



HAL
open science

Architectures moléculaires complexes pour la reconnaissance de bio(macro)molécules d'intérêt thérapeutique

Benjamin Ourri

► **To cite this version:**

Benjamin Ourri. Architectures moléculaires complexes pour la reconnaissance de bio(macro)molécules d'intérêt thérapeutique. Other. Université de Lyon, 2020. English. NNT : 2020LYSE1001 . tel-02457391

HAL Id: tel-02457391

<https://theses.hal.science/tel-02457391>

Submitted on 28 Jan 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT : xxx

THESE de DOCTORAT DE L'UNIVERSITE DE LYON

opérée au sein de
l'Université Claude Bernard Lyon 1

Ecole Doctorale N° 206
Chimie, Procédés, Environnement

Spécialité de doctorat : Chimie
Discipline : Chimie Organique

Soutenue publiquement le 06/01/2020, par :
Benjamin Ourri

Complex molecular architectures for the recognition of therapeutic bio(macro)molecules

Devant le jury composé de :

Heck, Marie-Pierre	Directrice de recherche, CEA, Université Paris Saclay, Rapporteur
Ulrich, Sébastien	Chargée de recherche CNRS, Université de Montpellier, Rapporteur
Metay, Estelle	Chargée de recherche CNRS, Université de Lyon, Examinatrice
Morell, Christophe	Professeur, Université de Lyon, Examineur
Hunter, Christopher	Professeur, Université de Cambridge, Examineur
Leclaire, Julien	Professeur, Université de Lyon, Directeur de thèse
O'Hara, John	Chercheur, UCB-Pharma, Tuteur entreprise
Dumont, Elise	Professeur, ENS Lyon, Membre invité

UNIVERSITE CLAUDE BERNARD - LYON 1

Président de l'Université	M. le Professeur Frédéric FLEURY
Président du Conseil Académique	M. le Professeur Hamda BEN HADID
Vice-président du Conseil d'Administration	M. le Professeur Didier REVEL
Vice-président du Conseil Formation et Vie Universitaire	M. le Professeur Philippe CHEVALIER
Vice-président de la Commission Recherche	M. Fabrice VALLÉE
Directrice Générale des Services	Mme Dominique MARCHAND

COMPOSANTES SANTE

Faculté de Médecine Lyon Est – Claude Bernard	Directeur : M. le Professeur G.RODE
Faculté de Médecine et de Maïeutique Lyon Sud – Charles Mérieux	Directeur : Mme la Professeure C. BURILLON
Faculté d'Odontologie	Directeur : M. le Professeur D. BOURGEOIS
Institut des Sciences Pharmaceutiques et Biologiques	Directeur : Mme la Professeure C. VINCIGUERRA
Institut des Sciences et Techniques de la Réadaptation	Directeur : M. X. PERROT
Département de formation et Centre de Recherche en Biologie Humaine	Directeur : Mme la Professeure A-M. SCHOTT

COMPOSANTES ET DEPARTEMENTS DE SCIENCES ET TECHNOLOGIE

Faculté des Sciences et Technologies	Directeur : M. F. DE MARCHI
Département Biologie	Directeur : M. le Professeur F. THEVENARD
Département Chimie Biochimie	Directeur : Mme C. FELIX
Département GEP	Directeur : M. Hassan HAMMOURI
Département Informatique	Directeur : M. le Professeur S. AKKOUCHE
Département Mathématiques	Directeur : M. le Professeur G. TOMANOV
Département Mécanique	Directeur : M. le Professeur H. BEN HADID
Département Physique	Directeur : M. le Professeur J-C PLENET
UFR Sciences et Techniques des Activités Physiques et Sportives	Directeur : M. Y.VANPOULLE
Observatoire des Sciences de l'Univers de Lyon	Directeur : M. B. GUIDERDONI
Polytech Lyon	Directeur : M. le Professeur E.PERRIN
Ecole Supérieure de Chimie Physique Electronique	Directeur : M. G. PIGNAULT
Institut Universitaire de Technologie de Lyon 1	Directeur : M. le Professeur C. VITON
Ecole Supérieure du Professorat et de l'Education	Directeur : M. le Professeur A. MOUGNIOTTE
Institut de Science Financière et d'Assurances	Directeur : M. N. LEBOISNE

Summary	5
Résumé en français	10
Remerciements	14
Nomenclature and abbreviations	17
General introduction	21
Preliminary remarks	22
Chapter 1: Polycarboxylates disulfide bridge cyclophanes for fundamental study and biomolecular recognition	23
Abstract	24
1. Introduction	26
2. An aside for a brief presentation of the DCC and the associated computational tools	30
a. Presentation of dynamic combinatorial chemistry.....	30
b. Density functional theory: fundamentals.....	32
c. Essentials of molecular dynamic simulation.....	33
3. Self-assembly of 2,5-dimercaptoterephthalic acid I-1 through DCC and selective complexation of a corresponding disulfide cyclophane I-1₄ with polyamines in water	34
a. Self-assembly of 2,5-dimercaptoterephthalic acid I-1 through DCC and structural properties of its oligomer I-1 ₄	34
i. Synthesis and self-assembly of I-1.....	34
ii. Structural properties of I-1 ₄	35
iii. Heterocyclo-oligomerisation with I-1 and other 1,4-bisthiophenols.....	36
b. Factors governing the selectivity of the carboxylate receptor I-1 ₄ with a series of α,ω -alkyldiammonium ions of increasing chain length in water.....	38
c. Factors governing the selectivity of the carboxylate receptor I-1 ₄ with spermine and spermidine in water.....	43
i. Qualitative analysis of the topology of and internal mobility of the complex.....	44
ii. Thermodynamic of the association.....	48
iii. Study of the solvation energy of the complexes.....	52
d. Conclusion.....	54
4. Self-assembly of 2,5-bis(mercaptomethyl)terephthalic acid I-15 by iodine or dioxygen oxidation of thiols into disulfide macrocycles	55
b. Additive-directed self-assembly of polythiols and kinetic trap.....	56
c. Synthesis of benzylic building blocks.....	58
d. Self-assembly by aerobic oxidation.....	60
e. Switch between thio-Michael addition and disulfide exchange.....	65
f. Thiol oxidation of 2,5-bis(mercaptomethyl)terephthalic acid I-15 with iodine.....	68
g. Study of the dimer properties.....	69
i. Crystallographic structure.....	69
ii. Binding of lithium and amines with I-15 ₂	71
iii. Post functionalisation of I-15 ₂ by amide bond formation.....	73
h. Conclusion.....	74
5. Separation of the thiol and the carboxylate moieties in a dithiophenol tetra carboxylate building block	75

a.	Overview of London dispersive interactions	76
i.	Computational description.....	76
ii.	Applications in organic chemistry	78
b.	Synthesis of a pair of bulk diastereoisomeric I-106 and I-106 building blocks.....	80
c.	Self-assembly through DCC by oxidation with dioxygen	82
i.	DCLs of one building blocks.....	82
ii.	DCLs of two building blocks simultaneously	85
d.	Structural properties of I-105₄	87
i.	NMR analysis	88
ii.	Crystallographic study	89
iii.	pH titration	91
e.	DFT optimisation of the diastereoisomers	91
f.	Modification of the external part and of the functionalisation of the sterically crowded 1,4-bisthiophenol.....	94
g.	Conclusion	96
6.	Conclusion	96
	Experimental section	98
a.	Synthetic procedures and characterisations	98
i.	Materials and methods	98
ii.	Preparative procedures and characterisations	98
b.	DOSY NMR experiments	119
c.	DCC, HPLC and UPLC/HRMS analysis	120
i.	Library preparation	120
ii.	Antiparallel experiments	120
iii.	HPLC analysis.....	120
iv.	UPLC analysis.....	121
v.	UPLC/HRMS analysis	121
vi.	Semi-preparative chiral HPLC.....	121
d.	Crystallography.....	121
e.	pH titrations.....	122
f.	ITC ^{136,216}	122
i.	Standard procedure	122
ii.	GFPs / I-14	122
g.	Octet®	123
h.	μ-FFE	123
i.	DFT.....	123
j.	Molecular dynamic	123
	Appendices.....	124
a.	Appendix I-I: List of molecules.....	124
b.	Appendix I-II: Additional conditions of reactions	126
c.	Appendix I-III: Data of the molecular dynamic simulation between I-1 and diamines.....	129
d.	Appendix I-IV: Data of the molecular dynamic simulation between I-14 and polyamines I-8 and I-14 . 130	
e.	Appendix I-V: Additional data of the DCLs for 1-15	131
f.	Appendix I-VI: CheckCIFiles and crystallographic data for I-15₂	145
g.	Appendix I-VII: DFT data for I-15₂	147
h.	Appendix I-V-III: ITC thermograms of I-15₂ vs polyamines and LiCl	149
i.	Appendix I-IX: NMR experiments of I-105 / I-105₄	152
j.	Appendix I-XI: pH titration of I-105₄	158
k.	Appendix I-XII: DFT data for I-105₄	159

Chapter 2: Towards macrocycles dedicated to the selective recognition of therapeutic proteins	165
Abstract	166
1. Design of the macrocycles	168
a. Introduction and context	168
b. Molecular recognition of lysine methylation	172
c. Analysis of enzymes	176
d. Proposition of design	179
2. Synthesis of corona[n]arenes out of framework of DCC	183
a. Access and applications of corona[n]arenes	183
b. Synthesis of corona[n]arenes	187
i. Synthesis of corona[n]arene[n]tetrazines	187
ii. Synthesis of corona[n]arene[n]pyridazines	190
c. Theoretical investigation of the corona[n]arenes synthesis	192
i. Modelling of the S _N _A r reaction	192
ii. Modelling of the IEDDA sequence	194
d. Crystallographic structure of II-51	196
e. Conclusion	198
3. Introduction of a Lysine-tag on proteins and binding with II-1₄	198
4. Introduction of side chains on a carboxylate aromatic 1,4-bisthiophenol	200
a. Protection of thiols by the trityl group	200
b. Protection of thiols by the cyanopropyl group	203
c. Deadlock route to asymmetric monomers	205
d. Conclusion	206
5. Setbacks to the connection of the functionalised monomers into a principal chain	207
a. Formation of reversible disulfide linkage between 2,5-functionalised-1,4-bisthiophenols through DCC	207
b. Rationalisation of the oxidative NS coupling	219
i. Investigation of a radicalar mechanism by DFT	219
ii. Description of a recently proposed electrophilic mechanism regarding the formation of sulfenyl amides	222
c. Overcoming the NS coupling on N ₂ ,N ₅ -bis(aminoalkyl)-1,4-bisthiophenols	223
d. Exploration of the nucleophilicity of 1,4-bisthiophenols	231
e. Conclusion	232
6. Towards combination of disulfide and hydrazone/imines exchanges for the synthesis of cavitants	233
a. Synthesis of 1,4-bisthiophenol bearing amines in the 2,5 positions of the aromatic core	234
b. Synthesis of 1,4-bisthiophenol with aldehyde moieties in the ortho positions and subsequent formation of hydrazones	236
c. Conclusion	239
7. Conclusion	239
Experimental section	240
a. Synthetic procedures and characterisations	240
iii. Materials and methods	240
iv. Preparative procedures and characterisations	240
b. DCC, HPLC and UPLC/HRMS analysis	268
vii. Library preparation	268

viii.	HPLC analysis	268
ix.	UPLC analysis.....	269
x.	UPLC/HRMS analysis	269
c.	Crystallography.....	269
d.	DFT details	269
Appendices.....		271
a.	Appendix II-I: List of molecules.....	271
b.	Appendix II-II: Reaction conditions.....	274
c.	Appendix II-III: Data of the binding assays between II-14 and GFPs	277
d.	Appendix II-IV: Additional HRMS or LC/MS data	284
e.	Appendix II-V: CheckCIFiles and crystallographic data for II-51	290
f.	Appendix II-VI: DFT TS coordinates.....	294
g.	Appendix II-VII: Set of amplification data for the DCLs between II-69 and II-1 (4 + 4 mM) in DMF..	306
Chapter 3: Poly-dendrigrfts of lysine: flexible objects for various chemistry.....		311
Abstract.....		312
1. Introduction.....		314
2. Recent advances in the neutralisation of heparin		315
a.	Proteic/peptidic/pseudo-peptidic binders	316
b.	Self-assembled binders.....	318
c.	Polymeric/dendritic binders	320
d.	Conclusion	323
3. Overview of dendrigrfts of lysine		324
4. Exploration of DGL applications		328
a.	Neutralisation of heparin.....	328
b.	Detection of biomolecules.....	330
c.	G₃ acting like an hydrolase	333
e.	Conclusion	334
5. Understanding the DGL – protamine / heparin binding by molecular dynamic simulations		335
a.	Study of the binding of protamine with heparin	337
b.	Study of the binding of DGLs with heparin and comparison with protamine	343
c.	Conclusion	348
6. Conclusion		348
Experimental section		349
a.	Synthetic procedures and characterisations	349
v.	Materials and methods	349
vi.	Preparative procedures and characterisations	349
b.	HPLC analysis	356
i.	HPLC analysis.....	357
ii.	UPLC analysis.....	357
xi.	UPLC/HRMS analysis	358
c.	DGL / Azur A / heparin UV / absorbance titrations	358
d.	Fluorescence.....	360
i.	DGL / III-3_n titrations	360
ii.	DGL / III-3_n titrations G₄ / III-3_n / BL21(DE3) titrations.....	362
iii.	G₂CF / BL21(DE3) measurements.....	365
e.	pH titration	365

f. ITC.....	365
g. Molecular dynamic.....	366
Appendices.....	369
a. Appendix III-I: List of molecules.....	369
b. Appendix III-II: Fluorescence	370
c. Appendix III-III: pH titration.....	379
d. Appendix III-IV: ITC	381
e. Appendix III-V: Supporting information regarding the MD simulations of protamines with heparins 382	
f. Appendix III-VI: Supporting information regarding the MD simulations of DGLs with heparins.....	392
Appendix: Activities during the PhD	399
References.....	401
General conclusion.....	425

La reconnaissance de biomolécules dans des milieux biologiques complexes est un réel défi pour les chimistes et les biologistes, associé à des enjeux médicaux majeurs. Face à cette problématique, le chimiste peut choisir d'utiliser des molécules désignées par ses soins, ou encore de sélectionner et d'utiliser directement des structures commerciales ou naturelles.

Suivant cette dernière approche, les dendrigrafts de lysines (DGLs) ont montré une neutralisation des héparines de différentes tailles (UFH, enoxaparin et Fondaparinux) supérieures à l'action de la protamine dans le sang, le seul médicament autorisé en cas de surdosage de l'anticoagulant et ayant de nombreux effets secondaires.

Une étude par dynamique moléculaire a permis de mettre en avant le mécanisme d'interaction entre les héparines d'une part, et les DGLs et la protamine d'autre part (Figure 1). Il a notamment été montré que la flexibilité du dendrimère lui permet d'optimiser ses charges et les interactions électrostatiques qu'il développe avec les héparines. Par ailleurs, la plus grande affinité des héparines de longue taille avec la protamine est notamment due à des interactions électrostatiques supplémentaires de moyenne distance. Avant une potentielle commercialisation pour cette application, le dendrimère doit encore être testé sur des modèles animaux et son éventuelle toxicité évaluée. Par ailleurs, les DGLs ont été utilisés en combinaison avec des homopolymères d'acide aspartique et de carboxylfluorescéine. Leur complexation avec le DGL inhibe leur fluorescence par agrégation. Lorsqu'un objet dissocie le complexe entre le DGL et cette sonde, la fluorescence est rétablie. Ce système de détection a été testé sur des objets d'intérêts biomédicaux, telle que la spermine, des bactéries Gram négative et des protéines thérapeutiques avec des limites de détection intéressantes mais trop haute pour envisager des applications, nécessitant des améliorations telle que la combinaison avec des nanoparticules. Enfin, un DGL a montré d'intéressantes propriétés de mime d'enzyme de type estérase.

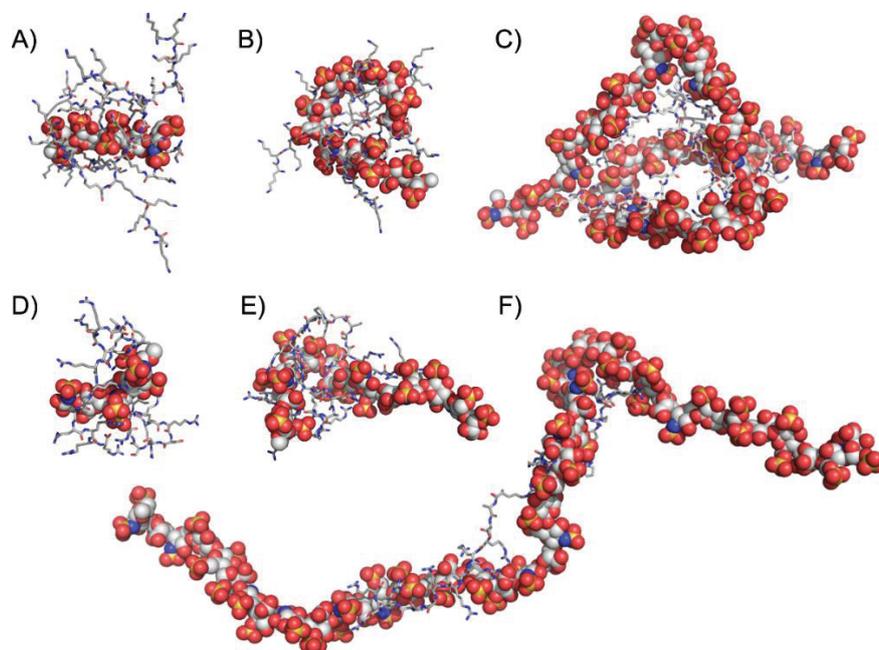
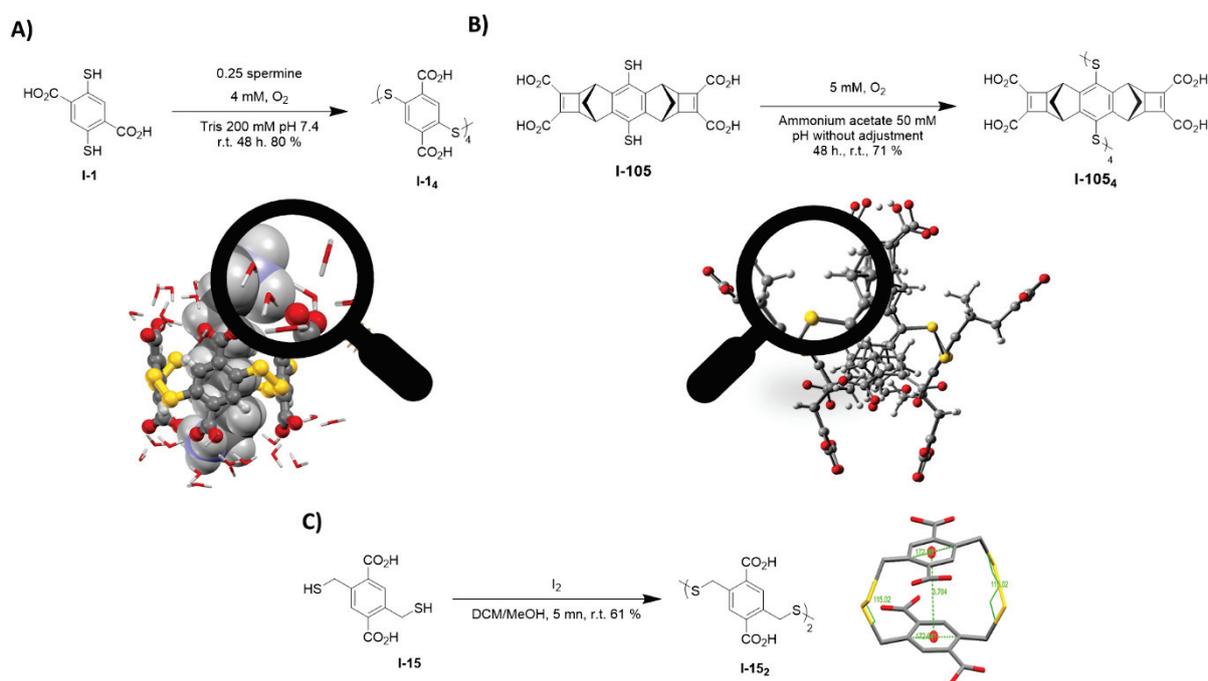


Figure 1 : Complexation entre des héparines de diverses tailles et ses inhibiteurs. Capture de simulations en dynamique moléculaire. Ball : heparines, stick : inhibiteurs. A) Fondaparinux et G_2 (DGL de génération 2), B) Héparine de taille médiane et G_2 , C) Héparine longue et G_2 , D) Fondaparinux et protamine, E) Héparine de taille médiane et protamine, F) Héparine longue et protamine.

Suivant la première approche évoquée de design et synthèse, nous avons utilisé la chimie combinatoire dynamique (CCD) pour obtenir des nouveaux récepteurs synthétiques à partir de briques moléculaires diverses de type 1,4-dithiophénols. Des études à la fois théorique, en DFT et dynamique moléculaire, et expérimentale, ont été menés pour comprendre les phénomènes régissant l'auto-assemblage de ces briques et la complexation des cavitands résultants correspondants avec des biomolécules d'intérêt.

L'acide 2,5-dimercaptoterephthalique **I-1** s'auto-assemble en solution aqueuse en un unique diastéréoisomère homochirale **I-1₄** par effet template de la spermine. La sélectivité de ce récepteur pour des α,ω -alkyldiammoniums est en partie régie par des effets de solvant, en l'occurrence une stabilisation enthalpique de solvation des complexes. Le récepteur a montré des constantes d'association exceptionnelles de l'ordre de 10^{11} M^{-1} avec la spermine et la spermidine (Scheme 1). La sélectivité observée pour la spermidine est en partie due à des effets de solvants, et à une plus forte mobilité translationnelle et rotationnelle de la spermidine dans son complexe d'inclusion.



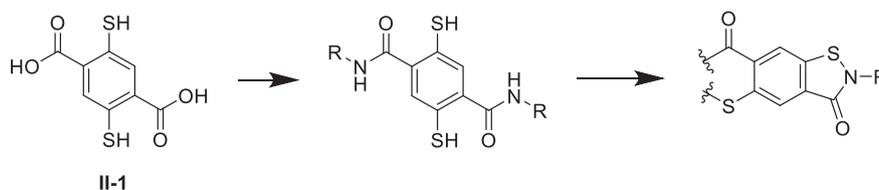
*Scheme 1 : Auto-assemblage de briques dithiols A) de la brique **I-1** en tétramère cyclique **I-1₄**. La sélectivité de ce récepteur pour des polyamines est partiellement régie par des effets de solvation. B) de la brique **I-105** en tétramère cyclique **I-105₄**, influencé par les effets stériques. C) de la brique **I-15** par le diode en milieu organique en un dimère cyclique fortement tendu au niveau des ponts disulfures.*

L'auto assemblage en solution aqueuse de l'acide 2,5-bis(mercaptométhyl)terephthalique **I-15** a conduit à la formation d'une série d'oligomères cycliques de taille variée (incorporant de 2 à 10 unités), dont la proportion durant les chimiothèques varient par effet template (Scheme 1). L'utilisation de diode en tant qu'oxydant a permis d'isoler un dimère cyclique **I-15₂** pour lequel la structure cristallographie a révélé une forte contrainte sur les ponts disulfures. Cet objet a pu servir de modèle de post-fonctionnalisation par liaison amide.

Une brique 1,4-dithiophénol encombré **I-105** a montré un phénomène spectaculaire de sélection par auto-assemblage en solution aqueuse lors d'une chimiothèque combinatoire dynamique (Scheme 1).. La brique s'auto-assemble en un unique diastéréoisomère (sur 24 possibilités) d'un tétramère cyclique **I-105₄**. Une série de modélisation en DFT et une étude RMN a pu montré que la configuration des ponts

disulfures est alternativement MPMP, et que les faces encombrés sont orientés à l'extérieur de la cavité du macrocycle. Lors de la cristallisation de cet objet, une métathèse de pont disulfure a conduit à une dimérisation en octamère avec inversion de la géométrie des ponts disulfures concernés, ce qui n'a encore jamais été décrit. L'auto-assemblage de **I-105** est régi par des effets de répulsions stériques. Il a été dans un premier temps postulé une sélection par les forces attractives de London, ce qui a été invalidé par la DFT.

Enfin, une voie de synthèse a été élaboré et a conduit à la synthèse d'une famille de 1,4-dithiophénols fonctionnalisés en position 2,5 par des liaisons amide avec des groupements chimiques variés (amines, acide aminés, azoture, alcools), permettant l'élaboration de brique sur mesures pour la reconnaissance de protéines thérapeutiques affectées par des modifications post-translationnelles. Toutefois, l'auto-assemblage de ses briques moléculaires en solutions a donné lieu un phénomène inattendu d'oxydation des thiols en amides sulfényle cycliques, contrecarrant notre stratégie de synthèse de cavitants. Des solutions partielles ont pu être proposés pour faciliter la formation souhaitée de pont disulfures à partir des thiols, ce qui a conduit à l'observation en chimiothèque combinatoire dynamique d'hétéro et homocyclooligomères issues de ces nouvelles briques. Leur application n'est toutefois pas encore possible et nécessite de nombreuses optimisations et une revisite de leur design pour éviter ce couplage parasite.



Scheme 2: Obtention de brique 1,4-dithiophénols fonctionnalisés en position 2,5 par liaison amide sur 11 exemples en 3 ou 4 étapes (R: alkyls varies, comportant des amines, alcools, azoture acide amines) par 1) protection des thiols avec les groupements trityls ou cyanopropyls, 2) couplage peptidique sur les positions 2,5; 3) déprotection des thiols en une ou deux étapes. L'auto-assemblage de ces briques souffrent principalement d'oxydation des thiols en amide sulfényle au lieu de ponts disulfures.

Ces efforts s'inscrivent dans une volonté de développer des récepteurs pour la reconnaissance de protéines thérapeutiques affectées par des modifications post-translationnelles. Dans ce but, une brique 2,5-dimercaptoterephthalaldehyde **II-138** n'a pas pu être mise à profit en chimiothèque combinatoire dynamique basée sur des échanges simultanées thiolate/disulfides et hydrazines/acylhydrazones. De même, la famille des corona[n]arènes a été exploré mais des difficultés synthétiques ont été rencontrés pour l'obtention des groupements fonctionnels désirés sur ses cavitants pour la reconnaissance souhaitée.

Mots-clés: Chimie combinatoire dynamique, briques 1,4-dithiophénol-2,5-fonctionnalisées, modifications post-translationnelles, reconnaissance moléculaire, neutralisation des heparines, dendrigrafte de lysine, dynamique moléculaire, DFT

"I almost wish we were butterflies and liv'd but three summer days - three such days with you I could fill with more delight than fifty common years could ever contain."

John Keats

Remerciements

J'exprime ici toute ma reconnaissance envers mon directeur de thèse, Julien Leclaire, pour son travail de supervision, son humour décapant et sa confiance dans mon travail. Je souhaite qu'il trouve dans ces courtes lignes une admiration pour sa réflexion scientifique et ses intuitions.

Je souhaite également remercier UCB Pharma pour le financement de ma thèse, à travers mes tuteurs entreprises, Edith Norrant et John O'Hara.

Ces trois années de formation à la recherche par la recherche ont été riches d'enseignements et d'expériences. J'ai pu bénéficier d'un environnement de travail idéal, au sein d'un institut qui m'a quasi intégralement formé en tant que chimiste de synthèse depuis mon « M1 ». Que tous ses membres, permanents ou non, reçoivent ici toute ma gratitude pour l'esprit d'échange et d'entraide qui règnent dans nos murs.

A travers ces remerciements, je voudrais saluer les encadrants qui ont successivement contribué à mon apprentissage et à me rendre encore plus rêveur du métier de chercheur que je ne pouvais déjà l'être avant mon doctorat. Parmi eux, Olivier Baudoin, Grégory Danoun, David Dailler, Ulrich Darbost, Christophe Morell. Mon passage dans la start-up Calixar m'a également beaucoup appris et m'a aidé à me développer en tant qu'élève ingénieur. Que Julien Dauvergne et Emmanuel Dejean en reçoivent ma reconnaissance. L'enseignement de toutes ces personnes passionnées par leur métier m'a armé pour me confronter à trois années de recherche intense au sein de l'équipe CSAp.

L'éphémère apporte une beauté incontestable à la vie. La composition transitoire des équipes de recherche en est une illustration que j'ai pleinement vécue pendant un peu plus de 3 années. Les allées et venues des doctorants, post-doctorants et stagiaires ont été un réel fil d'Ariane pour moi et l'occasion de belles rencontres. Je salue ici Geoffroy Germain, Yves Pellé, Emeric Jamet, Mélissa Dumartin, Marion Donnier-Maréchal, Jean Septavaux, Ousseynou N'Diaye, Marie-Noelle Poradowski Aurélie Rago, William Vasseur et Youssef Nassar.

L'excellence de notre recherche ne peut se concevoir sans un esprit de collaboration et de formation tourné vers les jeunes comme les doctorants. Lors de cette thèse, j'ai eu la chance de côtoyer Lionel Perrin, Elise Dumont et Gérald Monard qui sont une très belle illustration de cet esprit et que je remercie chaleureusement pour leur contribution à ma thèse et à mes connaissances.

Isabelle Bonnamour trouve ici une place particulière. Elle est le premier contact que j'ai eu lorsque j'étais lycéen (et chevelu) avec le monde de la recherche. Son investissement dans les Olympiades a participé à ma vocation de chercheur et à une belle aventure humaine par la suite. Je tiens à lui exprimer ici toute mon admiration pour sa carrière ainsi que mille merci pour ce qu'elle m'a appris en M2 et pour son soutien indéfectible ces dernières années.

Nathalie Perol et Florent Perret, maitres de conférences de mon équipe, m'ont été d'un grand soutien et m'ont prodigué de bons moments de discussion et de détente. Je remercie également Laurent Vial, avec qui j'ai découvert une thématique de recherche passionnante et une philosophie d'investigation qui m'a beaucoup plu. Cette dernière année à travailler avec lui fut agréable et pleine d'enseignements.

Le quotidien de ce doctorat n'aurait été le même sans la présence de mes collègues doctorants et amis Fanny Demontrond, Clara Tosi et Francesca Marocco Stuardi. Je leur souhaite une belle réussite dans leur projet et surtout leur adresse un grand merci pour les moments passés ensemble.

J'ai commencé mes études supérieures avec Alexandre Heloin, et j'ai eu la chance de voir nos parcours respectifs se croiser à plusieurs reprises pour finir en apothéose en thèse. Son sens de la rigueur scientifique n'a d'égal que sa gentillesse. Je lui suis reconnaissant pour ce qu'il m'a appris et lui exprime ici tout le bonheur que j'ai d'avoir partagé 1021 journées au laboratoire avec lui et bien plus encore.

En dehors de la sphère professionnel, les amis et la famille contribuent pleinement à la réussite des projets. A ce titre, j'aimerais mentionner mon amitié indéfectible et mon admiration sans faille pour Timothée Reymond.

La réalisation de mon projet d'études à travers mon cursus en classe préparatoires, puis à CPE Lyon et enfin en thèse à l'Université Lyon 1, n'aurait tout simplement pas été possible sans des rencontres décisives, qui m'ont façonné en tant qu'être humain, et qui ont contribué à mes rêves. Que Jacqueline et Yves ainsi que Christiane et Pierre, trouvent dans ces lignes une infinie reconnaissance pour leur soutien, leur générosité. Ils sont une source d'inspiration sans fin pour moi. J'espère qu'ils seront fiers de ce travail et de mon parcours. De même, j'ai une pensée pour Sonia Doyen et les Anciens Martins, dont je suis fier d'être membre et qui m'ont solidement aidé pendant de nombreuses années.

Les mots sont parfois insuffisants, alors je me contenterai simplement de finir ces remerciements par Bérengère. Elle n'a eu de cesse de m'aider dans mes choix et à m'aider à relativiser les déconvenues du quotidien à la paillassse. J'espère qu'elle trouvera autant de joie à poursuivre un doctorat que la mienne. Que ces mots résonnent en elle comme une preuve d'amour et d'admiration pour toutes ses qualités dont elle n'a pas conscience.

Et bien évidemment, j'embrasse mon frère et mes sœurs, mes parents, ma grand-mère, mes amis et ma filleule Lily !

Nomenclature and abbreviations

Abbreviations	Meaning
α	Dipole polarisability
δ	Chemical shift
ψ	Wavelength
λ	Molar extinction coefficient
γ	Collision frequency
ρ	Electronic density
η	viscosity
Δt	Difference of time
$\Delta_r G^\circ/\Delta G$	Standard free energy of reaction
$\Delta\Delta_r G^\circ$	Relative standard free energy of reaction
$\Delta_r H^\circ/\Delta H$	Standard enthalpy of reaction
$\Delta\Delta_r H^\circ$	Relative standard enthalpy of reaction
$\Delta_r S^\circ/\Delta S$	Standard entropy of reaction
$\Delta\Delta_r S^\circ$	Relative standard entropy of reaction
$^\circ\text{C}$	Celsius degree
ΔK	GFP mutant/error on Gibbs free energy
	Entropy
λ_{ex}	Excitation wavelength
λ_{em}	Emission wavelength
μ -FFE	Micro free flow electrophoresis
2D	2 dimensions
\AA	Angstrom
A	Absorbance
Ac	Acetyl
Ad	Adamantyl
AIBN	Azobisisobutyronitrile
APFD	Austin-Frisch-Petersson functional
APT	Attached Proton Test
APTMAC	N-acrylamidopropyl-N,N,N-trimethylammonium chloride
APTT	Activated partial thromboplastin time
AT	Anti-thrombin

Ar	Aryl
atm	Atmosphere (pression unit)
B3LYP	Becke, 3-parameter, Lee–Yang–Parr
BINOL	1,1'-bi-2-naphtol
Boc	<i>tert</i> -butoxycarbonyl
C_n^{AB}	Dispersion coefficient at order n for A-B
CBX7	Chromobox protein homolog 7
CCD	Cambridge Crystallographic Database
CETA	Cetyltrimethylammonium bromide
CF	5(6)-carboxyfluorescein
CIP	Centre d'ingénierie des protéines
CID	Collision-induced dissociation
COSY	Correlation spectroscopy
CPB	Cardiopulmonary bypass study
CSAp	Applied supramolecular chemistry
CSD	Cambridge Structural Database
CuAAC	Copper-Catalyzed Azide-Alkyne Cycloaddition
d	Day
D	Diffusion coefficient
dG	Local contribution to Gibbs free energy of solvation
dH	Local contribution to enthalpy of solvation
dS	Local contribution to entropy of solvation
Da	Dalton
DA	Diels-Alder
DABCO	1,4-diazabicyclo[2.2.2]octane
DCL	Dynamic combinatorial library
DCC	Dynamic combinatorial chemistry
DCM	Dichloromethane
Dea	Deamidation

DED	Dispersion energy donor
DETA	Diethylenetriamine
DEPT	Distortionless enhancement by polarization transfer
Dex	Dextran
DLS	Diffraction light scattering
DFT	Density functional theory
DGL	Dendrigraft of poly-Lysine
DHDF	Double hybrid density functionals
DLS	Dynamic light scattering
DMAD	Dimethyl acetylenedicarboxylate
DMTMM	4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride
Dn	DFT-D dispersion, version n
DNA	Deoxyribonucleic acid
D _n CF	Size n of polymers of aspartic acid with carboxyfluorescein
DIBAL-H	Diisobutylaluminium hydride
DIPEA	Diisopropylethylamine
DMA	Dimethylacetamide
DMF	Dimethylformamide
DMSO	Dimethylsulfoxide
DTT	DL-dithiotreitol
E	energy
E _{xc}	Exchange correlation energy
E _c	Exchange energy
E _x	Correlation energy
EDC/EDCI	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
ENS	Ecole normale supérieure
eq	Equivalent / équation
ESI	Electrospray ionisation
Et	Ethyl
ETD	Electron-transfer dissociation
F	Spatial direction/Functional
FDA	Food drug administration
f _{damp}	Becke-Johnson damping function
Fmoc	Fluorenylmethoxycarbonyl
g	Gram

G	Gibbs free energy
G _n	Size n of dendrigrafts of poly-L-lysine
GABA	γ-aminobutyric acid
GAG	Glycosaminoglycan
GB	Generalised Born
GCD	γ-cyclodextrin
GGA	Approximation of semi local generalised gradient
GIST	Grid inhomogeneous solvation theory
GFP	Green fluorescent protein
Glu	Glutamic acid
GPU	Graphics Processing Unit
GTMAC	Glycidyltrimethyl-ammonium chloride
h	Hour
H	Hamiltonian
HBG	Heparin binding group
HCT	High Capacity Ion Trap
HCTU	O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMPT	Hexamethylphosphoramide
HOBt	Benzohydroxytriazole
HPLC	High performance liquid chromatography
HRMS	High resolution mass spectrometry
HMBC	Heteronuclear multiple-bond correlation spectroscopy
HPG	Hyperbranched polyglycerol
HSQC	Heteronuclear single-quantum correlation spectroscopy
Hz	Hertz
IMAC	Immobilised metal ion affinity chromatography
I	Intensity / First ionization potential
IBCP	Institut de biochimie et chimie des protéines

IC ₅₀	Half maximal inhibitory concentration
ICBMS	Institut de chimie et biochimie moléculaires et supramoléculaires
iEDDA	Inverse electronic demand Diels-Alder
ISA	Institut des sciences analytiques
ITC	Isothermal calorimetric titration
J	Joule
K / K _a	Thermodynamic constant of association
k _B	Boltzmann constant
K _d	Thermodynamic constant of dissociation
K _{men}	<i>n</i> methylation state of lysine
KISS	Keep it simple stupid
L	Litre
LIE	Linear interaction energy
LMWH	Low molecular weigh heparin
LOD	Limit of detection
m	mass
M	mol/L
<i>m</i> -CPBA	Meta-chloroperoxybenzoic acid
m	Meter
MAA	GFP mutant
MALDI	Matrix Assisted Laser Desorption Ionisation
MalB	Mallard Blue
MB	Methylene Blue
MD	Molecular dynamic
Me	Methyl
MetOx	Methionine oxidation
MKQ	GFP mutant
MM-PBSA/GBSA	Molecular mechanics energies combined with the Poisson-Boltzmann or generalized Born and surface area continuum solvation
mol	Mole
mp	Melting point

MP2	Moller Plesset perturbation at order 2
mn / min	Minute
MS	Mass spectrometry
MW	Molecular weight
N	Binding number/stoichiometry
NBS	N-bromosuccinimide
NCA	α-aminoacid N-carboxyanhydride
NCI	Non-covalent interactions
NHC	N-heterocyclic carbene
NKR	Newman-Kwart rearrangement
NMP	N-methylpyrrolidine / N-methyl morpholine
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser Spectroscopy
PAH	Poly(allylamine hydrochloride)
PAMAM	Poly(amidoamine) dendrimers
PB	Phosphate buffer / Poisson Boltzmann
PBS	Phosphate-buffered saline
PCM	Polarisable continuum model
PDMAEMA	Poly(2-(dimethylamino)ethyl methacrylate)
PEG	Polyethylene glycol
pH	Hydrogen potential
pK _a	Acid base dissociation constant
pI	Ionisation potential
PMAPTAC	Poly(3-methacraloylamino)propyl trimethylammonium chloride
ppm	Partie per million
PPV	Polyphenylenevinylene
PSMN	Pôle scientifique de modelling numérique
PTM	Post translational modification
Pul	Pullulan
PyBOP	Benzotriazole-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate

PyBrOP	Bromotripyrrolidinophosphonium hexafluorophosphate
QTOF	Quadripolar time of flight
r	Intermolecular distance
r.t.	Room temperature
R _H	Hydrodynamic radius
riAT	Recombinant-inactivated antithrombin
R _{men}	<i>n</i> methylation state of arginine
RMSD	Root mean square deviation
RNA	Ribonucleic acid
ROM	Ring opening metathesis
SASA	Solvent accessible surface area
SAMul	Self-assembled multivalent micelles
S _n	global scale factor
SDS-PAGE	Sodium dodecyl sulfate–polyacrylamide gel electrophoresis
SP	Single point
SPAM	Maps in reverse
T	Temperature/Kinetic energy operator
<i>t</i> Bu	<i>tert</i> -butyl
TCCA	Trichloroisocyanuric acid
TD-DFT	Time dependant density functional theory
T _{eb}	Boiling point
TNB	2-nitro-5-thiobenzoate

TOF	Time of flight
TFA	Trifluoroacetic acid
T _{fus}	Fusion point
TIPS	Triisopropylsilane
TLC	Thin layer chromatography
TMEDA	<i>N,N,N',N'</i> -tetramethylethane-1,2-diamine
TMS	Trimethylsilane
TMSP	Trimethylsilylpropanoic acid
t _r	Retention time
tr	Tour
Tr	Trityl
Tris	Tris(hydroxymethyl)aminomethane
TS	Transition State
U	Tension
UA	Arbitrary unit
UCBL	Université Claude Bernard de Lyon
UFH	Unfractionated heparin
UHRA	Universal heparin reversal agent
UPLC	Ultra-performance liquid chromatography
UV	Ultraviolet
V	Potential energy operator
VdW	Van der Waals
VLP	Virus-like particle
z	charge

Natural molecules, such as enzymes and proteins, perform highly specific molecular recognition in water, for instance in the formation of peptides from simple nucleotides during the translation of DNA¹. Molecular recognition refers to specific interactions between several molecules through noncovalent bonds. Nature inspires chemists in their design of multimolecular supramolecular architectures of similar complexity and functionality that are destined to recognise biomolecules in a complex biological media. The obtention of such efficient receptors is conditioned by their selectivity towards a target and their specificity in a competitive media such as a water.

The chemists can address this issue following two approaches: either they design and synthesise molecules or they select commercially available or natural molecules. In both cases, the comprehension of the fundamental intermolecular forces are essential to correctly understand 1) how the natural molecules work, 2) how the chemist's molecule will interact with the target. In this context, we are interested in two classes of synthetic receptors that are adapted to study fundamental interactions in water. These receptors have two different topologies and cavities, ramified and cylindric, allowing various mechanisms of interaction with guests. The cylindric receptors which are investigated in this thesis are linked to the family of cyclophanes. These molecules were widely employed to study fundamental principles such as the π -cation interactions² or the hydrophobic effects³. The ramified molecules are commercially available dendrigrafts of lysine (DGL). These dendrigrafts have various applications in water thanks to their high cationic charge density and particular shape⁴.

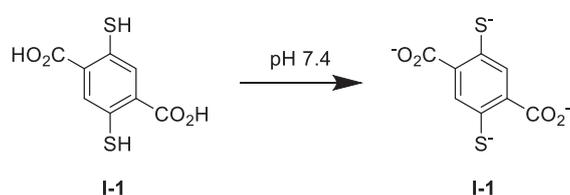
In the first part, the cylindric receptors will be synthesised by self-assembly of polythiols into disulfide cyclic oligomers through dynamic covalent chemistry (DCC). These polythiols and macrocycles will be used in water to explore the driving forces of complexation with various biogenic polyamines, as well as the influence of steric effects in self-assembling. The second part will be dedicated to the recognition of modified therapeutic proteins that are affected by post-translational modifications. The same family of disulfide bridge cyclophanes will be used, following a molecular design that is inspired by the analysis of natural systems. This part is linked to an industrial partnership with UCB pharma. Then, the dendrigrafts will be benchmarked for a variety of applications in water, focussing on detection of bio(macro)molecules and the neutralisation of heparin in replacement of protamine, the only FDA-approved drug in case of heparin overdose. The mechanism of association between, on the one hand, protamine or DGL, and, on the other hand, heparin of various sizes will be particularly studied.

Our strategy relies on a combination of experimental and computational studies that aim to get a better understanding of the fundamental forces which drive the non-covalent molecular associations in water.

Preliminary remarks

Following the idea of Dr. Laurent Vial and the suggestion of Wolle and Muller⁵, an examination of 3D structures is proposed in this thesis in order to enable the reader to validate or invalidate conclusions that are drawn. The reader is invited to download on their smartphones/tablets the free « augment » app and to scan the QR codes that are displayed in the text with the app. A 3D model in augmented reality will be proposed. The option « 3D view » gives a better experience. Following the smartphone or the tablet, the colours might be missing.

This thesis explores host guest chemistry in aqueous solutions. The protonation state of the molecules that are involved in aqueous solutions are most of the time neutral for clarity. For instance, the following molecules should be a tetraanion in physiological conditions but is always represented neutral.

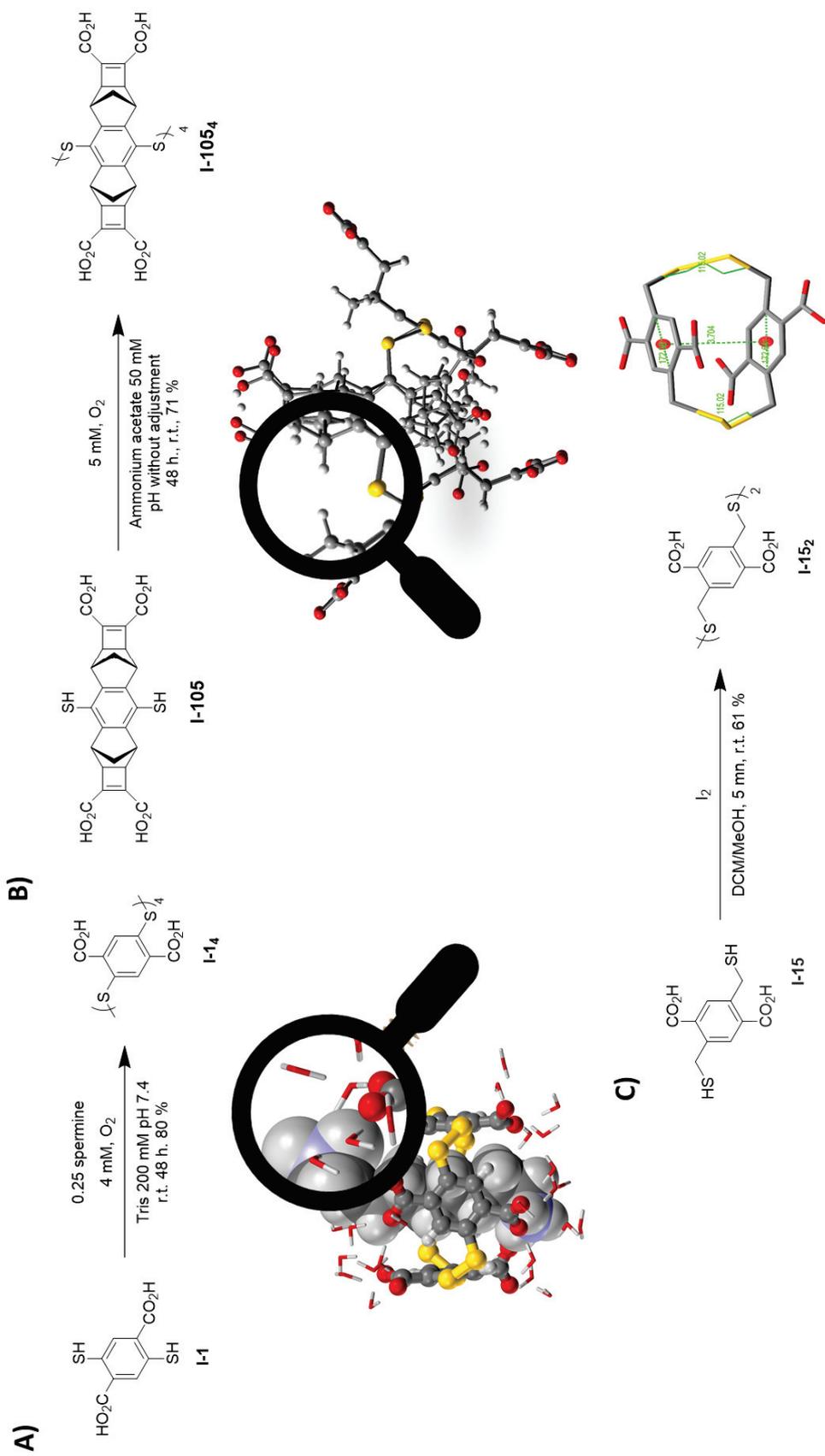


Scheme 3 : Dissolution of I-1 in an aqueous buffer gives a tetraanion in solution

Chapter 1: Polycarboxylates disulfide bridge cyclophanes for fundamental study and biomolecular recognition

Abstract

Cyclophanes are a venerable class of compounds, containing aromatic rings bridged by aliphatic chains in two non-adjacent positions⁶. These molecules were widely employed to study fundamental principles such as the π -cation interactions² or the hydrophobic effects³. Recently, cucurbiturils revealed the complexity of the water network and the solvation effects inside a cavitand⁷. Cyclophanes are a sub-category of the family of cavitand, which are molecules containing enforced cavities large enough to complex and even surround organic guest compounds⁸. Cyclophanes also have the features allowing many applications, from biomolecular recognition⁹ to material science¹⁰. Particularly, disulfide-bridged polycarboxylate cyclophanes obtained by dynamic combinatorial chemistry presents well-defined apolar cavity with carboxylate groups dedicated to solubility and electrostatic interactions. Nevertheless, they have never been used to study fundamental concepts such as the influence of solvation effects in supramolecular chemistry and steric effects in self-assembly. The 1,4-bisthiolphenol **I-1** forms a homocyclo-oligomers **I-1₄** by template effect with spermine through dynamic covalent chemistry (DCC). This polyanionic macrocycle showed high affinity for polyamines in aqueous solution. By a combination of experimental and computational studies, it was shown that solvation effects governed the selectivity of the associations. In fact, the enthalpy of complex solvation and of guest desolvation were determinant regarding a series of diamines with increasing chain lengths. Regarding spermine and spermidine, the solvation also ruled the selectivity, but the exact scenario is not fully elucidated yet. Several modifications such as adding one carbon between thiols and the aromatic core were obtained on **I-1** and the corresponding homo/heterocyclo-oligomers were assembled and characterised. The use of iodine as oxidant of the thiols into disulfide bonds highlighted the difference between thermodynamic and kinetic DCLs thanks to a methodology without additives as reported and allowed an expansion of the diameter of the cavity. Nevertheless, in order to get more rigid macrocycles, a further modification turning disulfide bonds into thioether was attempted. Another modification consisted in the increase of the height of the cavity. That resulted in a spectacular selection of an homocyclo-tetramer **I-105₄** of configuration In_In_In_In_MPMP in DCL (In referring to the orientation of the face in or out the cavity, M/P standing for the disulfide bond stereogenic configuration). Under some conditions of **I-105₄** crystallisation, an unexpected metathesis of disulfide bonds converted two **I-105₄** into an octamer with a disulfide configuration MPPPMPPP. The London forces were postulated to drive the self-assembling of **I-105**, but a DFT study showed that the self-assembly was governed by repulsive steric effects.

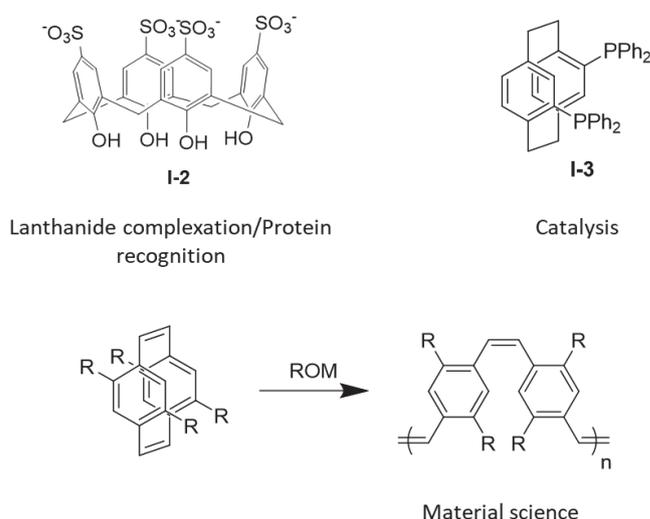


Scheme 4 : Self-assembly of 1,4-bis(2-mercapto-5-sulfanylphenyl)acetic acid (I-1) into I-14 by templating with spermine. The corresponding receptor had high affinity for various polyamines including spermine and spermidine. The selectivity of the complexation over polyamines was influenced by solvent effects. B) I-105 into I-105₄ by a spectacular selection over 24 diastereoisomers of cyclic tetramer influenced by steric effects. C) I-15 into I-15₂ by oxidation of thiols with iodine, leading a strain cyclic dimer..

1. Introduction

The objectives of this chapter were to use disulfide-bridged cyclophanes obtained by dynamic combinatorial chemistry (DCC) in order to study fundamental concepts such as the influence of solvation and steric effects in molecular recognition and self-assembling. The aim was also to expand our family of disulfide-based polycarboxylate cyclophanes named dyn[a]arenes.

Cyclophanes are organic compounds containing aromatic rings bridged by aliphatic chains in two non-adjacent positions⁶. Following the positions of connection, the cyclophanes are [n]metacyclophanes, [n]paracyclophanes or [n,n']cyclophanes. Meta and para correspond to the nomenclature of aromatic substitution while n corresponds to the number of atoms forming the bridge. Their first synthesis was reported in 1899¹¹ and 1951¹². These macrocycles feature unusual properties and have focused attention for decades. Among cyclophanes, well-known family emerged, such as calixarenes (metacyclophane)¹³ or pillararenes (paracyclophane)¹⁴. The cyclophanes can adopt a specific topology and original conformations when aromatic rings are distorted because of the large strain from the short aliphatic bridge connecting the two phenylene groups¹⁵. These specific features enabled many fundamental studies and applications (Scheme 5).

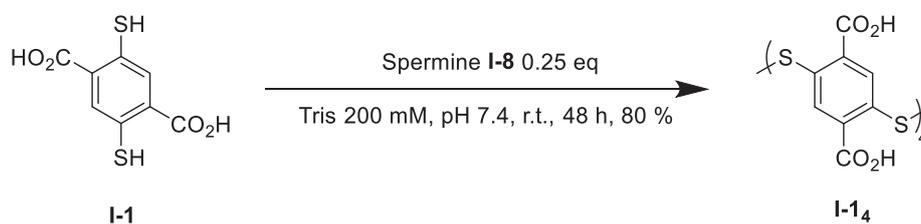


Scheme 5 : Examples of cyclophanes and their applications^{10,8a,9,16}. Top left, calixarenes, right, phosphine cyclophane, bottom, strained cyclophanes, R various groups.

Among them, many concepts linked to (anti)aromaticity were developed thanks to cyclophanes^{12b,17}, such as the particular spatial arrangement between (anti)aromatic rings¹⁸. Besides, the molecular recognition by cyclophanes was widely explored over the last decades¹⁹. The cyclophanes are efficient receptors in aqueous solutions for a variety of substrates such as arenes²⁰, steroids²¹, adenosine and amino acids, equalling enzymes and other proteins⁹. The studies of cyclophanes led to the discovery of important notions in host guest chemistry, such as the cation/ π interactions², which is a noncovalent molecular interaction between the face of an electron-rich π system and an adjacent cation²². This discovery enabled many applications in biochemistry and biology²³. Thus, cyclophanes are good candidates to study fundamental concepts such as solvation and steric effects²⁴. Steric effects are nonbonding interactions that influence the shape and reactivity of ions and molecules, complementing electrostatic interactions. The solvation refers to the organisation of solute and solvent into solvation spheres, including bonds, H-bonds, and Van der Waals forces²⁴. For instance, the study of water soluble cyclophanes revealed that a close apolar complexation was enthalpy-driven^{3a-c}. Nowadays, this notion

is named the non-classical hydrophobic effect^{3d}. Specific conformation and configuration of the cyclophane structure is a source of chirality²⁵ and helical structures²⁶, leading to asymmetric catalysis²⁷ for instance^{27b}. Such architectures are used in optical or electronic materials¹⁰, metal capture²⁸, biology²⁹ and polymer chemistry^{10a,30}. The extensive use of cyclophanes for such theoretical investigations and applications validates our choice to their use for the exploration of the influence of solvent and steric effects in self-assembling.

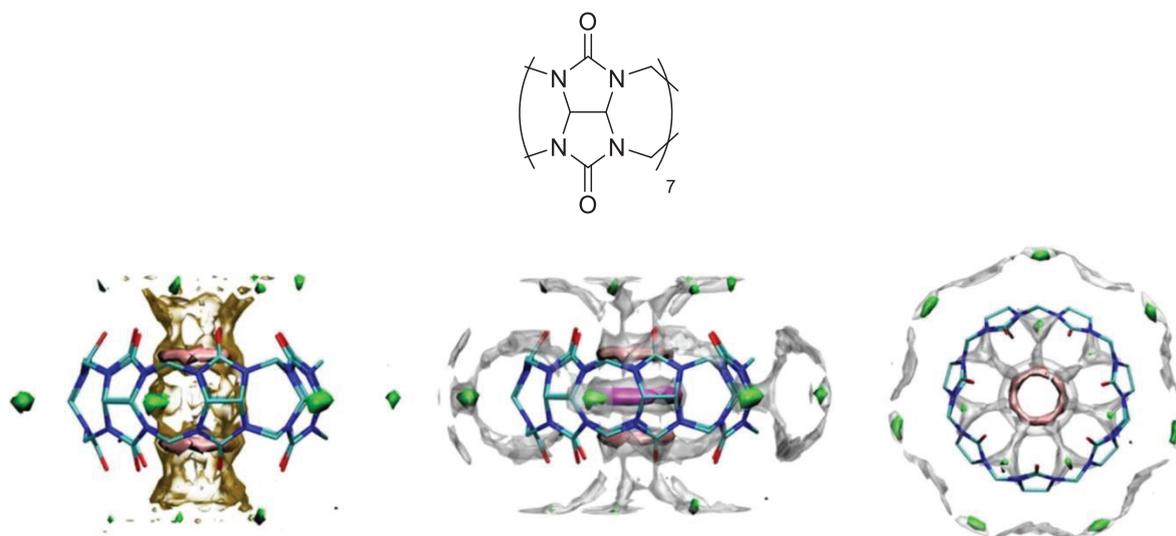
However, their synthesis is challenging given their complexity. The reactions involved had usually a poor functional tolerance and extensive purification steps³¹. In fact, the old route to cyclophane is based on intra and intermolecular macrocyclisations, occurring in harsh conditions³². Most of the time, a complex mixture of oligomers and polymers are obtained³³. The reaction leading to cyclophanes are for instance addition (Mannich reaction, Michael's addition) rearrangement (Beckmann, Claisen), coupling (Glaser-Eglinton, Castro-Stephens, Sonogashira), Friedel Craft reaction, metathesis, cycloaddition (4 + 2, Dotz benzannulation) and elimination (Hoffman, Wittig)³¹. They are all limited by unfavourable cyclisation steps. For instance, the Wurtz and Mc Murry couplings can help avoid an uncontrolled polymerisation and the Wittig reaction can give mixte heterocycles³¹. As an alternative route, a modern self-assembly and dynamic covalent chemistry under thermodynamic control has been recently developed to access such organic cages³⁴, tailored for specific applications³⁵. Dynamic combinatorial chemistry is a method of synthesis and screening of chemical compounds, based on self-assembly. Self-assembling is the spontaneous and reversible organisation of multiple building blocks into a single superstructure. In the field of supramolecular chemistry, DCC has led to a diversity of complex architectures³⁶, starting from building blocks and using reversible chemical bonds³⁷. The formation of products occurs under thermodynamic control, and the distribution depends on their relative stability³⁷. This strategy overcomes the issues of the traditional approach and extend the cyclophane design to 3D ones for instance³⁸. Dynamic combinatorial libraries (DCLs) of dithiol building blocks can respond to the stimuli of a guest to amplify the formation of disulfide bridged cyclophanes^{37,39}. For instance, the polycarboxylate disulfide-bridged paracyclophane **I-14** is synthesised by DCC thanks to the template effect of spermine (Scheme 6)⁴⁰. That macrocycle, widely used in our team, binds with various lysine derivatives⁴¹. The accessible molecular diversity and the selection/amplification of superstructures by self-assembling through DCC validates this synthetic plan to obtain synthetic receptors.



Scheme 6 : Amplification of I-14 by the template spermine I-6 from the I-1 building block in Tris buffer

The purpose of this chapter was to explore the modification of chains or chemical functions on that macrocycle, with all possible freedom degrees, and then to use the family of disulfide cyclophanes to study fundamental concepts (Scheme 8). Compared to the Dougherty's cyclophanes, our objects are polycarboxylate and dynamic, which can complexify the analysis. Solvation and hydrophobic effects are important phenomena, relatively misunderstood or underestimated in supramolecular and host guest chemistry. The solvation and desolvation processes play underexploited determinant roles in the

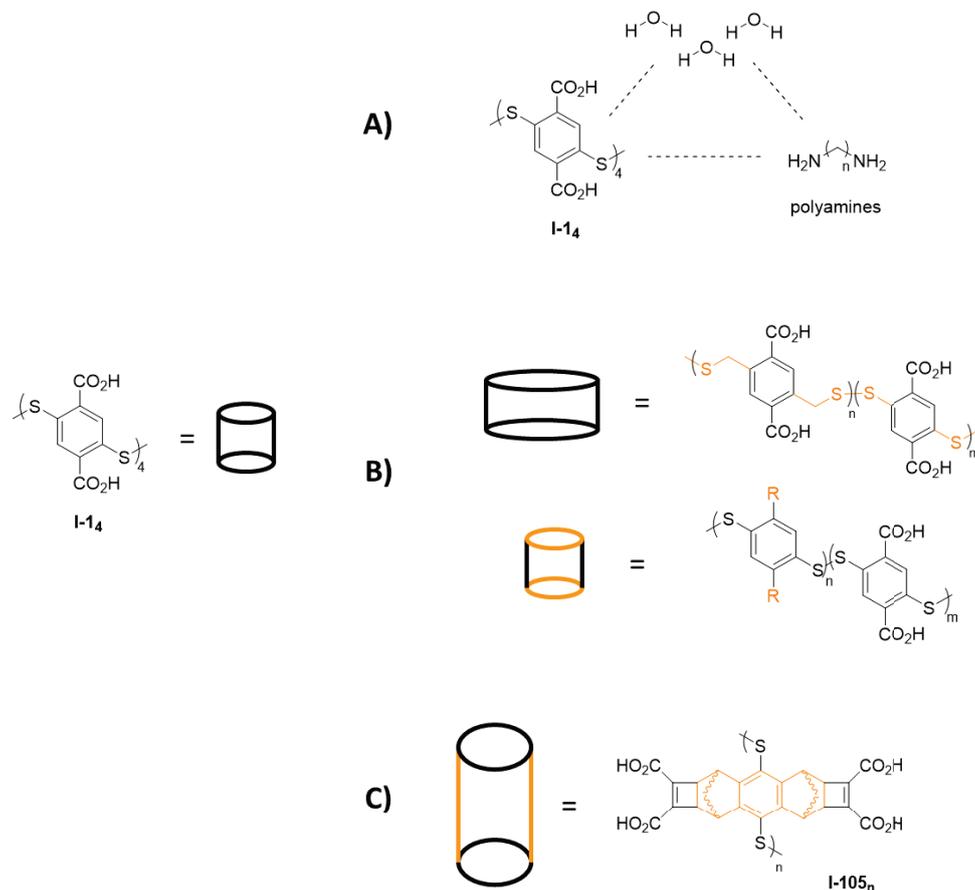
thermodynamics of the complexation between cavitands such as a cyclophane and guests. It is especially the case when disturbed or particularly stable water molecules are expelled upon complexation by the guest. Cyclophanes contributed to the investigation of solvent effects, for instance with the cucurbituril family⁷. Molecular dynamic underscored a toroidal region of high water density at the centre of a nonpolar cucurbit[7]uril⁴². The presence of water in this region is disfavoured energetically and entropically and contribute to the known ability of cucurbituril to bind guest molecules with unusually high affinities (Scheme 7). Maps of the enthalpy and entropy of solvation in and around the cavitand revealed unfavourable contributions to the solvation at the entry of the cavity, relatively to the bulk water^{7c}. By a combination of experimental and computational studies in aqueous solution, the release of perturbed “high-energy” water molecules from a cavity of a cucurbit[8]uril containing dicationic amines favoured the formation of an inclusion complex with electron-rich aromatic guests⁴³. That highlighted the choice of cyclophanes to study solvation relative effects in supramolecular chemistry. Cyclophanes usually contains a water network inside their cavity. Upon complexation by inclusion, the release of water molecules and the overall processes of binding can be entropically (classical hydrophobic effect) or enthalpically (non-classical hydrophobic effect)²⁴ driven, that will be investigated with **I-14**.



Scheme 7 : Spatial maps of second order translational entropy density of water, relative to bulk, in and around a cucurbit[7]uril, from molecular dynamic. Left: Entropy contoured at -0.02 kcal/mol/Å³ (pink), -0.01 kcal/mol/Å³ (transparent orange), and $+0.005$ kcal/mol/Å³ (green). Middle: same, omitting the contour at -0.01 kcal/mol/Å³ and additionally showing water number density contoured at 0.1 waters/Å³ (gray) and 0.175 waters/Å³ (magenta). Right: rotated view, showing only entropy density contours at -0.02 and 0.005 kcal/mol/Å³ and number density contour at 0.01 waters/Å³. Entropies are multiplied by temperature to convert to kcal/mol. Extracted from reference^{7c}.

Few variations on a **I-14** and/or its ligands combined with an experimental and computational study will reveal the key contributions of solvation on the association of **I-14** with polyamines in water. Those associations take place between cationic amines and anionic carboxylates and are mainly electrostatic. A modification of the chemical functions as well as the topology of the cavity of **I-14** will be relevant to study the inclusion complex and the influence of water molecules around and inside the empty cavity. A chain elongation will expand the diameter of the cavity, changing the flexibility of the complex. The use of heterocyclooligomers with different chemical functions will affect the electrostatic interactions. However, an increase of the complexity of the macrocycle will result in more complex

analysis and deconvolution of the contributions. The use of structurally near macrocycles compared to **I-14** paves the way for the templating with similar polyamines and the study of the resulting differences. In addition, the expansion of the cavity height and the separation of the thiol and the carboxylate by a bulk group will favour the steric effects while the electrostatic contribution will decrease. Such cyclophanes are adapted to study the influence of steric effects in self-assembly.



*Scheme 8: Goal and methodology in the chapter 1. The aim was A) study of the affinity for diamines of **I-14** and the solvation effects in its binding with polyamines; B) synthesis of heterocyclo-oligomers derived from **I-1** and others 1,4-bis(2,5-dithiophenyl)benzenes, as well as their homocyclo-oligomers, through DCC in order to modify the size of our cavitands and the chemical functions on the rims (modification in orange, R any chemical group); C) self-assembly of a bulk 1,4-bis(2,5-dithiophenyl)benzene **I-105** to obtain different cavity compared to **I-14** and study of the influence of steric effects in the process. The methodology combines synthetic work, physical chemistry and computational chemistry.*

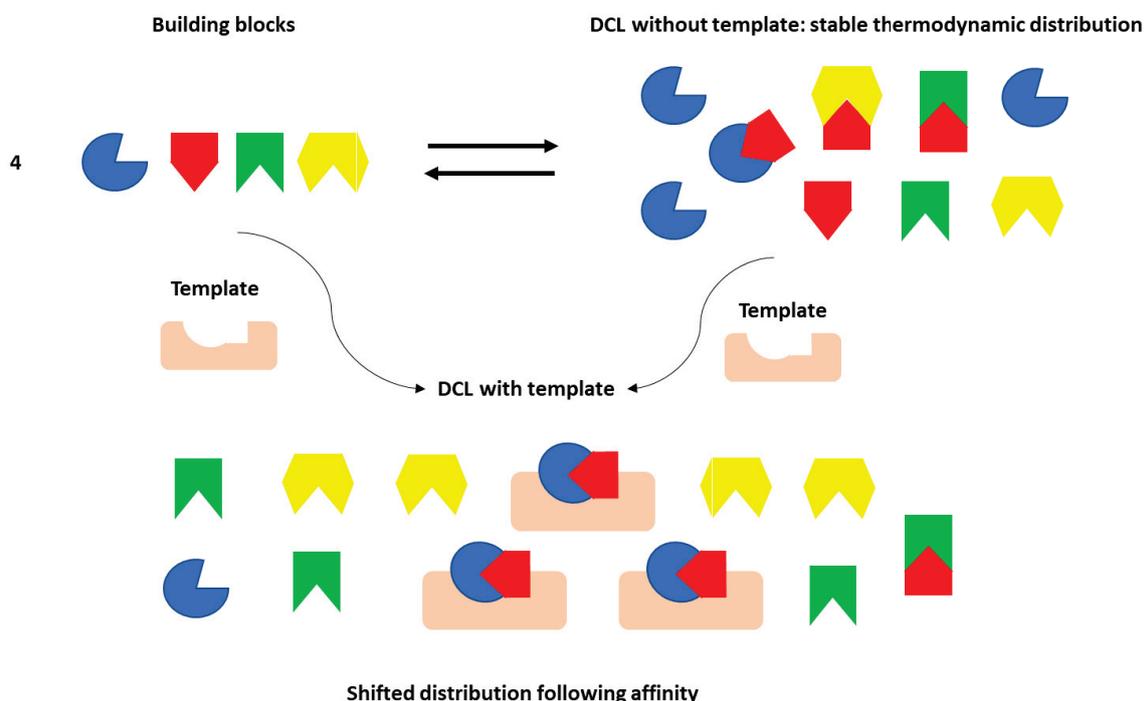
The recent progress made by computational chemistry and molecular dynamics can now provide a further insight into the thermodynamic of the binding mechanisms in host guest chemistry^{7a,43-44}. The synthesis of a previously known **I-1** and new building blocks (**I-15**, **I-105**) will be presented. Then, oxidation of their thiols into disulfide bonds through DCC will enable to study their self-assembly and the subtle change induced by the backbone modifications between the building blocks. Finally, state of the art computational methods such as density functional theory (DFT) and molecular dynamic (MD) will be used to investigate the self-assembling of **I-105** and the binding properties of **I-14** with various polyamines, compared to the experimental data.

2. An aside for a brief presentation of the DCC and the associated computational tools

An important part of the work presented in this manuscript is linked to dynamic covalent chemistry and in a certain extent to computational chemistry. The thermodynamic aspects of the DCLs which were studied experimentally were modelled by computational chemistry. Therefore, the DCC is briefly presented as well as density functional theory and molecular dynamic, which are the two computational tools that are the most used in the community and applied here.

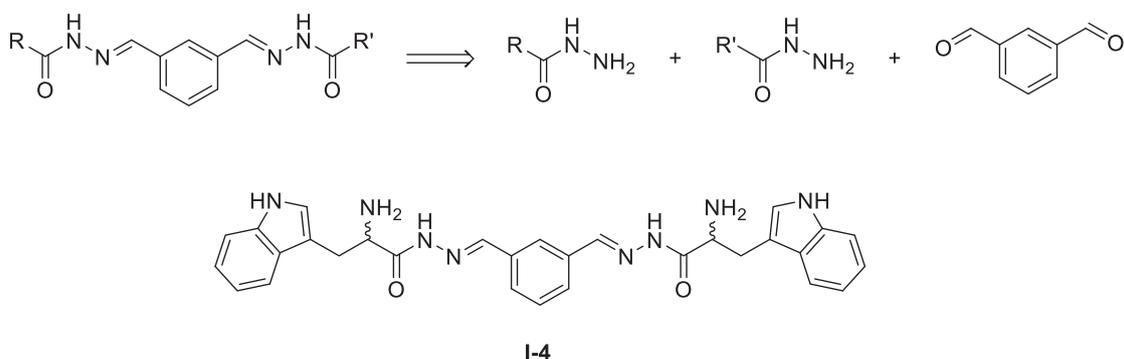
a. Presentation of dynamic combinatorial chemistry

Dynamic combinatorial chemistry is a method of synthesis and screening of chemical compounds, based on self-assembly⁴⁵. Self-assembly is the spontaneous and reversible organisation of multiple simple building blocks into libraries of complex structures⁴⁶ (Scheme 9). In a dynamic combinatorial library (DCL, or library), complex structures from the building blocks are continuously interconverting and virtually explore all the combinations⁴⁷. That is why DCC is a method of synthesis, at the opposite of classical synthetic chemistry⁴⁸ and combinatorial screening⁴⁹. The DCLs reach a stable thermodynamic equilibrium for a given set of starting conditions. The proportion of the member of DCLs are influenced by external stimuli⁵⁰ (redox potential, presence of a molecule for instance (Scheme 9) modifying the relative thermodynamic stability of the library members⁵¹, therefore changing their distributions by enrichment of high-affinity guests for instance. The amplification is linked to the association constants with the external stimulus in the case of a molecule⁵², that is why DCC is a method of screening. The molecule acting as an external stimulus is named template and its concentration dictates the selection/correction process⁵³.



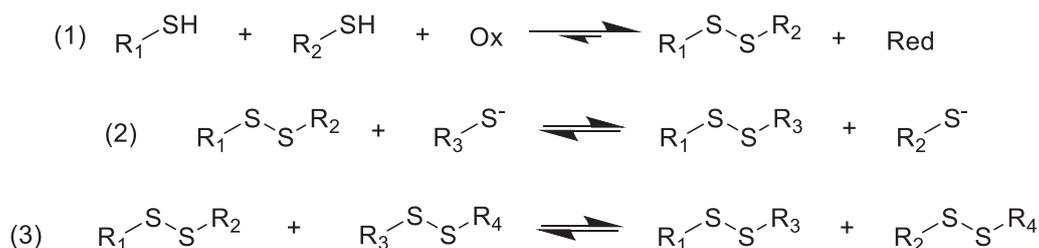
Scheme 9 : Principle of DCC. Different building blocks (left, top) self-assembles into a library of different structures (right, top) by a reversible process (covalent association or non-covalent interactions), reaching a stable thermodynamic distribution. Upon an external stimulus, here a template, the distribution is shifted to a ligand with the highest affinity to the template (bottom)

The application of this method of synthesis and selection are wide⁵⁴, for instance catalysis^{39j}, identification of lead compounds for drug discovery^{49,55} or design of functional materials^{35,56}. DCC is also employed in order to discover inhibitors of enzymes⁵⁷. For instance, a DCL of bis-acylhydrazones identified inhibitors of the aspartic protease endothiapepsin⁵⁸. The most efficient inhibitor **I-4** exhibited an IC₅₀ value of 54 nM, improving 240-fold in potency compared to the parent bis-acylhydrazone hits with a IC₅₀ value of 2.1 μM. The IC₅₀ were calculated thanks to fluorescence increase of the inhibitors separated from the DCLs, because of hydrolysis by endothiapepsin.



Scheme 10: Retrosynthetic analysis of the designed bis-acylhydrazone inhibitors and structure of the most potent inhibitor of the aspartic protease endothiapepsin (IC₅₀ = 54.5 nM, 250-fold improvement compared to the parent hit). R, R' various chemical groups.

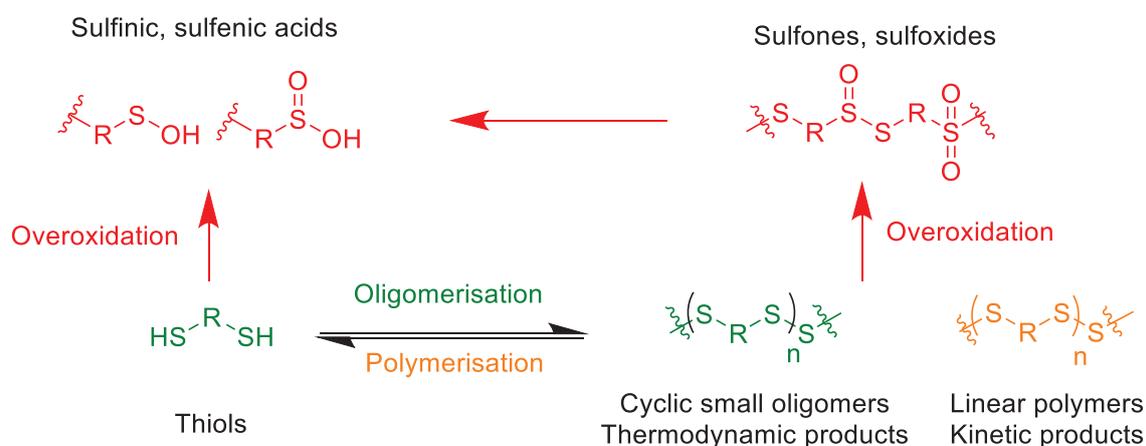
Usually, the DCL are conducted in small sizes aqueous solutions and analysed by LC/MS⁵⁹. Numerous cavitations were formed and amplified by DCC, enlarging the possibilities of structures and improving the yields of synthesis compared to synthetic chemistry^{38a,60}. The reversible organisation of the building blocks is mainly based on covalent links obtained by reversible reactions. The most frequently used reversible reactions in this thesis is the thiol disulfide exchange reaction^{39d} (Scheme 11). The mechanism is given below and requires a thiolate. Consequently, the reaction is pH sensitive (optimal 7-9) and can be stopped at low pH.



Scheme 11: Thiol disulfide exchange reaction, used in DCC. The thiolate concentration is optimum at pH 7-9. Reversibility thanks to (1) connection/disconnection of disulfide bond, (2) substitution, (3) metathesis

Starting from a DCL without preformed disulfide, the thiols are oxidised into a disulfide bond. The reversibility comes from three simultaneous equilibriums: (1) connection and disconnection of thiols into disulfide bond, (2) thiol-disulfide exchange and (3) metathesis of disulfide bonds. The metathesis of disulfide bonds is generally induced by a combination of a reducing agent and base or by irradiation with ultraviolet light⁶¹. In the conditions used for DCL (aqueous buffers at pH 7.4, mildly basic without reducing agent), this third reversible reaction is supposed to be limited.

However, the oxidation of thiols is not limited to disulfide bond. As a matter of fact, the thiols can be oxidised into sulfenic and sulfinic acids and the disulfide bond can also be oxidised in sulfones or sulfoxides for instance⁶². Those oxidations are not desired because they limit the reversible reaction. That is why they are called “overoxidation” in this thesis and must be avoided as much as possible. The building blocks used in this thesis are 1,4-bisthiophenols. The disulfide bond formation can lead to small cyclo-oligomers, which is desired, or to linear long polymers, which is not desired in order to obtain a correct exploration of the different structures and selection generated during a DCL. We are interested by obtaining macrocycles of small sizes. In this case, the disulfide bond formation is called “oligomerisation” in contrast with “polymerisation” (Scheme 12). The synthesis of cyclophanes such as pillararenes depends on the conditions⁶³. The cyclo-oligomerisation is a thermodynamic process, in contrast with the polymerisation, which is a kinetic process. The balance between both depends on the concentration, the temperature and all the reactions conditions.



Scheme 12: Desired reactivity of thiols for DCC based on thiol disulfide exchange (green, oligomerisation in small cyclic oligomers upon thermodynamic control, orange polymerisation upon kinetic control) and undesired reactivity (overoxidation into sulfenic, sulfenic acids, sulfones and sulfoxides), stopping the reversibility.

Due to the accessibility of thiols and the mild conditions of the reversible reactions, disulfide based DCLs are widely used and applied for the assembling of cavitants^{16,34a,37,38b,39a,b,64}. The oxidant that is mainly used in this thesis is molecular oxygen O₂. Iodine I₂ is also employed. In the literature and in this work, the DCL obtained by I₂ are said “kinetic”, while the DCL obtained by O₂ are said “thermodynamic”. The dissolution of oxygen in a solution is a slow process compared to the rate of the thiol-disulfide exchange, explaining the term thermodynamic. The equilibration can take place while the oxidation of thiols into disulfide bonds does not surpass the equilibration. At the opposite, iodine is added in solution in the DCL. Its dissolution and reaction by oxidation of thiols in the media are quick compared to the disulfide exchange rate, explaining the term kinetic.

b. Density functional theory: fundamentals

The DFT is an *ab initio* approach, which means that the calculation is based on theoretical principles and not on experimental data. The mathematical approximations defined the computational methods. One of the most famous is the Born Oppenheimer approximation⁶⁵, stating that the motion of atomic nuclei and electrons in a molecule can be separated (eq 1). The corresponding electronic hamiltonian can be written as followed for stationary nuclei (H hamiltonian, T kinetic energy operator, V potential energy operator) (eq 2).

$$\psi_{tot} = \psi_{electrons} \otimes \psi_{nuclear} \quad (1)$$

$$H_{electronic} = T_{electrons} + V_{electron/nuclear} + V_{electron/electron} \quad (2)$$

The DFT origin lies in the Thomas-Fermi model of the atom⁶⁶, which was improved later by Honenberg and Kohn. In short, the energy of an atom was approximated by the equation (3) (v nuclear potential, ρ electronic density).

$$E[\rho(r)] \approx 2,871 \int \rho(r)^{5/3} dr + \int \rho(r)v(r)dr + \iint \frac{\rho(r)\rho(r')}{|r-r'|} drdr' \quad (3)$$

The DFT energy is based on the electronic density, determining the whole system, while Hartree Fock methods are based on wavefunctions. In fact, the Honenberg and Kohn theorems state that the fundamental energy of a system is completely described by its electronic density⁶⁷. The fundamental energy is hence a functional¹ of the electronic density and a variational principle can be applied (eq 4).

$$E[\rho(r)] = \int \rho(r)v(r)dr + F[\rho(r)] \quad (4)$$

Nevertheless, there is no exact analytical formula giving the energy from the electronic density. The energy corresponding to the previous hamiltonian can be written as followed⁶⁸ (eq 5).

$$E = T + E_{electron|nuclei} + E_{electron|electron} \quad (5)$$

The electronic interactions are composed of a coulombic contribution J , and an exchange correlation term E_{xc} ⁶⁹. The last is due to the electronic indiscernibility. Thus, the energy of interactions between the electrons could be written as followed⁶⁸ (eq 6).

$$E = T + E_{electron|nuclei} + J + E_{xc} \quad (6)$$

The accuracy of DFT lies in the E_{xc} term. The other ones can be exactly determined. All the current developments intend to propose more accurate functionals accounting for this exchange/correlation⁷⁰. The current functionals are based on a generalised gradient approximation (GGA) rather than local density approximation (LDA). The electronic density is considered as a non-uniform electronic gaz. Hence, the functional depends on the electronic density and its first derivatives⁷¹. Meta-GGA functionals also depends on second derivatives. The most frequently used functional at the moment and in this work is B3LYP. It is an hybrid one, combining several other functionals⁷² (eq 7).

$$E_{xc}^{B3LYP} = E_{xc}^{LDA} + a_1(E_x^{HF} - E_x^{LDA}) + a_2(E_x^{GGA} - E_x^{LDA}) + a_3(E_c^{GGA} - E_c^{LDA}) \quad (7)$$

The DFT methods benefit from the correlation/exchange in its theoretical developments but suffer from the accurate determination of this contribution. The DFT is widely used in order to obtain optimised geometries, excited states, energies for instance⁷³. The calculation cost is reasonable with a satisfying accuracy compared to other computational methods. Finally, an attention should be paid for system highly influenced by dispersion forces (Computational description).

c. Essentials of molecular dynamic simulation

The molecular dynamic is a computational method using classical mechanics. The evolution of a molecular system during a given time is obtained through the Newton's law. This is a time dependent

¹ A functional is an application of a vector space to its scalar body

and an empirical approach. In fact, the bonds are not calculated but are parameters thanks to force fields. A force field is a list of parameters and potentials describing the potential energy of a system⁷⁴. The total energy of the system is written as followed (eq 8). For instance, the bond energy can be modelled by a spring potential (eq 9).

$$E_{total} = E_{bond} + E_{angle} + E_{dihedral} + E_{VDW} + E_{Electrostatic} \quad (8)$$

$$E_{l_{ij}} = \frac{1}{2} k_{ij} (r_{ij} - r_{ij}^0)^2 \quad (9)$$

The forces are calculated thanks to the force fields and then the Newton's law (eq 10) is integrated by a Taylor development (eq 11) for instance, giving the atomic positions at the next time step δt ⁷⁵. The natural motion of the molecular system is modelled with fixed conditions such as solvent box, temperature, pressure and boundary periodic limits⁷⁶. The given energy enables the atoms to move and to interact with their neighbours. The molecular dynamic is widely used in material science and medicinal chemistry⁷⁷.

$$\sum_i F_i = m \vec{a}_i \quad (10)$$

$$x(t + \delta t) = x(t) + v(t)\delta t + \frac{1}{2} a(t) \delta t^2 + \dots \quad (11)$$

These two computational methods could be applied in the context of dynamic covalent chemistry in order to explain the driving forces of self-assembly and the stability of some macrocycles.

3. Self-assembly of 2,5-dimercaptoterephthalic acid **I-1** through DCC and selective complexation of a corresponding disulfide cyclophane **I-1₄** with polyamines in water

Dynamic combinatorial chemistry is a supramolecular approach using self-assembly to generate a variety of compounds starting from building blocks. The distribution between the member of a DCL can change upon templating and can amplify a specific component. Such a phenomenon of selection can be used to selectively synthesised and purified synthetic receptors through DCC. Thanks to the process of correction and selection, these synthetic receptors approaches natural systems in terms of binding affinities with their templates or relative guests^{45,78}. Self-assembly^{39d,i,k,64g,7939d,i,k,64g,7939d,i,k,64g,79} of dithiol building block through DCC have been widely explored for the elaboration of new host molecules^{33d,i,k,57g,72}. In this context, this subchapter presents the self-assembly of a bithiol building block **I-1**, which is widely employed in our team, into a disulfide cyclotetramer **I-1₄** and the complexation of this receptor with polyamines.

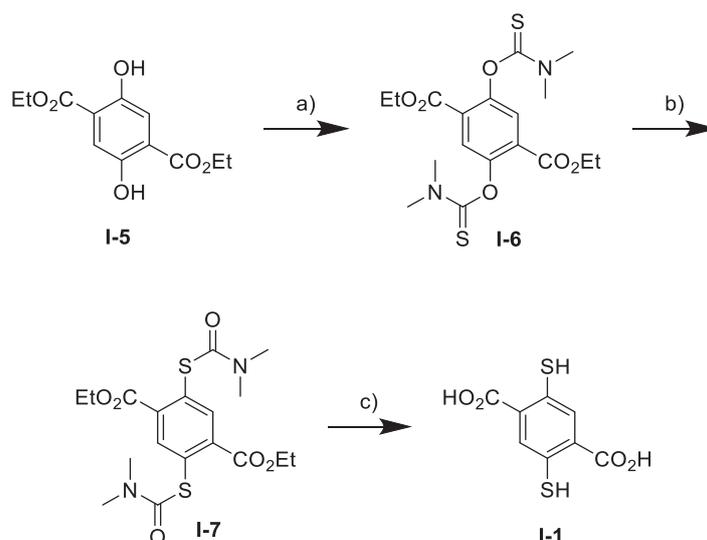
a. Self-assembly of 2,5-dimercaptoterephthalic acid **I-1** through DCC and structural properties of its oligomer **I-1₄**.

In this part, the aim was to synthesise the targeted building block and the derived cyclotetramer **I-1₄**.

i. Synthesis and self-assembly of **I-1**

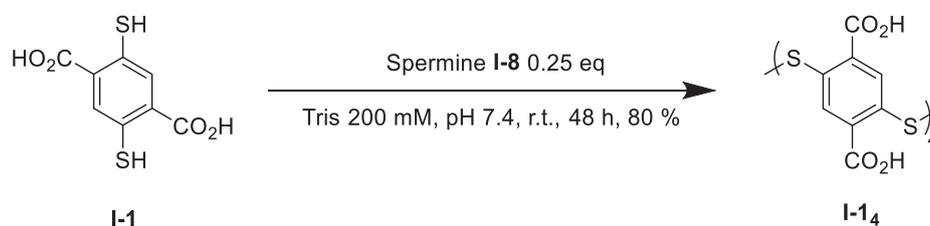
The building block **I-1** was used by our team and different research groups interested in thiol exchange reaction and DCC^{79b,80}. The bithiol was synthesised in 3 steps from **I-5** (*Scheme 13*). The first step consisted of the introduction of the O-thiocarbamate group on the alcohol position¹⁶ with a satisfying yield of 99 %. Then, a neat and thermal rearrangement of Newman-Kwart converted the O-thiocarbamate **I-6** into a S-thiocarbamate group¹⁶ **I-7** with an excellent yield of 82 % compared to other similar rearrangements⁸¹. Finally, a basic deprotection cleaved both S-thiocarbamate and ester groups

into thiols and carboxylic acids¹⁶ with a satisfying yield of 95 %. The overall yield was 77 % in 3 steps. The synthesis required no chromatographic purification and could be performed at a multigram scale.



Scheme 13: Synthesis of I-1 a) dimethylcarbamothioic chloride, DABCO, DMA, r.t., 24 h., 99 %; b) neat, 210°C, 82 %, c) KOH, EtOH/H₂O, 60°C, 5 h., 95 %

The dissolution of **I-1** at 4 mM in a Tris buffer (200 mM pH 7.4) led to the formation of three discrete macrocycles after 48 hours of stirring (LC-MS and ¹H NMR analysis): a trimer **I-1₃** (23 %), a tetramer **I-1₄** (74 %) and a pentamer **I-1₅** (3 %). The same DCL, templated by 1 mM of spermine, led to a spectacular amplification of 99 % of the tetramer **I-1₄**¹⁶. A precipitation induced template removal and separation of **I-1₄** with a yield of 80 % and a purity of 99 % (NMR).



Scheme 14 : Amplification of I-1₄ by the template spermine I-6 from the I-1 building block in Tris buffer

The macrocycle **I-1₄** was widely studied in the group over the last years^{40a,41,82}. It was named as a member of the family of dyn[n]arenes, with other similar disulfide macrocycles. The term “dyn” accounts for the dynamic aspect of the DCC.

ii. Structural properties of **I-1₄**

The formation of disulfide bonds from thiols was a source a stereogenic elements regarding the structure of **I-1₄**. There was not only one cyclic tetramer that could be formed through the self-assembly, but several because of specific configurations and conformations. As a matter of fact, each disulfide bonds could adopt a specific conformation⁸³, P or M referring to the nomenclature⁸⁴. This comes from the ideal torsion angle of 90° of the disulfide bonds⁸⁵. Consequently, 4 different stereoisomers of conformations were possible: PPPP, MMMP, MMPP and MPMP (Figure 2).

In addition, the monomeric unit is a 1,4-bisthiophenol with carboxylic acid in the 2,5 positions of the aromatic core. Given that the rotation around the axis of the sulfur-sulfur bonds are blocked in the macrocycles, a stereodescriptor of configuration prochiral R/S could be attributed to each disulfide bond⁸⁶. That resulted in different homo- or hetero-chiral **I-14** (Figure 2). **I-14** obtained by templating with spermine showed a sharp peak in ¹H NMR. A DFT modelling suggested an homochiral product (pS)₄/(pR)₄, minimising the repulsive Coulombic interactions between the carboxylate groups around the polar rims. The ¹³C NMR revealed a unique C–H signal. Given that the inversions of the disulfide bridge's dihedral angle was estimated to be around 75 kJ.mol⁻¹, the conformers were exchanged at the NMR time scale and could not be seen by NMR. That self-assembly of **I-1** into a homochiral **I-14**, templated by spermine, illustrated the isomolecular diversity from the oxidation of thiols into disulfide bonds through DCC. The configurational and conformation diversity is a common feature in this thesis regarding all the DCL involving 1,4-bisthiophenols.

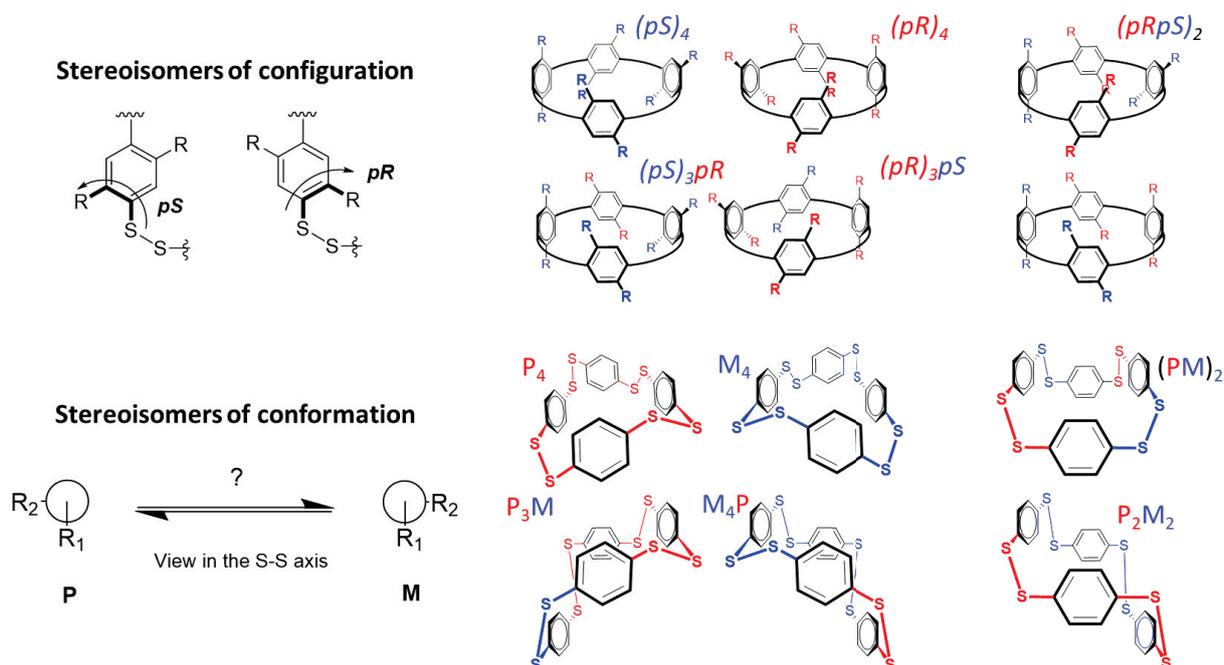
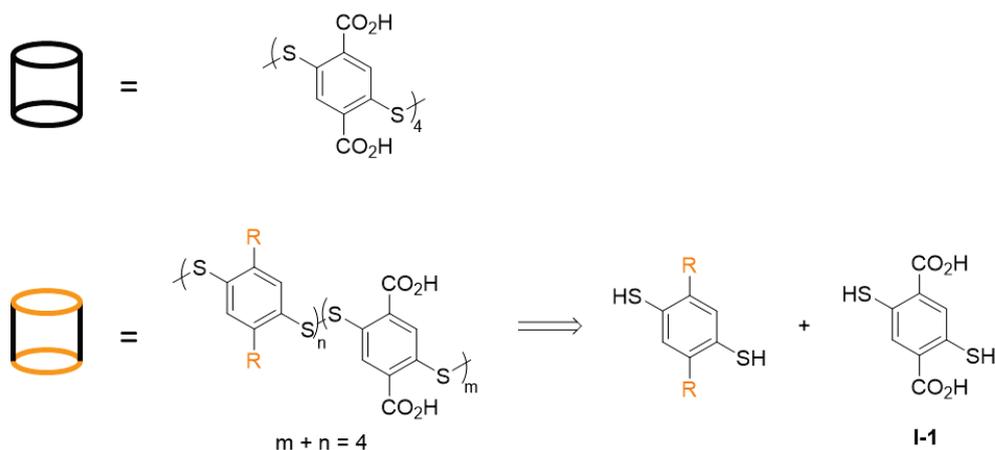


Figure 2 : Isomolecular diversity in the self-assembly of **I-1** into **I-14** regarding conformation and configuration concerns. The carboxylic acids are omitted for clarity for the conformation analysis

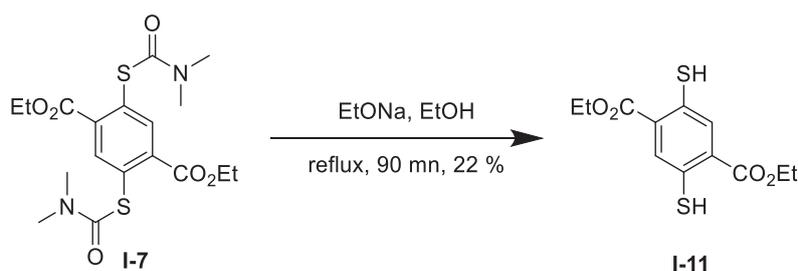
iii. Heterocyclo-oligomerisation with **I-1** and other 1,4-bisthiophenols

The aim of this subpart was to synthesise heterocyclo-oligomers, especially tetramer, of **I-1** and other 1,4-bisthiophenols through DCC in order to obtain modifications on the chemical functions available on the rims of the macrocycles.



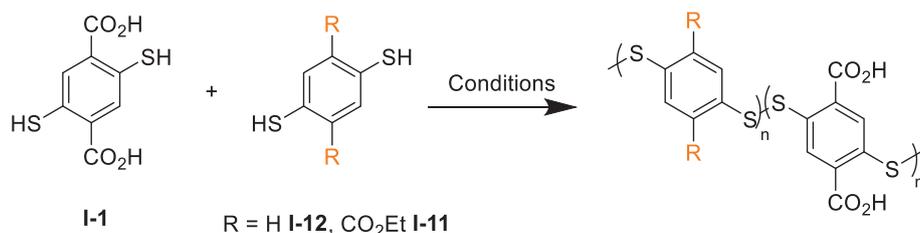
Scheme 15: Objectives of the subpart: synthesis of heterocyclooligomers from I-1 and 1,4-bisthiophenols through DCC (orange, R any chemical groups) in order to obtain a similar cavity than I-1₄ with different chemical functions on the rims

Two building blocks were combined with I-1 in DCLs: diethyl 2,5-dimercaptoterephthalate I-11 and dithiophenol I-12. I-12 was the product of a partial basic deprotection of S-thiocarbamate groups^{82a} with a poor yield of 22 % (Scheme 16).



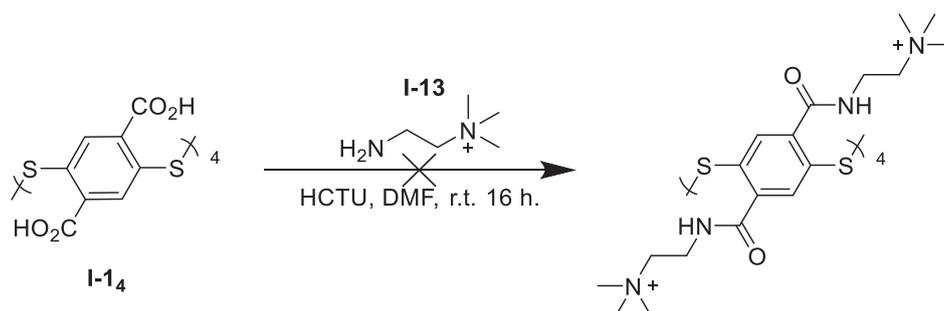
Scheme 16: Synthesis of I-11

The building blocks I-1 and I-11 or dithiophenol I-12 were dissolved at 2 mM respectively in solvent (Scheme 17). The DCLs were performed in a mixture 1:1 of Tris buffer 200 mM pH 7.4 and organic solvent (MeCN or DMSO). The mixtures were analysed by LC-MS and/or ESI-MS analysis. For both combinations of I-1 with I-11/12, linear, homo- and hetero-oligomers were observed in ESI-MS analysis. Nevertheless, no longer oligomers were detected. Precipitation occurred during the libraries and might result from a polymerisation or a poor solubility of the oligomers containing I-12 and I-11 units. The screening of the solvents, the stoichiometry and the templating by spermine did not led to the detection of desired heterocyclo-oligomers by LC-MS or ESI-MS. The conditions of DCLs could be optimised in the future.



Scheme 17: Building blocks for DCLs attempting to create a mutant of I-1₄

Alternatively, post-assembling modification of **I-1₄** was attempted by peptide-like coupling with **I-13** (Scheme 18). The amide bond formation involved HCTU^{79e,87} and gave a complex crude mixture, which could not be purified.

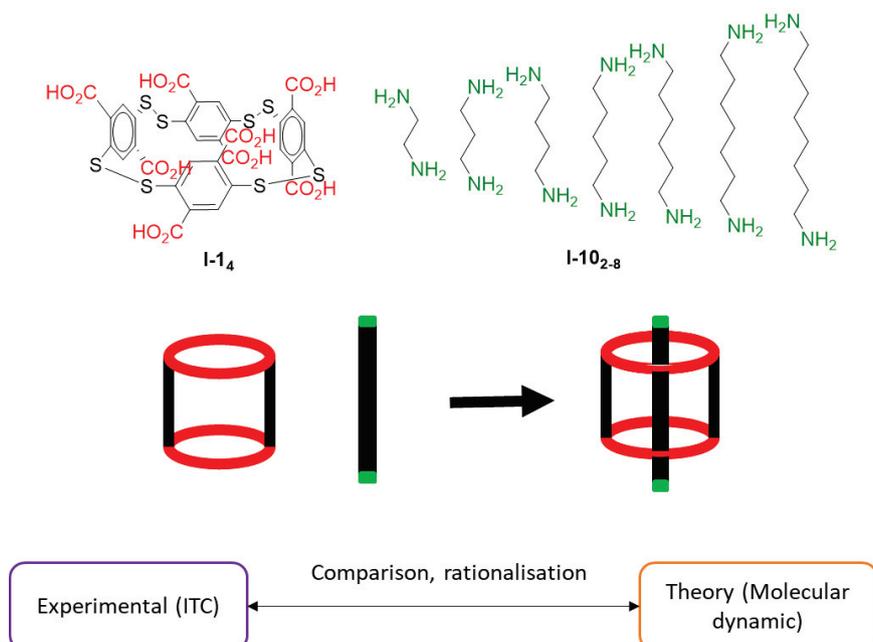


Scheme 18: Modification attempt of **I-1₄** by amide bond formation with cationic amine **I-13**

The cyclic tetramer **I-1₄** can be amplified through a DCL of **I-1** thanks to the spermine **I-8** template. It was also accessible by oxidation with iodine. No mutants or modifications could be performed on the macrocycle.

- b. Factors governing the selectivity of the carboxylate receptor **I-1₄** with a series of α,ω -alkyldiammonium ions of increasing chain length in water

The aim of this part explored the factors governing the association between the macrocycle **I-1** with various polyamines **I-10_n** (Scheme 19). Polyamines play significant roles in mammalian physiology. They regulate many metabolism in cells and act as antioxidants⁸⁸. for instance Putrescine promote cell division and ethylenediamine is an agonist of GABA⁸⁹. Besides, abnormal concentrations of polyamines are linked with many disease states, such as a defect in lysine metabolism⁹⁰. Polyamines can modulate the functions of DNA and RNA⁹¹. In fact, most of polyamines form a complex with these two biomolecules.



Scheme 19: Formation of pseudo rotaxanes with **I-1** and a series of diamines **I-10** with increasing chain lengths. The goal of this part was to compare a series of experimental data from ITC with computational data from

molecular dynamic in order to understand the driving forces of the complexation as well as the selectivity for **I-10₆**

Historically, the receptor **I-1₄** was discovered in a work linked to DNA¹⁶. **I-1₄** bound spermine with a K_d of 22 nM, enough to remove it from its natural host DNA and reverse some of the biological effects of spermine on the nucleic acids. In fact, low concentrations of spermine induces the formation of left-handed DNA, but upon addition of the receptor, the DNA reverted to its right-handed form. The macrocycle **I-1₄** is fully deprotonated in solution (determined by pH titration⁹²), bearing 8 negative charges, and its formation is templated by spermine **I-7**. Hence, it is a relevant receptor for polyammonium ions thanks to salt bridges with the anionic carboxylates. In addition, the hydrophobic cavity can interact with alkyl chain following the chain length. The binding strength between a series of terminal alkyl diamines **I-10** and **I-1₄** was investigated in order to understand the driving force of the complexation and the role of solvation². ITC experiments³ were performed in order to provide a set of thermodynamic data (Δ_rH° , Δ_rS° , Δ_rG° , K_a) which were compared to their theoretical counterpart obtained by molecular dynamic simulations. The computational resources were provided by the PSMN of the ENS de Lyon. The computational details are given in the experimental section (Molecular dynamic). **I-1₄** and the series of diamines form pseudorotaxanes upon complexation (Scheme 19). The experimental thermodynamic data of the association are presented in the following table, reaching an optimal value for hexanediamine **I-10₆** (Table 1).

Guest	K_a ($10^6 M^{-1}$)	Δ_rH° (kJ.mol ⁻¹)	$-T\Delta_rS^\circ$ (kJ.mol ⁻¹)
I-10₂	0.03 +/- 0.00	-5.04 +/- 1.05	-20.8 +/- 0.36
I-10₃	0.21 +/- 0.03	-5.48 +/- 0.82	-24.7 +/- 0.58
I-10₄	9.25 +/- 1.20	-24.7 +/- 0.32	-16.5 +/- 1.53
I-10₅	28.4 +/- 3.7	-26.5 +/- 1.06	-15.6 +/- 1.01
I-10₆	46.2 +/- 3.4	-27.5 +/- 0.07	-16.2 +/- 0.29
I-10₇	42.1 +/- 6.2	-33.3 +/- 0.20	-9.47 +/- 0.34
I-10₈	32.8 +/- 5.6	-37.7 +/- 1.42	-4.96 +/- 2.31

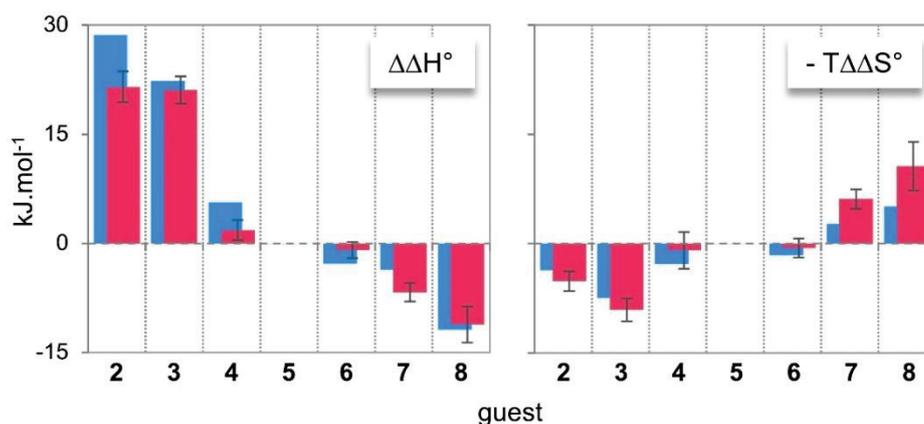
Table 1: Thermodynamic data of the association between **I-1₄** and the diamine series **I-10**

Computed energies of the association were obtained by MM-PBSA calculations, which is a common method to estimate the free energy of the binding of small molecules with biomacromolecules⁹³. The only valid comparison between computed data with the experimental data must be drawn on the relative thermodynamic parameters ($\Delta\Delta_rH^\circ$, $\Delta\Delta_rS^\circ$, $\Delta\Delta_rG^\circ$) because MMPBSA provides inaccurate absolute energies of complex formation (Scheme 17). The two sets of data followed the same trends. The relative enthalpy of association $\Delta\Delta_rH^\circ$ become more and more favourable as the chain length increased while $\Delta\Delta_rS^\circ$ was more and more unfavourable. This agreement suggested that both entropy from water molecules and energetic penalties associated with conformational fit upon complexation remained constant along the series.

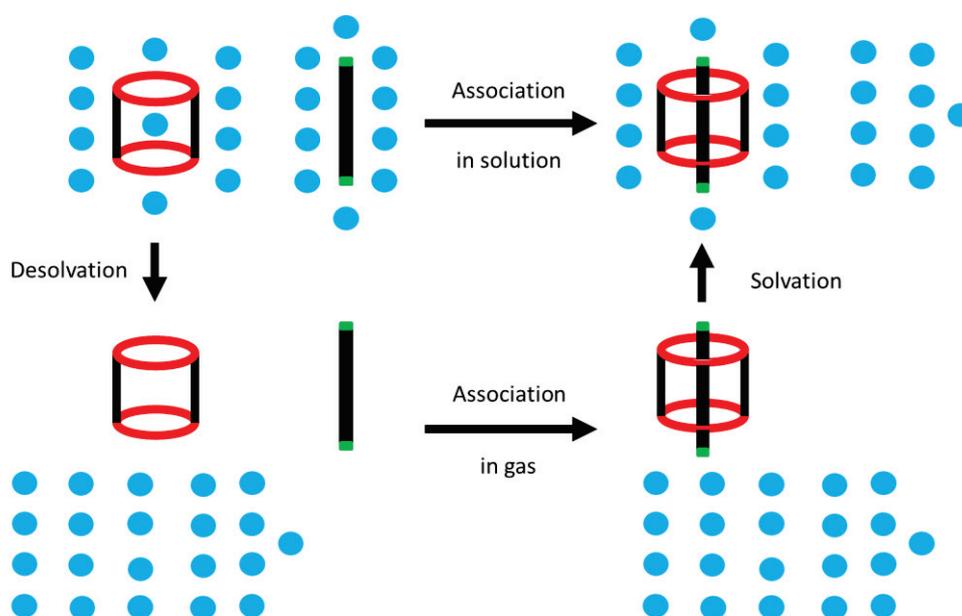
²Many people contributed to the work presented in this paragraph. My contribution is minor and mainly consisted in the **I-1₄** synthesis and a small part of the molecular dynamics. The reader should consider this subpart regarding polyamines as a bibliographic paragraph regarding rather than a personal contribution to the field. Nevertheless, the following subpart regarding spermine and spermidine is my full contribution to the theoretical investigation of the driving forces.

³ The ITC experiments were performed by Dr. Heloin

Firstly, a lock and key model could not explain the selectivity. As a matter of fact, the distance between plus and minus charges, which encompassed the electrostatic complementarity, and the SASA, which encompassed the geometric complementarity between host and guest, predicted that **I-105** must have the more favourable enthalpy of association $\Delta_r H^\circ$. However, **I-106** formed a more stable complex, even though its enthalpy of association $\Delta_r H^\circ$ was higher by 13 kJ.mol⁻¹ compared to the longest chain **I-108**. In addition, the classical hydrophobic effect was constant along the series and did not discriminate the diamines. Consequently, the lock and key theory was not adapted to the system. That is why a finer analysis involving water was explored thanks to a thermodynamic deconvolution (Scheme 21).



Scheme 20: Comparison between the experimental (red) and computed (blue) relative energies of binding between host **I-14** and guests **I-10**, with **I-105** as a reference.



Scheme 21: Thermodynamic deconvolution of the binding process between **I-14** and diamines into a desolvation step, an association in gas phase and a solvation of the complex. Each step is associated with an enthalpic and entropic contribution. Blue, water molecules.

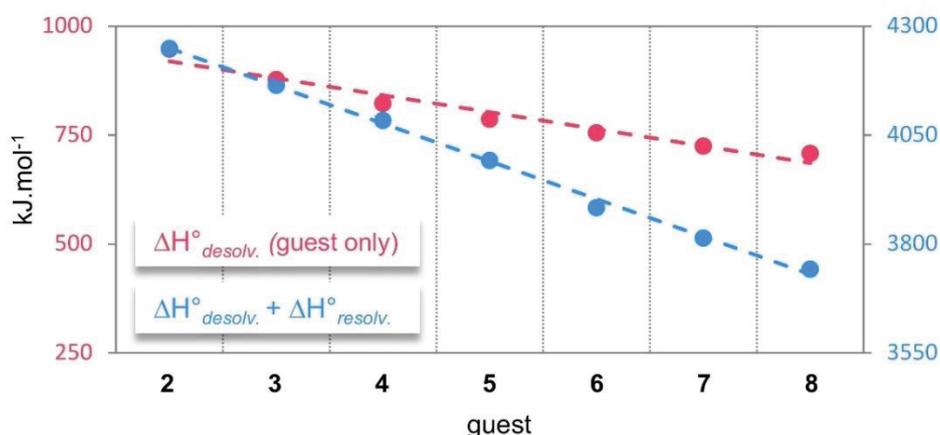
The process was formally divided into a desolvation of the partners (eq), an association of the receptor and the diamine in gas phase (eq) and finally a solvation of the complex (eq). To each step was associated an enthalpic $\Delta_r H^\circ$ and entropic contribution $\Delta_r S^\circ$, partially accessible thanks to MM-PBSA (eq 11 and 12).

$$\Delta_r H^\circ = \Delta_r H^\circ_{desolvation} + \Delta_r H^\circ_{association,gas} + \Delta_r H^\circ_{resolvation} \quad (11)$$

$$-T\Delta_r S^\circ = -T\Delta_r S^\circ_{desolvation} - T\Delta_r S^\circ_{association,gas} - T\Delta_r S^\circ_{resolvation} \quad (12)$$

The principal entropic contributions to an association process are due to changes in translational, rotational and vibrational entropy. Firstly, the relative entropic association in gas term $-T\Delta_r S^\circ_{association,gas}$ was less and less favourable along the series because of the increasing loss of freedom degrees for the diamine upon binding (ie translational and internal rotation freedom). The correspondence observed between the experimental relative entropy of complex formation in solution and the computed relative $-T\Delta_r S^\circ_{gaz,association}$ indicated that the sum of the solvation $-T\Delta_r S^\circ_{solvation}$ and desolvation $-T\Delta_r S^\circ_{desolvation}$ term, encompassing the classical hydrophobic effect and the release of water molecules from the polar areas remained constant along the series of guests.

Regarding the enthalpy $\Delta_r H^\circ$, the association in gas $\Delta_r H^\circ_{association,gas}$ was purely of electrostatic nature. The term was less and less favourable along the series because of the increasing inductive effect of the aliphatic chain. However, that did not follow the experimental trend. Consequently, the resolution of the complex $\Delta_r H^\circ_{resolvation}$ and desolvation of host and guest $\Delta_r H^\circ_{desolvation}$ terms were determinants. In fact, the experimental trend was similar to the sum of the enthalpic desolvation and resolution $\Delta_r H^\circ_{resolvation} + \Delta_r H^\circ_{desolvation}$. These computed values increased with the chain length and were dominated by polar contributions corresponding to the desolvation and resolution of the polar areas of the host and the guest.

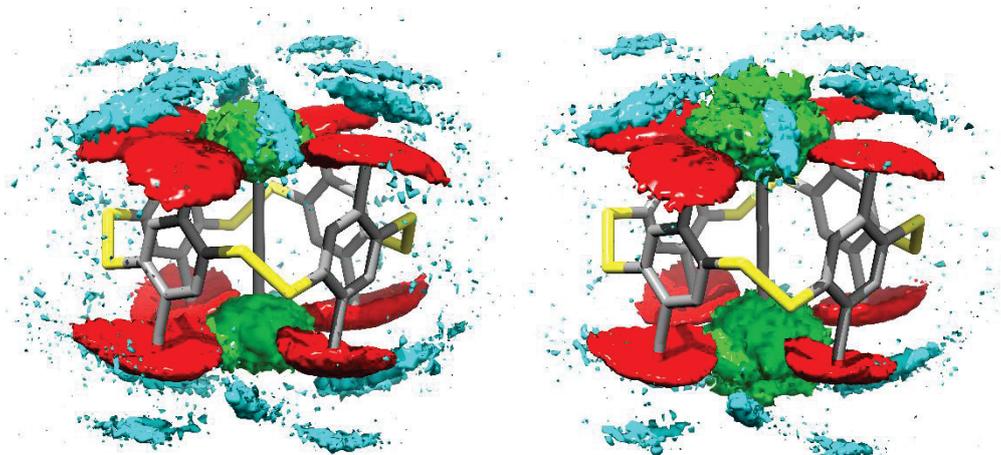


Scheme 22: Evolution of the computed enthalpy of desolvation of guests I-10_n, and of the corresponding relative enthalpy of solvation between bound and unbound states with host I-1.

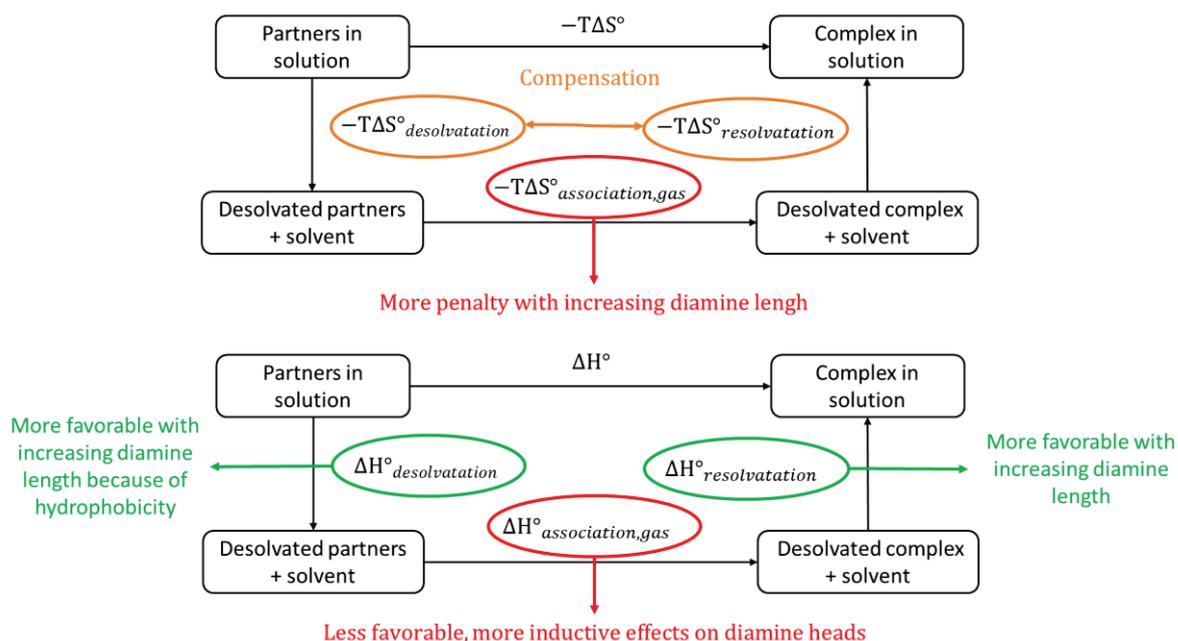
The contribution of the receptor was the same along the series. Nevertheless, the desolvation cost decreased with the chain length because the diamine became more and more hydrophobic (Scheme 84). Mathematically, the resolution of the complex $\Delta_r H^\circ_{resolvation}$ became more and more favourable with the guest's chain length despite its progressive protrusion from the cavity and the exposure to the solvent of its hydrophobic surface.

That more and more favourable $\Delta_r H^\circ_{resolvation}$ was attested by a higher water density at the polar heads of the carboxylates along the series (Scheme 23). The increased density of water around the ionic heads of the complexes resulted in both stronger and more numerous solute-solvent and solvent-solvent interactions, related to $\Delta_r H^\circ_{resolvation}$. It also led to an increased entropic cost of solvent immobilisation,

but the entropic benefit of desolvation of the diamine counterbalanced the effect, which was coherent with the constant sum of the entropic resolution and desolvation terms.



Scheme 23 : Superposition of the average position of both receptor **I-14** and guests with the density of presence
Blue : water (64.4 %), Red : COO⁻ (76.7 %), Green : NH₂⁺/NH₃⁺ (99.5 %). Left: putrescine **I-10₃**, right: cadaverine **I-10₄**.



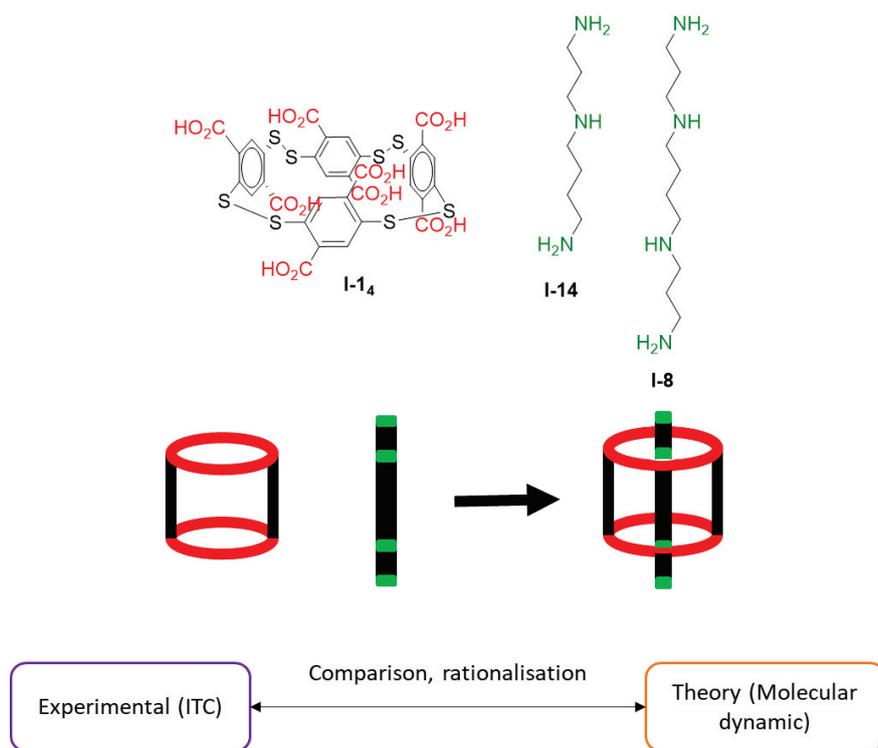
Scheme 24: Contribution of each enthalpic and entropic computed terms to the global picture of binding between **I-14** and the series of diamines, explaining the experimental trend.

To sum up, the association between **I-14** and diamines were dictated, as the chain length increased: 1) entropically, by the unfavourable loss of freedom degrees of the guest related to $\Delta_r S^\circ_{\text{association,gas}}$, 2) enthalpically, by two solvent related phenomena: the decrease of the enthalpic cost of guest desolvation $\Delta_r H^\circ_{\text{desolvation}}$ and the increase of the enthalpic benefit of complex resolution $\Delta_r H^\circ_{\text{resolution}}$ (Scheme 24).

The binding process was enthalpically highly influenced by two solvent related phenomena, rationalising the observed selectivity.

- c. Factors governing the selectivity of the carboxylate receptor **I-14** with spermine and spermidine in water

The aim was to explain the selectivity of the macrocycle **I-14** with two biogenic amines: spermine **I-8** and spermidine **I-14** (Scheme 25) and to study in details the influence of solvation in the process. These polyamines are essential regulators of cellular proliferation^{91,94} and hence are biological targets for their molecular recognition by synthetic receptors. ITC experiments were performed in order to provide a set of thermodynamic data regarding the complexation of the receptor with the polyamines and the stoichiometry of the complexes. The data was compared to their theoretical counterpart obtained by molecular dynamic simulations. The computational resources were provided by the PSMN of the ENS de Lyon. The computational details are given in the experimental section (Molecular dynamic).

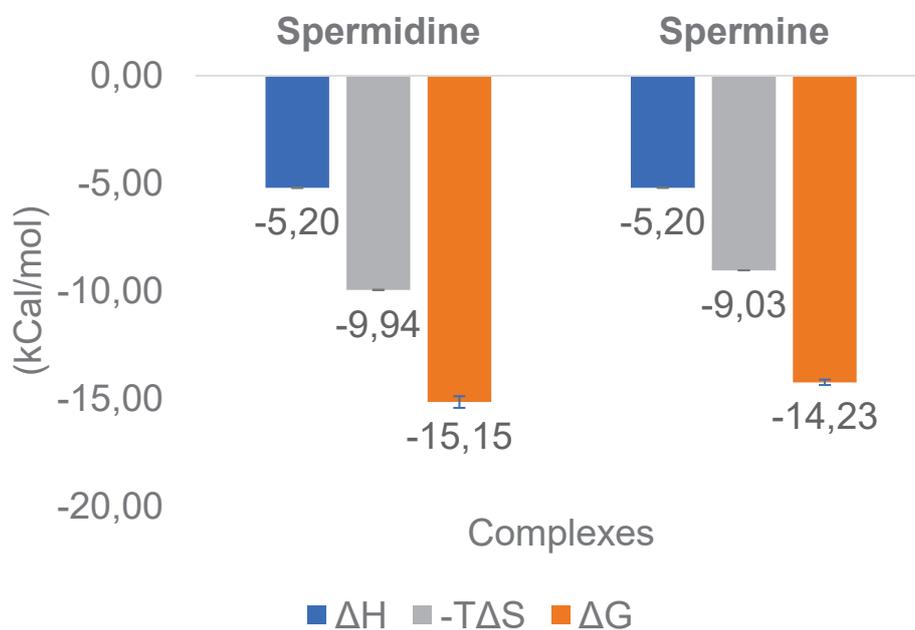


Scheme 25: Formation of pseudo rotaxanes with **I-14** and two biological amines **I-8** and **I-14**. The goal of this part was to compare a series of experimental data from ITC with computational data from molecular dynamic in order to understand the driving forces of the complexation as well as the selectivity for **I-14**

I-14 and the two amines formed pseudorotaxanes upon complexation (Scheme 25). The constants of the association were determined by means of ITC⁴, reaching an optimal value for spermidine **I-14** (Scheme 26). That was counterintuitive, because spermine **I-8** is a tetracation while spermidine **I-14** is a trication at physiological pH. At a first glance, a stronger association because of more electrostatic interactions was expected with spermine **I-8**. The constants of association reached 10^{10} M^{-1} for spermidine and 10^{11} M^{-1} for spermine. They are the highest described constants of association for a synthetic receptor in water. The difference between the two amines came from the entropy, more

⁴ The ITC experiments were performed by Dr. Heloin

favourable for spermidine **I-14**. The aim of molecular dynamic simulations was to explain that selectivity. Three systems were simulated: **I-14** alone, and the inclusion complexes with the two polyamines, during 100 ns in a box of water molecules. The cpptraj⁹⁵ and Few⁹⁶ packages were employed.



*Scheme 26 : Experimental thermodynamic data of the association between **I-14** and spermine **I-8** /spermidine **I-14**, obtained by ITC titrations. The experiments were triplicated. Errors were calculated by standard error calculations.*

i. Qualitative analysis of the topology of and internal mobility of the complex

The average position on the atoms during the simulations were superposed with the 3D representation of the probability of presence in order to evaluate the mobility of the guests and the host (Figure 3)⁵. A probability of presence of n % for an atom/molecule means that n % of the frames contains the atom/molecule within the coloured volume, illustrating the mobility of the atoms during the simulation.

⁵ Grid and average command in cpptraj

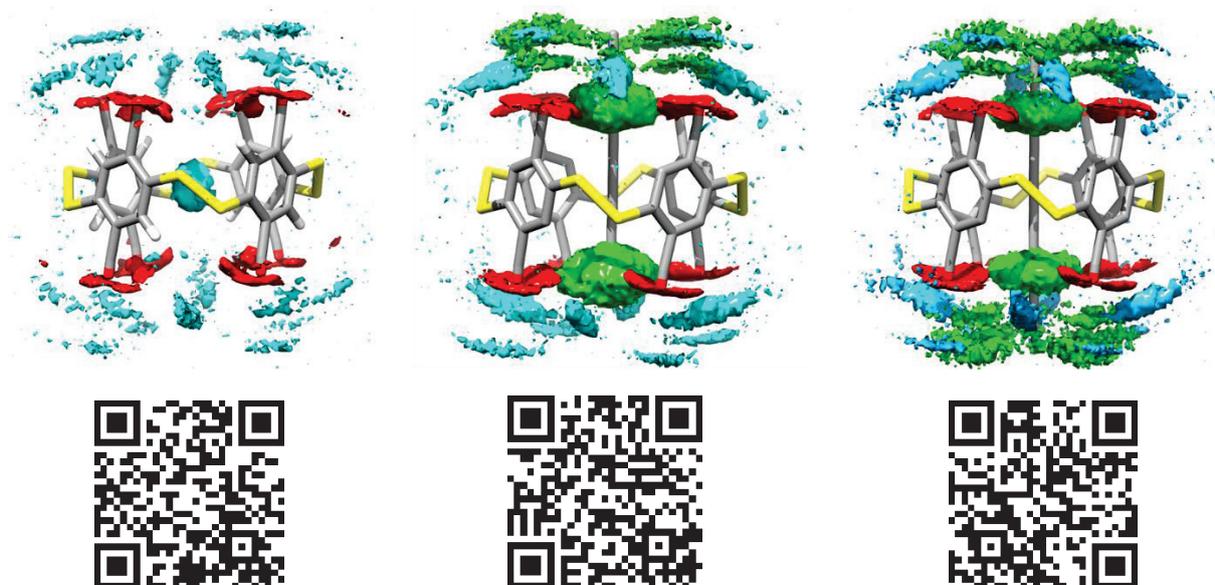


Figure 3 : Superposition of the average position of receptor **I-14** and guests during the simulation, with the density of presence Blue : water (64.4 %), Red : COO⁻ (76.7 %), Green : NH₂⁺/NH₃⁺ (99.5 %). Left: host alone, middle complex with spermidine **I-8**, right: spermidine **I-14**.

The different freedom degrees of the host and the guests are represented below. They are composed of translational and rotational degrees.

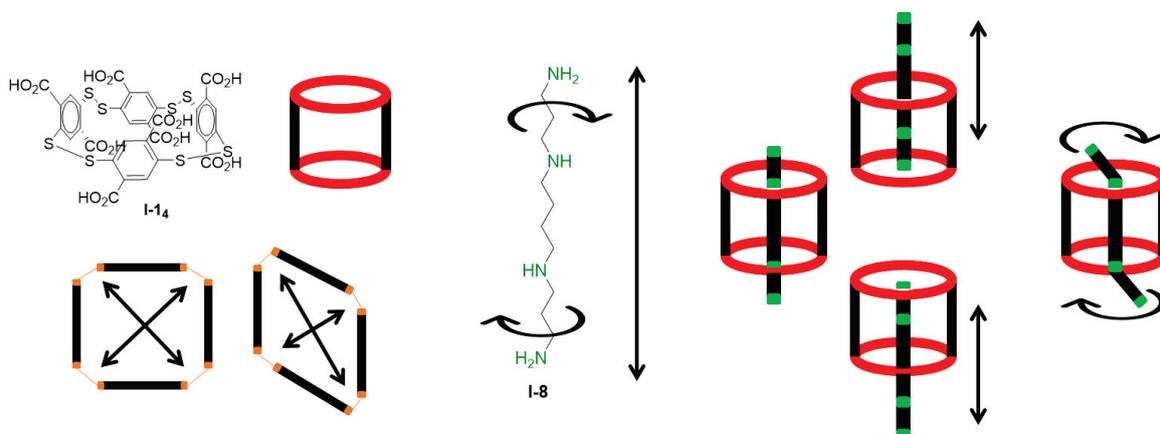


Figure 4 : Freedom degrees of the receptor **I-14** and the polyamines in the inclusion complexes with **I-14**. The faces of the receptor could modify the shape of the cavity because of the disulfide bridges while the polyamines could translate inside the cavity, and the external amines could spin around the rims of the cavity.

The average position of the host atoms was well defined, even for the simulation of the host alone. The host was poorly mobile, which was coherent with the strain induced by the short disulfide connection between the aromatic units. The guests had a translational mobility inside the cavity of the complexes, as attested by the unwell defined average positions of their atoms and the large green grid of NH₂⁺/NH₃⁺ probability of presence. The pictures revealed that the translational mobility was qualitatively higher for spermidine **I-14** compared to spermine **I-8**. The two internal amine moieties of spermidine **I-14** seemed to have a slightly different vertical mobility one over the other, which was coherent with the asymmetry of the guest. The external amines turned around the carboxylate rims, interacting without discrimination with all the carboxylates. The presence grid gave an interesting

picture of the solvation. A high density of water was found inside the empty cavity. That presence of water molecules was not a continuity to the outside of the cavity. This density will be important for the binding process. Indeed, the water molecule will be ejected from the cavity. Following its entropic and enthalpic stability, the influence on the binding is crucial. Regarding the complexes, the grids revealed that the solvation was denser at the rims of the complexes compared to the host alone, while they were no obvious differences between the two complexes at this stage.

Logically, the average number of contacts per frame between the host and the guest was higher with spermine **I-8** (7.9 vs 6.6 for spermidine **I-14**). The contacts⁶ were defined by a distance cut-off of 3.5 Å. The contacts are mainly due to salt-bridges between the anionic carboxylate groups and the cationic ammoniums. Spermine **I-8** has one additional ammonium (4) compared to spermidine **I-14** (3), increasing the number of contacts. The average value and snapshots indicated that 2 carboxylates established simultaneous contact with one ammonium when an external ammonium was engaged in a contact. Regarding internal ammoniums, two or only one carboxylate interacted with the ammonium. That explained the average value of 7.9 and 6.6 (for instance with spermine **I-8**, for a representative frame $2 \times 2 \text{ COO}^-/\text{NH}_3^+ + 1 \text{ COO}^-/\text{NH}_2^+_{\text{intern 1}} + 2 \text{ COO}^-/\text{NH}_2^+_{\text{intern 2}} = 8$). The details of the contacts between the host and the guests are detailed below (Figure 5). The map represents the proportion of contacts between each carboxylates and ammoniums for the two complexes. Differences emerged. Regarding the symmetric spermine **I-8**, the internal ammoniums were less in contact compared with the external ammoniums. The scenario was different for spermidine **I-14**. In fact, the triamine was locked in the cavity by the internal terminal ammonium, which was at nearly 80 % in contacts with all the carboxylates of the inferior rims during the simulations. That led to a more mobile external ammonium, which was less in contact with the superior rim compared to the external ammonium of spermine **I-8**. That mobility scenario was coherent with the grid of presence that is discussed above. The lifetime of a contact was short, on average 5 ps for each sort of contact. The maximum lifetime reached 100 ps for internal ammoniums, slightly higher compared to external ammoniums (Figure 6). That illustrated the mobility of the amines, due to the quick jumps of the ammoniums around the rims of the host.

⁶ Command Native contact in cpptraj.

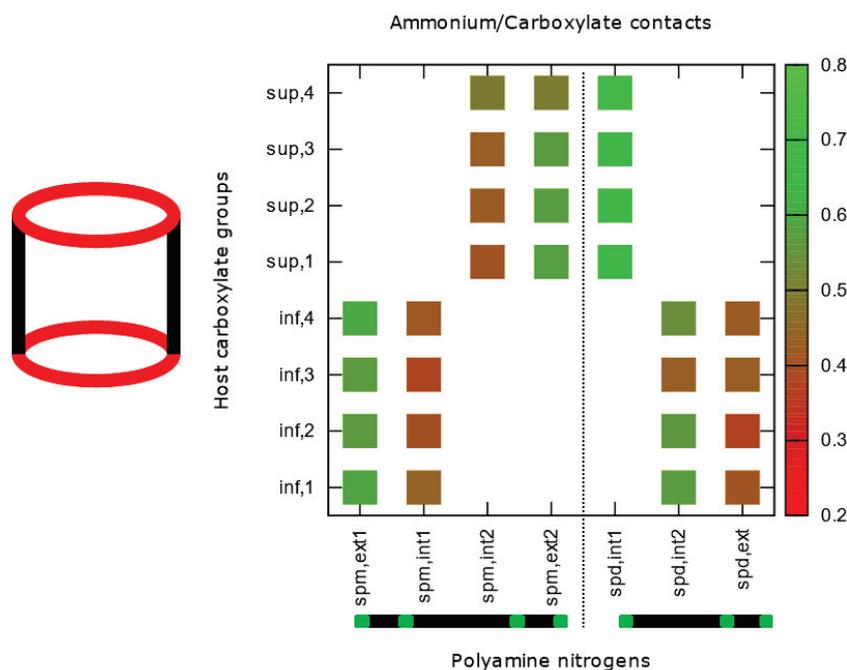


Figure 5: Proportion of contacts between the ammoniums of the guests (spermine I-8, spermidine I-14) and the carboxylate groups of the host I-14. The host has 4 carboxylates on each inferior and superior rims. The amines have internal and external ammonium groups. The contacts are normalised to 1. 0.8 means that the ammoniums are in contact with the carboxylates during 80 % of the simulation. A contact exists if the distance between the groups are equal or less than 3.5 Å



Figure 6 : Average and maximum lifetimes of the contacts between the carboxylate groups of the host rims and the ammonium groups of the guests spermine I-8 (spm) and spermidine I-14 (spd).

The root mean square deviation (RMSD) measures the deviation of a set of coordinates to a reference set of coordinates. A RMSD of 0.0 indicates a perfect overlap. The deviation of the host was similar (1,4) for each complex (Figure 7). The RMSD⁷ of spermidine I-14 in its complex was higher compared to spermine I-8 in its complex, resulting in a higher overall deviation for the complex of I-14 with spermidine I-14. That was coherent with the contacts analysis and proved that the spermidine I-14 had more mobility compared to spermine I-8 in the association.

⁷ Command RMSD with cpptraj.

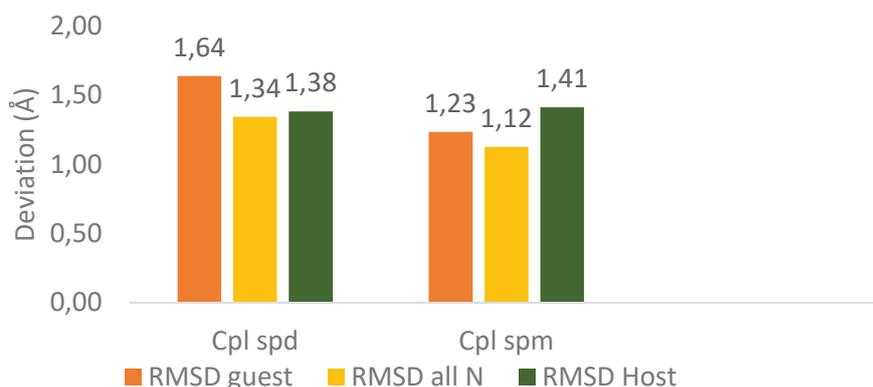


Figure 7 : RMSD of the host **I-14** and guests spermidine **I-14 (spd)** and spermine **I-8 (spm)** in their complexes.

The experimental thermodynamic data indicated a more favourable entropy of association for spermidine **I-14**. That analysis indicated that the mobility, thus the freedom degrees, of the two guests were different. That could play a role in the entropy of association. The analysis does not have a tendency and does not account for the solvation but starts explaining the entropic difference.

*There were more salt bridges in the association between spermine **I-14** and **I-14**, but the polyamine was less mobile compared to spermidine **I-14** in its complex. The guests showed vertical and spinning horizontal mobility, especially for external ammoniums. The empty host contained a high density of water molecules. The complexes revealed an increased density of water at the rims compared to the empty host.*

ii. Thermodynamic of the association

By the same comparative approach explored with diamines, the experimental and computational data extracted from ITC and MMPBSA were compared. The only valid comparison between such computed data with the experimental data must be drawn on the relative thermodynamic parameters ($\Delta\Delta_rH^\circ$, $\Delta\Delta_rS^\circ$, $\Delta\Delta_rG^\circ$) because MMPBSA provides inaccurate absolute energies of complex formation. The data were analysed by difference (Figure 8).

As a matter of fact, the MMPBSA calculation indicated a more favourable relative enthalpy of complexation $\Delta\Delta_rH^\circ$ for spermine **I-8**, while ITC experiments showed similar enthalpy of association. Computed $\Delta\Delta_rS^\circ$ relative entropy of complexation for spermine **I-8** was less favourable compared to spermidine **I-14**, similarly to the experimental trend. That resulted in a contradictory relative free energy of association $\Delta\Delta_rG^\circ$, experimentally slightly better for the spermidine **I-14** by nearly 1 kcal.mol⁻¹ and computationally more favourable for spermine.

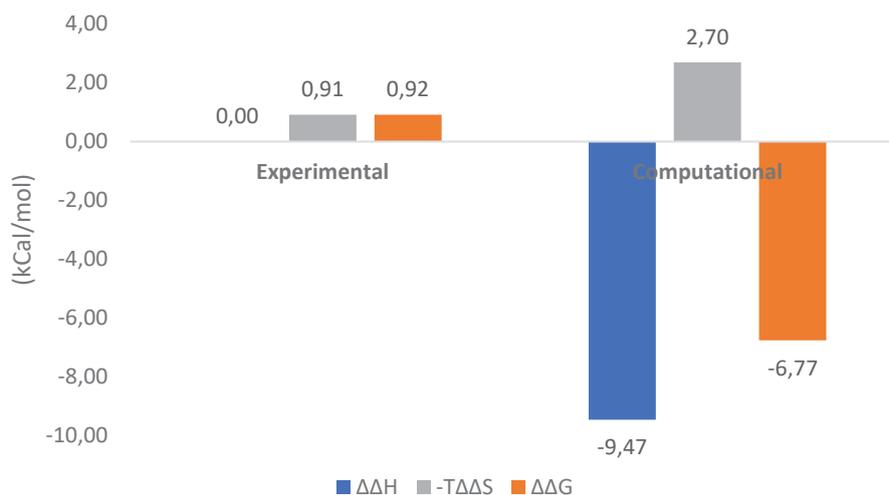
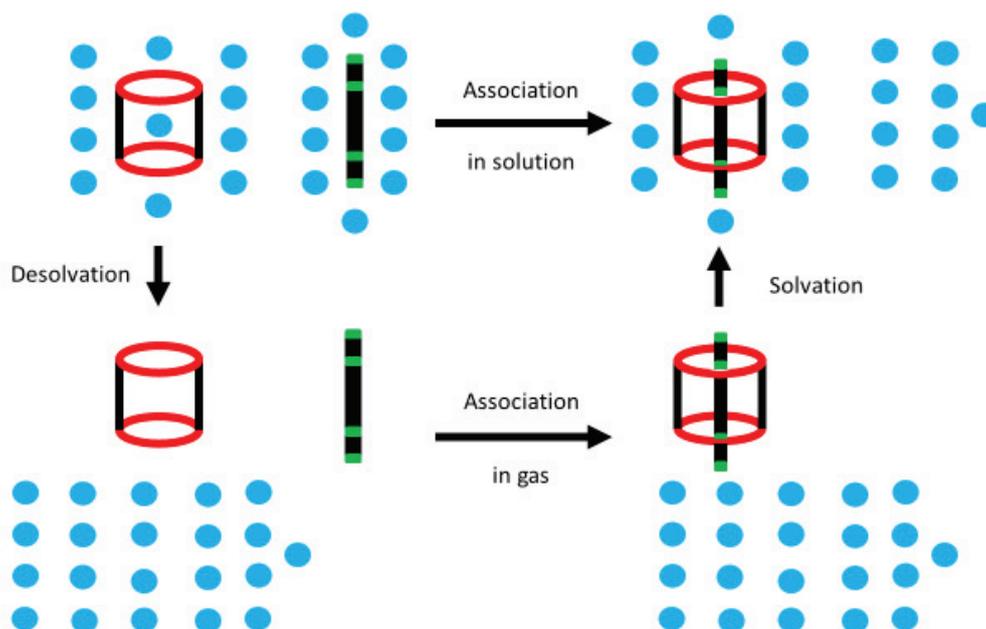


Figure 8 : Comparison between experimental (ITC) and computational data (MMPBSA) of the association between I-14 and spermine I-8 compared to spermidine I-14.

The disagreement between experimental and computed parameters ($\Delta\Delta_r H^\circ$, $\Delta\Delta_r S^\circ$, $\Delta\Delta_r G^\circ$) suggested that the entropy from water molecules did not remain constant with two polyamines, given that the snapshots of the two complexes showed a similar conformational fit upon complexation. Like the precedent study with diamines, a finer analysis involving water was explored thanks to a thermodynamic deconvolution (Scheme 21). The process was formally divided into a desolvation of the partners (eq), an association of the receptor and the diamine in gas phase (eq) and finally a solvation of the complex (eq). To each step was associated an enthalpic $\Delta_r H^\circ$ and entropic contribution $\Delta_r S^\circ$, partially accessible thanks to MM-PBSA (eq 11 and 12).



Scheme 27 : Thermodynamic deconvolution of the binding process between I-14 and diamines into a desolvation step, an association in gas phase and a solvation of the complex. Each step is associated with an enthalpic and entropic contribution. Blue, water molecules.

$$\Delta_r H^\circ = \Delta_r H^\circ_{desolvation} + \Delta_r H^\circ_{association,gas} + \Delta_r H^\circ_{resolvation} \quad (11)$$

$$-T\Delta_r S^\circ = -T\Delta_r S^\circ_{desolvation} - T\Delta_r S^\circ_{association,gas} - T\Delta_r S^\circ_{resolvation} \quad (12)$$

$\Delta_r H^\circ$ was first analysed. The desolvation $\Delta_r H^\circ_{desolvation}$ of the receptor was similar for the two complexes. The desolvation $\Delta_r H^\circ_{desolvation}$ of the guests was less favourable for spermine **I-8** for + 190 kcal.mol⁻¹. In fact, spermine had one more cationic charge with only 2 additional carbons. That resulted in more interaction between spermine **I-8** and the solvent to be disrupted during the complexation compared to spermidine **I-14**. Regarding the association in gas $\Delta_r H^\circ_{gas,association}$, the term was more favourable for the complex with spermine **I-8** by -387 kcal.mol⁻¹. That was coherent given that there was one more cationic charge in the inclusion complex interacting with the carboxylate rims of the receptor. A linear interaction energy analysis⁸ validated the result (Figure 9). This illustrated the increasing energy resulting from electrostatic interactions (resulting in a more favourable $\Delta_r H^\circ_{gas,association}$) between the host and spermine **I-8** compared to spermidine **I-14** (by -180 kcal.mol⁻¹), and that Van der Waals contributions were weak.

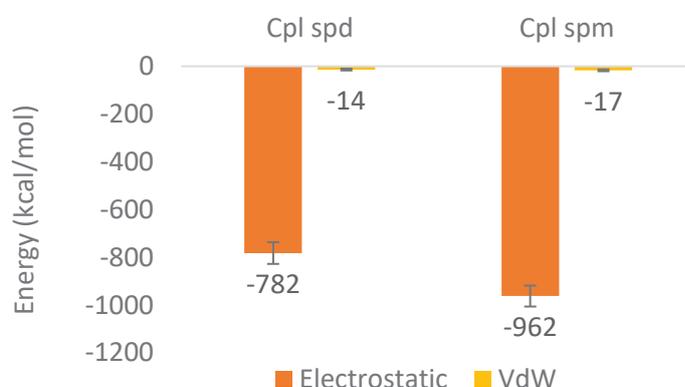


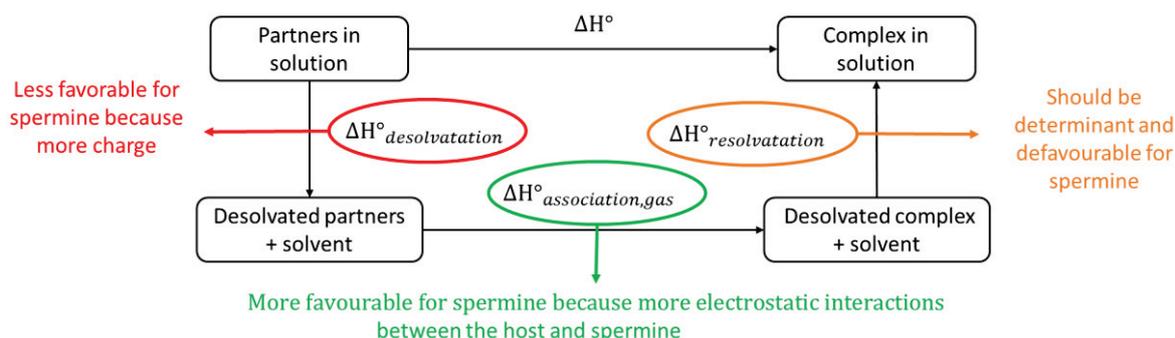
Figure 9 : Linear interaction energy analysis of the complex between **I-14** and spermine **I-8** (spm) / spermidine **I-14** (spd).

Finally, $\Delta_r H^\circ_{resolvation}$ was less favourable for the complex with spermine **I-8** by 185 kcal.mol⁻¹, which was at the opposite of the $\Delta_r H^\circ_{resolvation}$ of the diamines, more favourable for increasing chain lengths. The difference came from the presence of the internal amines, which depleted the accessible negative charge density of the carboxylate by salt bridges for the water molecules.

Therefore, $\Delta_r H^\circ_{resolvation}$ was less favourable in the case of spermine **I-8** at the opposite of $\Delta_r H^\circ_{desolvation}$, cancelling their contribution to the selectivity of the complexation. Therefore, the resulting enthalpy of association $\Delta_r H^\circ$ should followed the contribution $\Delta_r H^\circ_{gas,association}$, which does not match the experimental trend. Given that $\Delta_r H^\circ_{gas,association}$ was the more reliable and that the tendency of the desolvation $\Delta_r H^\circ_{desolvation}$ was logical, we hypothesised that the solvation term $\Delta_r H^\circ_{resolvation}$ was the less accurate in MMPBSA and should explain the equal experimental enthalpy of association (Scheme 28). We supposed that the trend (i.e. less favourable solvation of spermine **I-8** complex compared to spermidine **I-14** complex) was correct but that the values were not correctly estimated. In fact, the grid density of presence of water (Figure 3) showed an increased density of water molecules at the

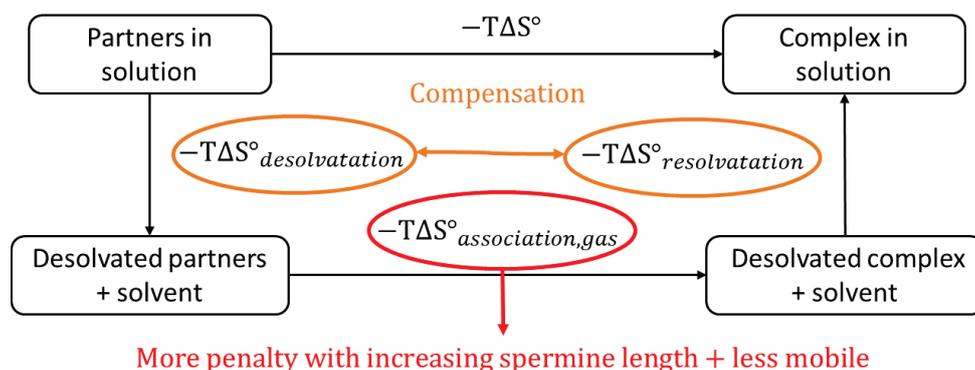
⁸ Command LIE, cpptraj

lower rims of the spermidine **I-14** complex. The water density is a direct mark of a better solvation and hence a more favourable enthalpy of solvation.



Scheme 28 : Contribution of each enthalpic computed terms to the global picture of binding between I-1₄ and the two polyamines in order to explain the experimental trend which was similar enthalpy of association for the guests.

Then, the entropy $-T\Delta S^\circ$ was analysed. The term of association in gas $-T\Delta_r S^\circ_{\text{gaz,association}}$ was found more unfavourable for spermine **I-8** by $2.70 \text{ kcal.mol}^{-1}$. That was coherent with the previous analysis of the topology and indicated a loss of freedom degrees upon binding (ie translational and internal rotational freedom). The receptor was similarly mobile for both guests, while spermine **I-8** was less mobile compared to spermidine **I-14**. That can explain the difference. The correspondence observed between the experimental relative entropy of complex formation in solution and the computed relative $-T\Delta_r S^\circ_{\text{gaz,association}}$ seemed to indicate that the sum of the solvation $-T\Delta_r S^\circ_{\text{solvation}}$ and desolvation $-T\Delta_r S^\circ_{\text{desolvation}}$ term, encompassing the classical hydrophobic effect and the release of water molecules from the polar areas remained constant along the series of guests. At this point, the MMPBSA explained the entropic scenario of the association (Scheme 29).



Scheme 29 : Contribution of each entropic computed terms to the global picture of binding between I-1₄ and the two polyamines in order to explain the experimental trend which was less favourable entropy of association for spermine I-8.

Consequently, the remaining question was the enthalpic contribution of solvation $\Delta_r H^\circ_{\text{resolvation}}$ of the complexes. A quantitative investigation of the interactions between water molecules and the complex was made thanks to the molecular dynamic simulations and cpptraj.

The increased loss of freedom degrees of spermine I-8 compared to spermidine seem to explain the difference of entropy between the associations. Regarding the enthalpy, the association in gas phase

was in favour of the spermine I-8 complex, while the desolvation of spermine I-8 was less favourable compared to spermidine I-14.

iii. Study of the solvation energy of the complexes

Different tools were employed in order to study the influence of water in the associations. As previously discussed, the grid density of water presence (Figure 3) showed an increased density of at the lower rims of the spermidine I-14 complex, accounting for a more favourable enthalpy of solvation. That could be explained by the disruption of water networks at the lower rim by the external ammonium of spermine I-8. As a matter of fact, the spinning move and the jumping salt bridges prevented from a strong organisation of the solvation spheres in this part of the complex, resulting in a less favourable enthalpy of solvation for spermine I-8 complex. It has been shown that solvent organisation is a key contributor to the thermodynamics of protein–ligand recognition⁹⁷. The solvation spheres of the complex with spermine I-8 was slightly more important than spermidine I-14, but the network was probably less stable in the lower rim area (70 vs 66 water molecules at 3.4 Å, 143 vs 136 at 5 Å⁹. Defined as the count of waters within a certain distance of the atoms of the complex). The slightly higher number of water molecules was coherent with the slightly higher solvent accessible¹⁰ surface area for spermine I-8 complex (622 vs 607 Å²).

Several packages were developed by the community in order to analyse the thermodynamics of hydration sites of proteins and ligands⁹⁸. The decomposition of the global solvation process into local contributions is a widely employed method to understand the stability of biomolecular complexes⁹⁹. Among the available packages, two are based on different approaches and are implemented in cpptraj tools and were used: SPAM¹⁰⁰ and GIST^{7c,42}.

SPAM¹¹ (maps spelled in reverse) finds discrete hydration sites at the water–host interface and computes a local free energy from the distribution of interaction energies between water and the environment. SPAM can qualitatively estimate the thermodynamic profile of bound water molecules and can highlight “hot spots”. The hot spots were determined for the empty receptor I-14. The pictures highlighted the bound water molecules and particularly showed the presence of a discrete high density of water molecules inside the cavity. That was coherent with the grid of water presence obtained previously by grid and average command of cpptraj (Figure 3).

⁹ Command watershell, cpptraj.

¹⁰ Command molsurf, cpptraj.

¹¹ Command volmap and SPAM, cpptraj.

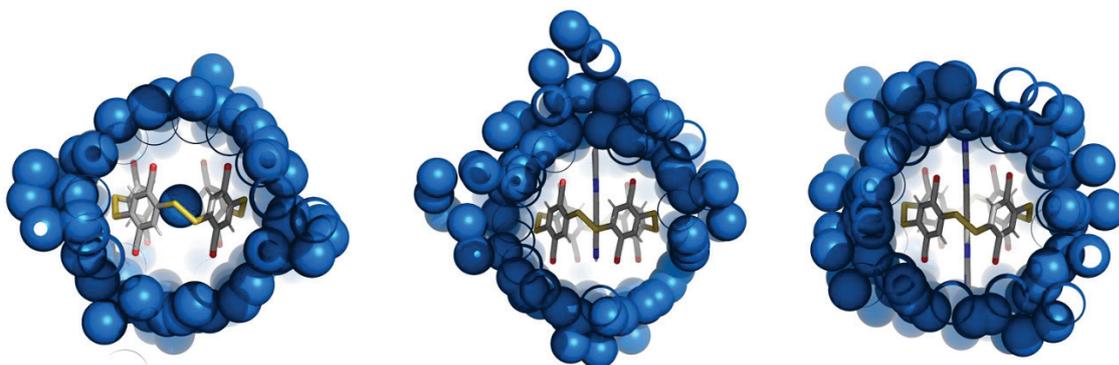


Figure 10 : Hot spots (blue) of bound water molecules thank to the SPAM analysis. Left: receptor alone, middle: complex with spermidine I-14, right: complex with spermine I-8.

The local free energy profile is computed in a specific position if there is a water molecule in this position at each frame of the simulations. The profiles are obtained by the calculation of the energetic distribution between water molecules and their environment thanks to a force- and energy electrostatic term assumed for the bulk. The profiles were expected to give information on the solvation of the complexes. All the bound water molecules contributed positively and similarly to the solvation energy of the complexes and empty host (Table 2). It was impossible to discriminate the complex with spermine I-8 from the complex with spermidine I-14 regarding the hot spots at the lower rims of the complexes in order to highlight a difference of solvation by the SPAM package.

Local energetic parameters	dG (kcal.mol ⁻¹)	dH (kcal.mol ⁻¹)	-TdS (kcal.mol ⁻¹)
Guest	-11.17 ; -0.66	-2.28 ; -1.64	-8.985 ; 1.17
Spermidine I-14	-9.55 ; -1.64	-2.43 ; -1.72	-7.26 ; 0.14
Spermine I-8	-16.87 ; -1.38	-3.68 ; 14.37	-20.83 ; 0.53

Table 2 : Range of the energetic profiles of bound water molecules relative to the bulk around host I-14 and its complexes with spermine I-8 and spermidine I-14, by SPAM package.

Nevertheless, SPAM provided information on the hot spot found inside the cavity of the host alone. The hot spot had an unexpected profile (relatively to the bulk): dG: -6.5 kcal.mol⁻¹; dH: 14.4 kcal.mol⁻¹, -TdS: -20.8 kcal.mol⁻¹. The favourable entropy -TdS of a water molecule inside the hydrophobic environment of the cavity relatively to the bulk was counterintuitive. As a matter of fact, the water molecules are expected to be more structured in that kind of environment and become “frustrated”. Similar situations were found in the literature. In particular, the water filling carbon nanotubes gained entropy compared to the bulk¹⁰¹. The authors of SPAM explained that, in that case of favourable solvent entropic term relative to the bulk-TdS, the water is trapped in a homogeneous hydrophobic environment where no orientation preference was imposed, explaining the gain in entropy for our hot spot compared to the bulk. The enthalpic contribution dH revealed the water molecule was enthalpically disfavoured in the cavity. That was important to understand the driving forces of the association between our polycarboxylate disulfide bridge macrocycles with polyamines. That referred to the definition of non-classical hydrophobic effect.

SPAM is a statistical-mechanics based approach, considering the distribution of interaction energies between water molecules and its environment at a given site, and the perturbation to this distribution

upon binding without water–water correlation. Each hot spot is equivalent to a single water molecule with a distribution function specific to that site but independent of others. At the opposite, GRID is based on the inhomogeneous solvation theory¹⁰². In this theory, the water is described in terms of the translational and rotational freedom degrees. Then, the excess entropy of bound water due to the translational and rotational restriction exerted by the local environment are extracted. The solute generates an external field that creates fluctuations around. That formalism led to contrasted entropy compared to SPAM calculation because of the hypothesis of a uniform translational and orientational distribution for water in the bulk. For instance, the entropy of solvation of the density inside the empty host was not described as more favourable compared to the bulk but slightly more unfavourable (Figure 11, left, red). GIST was computed on the empty host and the two complexes. GIST provides a lot of information, especially the enthalpic solute/solute and water/solute interactions as well as the entropic orientational and translational contributions on a 3D grid divided in small voxels. Additions thanks to the GISTPP software^{7b} gave the overall map of enthalpy and entropy of solvation (Figure 11).

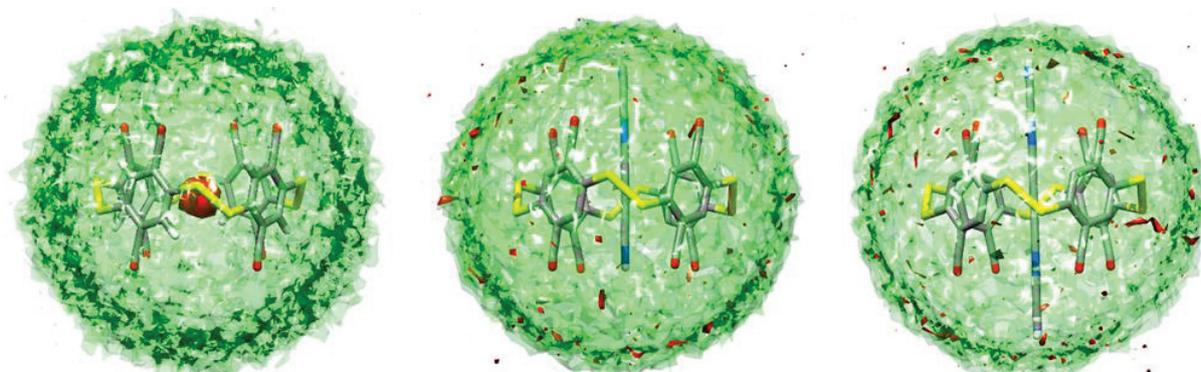


Figure 11: Maps of the enthalpy and entropy of solvation obtained by GIST. Green, $E_{tot} = -0.00862 \text{ kcal/mol.A}^3$, Red, $-TdS_{tot} = 0.000585 \text{ kcal/mol.A}^3$

Integration of the voxel values over the whole system and combination of the thermodynamic parameters with the GIST analysis on guests alone enabled an “end of state analysis” i.e. a calculation of a free energy of solvation for the complexes. The solvation free energy of spermine **I-8** was found more favourable by $-256 \text{ kcal.mol}^{-1}$. That data was not helpful in order to solve the scenario of the association between **I-14** and polyamines. Given the qualitative results with GIST, it is still not clear if the expulsion of central water density during the binding with polyamines referred to a classical or non-classical hydrophobic effect.

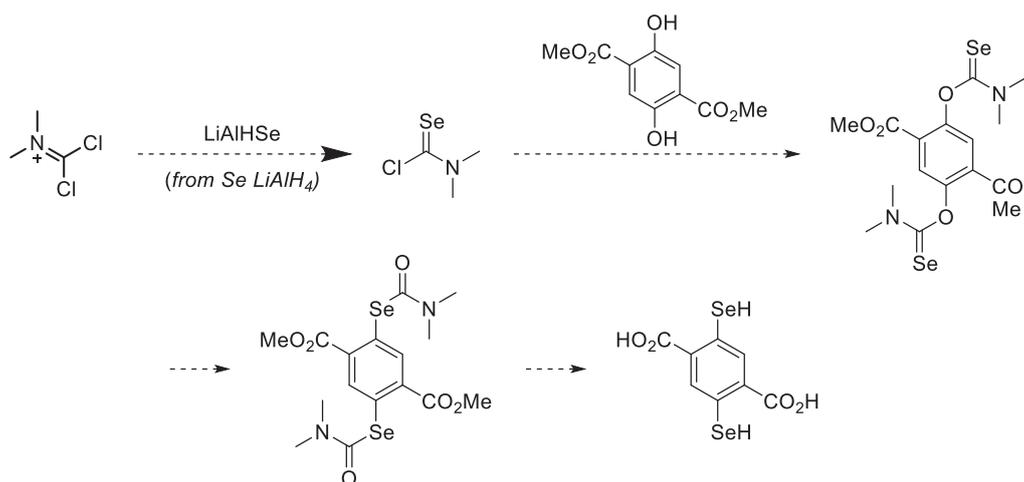
The use of these two different tools should be prudent given that their theoretical basis is different. They could not complement each other but at least give qualitative information. A major improvement will consist in the use of thermodynamic integration with the FEW utility⁹⁶. Compared to MMPBSA, the thermodynamic values are calculated thanks to explicit solvent and includes entropic contributions. That can overcome the limitations and the incoherence with experimental data of MMPBSA. It could also be interesting to estimate the residential time of the water molecules inside the hydrophobic cavity.

It is not clear if the water density inside the cavity led to classical or non-classical hydrophobic effects. The two tools GIST and SPAM did not help to discriminate the two complexes on their enthalpy of solvation, which remained appreciated thanks to the grid of water presence.

d. Conclusion

The polycarboxylate disulfide bridge cyclophane **I-14** showed high affinity for polyammoniums in solutions, with K_a up to 10^{11} M^{-1} . The selectivity of the complexation for a series of α,ω -alkyl diammoniums was governed by the solvation of the complexes and by the desolvation of the guests. Regarding the binding of **I-14** with spermine **I-8** and spermidine **I-14**, the solvation effects are also discriminant in the selectivity for spermidine **I-14**, even though the exact picture is not yet depicted. Spermidine is less charged but was more mobile in the complex, which contributed to the selectivity. The investigation highlighted the solvophilic effects in the binding process, which is more and more considered in drug discovery¹⁰³. The unprecedented affinity of the receptor for spermine **I-8** led to biological applications *in cellulo*, currently under investigation. The formation of heterocyclo-oligomers that changes the chemical functions of **I-14**, for both fundamental and applied studies encountered synthetic issues-during the assembly of the corresponding 1,4-bisthiophenols.

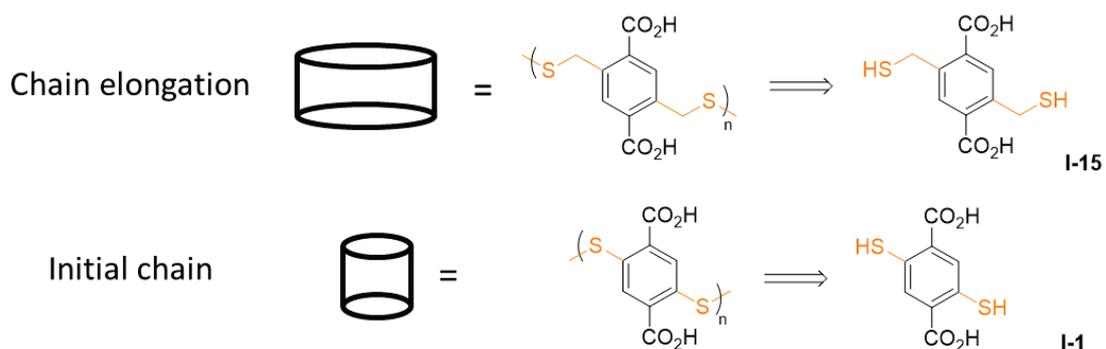
Another modification involving the replacement of disulfide bond into diselenide bonds could find applications in bioimaging. The diselenide bond is also reversible and can be formed thanks to irradiation and used in DCC^{39e,104}. The synthesis can be done in the future (Scheme 30). The synthesis is similar with **I-14**. The steps are described (selenocarbamoyl chloride formation¹⁰⁵, O-selenocarbamate formation¹⁰⁶, NKR rearrangement¹⁰⁷ and deprotection¹⁰⁷) for similar compounds until the building blocks. A self-assembly could then lead to a polycarboxylate diselenide bridge cyclophane.



Scheme 30 : Hypothetic synthetic plan to a diselenol building block

4. Self-assembly of 2,5-bis(mercaptomethyl)terephthalic acid **I-15** by iodine or dioxygen oxidation of thiols into disulfide macrocycles

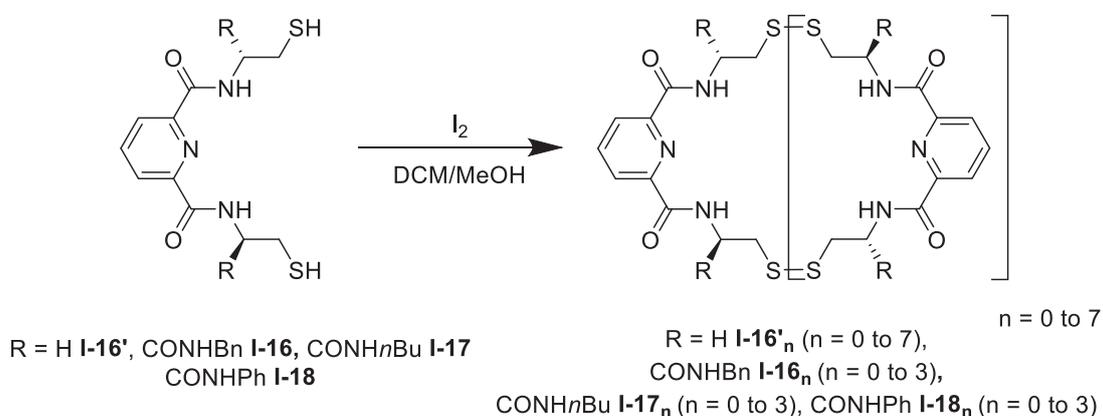
The aim of this subchapter was to lengthen the bridges between the units of polycarboxylate disulfide oligomers, expanding the diameter of the cavity of **I-14** for instance. The macrocycles will be obtained by the self-assembly through DCC of 2,5-bis(mercaptomethyl)terephthalic acid **I-15**, with an additional carbon between the aromatic core and the thiol compared to **I-1** (Scheme 31). Two different oxidants will be used, oxygen, and iodine and their influence on the self-assembly of **I-15** will be studied.



Scheme 31 : Desired chain modification of polycarboxylate disulfide cyclophanes obtained through self-assembly of polythiols. Bottom: initial situation, top, chain elongation with a building block I-15 incorporating an additional carbon between the aromatic core and the thiol function compared to I-1.

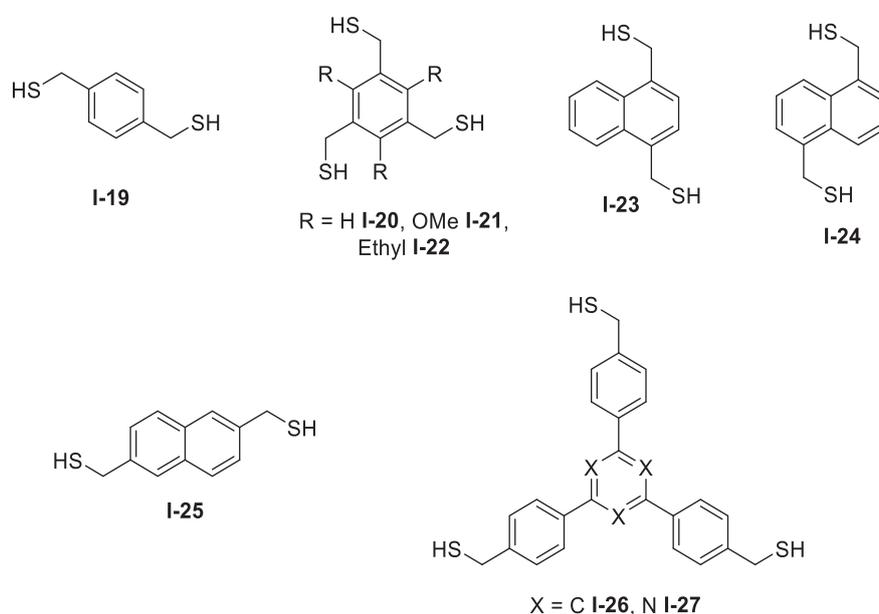
b. Additive-directed self-assembly of polythiols and kinetic trap

This bibliographic subpart presents the synthesis of disulfide-based cyclophanes using iodine as an oxidizing agent and pnictogen additives, nucleating the self-assembling of the polythiols. This synthetic methodology was the basis of the following experimental work. Iodine is known to be an efficient oxidant for the disulfide bond formation from thiols¹⁰⁸. The reaction is fast and proceeds through the formation of a sulfenyl iodide group as an intermediate¹⁰⁹. A catalytic version was proposed with oxygen peroxide through the same mechanism for the obtention of disulfide bond from thiols with similar rate¹¹⁰. Nevertheless, sulfenic or sulfinic acids can also be formed under these condition depending on the thiol structures, which is the main drawback of the method¹¹¹. In particular, the presence of carboxyl groups at the ortho position of the aromatic thiol ease the overoxidation¹¹¹. The iodine oxidation of thiols can be applied to the dynamic covalent chemistry in order to form the disulfide bonds faster than the thiol-disulfide equilibration, or to obtain kinetically controlled DCLs (Scheme 32). For instance, a family of receptors I-16-18_n of carboxylates based on bis-thiol dipicolinic amides was obtained by oxidation of thiols to disulfides by iodine in a mixture of DCM and MeOH¹¹². The proportion between the different cyclo-oligomers I-16-18_n could be controlled following the experimental procedure. Indeed, an addition of iodine into a concentrated solution of I-16-18 was in favour of larger oligomers (n = 3 – 7), while an addition of the building block into an iodine solution gave small oligomers (n = 0, 1, 2). The distribution could then be controlled thermodynamically by virtue of the disulfide bond exchange induced by a catalytic amount of thiolate coming from basic conditions¹¹³.



Scheme 32: Oligomerisation obtained by oxidation of thiols I-16-18 with iodine into disulfide bonds I-16-18_n.

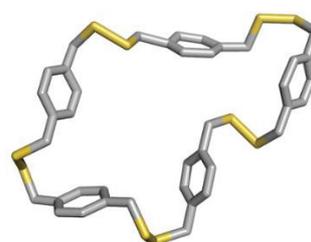
Recently, a pnictogen and metal-assisted self-assembly of benzylic bis/trithiols with iodine oxidation of thiols efficiently promote the isolation of a family of disulfide cyclophanes (Scheme 33)¹¹⁴. This method led to a variety of large organic cages and to highly strained cyclophanes, which was remarkable and difficult to obtain through conventional synthesis (Scheme 34). In fact, the yields after isolation were very decent. The small oligomers as **I-19**₃ could be easily obtained with a yield around 30 %. The dimer **I-19**₂ could be isolated with a yield up to nearly 70 %. The larger oligomers like **I-19**₅ had lower yields, above 10 %. The oxidation of thiols with iodine took place in an organic media and in presence of pnictogen additives such as AsCl₃ and SbCl₃^{114a,b,d}. The additives prevented the formation of undesired polymers by nucleating the cyclo-oligomers by cation- π interactions¹¹⁵. These toxic salts could be replaced only by Cu(II) salts¹¹⁶. Many other additives such as Zn(II) or Ca(II) salts were unable to promote the formation cyclophanes^{114d}. In another article, the dimer of benzylic thiols was obtained similarly without the use of any salt^{108b}. Other chemical oxidants such as NBS and *m*-CPBA mainly produced polymers and insoluble species or oxidised the thiols beyond the disulfides^{114d}. The minimisation of the distortion of the resulting disulfide bonds influenced the macrocyclic distribution during the self-assembly of these polythiols.



Scheme 33: Example of building blocks used for the self-assembly by iodine, directed by Pnictogen additives

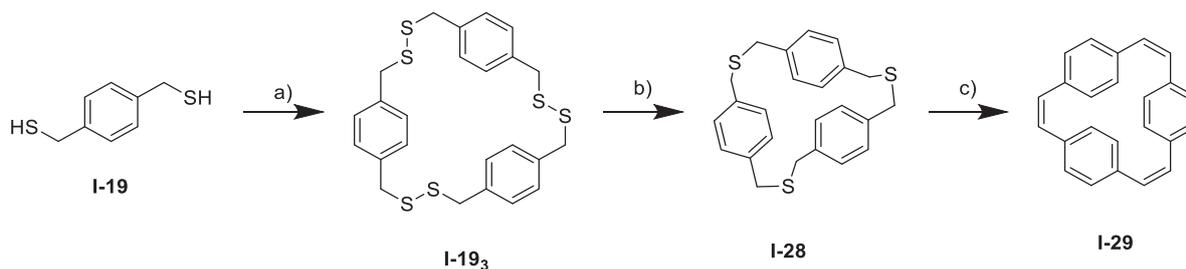


I-19₃, AsCl₃, I₂, toluene,
DCM from **I-19**, 30 %

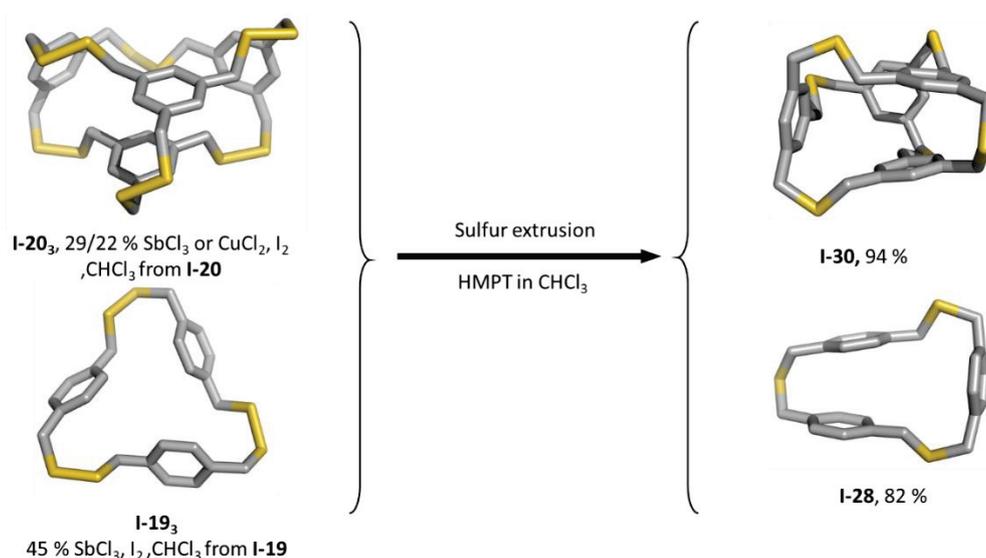


I-19₅, SbCl₃ (6 %)
or CuCl₂ (4 %), I₂, CHCl₃ from **I-19**

Scheme 34: Example of self-assembled cyclophane by iodine, directed by pnictogen additives. Structures obtained thanks to the CCDC database¹¹⁷ and Chimera software¹¹⁸.



Scheme 35: Examples of a sequence including self-assembly with iodine and pnictogen additives, sulfur extrusion and desulfurization a) I_2 , $SbCl_3$, $CHCl_3$, r.t., 45 %; b) HMPT, $CHCl_3$, r.t., 82 %; c) MeI, base, r.t., no yield provided^{114b}



Scheme 36 : Examples of sulfur extrusion from self-assembled cyclophanes obtained by iodine and pnictogen additives. Structures obtained thanks to the CCDC database¹¹⁷ and Chimera software¹¹⁸.

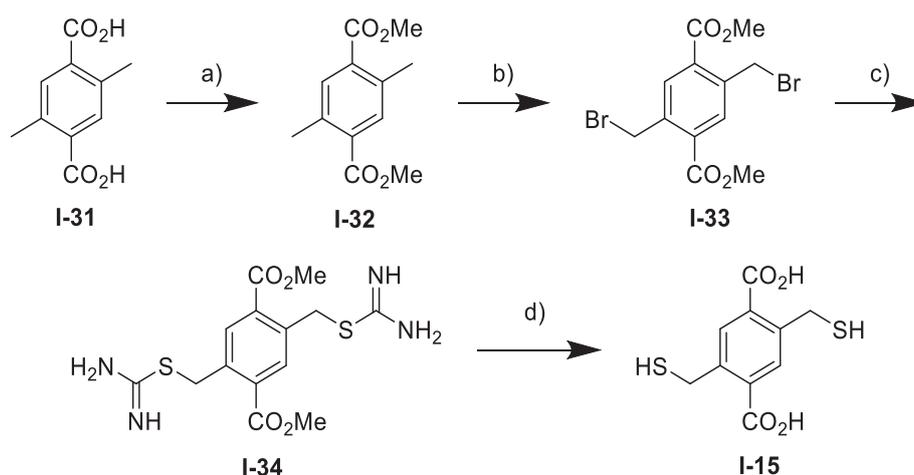
Furthermore, these disulfide cyclophanes could be covalently trapped by a non-reversible sulfur extrusion process^{114b,c,116}. The reaction occurred at room temperature, giving thioether cyclophanes with high yields thanks to the desulfurizing agent HMPT (Scheme 36). The phosphoramidite removed a sulfur atom from the disulfide bond and therefore decreased the strain in the cyclophane. Indeed, the thioether linkages had more degrees of freedom compared to the oriented disulfide bond which adopts an ideal angle of 90° in order to minimise the repulsion between the sulfur electronic lone pairs¹¹⁹. The method was illustrated with the self-assembly of **I-19** and its sulfur extrusion (**I-28**). Besides, the thioether linkages could be further desulfurized (**I-29**) thanks to methylating agent and a base, or a sequence with butyl lithium and Borch's reagent^{114b}.

A variety of disulfide macrocycles could be obtained by oxidation with iodine from benzylic thiols. These cyclophanes could then undergo a desulfurization process.

c. Synthesis of benzylic building blocks

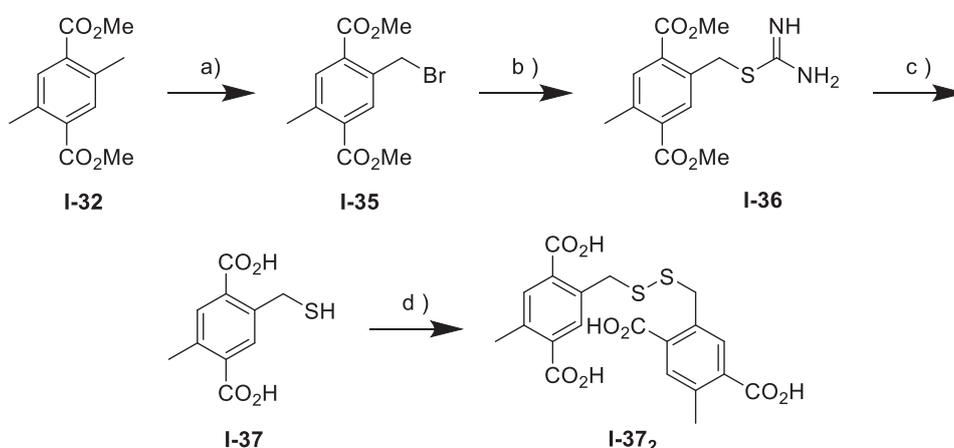
This subpart presents the synthesis of the 2,5-bis(mercaptomethyl)terephthalic acid **I-15**, which is the elongated version of 1,4-bisthiophenol **I-1**. A 4-step synthesis starting from commercial **I-31** led to the

I-35 analogue of **I-1** (Scheme 37). Firstly, an esterification¹²⁰ enabled to protect the carboxylic acids into methyl esters in **I-32** with a satisfying yield of 99 %. Then, a Wohl-Ziegler radicalar bromination¹²¹ gave **I-33** with a yield of 87 %, which was coherent with the reported 85 %. This reaction usually involves enriched aromatics¹²² and hence was herein slow. A nucleophilic substitution with thiourea led to **I-34** quantitatively¹²³. **I-33** should be introduced slowly in the presence of thiourea, because a nucleophilic substitution between the intermediate and thiourea took place. Finally, the basic hydrolysis of the thiuronium salt¹²⁴ **I-34** gave the desired product **I-15** with a moderate yield of 57 %. The 4-step synthesis was achieved with a satisfying yield of 48 % at a gram scale. Furthermore, an analogue **I-37** of **I-15** and its linear dimer **I-37₂** were synthesised through the same procedure in order to explore the optimal conditions of sulfur extrusion (Scheme 38).



Scheme 37: Synthesis of **I-35** a) SOCl_2 , MeOH, reflux, 16 h., 99 %; b) NBS, AIBN, DCM, reflux, 16 h, 86 %; c) thiourea, MeOH, r.t., 16h., 99 %; d) KOH, MeOH/H₂O, reflux, 2 h., 57 %

The bromination¹²¹ of **I-32** gave **I-35** with an excellent yield of 92 %. The thiourea moiety was then successfully introduced in **I-36** thanks to a nucleophilic substitution with thiourea¹²³ with a moderate yield of 68 %. Finally, **I-37** was obtained in basic conditions¹²⁴. The global yield is 39 % in 4 steps and was comparable to the synthesis of **I-15**. Then, oxidation of thiols with iodine in methanol¹¹² gave the disulfide dimer **I-37₂** with a satisfying yield of 70 %.



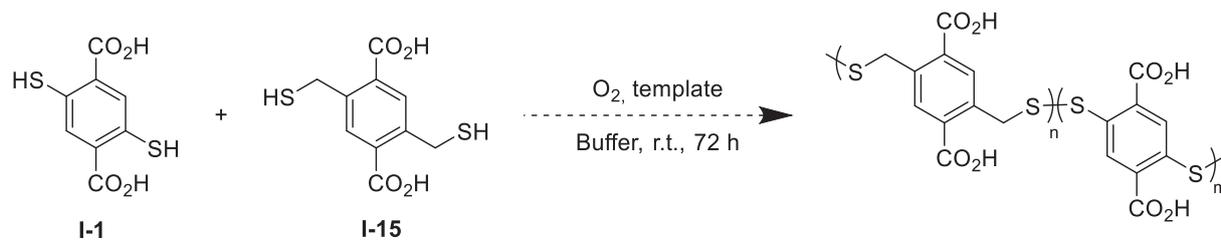
Scheme 38: Synthesis of **I-37₂** a) NBS, AIBN, DCM, reflux, 16 h, 92 %; b) thiourea, MeOH, r.t., 16 h., 68 %; c) KOH, MeOH/H₂O, reflux, 2 h., 67 %; d) I₂, MeOH, 5 mn, 70 %

The polycarboxylate disulfide-bridge cyclo-oligomers are characterised by a relative stability in aqueous media in the absence of template. Their disulfide bonds can be cleaved or oxidised even in mild conditions¹²⁵. Their transformation into a thioether bond by sulfur extrusion was envisioned to our family of polycarboxylate disulfide cyclophanes. This sulfur extrusion was studied on this dimer by another PhD student^{82d}, with contrasted results in the conditions tested. The literature reports that the sulfur extrusion could be driven by a strain release^{32b,114b}, which was not the case from our unconstrained linear substrate **I-37**₂. A cyclic model should be used to find optimal conditions.

Two new building blocks were synthesised. They were elongated analogues of I-1, bearing two carboxylates and benzylic thiols instead of thiophenols.

d. Self-assembly by aerobic oxidation

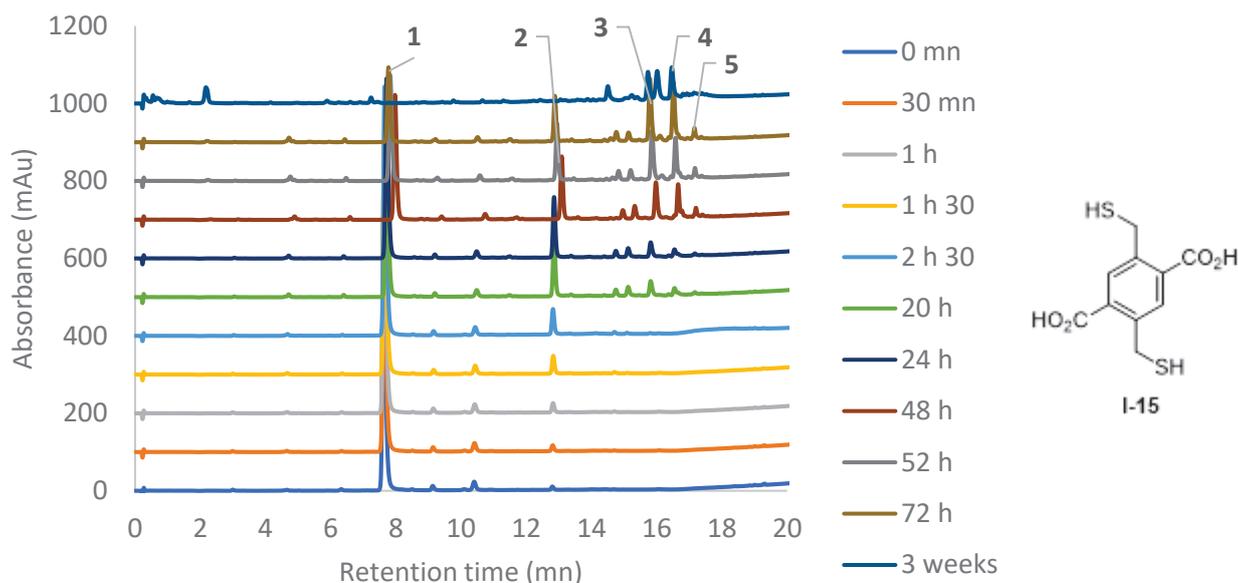
The aim of this subpart was to study the self-assembly of **I-15** with or without **I-1** in aerobic conditions, in order to form disulfide hetero- and homo-oligomers. The building block **I-15** was dissolved between 2-10 mM in aqueous buffers (ammonium acetate or Tris 200 mM, pH 7.4-10). The mixtures were left open, stirred and analysed by UPLC and LC/MS. The details of the conditions are given in appendix (Appendix I-II: Additional conditions of reactions).



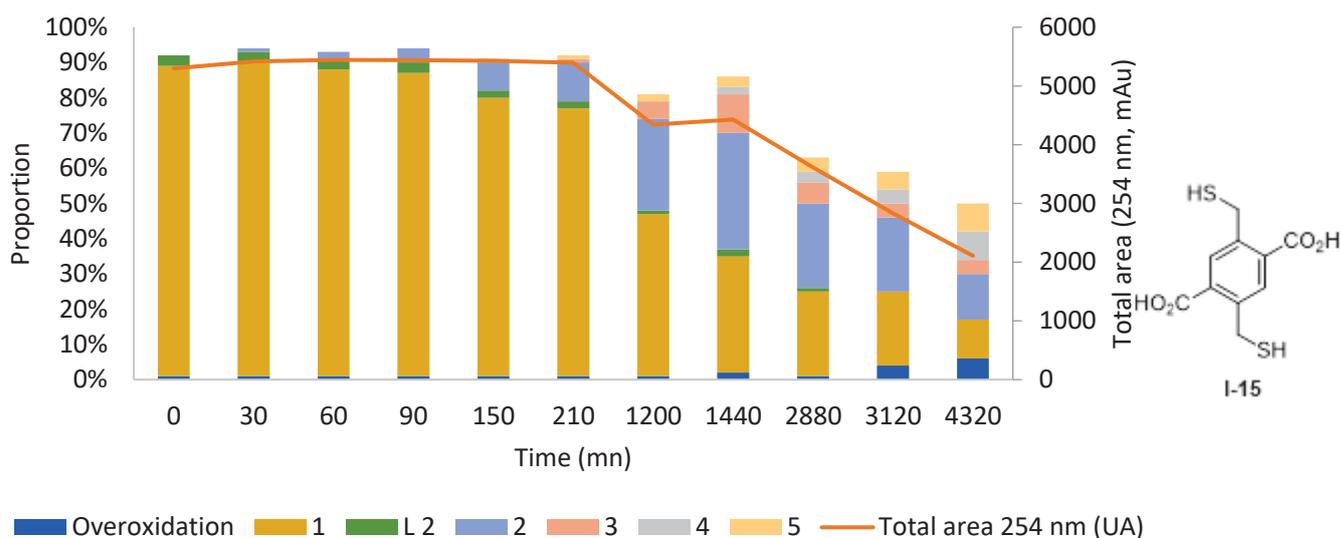
Scheme 39: Desired formation of disulfide bonds during DCLs between I-1 and I-15 with templates and oxygen as oxidants

The total area usually was a metric of the quantity of matter of polythiols and their corresponding oligomers in our DCLs, because of the additive Beer-Lambert law, linking the intensity of the absorbance with the quantity of matter of each compounds following their molar extinction coefficient λ^{126} . In a DCL composed of one building block, we applied the hypothesis that each compound had the same λ given that their chemical moieties involved in the absorbance were not changed by the oligomerisation. In the case of a DCL involving **I-1** and **I-15**, the concentrations of the macrocyclic products formed were obtained from their corresponding HPLC-UV peak areas after determining the effective absorption coefficients of the two building blocks.

Firstly, the dissolution of **I-15** at 4 mM in Tris buffer (200 mM, pH 7.4) led to a variety of disulfide cyclophanes, including up to 10 monomeric units after 72 hours with different diastereoisomers of the same sizes (Scheme 40, Scheme 41). The kinetic of disulfide bonds formation was slow compared to **I-1** in the same conditions^{16,40a}. In fact, full consumption of the starting material required one week, compared to one day for **I-1**. During the first hours, a cyclic dimer was formed, followed by a cyclic trimer and then by higher sizes of cyclophanes.



Scheme 40 : Evolution of **I-15** (4 mM) in Tris (200 mM, pH 7.4), method B. n = disulfide macrocycle containing n units, 1 = starting building block

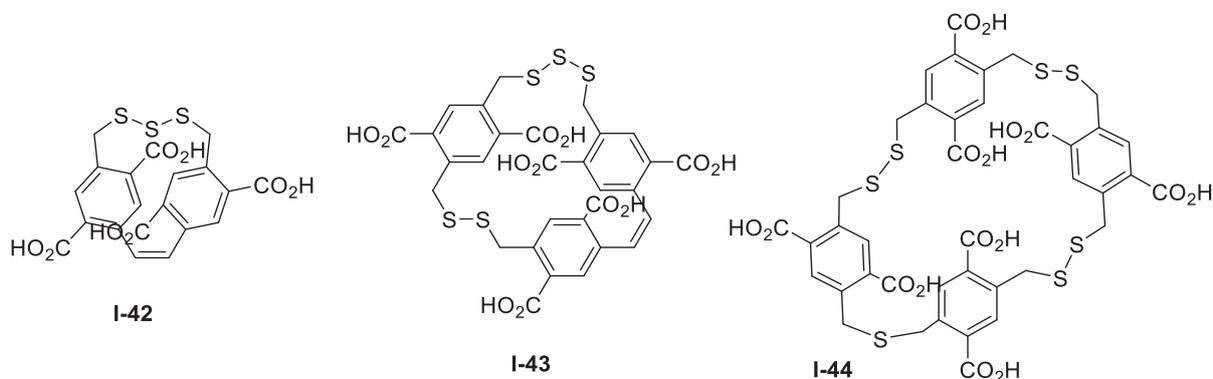


Scheme 41: Building block and macrocycle evolution of **I-15** (4 mM) in Tris (200 mM, pH 7.4), method B. n = disulfide macrocycle containing n units, 1 starting building block, L linear.

These species were stable in the media and are not degraded after one week. Nevertheless, a precipitate and a progressive loss of UV signal in HPLC was progressively observed over 72 hours. This might be attributed to the formation of linear polymers which was reported for similar benzylic bithiols oxidised with oxygen^{114a}.

In addition, other minor constituents of the DCL were identified. Sulfones, sulfenic and sulfinic acids were detected. The overoxidation of thiols occurred on the building blocks but also on small linear oligomers. Furthermore, unexpected small species bearing trisulfide, thioether and alkene bonds were also identified (Scheme 42). They could be detected only in UPLC because of their very low amounts. The sulfur extrusion has never been reported as a spontaneous process and required reagent such as

HMPT with similar bithiols^{114b,c,116}. Cyclic oligomers suffering from oxidation of their disulfide bridges were observed. This overoxidation might facilitate the sulfur extrusion by a SO₂ release, as described in the literature¹²⁷. This DCL illustrated the accessible molecular diversity from a simple building block by DCC.

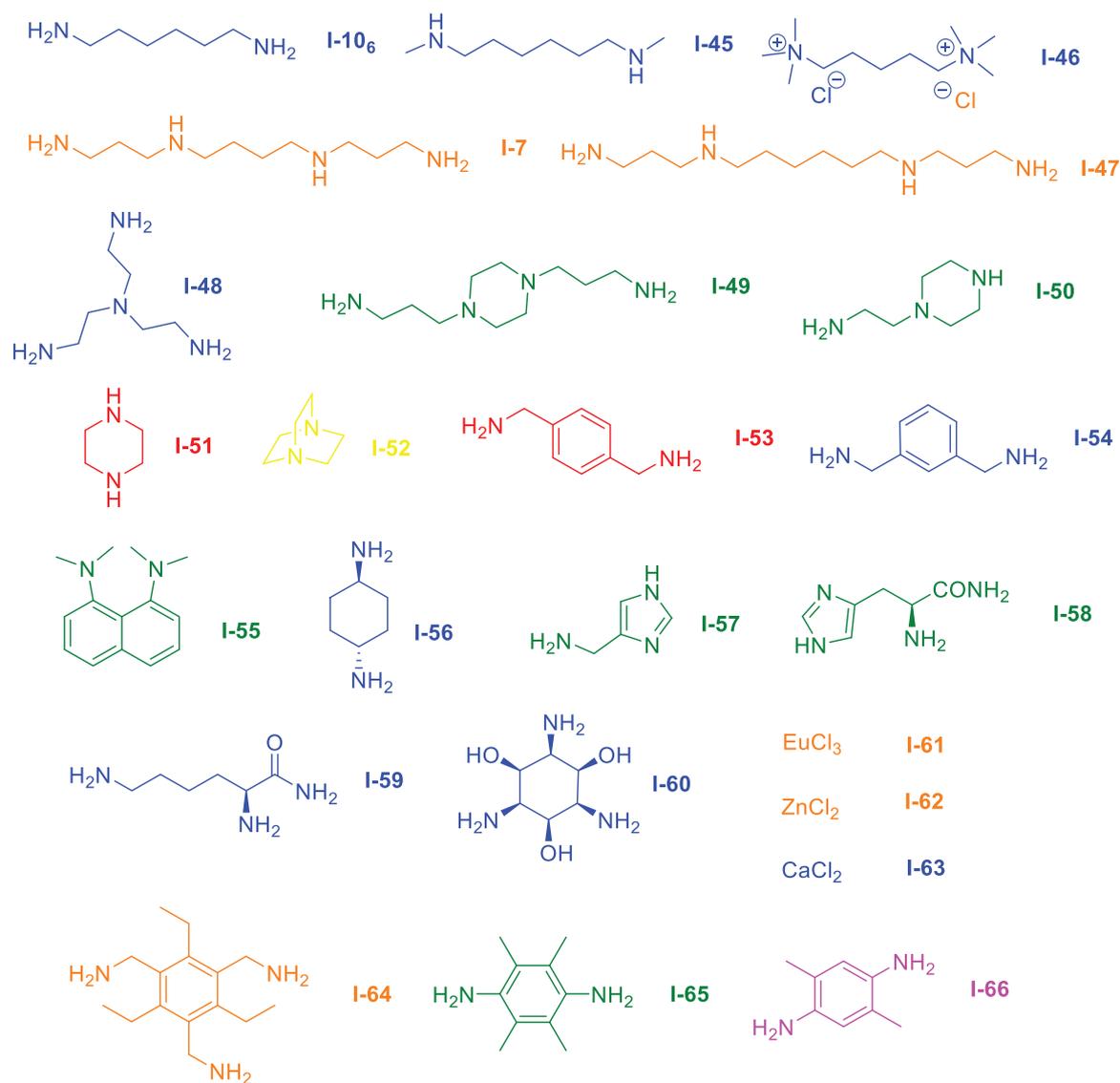


Scheme 42: Examples of trisulfide I-42, thioether I-44, and alkene bonds I-42 found in small oligomers of I-15 during the experiment of DCC at 4 mM in Tris 200 mM pH 7.4, detected by LC/MS

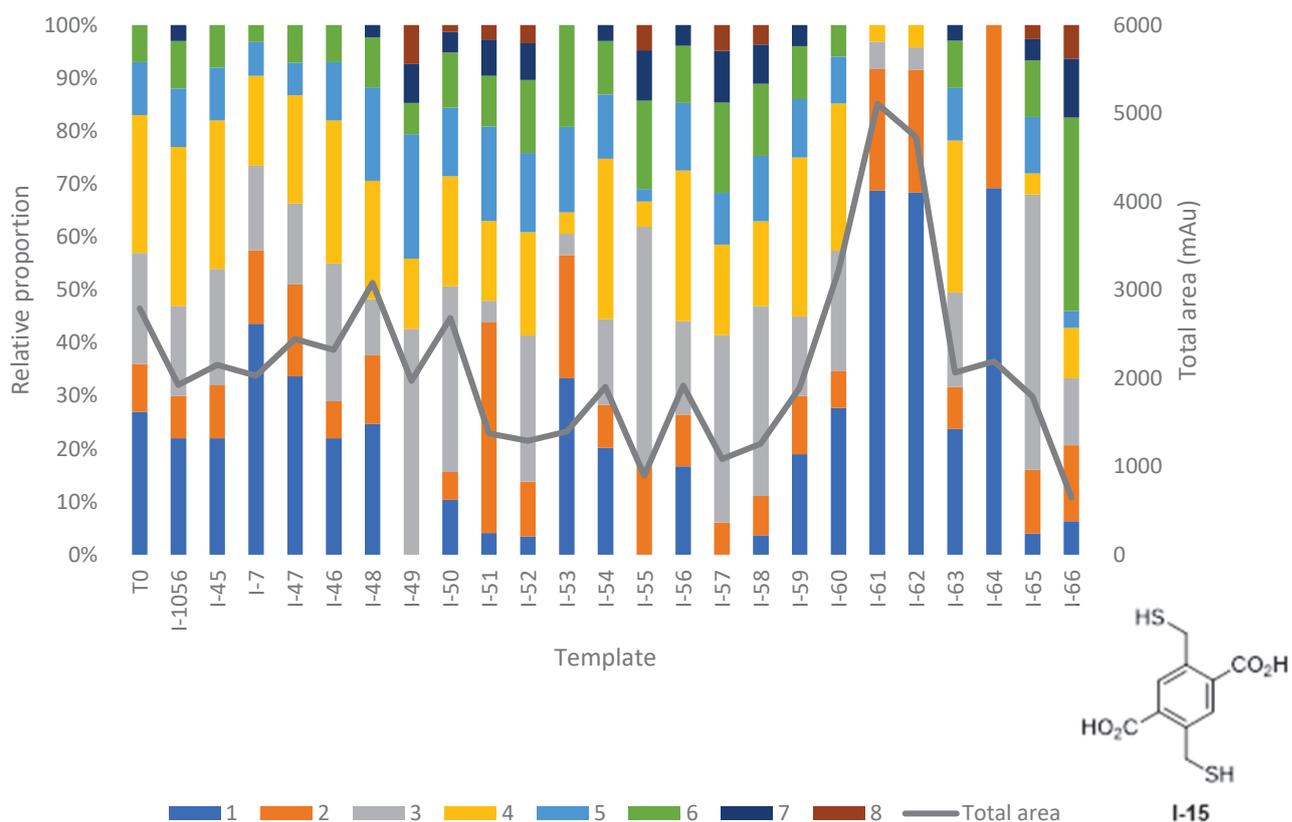
Templated libraries were generated by the dissolution of 0.25 equivalents (1 mM) of I-15 at 4 mM in Tris buffer 200 mM pH 7.4. Polyamines were screened as well as some metallic salts (Scheme 43). The results are summarised in the next figure (Scheme 44). The analysis was done after one week of stirring, in order to observe the influence of the templates on the conversion of I-15 as well as the amplification of macrocyclic products.

The kinetic of the templated DCLs was similar to the DCL without template. The details are available in appendix (Appendix I-V: Additional data of the DCLs for 1-15). A variety of relative proportions between macrocycles were observed. Some templates gave a similar distribution than the DCL without templating: I-105₆, I-45, I-46, I-48, I-54, I-56, I-59, I-60 and I-63. Other templates decreased slightly the kinetic of consumption of the starting building block I-15 without change in the distribution of macrocycles: I-7 I-47, I-53 and especially I-61, I-62 and I-64. The last three led to an impressive decrease of the consumption of I-15 of only 30 % after one week. Some templates amplified a cyclic trimer: I-49, I-50, I-55, I-57, I-58 and especially I-65. Some templates amplified a cyclic dimer: I-51 and I-53. I-52 increase the kinetic of consumption of the starting building block I-15 without change in the distribution. I-66 amplified a cyclic hexamer. An important number of templates led to a loss of UV signal in HPLC, which might be due to a polymerisation.

These amplifications were difficult to rationalise. The metallic salts could prevent from an oligomerisation by a binding with I-15. Regarding the templates leading to no amplification, there were all cationic in aqueous solutions. It was highly probable that they interacted with I-15 and its oligomers by electrostatic interactions or salt bridges without discriminant stabilising interactions. The amplification of a cyclic dimer might result from a stabilisation of the dimer not by an inclusion complex given that the dimer was too small to fit a guest. Regarding I-66, a cyclic hexamer might be particularly adapted to accommodate one or several template molecules. Similarly, the amplification of cyclic tetramer might result from a particularly stable inclusion complex with the templates. Nevertheless, there was no proofs.



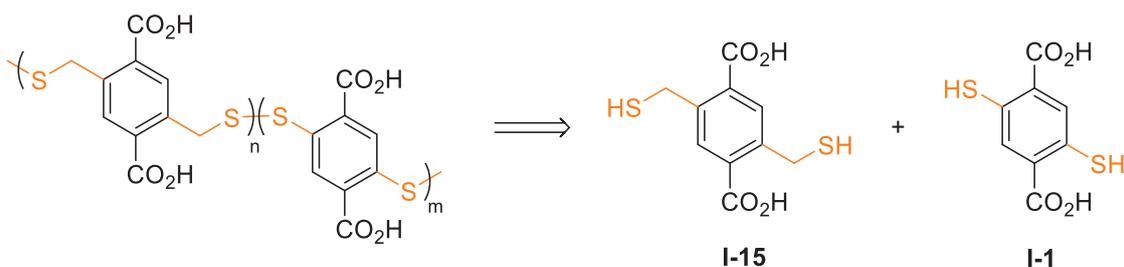
Scheme 43: Templates used for DCLs of I-15. Blue: no amplification of macrocyclic products. Orange: decrease of the rate of I-15 consumption without change in the distribution. Green: amplification of a cyclic trimer. Red amplification of a cyclic dimer. Yellow, increase of the rate of I-15 consumption without change in the distribution. Purple: amplification of a cyclic hexamer



Scheme 44: Evolution of the macrocycle proportions by templating the DCLs of **I-15** (4 mM, Tris 200 mM, pH 7.4, template at 0.25 eq), after one week. 1: building block, *n* cyclic oligomer with *n* units.

The DCLs were analysed at one week. However, it was clear that there was still **I-15** in a majority of the DCLs. Therefore, the analysis should be done later, around two weeks in order to better appreciate the amplification.

The influence of the building block concentration was studied in the case of a DCL formed from **I-15** templated by **I-7**. The kinetics of **I-15** disappearance was not increased by higher concentration of **I-15** from 2 to 10 mM (Appendix I-V: Additional data of the DCLs for **1-15**). Regarding the influence of the pH, **I-15** was slightly quicker consumed at pH 9 and 10 rather than at pH 7.4 in a DCL with **I-7** (see appendix). This was expected as the disulfide bond formation is catalysed in basic conditions¹²⁸. Another buffer, ammonium acetate (200 mM, pH 7.4) decreased the kinetic of **I-15** consumption. The linear dimers and trimers could be observed, and the UV intensity also decreased during the DCL. Thus, this buffer did not help to prevent from the supposed polymerisation. Finally, the dissolution of **I-15** at 2 mM with **I-1** at 2 mM in aqueous buffers was explored in order to create hetero-cyclophanes (Scheme 45).



Scheme 45 : Self-assembly of **I-15** and **I-1** through DCC by disulfide bond formation in order to obtain heterocyclo-oligomers

In both Tris and ammonium acetate 200 mM pH 7.4, without template, the HPLC peak areas of both building blocks vanished in 24 hours without an apparition of other signals (Appendix I-V: Additional data of the DCLs for **I-15**, Figure 17). The same conditions with templating by 1 mM of spermine **I-8** avoided the quick consumption of **I-15**. After 2 days, the only macrocyclic products were homocyclo-oligomers of **I-15** (**I-15₃**, **I-15₄** and **I-15₅**). We hypothesised that these homocyclo-oligomers might form stable inclusion complex with spermine **I-8**. It was unexpected, given that **I-1₄** had a high affinity for spermine **I-8**. A control of the DCL by ¹H NMR might show if homo-oligomers of **I-1** are formed during the DCL.

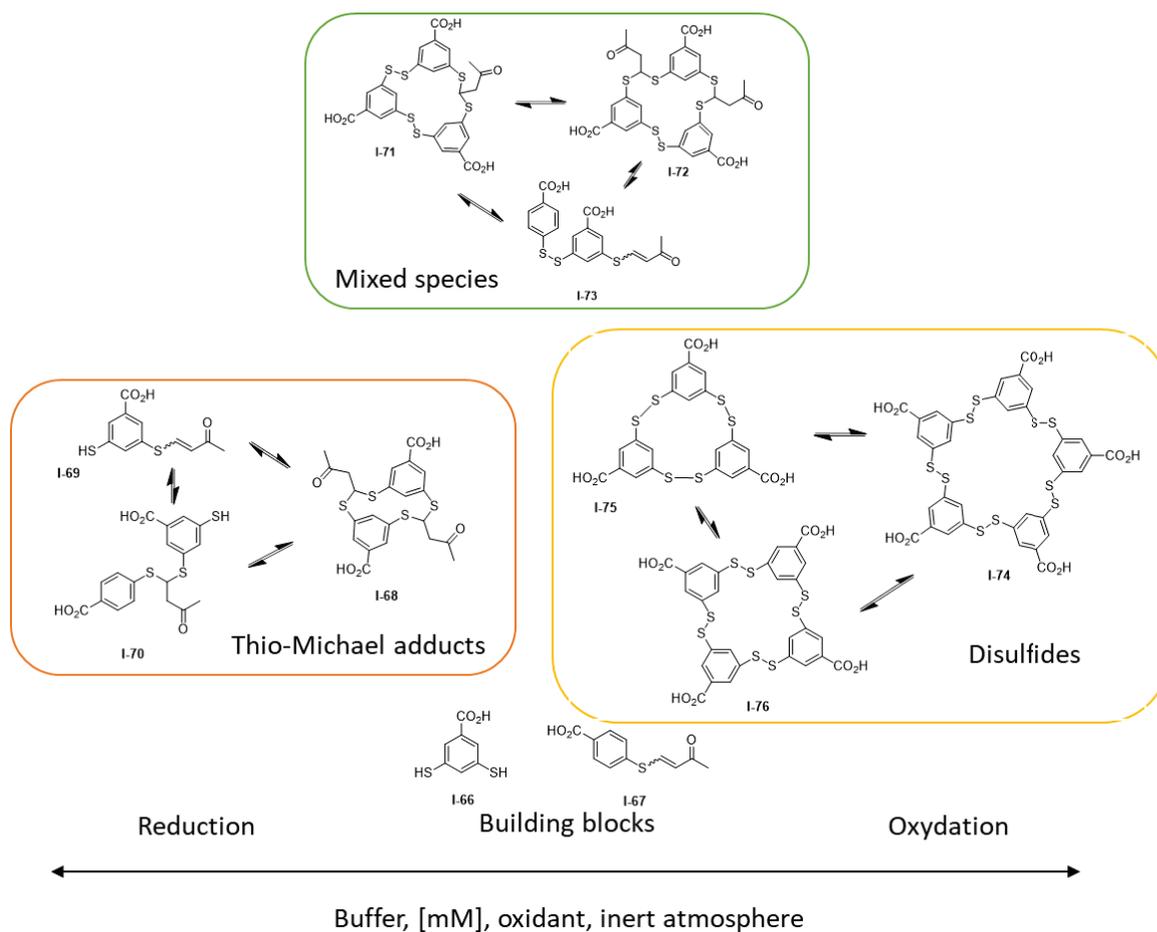
*The dissolution of **I-15** at 4 mM in Tris buffer 200 mM pH 7.4 led a variety of homocyclo-oligomers. The presence of templates at 1 mM led to important variation on the distribution of macrocyclic products. A decrease of the total area under peaks was probably due to a competition of the oligomerisation with polymerisation. The formation of heterocyclo-oligomers between **I-15** and **I-1** was not obtained.*

e. Switch between thio-Michael addition and disulfide exchange

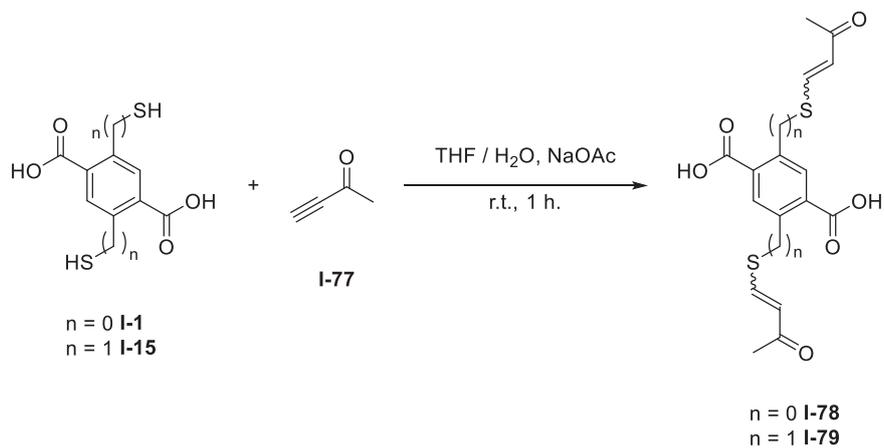
The aim of this subpart was to turn from oxidative coupling between polythiols to coupling between dinucleophiles and dielectrophiles, using thio-Michael addition and disulfide exchange in order to obtain alternated hetero-oligomers.

Antiparallel chemistry consisted of the channelling of one functional group into one of two reversible reactions following a controlled parameter¹²⁹. The disulfide exchange and thio-Michael addition were the chosen reversible reaction while the oxidant quantity was the parameter enabling the reversible switch between the products of the two reactions. The thiol was the common functional group. For instance, the reported system (Scheme 46) can be reversibly converted thio-Michael adducts, mixed species and disulfides¹²⁹ by means of oxidation and reduction.

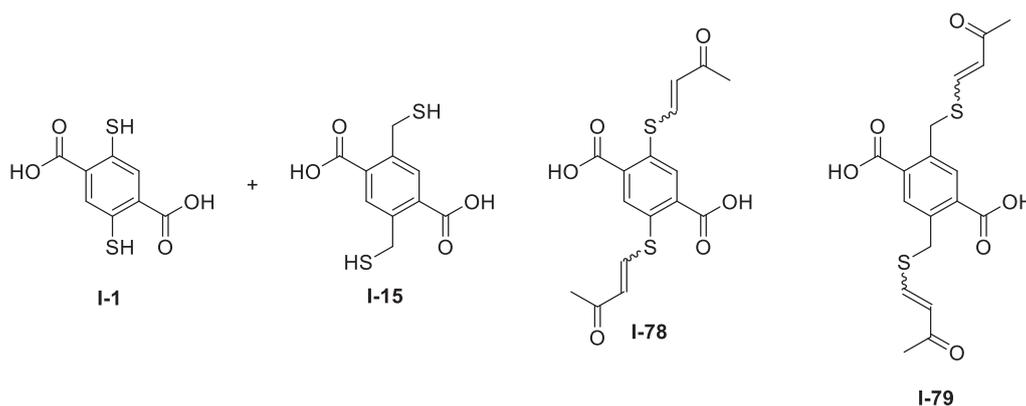
Inspired by the previous examples, the method was envisioned for **I-1** and **I-15**. Thio-Michael adducts **I-78** and **I-79** were obtained by a nucleophilic attack of thiols on but-3-yn-2-one **I-77**¹²⁹⁻¹³⁰ with satisfying yields of 76 % for **I-1** (**I-78**) and 40 % for **I-15** (**I-79**) (Scheme 47). Then, the thio-Michael adducts were combined with the bistiols (Scheme 48, Table 3).



Scheme 46: Accessible species from two building blocks by disulfide exchange and thio-Michael addition upon controlled quantity of sodium borate



Scheme 47: Synthesis of vinyl sulfides. $n = 0$, I-78 76 %; $n = 1$, I-79 40 %



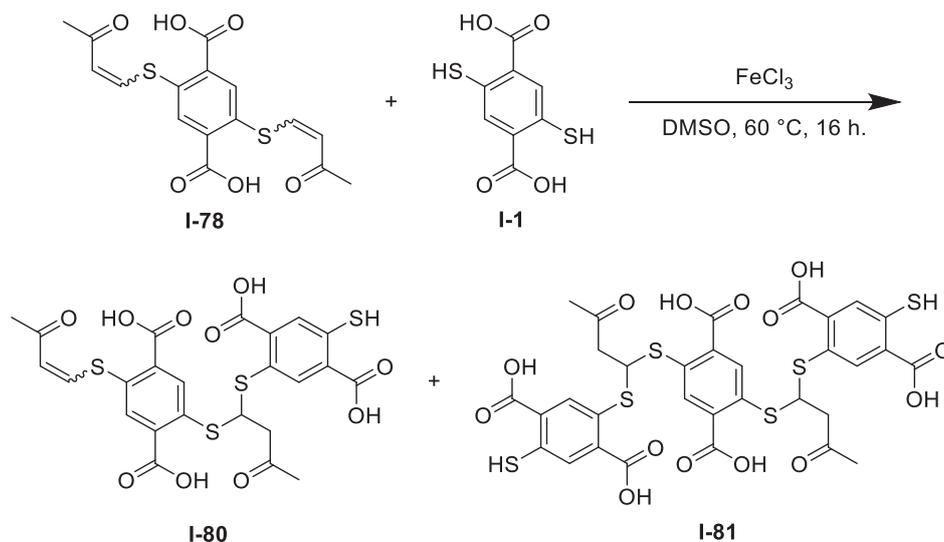
Scheme 48 : Members of the assays for dynamic thiol exchange with thio-Michael addition

Entry	I-1	I-15	I-78	I-79	Conditions
1	5 mM	/	5 mM	/	0 % oxidation
2					50 % oxidation
3					100 % oxidation
4	/	5 mM	/	5 mM	0 % oxidation
5					50 % oxidation
6					100 % oxidation
7	2.5 mM				0 % oxidation
8					50 % oxidation
9					100 % oxidation
10	10 μ M	/	10 μ M	/	DMSO-d ⁶ , Et ₃ N
11	/	10 μ M	/	10 μ M	
12	5 μ M				

Table 3: Set of conditions for the assays of antiparallel chemistry involving two bisthiols and their corresponding thio-Michael adduct with I-77. The oxidation level was obtained by mixing an adequate degassed solution of sodium borate (oxidant) with degassed DMSO and degassed stock solutions of building blocks.

The experiments were first run in a degassed phosphate buffer (100 mM, 7.8) but due to a poor solubility of the mixture, degassed DMSO was added at a volume ratio of 1:1. The oxidation levels were controlled by the concentration of oxidant sodium borate dissolved in stock solutions of phosphate buffer. All the solutions were degassed, avoiding the presence of oxygen. The only oxidant was thus sodium borate. The mixture was analysed by ESI-MS analysis and HPLC during 72 h. For oxidation levels set at 0 % (ie no sodium borate) (entries 1, 4 and 7 Table 3), no disulfide bridge was formed, which was logical. Nevertheless, only the building blocks were detected. The formation of dithianes by thiol-Michael addition was not observed, that was unexpected because of the supposed strong nucleophilicity properties of the thiols. Nevertheless, the presence of carboxylates at the ortho position of the thiol might prevent from the nucleophilic addition, compared to the meta positions of the reported example. Regarding the oxidation level set at 50 and 100 % (ie half stoichiometric and stoichiometric amount of sodium borate regarding thiols) (entries 2, 3, 5, 6, 8, and 9 Table 3), the results were also unexpected. For I-1, no mixed species were observed and only I-1₄ and I-1₃ were detected (ESI-MS analysis and HPLC). No overoxidation of thiols into sulfinic or sulfenic acids were detected at 100 % of oxidation level. The conditions might be optimised, for instance with higher concentrations of species and in pure organic media.

Another condition of reversible covalent bond exchange between β -sulfido- α,β -unsaturated carbonyls and dithianes, involving DMSO and DIPEA as catalyst, was used¹³⁰. The composition of libraries were monitored by ESI-MS analysis, HPLC and ¹H NMR. Homo dithianes of **I-1** and **I-15** were observed by ESI-MS analysis, at the opposite of hetero and mixed species. Referent dithianes of **I-1** was made by nucleophilic attack of **I-1** with **I-78** in DMSO and by activation with Fe(III)¹³¹ and was not isolated (Scheme 49). At this point, the conditions were not further optimised to obtain an efficient coupling between dielectrophiles and dinucleophiles into alternated hetero-oligomers. The concentrations might be too low (μ M) and should be increased.

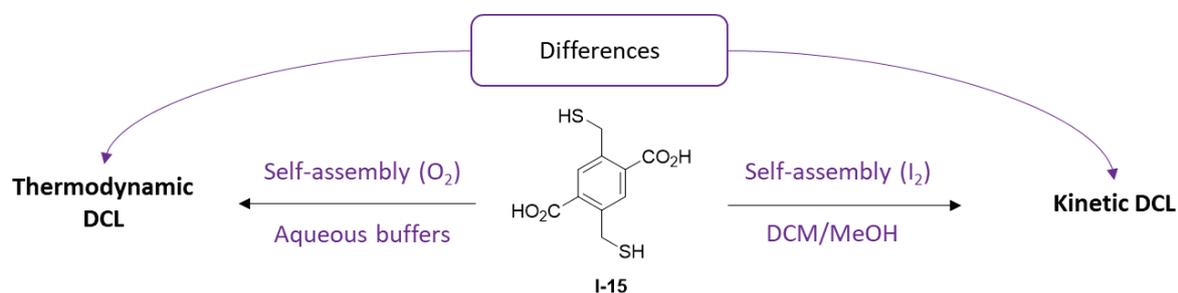


Scheme 49 : Formation of referent homo dithianes from **I-78** and **I-1**

Thio-Michael adducts of **I-1** and **I-15** and but-3-yn-2-one were efficiently prepared. The nucleophilic attack of **I-1** and **I-15** with dielectrophiles such as these adducts was surprisingly not spontaneous in aqueous mixtures at millimolar concentrations.

f. Thiol oxidation of 2,5-bis(mercaptomethyl)terephthalic acid **I-15** with iodine

The aim of this part was to study to apply the reported thiol oxidation by iodine (Scheme 50).

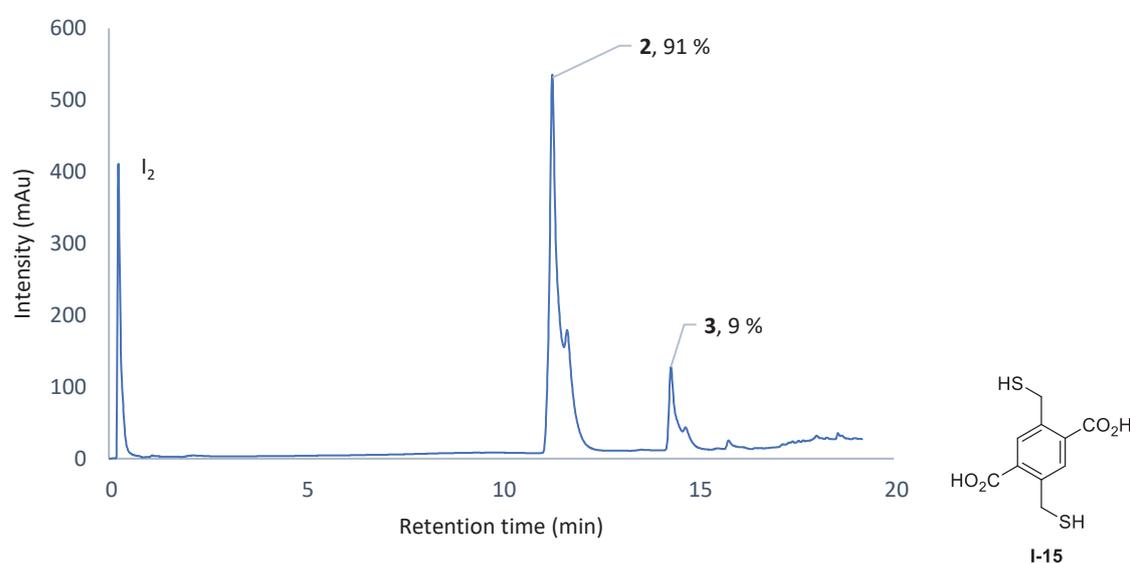


Scheme 50: Application of the reported method of thiol oxidation to disulfide bond by iodine compared to oxidation with oxygen

The self-assembly of 2,5-carboxylate-1,4-phenylenedimethanethiol **I-15** into disulfide bridge cyclophanes by oxidation of thiols into disulfide bonds by iodine was explored. The aim was to compare the composition of a DCL of **I-15** obtained by slow oxygen oxidation with a DCL obtained by oxidation

with iodine. Discrete disulfide cyclophanes were formed in a mixture of DCM/MeOH^{108b,112,114a,d}, after 5 mn of slow addition of a concentrate solution of iodine into a 20 mM solution of **1-15**. LC/MS analysis revealed the identity of the two formed macrocycles. A cyclic dimer and a cyclic trimer represented respectively 91 % and 9 % of the crude mixture (HPLC) (Scheme 51). The addition of Cu(II), Eu(III) and Zn(II) salts at 0.5 eq did not modify the proportion of trimer **I-15₃** and dimer **I-15₂**. Other solvents such as aqueous buffer and pure methanol or ethanol led to complex mixtures.

The difference between the DCL with iodine in DCM/MeOH and the DCLs with slow air oxidation in aqueous buffers (containing cyclic-oligomers with up to 10 units) was remarkable. That highlighted the role of the rate determining step. In fact, regarding oxidation with iodine, the rate determining step was the thiol-disulfide exchange while it was the oxidation of thiols into disulfide bond in the case of oxygen. Given that iodine was added in excess, the members of the DCL could equilibrate for a very short time until full consumption of the thiols.



Scheme 51: Mixture composition of the DCL involving I₂ and I-15, after 5 mn, method B, 2 cyclic dimer, 3 cyclic trimer

The macrocycle **I-15₂** was then isolated by a template-free spontaneous precipitation without further purification with a satisfying yield of 61 % compared to **I-1₄** obtained by templating followed by precipitation induced template removal. **I-15₃** could not be isolated. A separation by preparative HPLC could help to isolate the trimer. **I-15₂** was stable over two weeks (longer times were not analysed) in aqueous solution and in organic media.

Oxidation with iodine gave discrete disulfide cyclophanes without salt additives. Among them, the cyclic dimer I-15₂ was isolated.

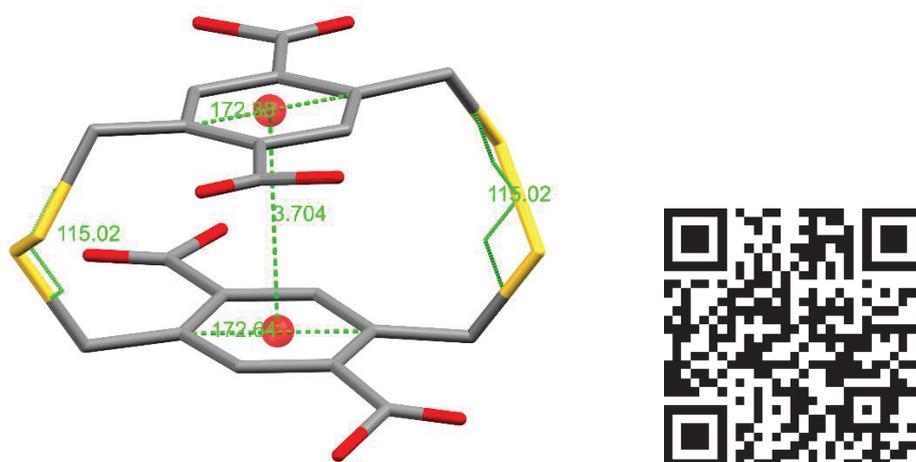
g. Study of the dimer properties

The dimer **I-15₂** was studied by crystallography and evaluated for its host guest chemistry in this subpart.

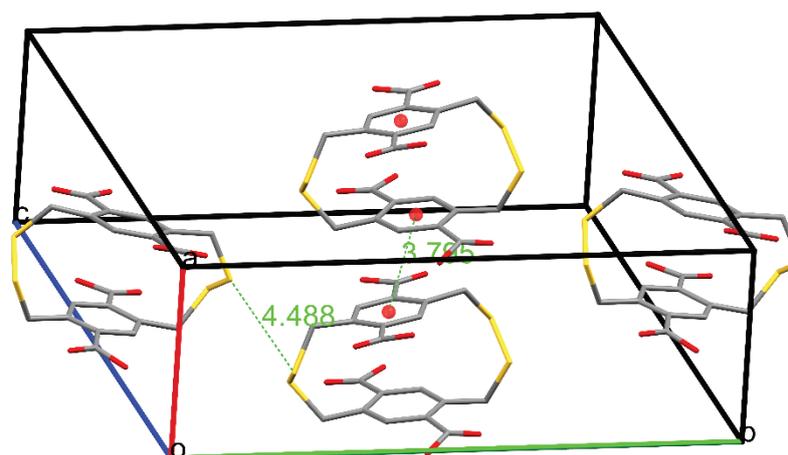
i. Crystallographic structure

A suitable crystal for X-ray diffraction was obtained by slow diffusion of CHCl₃ in a solution of **I-15₂** in DMSO (Scheme 52). The dimer crystallised in a Pba₂ orthorhombic structure. The molecules stacked

through aromatic interactions in the crystal (Scheme 53). The distance between two nearest aromatic centroids was 3.795 Å, likewise to similar cyclophanes^{114a,d}. the overlap between two neighbouring sulfur atoms was 4.488 Å and hence too far to suggest an intermolecular S-S interaction. That was at the opposite of one example^{114d} and similarly to another one^{114a}.



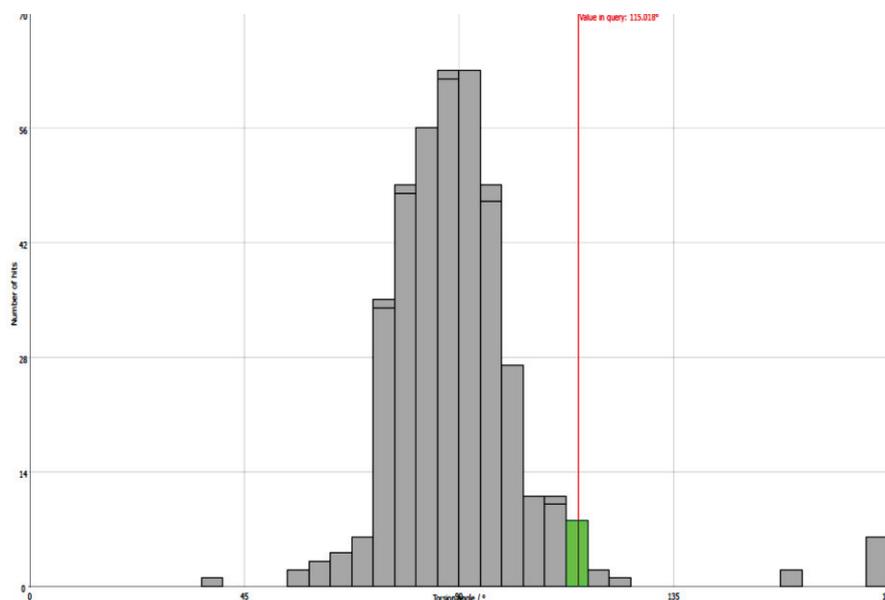
Scheme 52: X-ray structure for **I-15₂**, snapshot from Mercury software¹³²



Scheme 53: Packing unit for **I-15₂** crystallographic structure snapshot, from Mercury software¹³²

The solid state was driven by a porous packing arrangement despite the destabilising electrostatic interactions between the carboxylic acids. The dimer co-crystallised with a DMSO molecule, but not in the pores of the macrocycles. That led to an unfavourable void space formed by the extension of the pores along the crystal, explaining the high disorder observed for this X-ray structure. The size cavity was 3.704 Å. The C-S-S-C torsion angle of the disulfides bonds was 115.02° and substantially deviated from the theoretical angle (90°)¹¹⁹. The disulfide bond adopted a highly strained conformation. Such a strain could come from the electrostatic repulsion between the carboxylic moieties. In fact, a similar dimer without these carboxylic acids showed a deviation in a lesser extent^{114d}. Only 20 nearby structures were identified with torsional angles around 115° in the Cambridge Crystallographic Database (Scheme 54). Most of them appeared as bridging units in cyclic structures. Such important strain caused a slight deviation of the methylene carbon from the benzene plane (172.64° and 172.38°). The aromatic rings were co-planar and adopted a face-to-face orientation, which was common in paracyclophanes and decreased the strain¹³³. Indeed, arenes oriented perpendicular to the macrocycle plane and the methylene protons were located along the macrocyclic equator. The structure was quite

similar to ones obtained by the Johnson's group^{114a-c,114e}. According to the configuration and conformation descriptors^{84,86,119}, this dimer is (pR)₂ and PP. A more precise description of these descriptors are given in a next part (Structural properties of). **I-15₂** could be assimilated to a highly strained cyclophane, a subfamily showing particular structural and physical properties¹³⁴.



Scheme 54: Statistical repartition of torsion angle for C-S-S-C in CSD (snapshot from Mogul software¹³⁵, 14/03/18)

A DFT optimisation (DFT) gave similar energy without significant differences for all the possible conformers (Appendix I-VII: DFT data for **I-15₂**). The (pR)₂ PP diastereoisomer may be the unique to crystallise, and some others could remain in solution, or it could be the unique formed by the iodine oxidation. The conformational exchange barrier was 11.9 kcal.mol⁻¹ between the pR2 PP to a pR2 PM, which was a relatively low energy. The X-ray structure showed that the cavity was unambiguously short and did not contain a solvent molecule. Nevertheless, its properties of host guest chemistry were investigated.

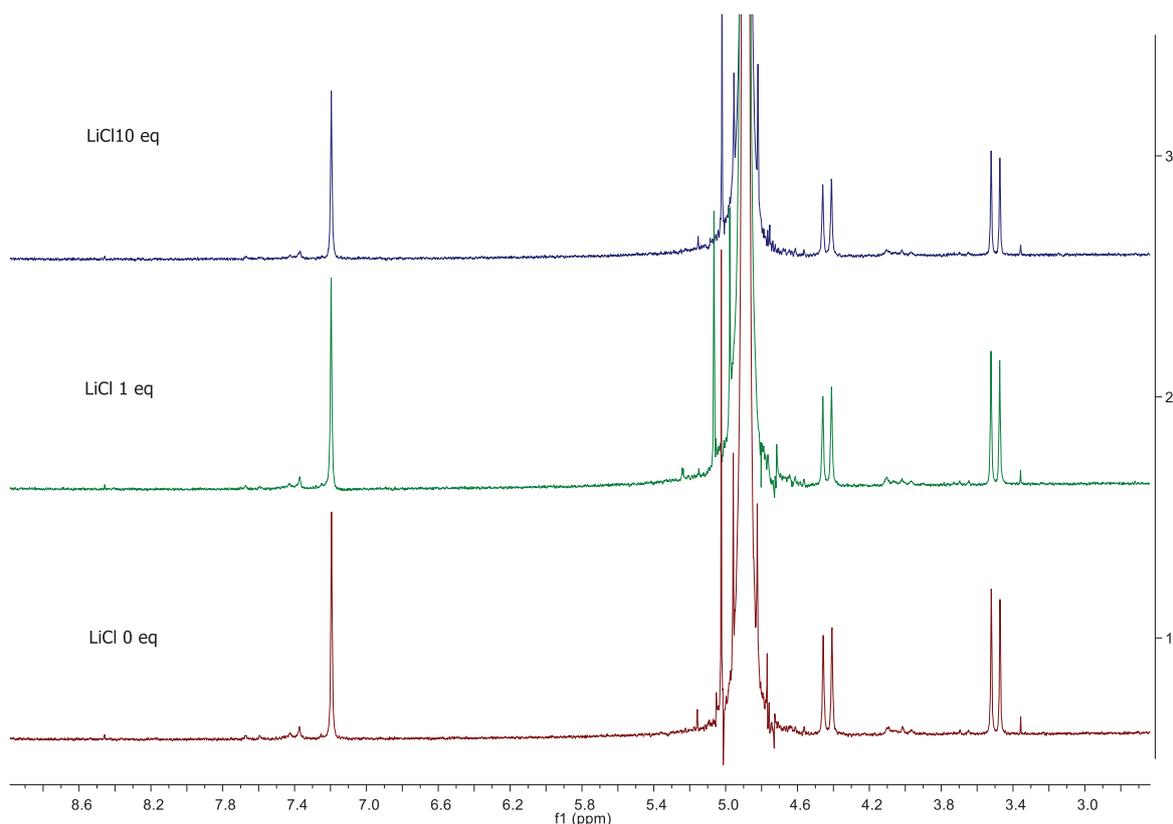
A crystallographic structure revealed a highly strained structure with an unusual C-S-S-C torsion angle in the solid state.

ii. Binding of lithium and amines with **I-15₂**

The binding of **I-15₂** with lithium ion and amines was investigated by means of ITC and ¹H NMR. The titration of a 1 mM solution of **I-15₂** by lithium chloride was conducted in water and followed by ¹H NMR (Scheme 55). No shift of the signals of aromatic protons were observed, proving that there was no deformation of the shape of **I-15₂** upon LiCl addition.

Regarding the titration of a 2.5 mM solution of **I-15₂** by lithium in Tris buffer 200 mM pH 7.4, the sigmoidal factor¹³⁶ was 4, under the acceptance criteria (5-1000) (Table 4). In addition, the error on the association constant was 34 %, superior to the commonly accepted value of 20 %. Thus, there was no clear complexation between Li⁺ and **I-15₂**¹².

¹² Titration experiment of 2.5 mM **I-15₂** with LiCl was carried at 298 K on a Malvern ITC200 microcalorimeter. Under these conditions, the quantification of a binding event is limited to affinities that are superior to 1/[(4-



Scheme 55: ^1H NMR spectrum of the addition of lithium chloride on a solution of **I-15₂** (1 mM, pH adjusted D_2O 7.4 with TMSP)

N (1:1 model)	0.916 +/- 0.066
K (L/mol)	$2.09 \cdot 10^3$ +/- 715
ΔH (cal/mol)	-16.57 +/- 1.766
ΔS (cal/mol/°)	15.1

Table 4 : ITC data from **I-15₂** (2.5 mM) and LiCl (25 mM) in Tris buffer 200 mM pH 7.4

Solution of **1-15₂** (1-2.5 mM) were titrated by various amines in Tris buffer 200 mM pH 7.4. Regarding butadiamine **I-10₄**, the sigmoidal factor was low (0.640) and the error on the association constant important (58 %) ¹³. The titration of **1-15₂** (1 mM) with spermine **I-8** had a low sigmoidal factor of 5 but a reliable association constant (19 % error). The titration of **1-15₂** (1 mM) with spermidine **I-14** had a low sigmoidal factor of 2 and a quasi-reliable association constant (21 % error). For both polyamines, the binding was weak and non-selective.

$8)_2] \sim 4.10^2 \text{ M}^{-1}$ ($K_d > 0.4 \text{ mM}$). See: <https://cmi.hms.harvard.edu/files/cmi/files/microcal-itc200-system-user-manual.pdf>

¹³ Titration experiments of 1/2.5 mM **I-15₂** with amines were carried at 298 K on a Malvern ITC200 microcalorimeter. Under these conditions, the quantification of a binding event is limited to affinities that are superior to $1/[\text{I-15}_2] \sim 10^3 / 4.10^2 \text{ M}^{-1}$ ($K_d > 1 / 0.4 \text{ mM}$). See: <https://cmi.hms.harvard.edu/files/cmi/files/microcal-itc200-system-user-manual.pdf>

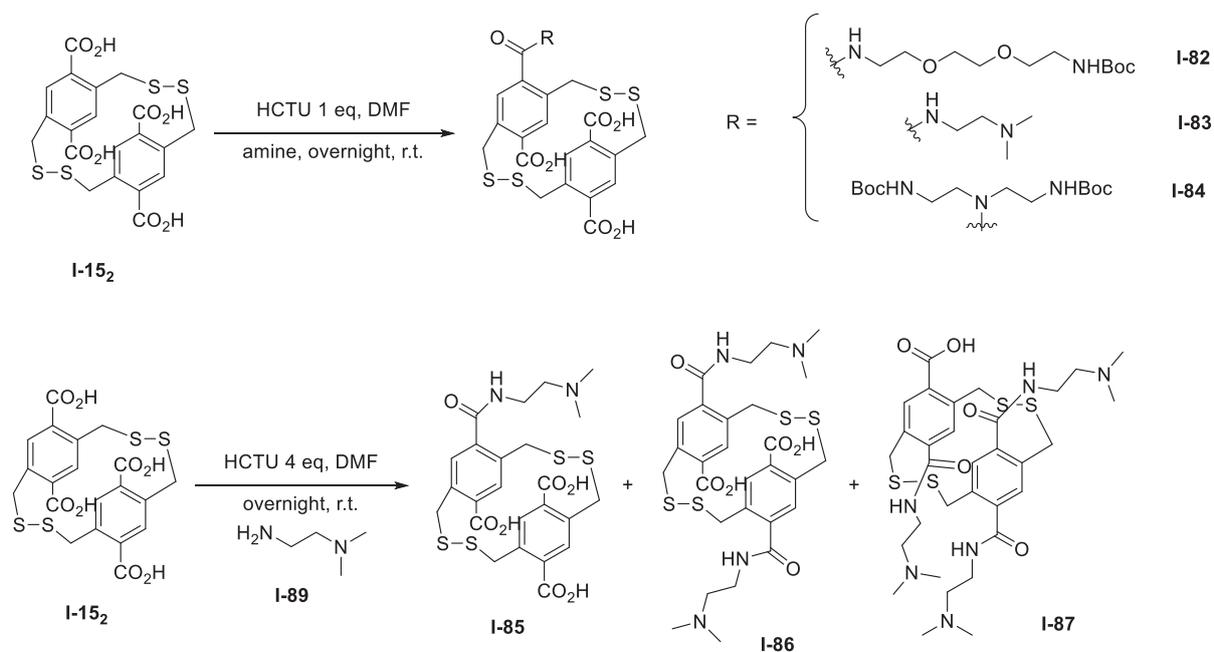
Ligand	Spermine I-8 (10 mM)	Spermine I-8 (25 mM)	Spermidine I-14 (10 mM)	Spermidine I-14 (25 mM)	Butanediamine I-10₄ (10 mM)
N (1:1 model)	0.902 +/- 0.04	0.900 +/- 0.04	0.938 +/- 0.051	0.618 +/- 0.06	0.360 +/- 0.248
K (L/mol)	5.02 10 ³ +/- 998	2.74 10 ³ +/- 541	5.16 10 ³ +/- 1.24 10 ³	1.36 10 ³ +/- 291	1.90 10 ³ +/- 1.11 10 ³
ΔH (cal/mol)	-181.1 +/- 11.6	-203.8 +/- 12.3	-129.7 +/- 10.7	-446.0 +/- 53.5	-82.2 +/- 66.7
ΔS (cal/mol/°)	16.3	15.0	16.6	12.8	14.7

Table 5: ITC data from **I-15₂** (1.0 or 2.5 mM) and various amines (10 or 25 mM) in Tris buffer 200 mM pH 7.4

A weak and non-selective binding was observed by ITC between spermidine **I-14**/spermine **I-8** and **I-15₂**.

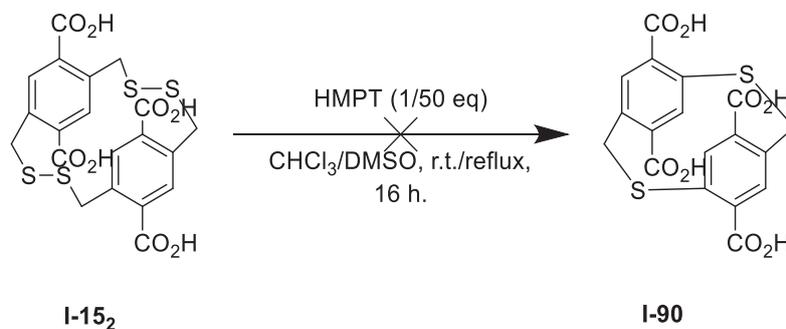
iii. Post functionalisation of **I-15₂** by amide bond formation

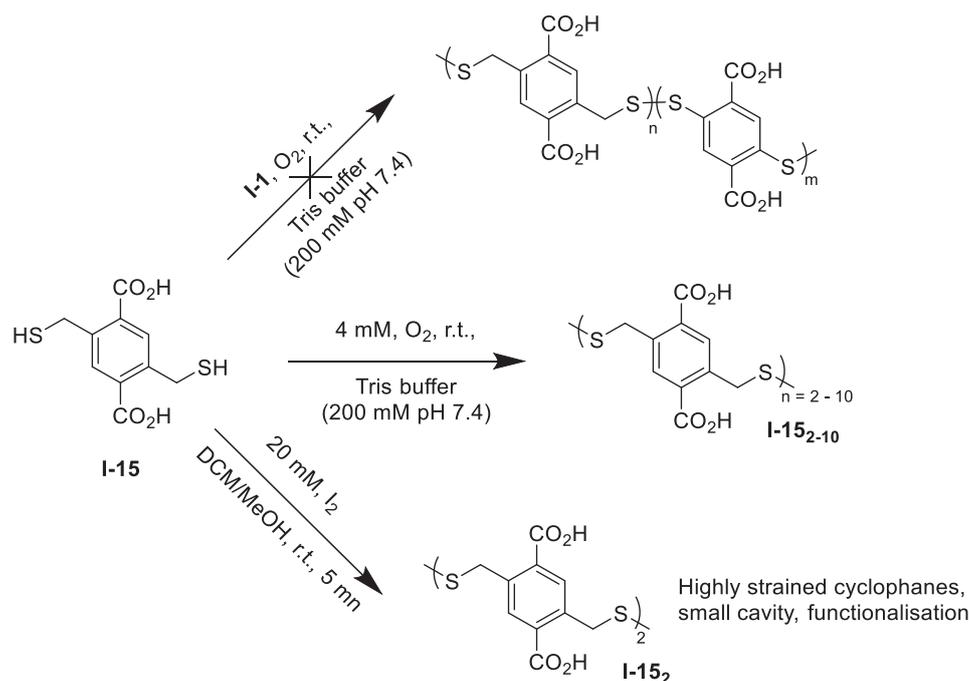
The reaction of the free carboxylic acids of the dimer **I-15₂** with various amines was explored in order to introduce new chemical functions on the cyclophane. Various amines were tested (Scheme 56). The amines were introduced with one or 4 equivalents of HCTU in DMF for 16 hours. Three examples of amide bond formation on 1 to the 4 carboxylic acid positions of **I-15₂** were observed by LC/MS with one equivalent of HCTU. In the case of N₁,N₁-dimethylethane-1,2-diamine **I-89**, increasing the equivalent of HCTU to 4 led to a mixture containing the species bearing one, two and three amide bond respectively, the conversion into tri-functionalised cyclophane **I-87** was 44 % (HPLC). A semi-preparative chromatographic purification failed to isolate the products for further characterisations. These results paved the way for an optimisation of the reaction and of the isolation of the functionalised macrocycles.



Scheme 56: Examples of side arm introduction by amide bond formation on **I-15₂**

The strain release drove the sulfur extrusion on similar cyclophanes^{114a,b} and was logically investigated on **I-15₂**. Even harsh conditions such as dissolution of **I-15** in HMPT and reflux during 16 hours^{114b} were unsuccessful to extrude sulfur atoms (Scheme 57). The mixture was analysed by ESI-MS analysis. No conversion was observed. The access to the disulfide bonds may be difficult for the reagent. In addition, the corresponding thio-ether could be too small and constrained. The benzylic ring will be in close contact. The Johnson's cyclophanes were less constrained and consequently could benefit from sulfur extrusion^{114a}.

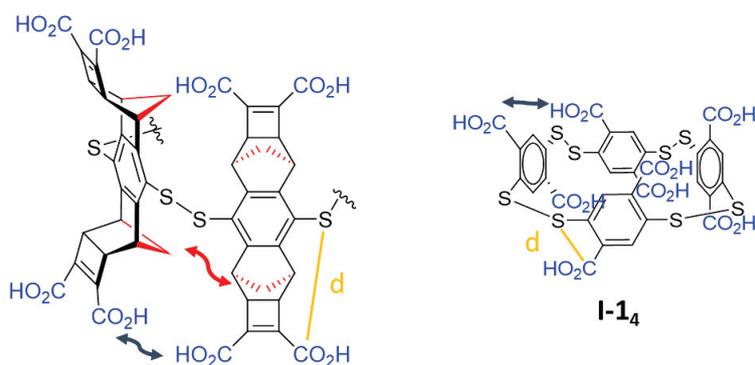




Scheme 58: Summarize of the self-assembly properties of I-15

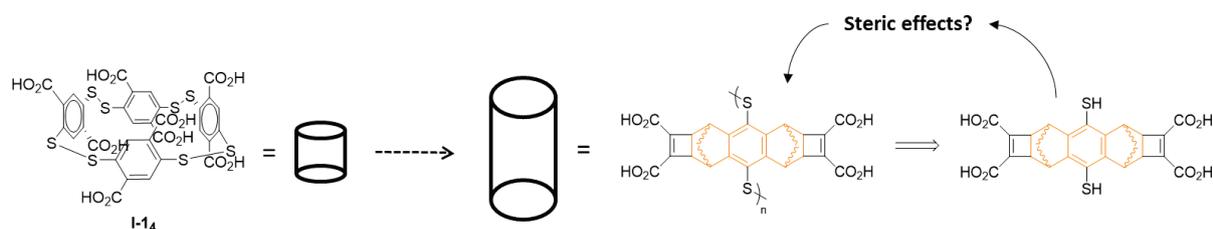
5. Separation of the thiol and the carboxylate moieties in a dithiophenol tetra carboxylate building block

The aim of this subchapter was to study the self-assembly through DCC of building block based on 1,4-bis(2,4,6-tricarboxylphenyl)ethane-2,5-dithiol including a spacing unit between the carboxylic acids and the thiol moieties. The building block **I-105** contained especially an octahydro-anthracene aromatic core, incorporating two methylene bridges. During the formation of cyclo-oligomers, the faces of **I-105** were expected to be closely in contact because of its bulk structures. That would result in a self-assembly that would be possibly influenced by the interactions of the methylene bridges inside the cavity of the corresponding oligomers.



*Figure 12 : Expected interactions between the monomers of self-assembled disulfide cyclophanes from 1,4-bis(2,4,6-tricarboxylphenyl)ethane-2,5-dithiols. Left, macrocycles from **I-105** (arbitrarily represented with two orientations inside the cavity regarding methylene bridges), with higher distance d between carboxylate and sulfur compared to **I-14** (right). The methylene bridges would be in close contact, involving a steric attraction or repulsion (red), while electrostatic interactions (blue) were predominant between the monomers of **I-14**.*

Compared to the self-assembly of **I-1**, highly influenced by the electrostatic contributions of the carboxylates, the DCL of **I-105** was supposed to be influenced by steric effects and non-covalent interactions such as Keesom, Debye and London forces, which are usually poorly considered for the design of receptors and in fundamental studies. Following the orientation of the methylene bridge inside or outside the cavity of **I-105** macrocyclic species, either the methylene bridges will be responsible for steric hindrance or they will stabilise the overall structures. As a matter of fact weak non covalent interactions such as London forces¹³⁸ can be responsible for the stability of crowded molecular architectures. As instance, they explain the boiling point of alkanes¹³⁹ and play a part in DNA stability¹⁴⁰. Regarding **I-14**, the desired corresponding macrocycles from **I-105** will differ from the height of their cavity. In order to study the influence of steric effects in the self-assembly of **I-105**, DCLs were studied by LC/MS and ¹H NMR, supported by modelling by DFT (Scheme 59).



Scheme 59: Goal and methodology in the subchapter. Synthesis of an aromatic and sterically crowded bisthiol containing carboxylic acids, self-assembly upon DCC and rationalisation of the process thanks to DFT.

a. Overview of London dispersive interactions

This bibliographic part provides key elements about London interactions. The non-covalent interactions could be divided into an electronic and a Van der Waals term¹⁴¹ (eq. 1). The latter is composed of the Keesom energy (orientational forces between two permanent dipoles), the Debye energy (interaction on a molecule induced a permanent dipole) and the London energy (dispersion forces between two induced dipoles created by a polarisation) (eq. 2).

$$E_{non-covalent} = E_{electrostatic} + E_{Van\ der\ Waals} \quad (1)$$

$$E_{Van\ der\ Waals} = E_{Keesom} + E_{Debye} + E_{London} \quad (2)$$

The total energy of these attractive forces varies in r^{-6} , r being the distance between two interacting molecules¹⁴¹. The London forces come from an electronic and long range correlation, increasing with the system size¹³⁸. Basically, they can be approximately written as followed, with α dipole polarizability, I first ionization potential and r intermolecular distance (eq. 3):

$$E_{AB}^{London} = -\frac{3}{2} \frac{I_A I_B}{I_A + I_B} \frac{\alpha_A \alpha_B}{r^6} \quad (3)$$

i. Computational description

Post ab initio methods can describe the electronic correlation. They consist in a first *ab initio* calculations, in which only the global effect of the electronic correlation is taken into account, for instance *Hartree Fock*. Then, a correction is applied in order to correctly describe the correlation. The correction methods can work by a configuration interaction. This is the case in *Coupled Cluster* approaches. A development by perturbation series can also correct the first calculation, for example with *Moller – Plesset*. However, they require an important calculation cost for large systems¹⁴². For instance, *Domain Based Local Pair Natural Orbital Coupled Cluster*¹⁴³ locally describes the correlation.

Besides, the coulombic integrals, responsible for the correlation, can be approximated by two centres integrals instead of three centres, for instance in *Resolution of Identity for Moller Plesset at the second order RI-MP2*¹⁴⁴. Even though the computational cost decreases, it remains important. The DFT methods offer a compromise between the cost and the accuracy. Nevertheless, most of the functionals do not describe the London forces properly¹⁴⁵. The electronic correlation is included in the density functional theory, but poorly approximated in the functional expression, particularly in the correlation exchange term (Density functional theory). 4 levels of accuracy were proposed in order to categorise the functional by their description of the dispersion interactions¹⁴⁶.

- Level 1, *parametric functionals*. Parametric functionals such as *meta-GGA*¹⁴⁷ are the lowest level. They do not describe long range interactions, but their parameters include a dispersion contribution. Some geometries of equilibrium were properly described¹⁴⁸ even though the contribution did not have a physical meaning^{145a}.
- Level 2, *DFT-D*. The total energy can be corrected after the DFT calculation by the dispersion energy (eq. 4). The term is parametric and can be written as followed for two atoms¹⁴⁹ (eq. 5), with C_n^{AB} the dispersion coefficient at order n for A-B, r their internuclear distance, s_n the global scale factor, dependent of the functional for adjustment and f_{damp} the Becke-Johnson damping function, which avoids a divergence for short r and a double counting of the correlation for medium distance¹⁴⁹⁻¹⁵⁰. Most of the time, the damping function and hence the dispersion energy are zero or a constant for short distance. This can create artificial repulsive forces that lead to an elongation of non-covalent interactions¹⁴⁹. Recently, an improvement of the short-range behaviour was proposed¹⁵¹.

$$E_{DFT-D} = E_{DFT} + E_{Dispersion} \quad (4)$$

$$E_{Dispersion} = -\frac{1}{2} \sum_{A \neq B} \sum_n s_n \frac{C_n^{AB}}{r_{AB}^n} f_{damp}(r) \quad (5)$$

The D_n correction is frequently used. The first versions D1 and D2 account for C_6 terms¹⁵². These dispersion coefficients are calculated by an empirical derivative interpolation. The D3 version adds the C_8 and C_{10} coefficients and improve their calculation by a physical basis¹⁵³. The atomic connectivity is considered for the polarizability calculation which is necessary for large and stacked π and σ systems¹⁵⁴. A recent D4 version includes a charge dependent polarizability and is mainly used for organometallic complexes¹⁵⁵. The London dispersion effects were accurately quantified in organometallic chemistry by this level¹⁴².

- Level 3, *VdW-DFT*. A non-local function can be added to the exchange correlation term, which avoids the use of external data for the C_n description. This approach is hence *ab initio* and named Van der Waals DFT. Several functionals were proposed between 2004 and 2010¹⁵⁶. The VV10 functional is based on the approximation of semi local generalised gradient with a non-local correlation term¹⁵⁷. This functional showed similar results than hybrid GGA functionals corrected with a D3 method, which are less time consuming¹⁵⁸. The *VdW-DFT* and *DFT-D* are recommended highlighting the importance of the dispersion energy in intramolecular cases¹⁵⁹. A recent study compared the accuracy of the three last levels, including the level 4¹⁶⁰.
- Level 4, *double hydride functional DHDF*. For the double hybrid functionals, the exchange correlation energy term is written as the sum of a non-local Fock exchange term E_x^{HF} and a non-local correlation term from MP2 E_c^{MP2} (eq. 6)¹⁵⁸.

$$E_{XC}^{DHDF} = (1 - \alpha_x)E_x^{DFT} + \alpha_x E_x^{HF} + (1 - \alpha_c)E_c^{DFT} + \alpha_c E_c^{MP2} \quad (6)$$

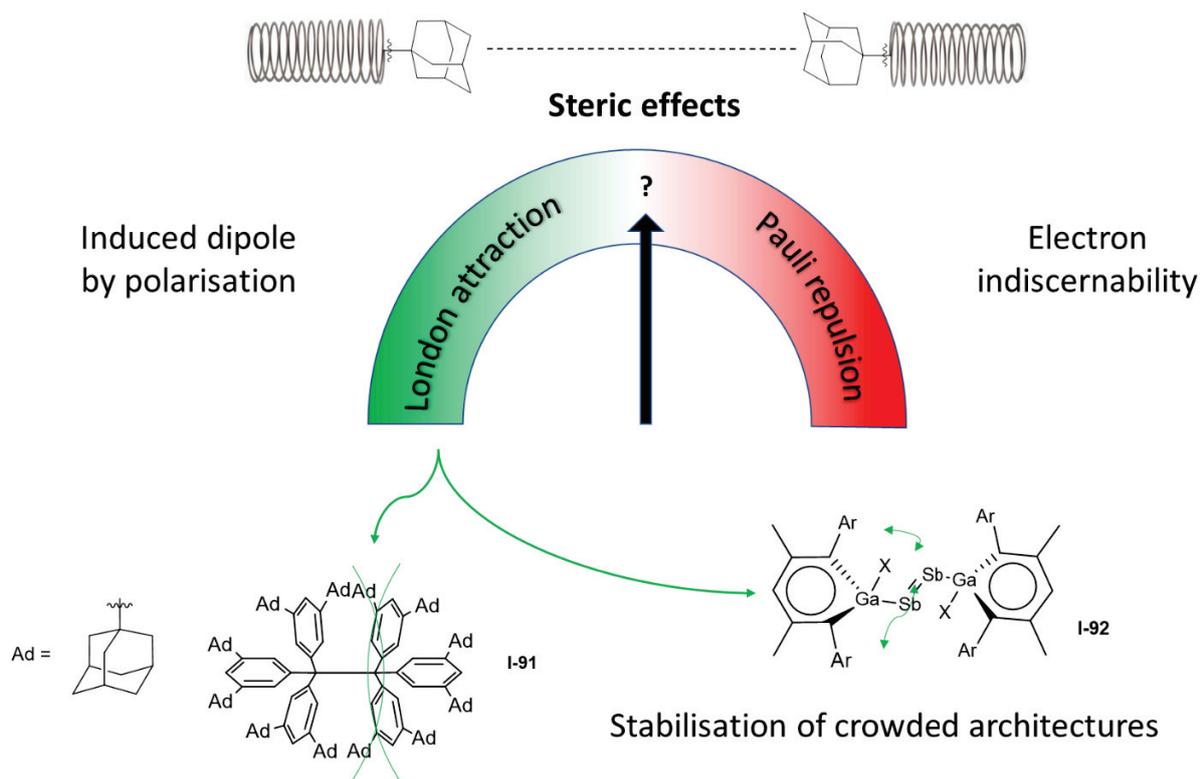
E_x^{DFT} and E_c^{DFT} are the exchange and correlation contribution from the DFT approximation. The first three terms are the expression of a hybrid GGA functional. The α parameters modulate the contributions. They are defined by a semi-empirical approach, based on thermodynamic data, or by a non-empirical approach with an adiabatic connection. The semi empirical is more accurate according to a recent study¹⁶¹. The *DHDF* functionals showed poor deviation and high accuracy on an organometallic benchmark but suffered from their calculation cost¹⁶².

Regarding the size of the further studied systems (DFT optimisation of the diastereoisomers), the *DHDF* method was not retained because of the calculation cost. In addition, *VdW-DFT* was not chosen, as they are not more accurate than a D3 method. Consequently, a Dn corrected B3LYP-D3BJ and a recent parametric M062X functionals were selected. These functionals were used for the study of the conformational isomerism of larger alkenes¹⁶³.

Several approaches in DFT enable to account for dispersion forces. Parametric functionals and dispersion correction DFT-D show correct results and low calculation costs.

ii. Applications in organic chemistry

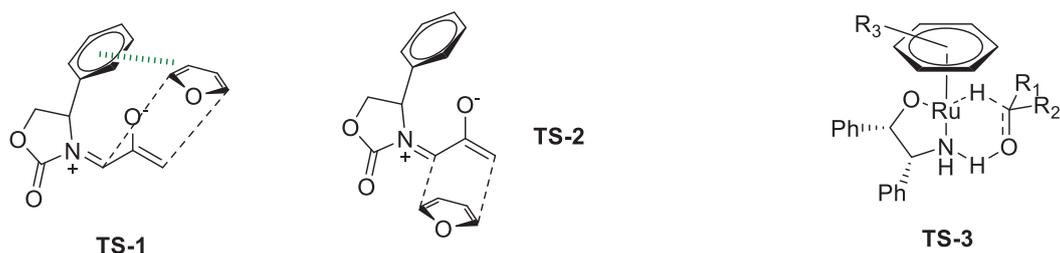
Usually, the energy resulting from dispersive interactions are underestimated compared to the electrostatic ones. For instance, the electrostatic energy between two charges at 5 Å is the same as the dispersion energy between two fragments of 100 atoms (C, H, N, O) at the same distance, i.e. 60 kcal.mol⁻¹¹⁶⁴. They influence the properties of surfaces¹⁶⁵ and various objects, such as conformational and rotational barriers¹⁶⁶ or agostic complexes¹⁶⁷. Furthermore, they play a part in the selectivity in synthetic chemistry because of their involvement in transition states. A review highlighted the importance of London forces and its misunderstanding in a lot of different rationalising models or concepts in synthetic chemistry¹³⁸. The review also described the computational modelling of representative cases in which dispersion effects are important. The concept of steric effect was clarified. The introduction of a bulky group is indeed usually assumed to hamper the reactivity of the system. The interactions between two non-polar fragments are divided into a Pauli exchange steric repulsion and a London dispersion attraction (Scheme 60). The Pauli repulsion is the result of the electron indiscernibility. Two electronic clouds in approach develop a repulsive force, counterbalanced by their induced polarizability. Consequently, the interaction between bulky groups can be attractive at a specific distance. This delicate balance can be in favour of the stabilisation of large and crowded structures, showing unusual long C-C bond, for instance with triphenylmethyl derivative **I-91**¹⁶⁸. Alkyl-substituted hexaphenylethane molecules were formed starting from their radicalar triphenylmethyl derivative monomers. These dimers contained a weak and long covalent bond C-C bond. The dimers were in equilibrium with their monomers. The determination of the equilibrium put in light the ability of the alkyl groups to stabilise the overall dimers. The equilibria were more in favour of the dimer for substituents such as adamantyl (Ad) and *tert*-butyl rather than cyclohexyl or phenyl. Such rigid groups can overcome the entropic penalty and the Pauli repulsion with a specific conformation maximising the London forces. That proved the role of stabilising London interactions in molecular stability.



Scheme 60: Steric effects, divided into an attractive London dispersion forces and a repulsive Pauli exclusion. Examples of unexpected stable structures thanks to dispersion forces¹⁶⁸⁻¹⁶⁹. Ar: 2,6-*i*Pr-C₆H₃, X: F, Cl, Br, I, NMe₂, NEt₂, OEt.

These steric effects coming from a specific moiety were considered as a tool named dispersion energy donors DEDs¹⁷⁰. This concept is widely used in organometallic chemistry. For example, the properties and the stability of Gadolinium substituted dipnictenes were proved to be strongly influenced by the London forces occurring between the substituents on the ligands and by pnictogen- π interactions between the metal and the ligand **I-92**¹⁶⁹. Besides, the London dispersion contributes to the stability of stilbinidenes¹⁷¹ and NHC compounds¹⁷². The dispersion effects are important in the solid state and are obviously weakened in solution, but in this case, intramolecular contributions remain strong. Particularly, molecular recognition and biological processes are influenced by London forces¹⁷³. In addition, enzymatic pockets can be protected from the solvent preventing from electrostatic dominant interactions. The coulombic forces can be completely overcome by the London dispersion. This is the case regarding the stability of the viologen cation radical pimer. The stacking led to attractive dispersion forces overcoming the Coulombic repulsion of bringing two cationic species into contact¹⁷⁴. Even in ionic liquids, the London forces compete with the hydrogen bonding¹⁷⁵. Regarding the hydrophobic effect, divided into a classic entropic part, and a non-classical enthalpic part, some studies suggested a contribution of the London interactions¹⁷⁶. Furthermore, they can accelerate the nucleophilic substitutions of substituted aromatic chlorides¹⁷⁷. The attraction dispersion was more important for the ortho substitution compared to the para position. It was correlated to the kinetic rates of reaction. Finally, London forces can stabilise transition states, which can create selectivity in the case of a preferential stabilisation. This is particularly the case in organometallic catalysis¹⁷⁸. In fact, small changes of the interactions between the different moieties in a transition state can lead to important selectivity. For instance, the aryl substituents influenced the TS stabilisation of carbonyl hydrogenations with Ruthenium (II) complexes (**TS-3**)¹⁷⁹. The enantiomeric excess of the resulting

alcohol was directly linked to the attractive CH/ π interactions between the η_6 -arene ligand and phenyl group, which was a manifestation of the London forces.



Scheme 61 : Favoured TS by London dispersion. **TS-1**¹⁸⁰ and **TS-3**¹⁷⁹ benefit from CH/ π interactions counterbalancing the Pauli repulsion and stabilising the TS.

The dispersion attraction was also observed for the [4 + 3] cycloaddition reactions between alkoxy allyl cations and furan. The most crowded product was formed thanks to the CH/ π interactions in the TS (**TS-2** preferred to **TS-3**)¹⁸⁰. The pericyclic transition states are highly influenced by London forces¹⁸¹. The London dispersion can be responsible for enantioselective catalysis, even in a polar media and a competitive electrostatic environment¹⁸². Besides, a study combining computational and spectral techniques proved the stabilisation effect of London forces for an example of enantioselective peptide catalysed acylation¹⁸³. However, there is currently a need for a tool quantifying the possible steric effects of a chemical group¹⁸⁴. It is indeed difficult to predict the behaviour of a moiety in terms of steric attraction or repulsion¹⁸⁵. Finally, the self-assembly in supramolecular chemistry can be driven by the London dispersion. For instance, they dictated the association between a copper (I) bimetallic molecular clip and an aliphatic linkers¹⁸⁶. In a nutshell, the dispersion forces are ubiquitous and can be dominant in the stabilisation of molecular architectures. The design of molecules can benefit from these forces, even though there is a lack of predictive tools.

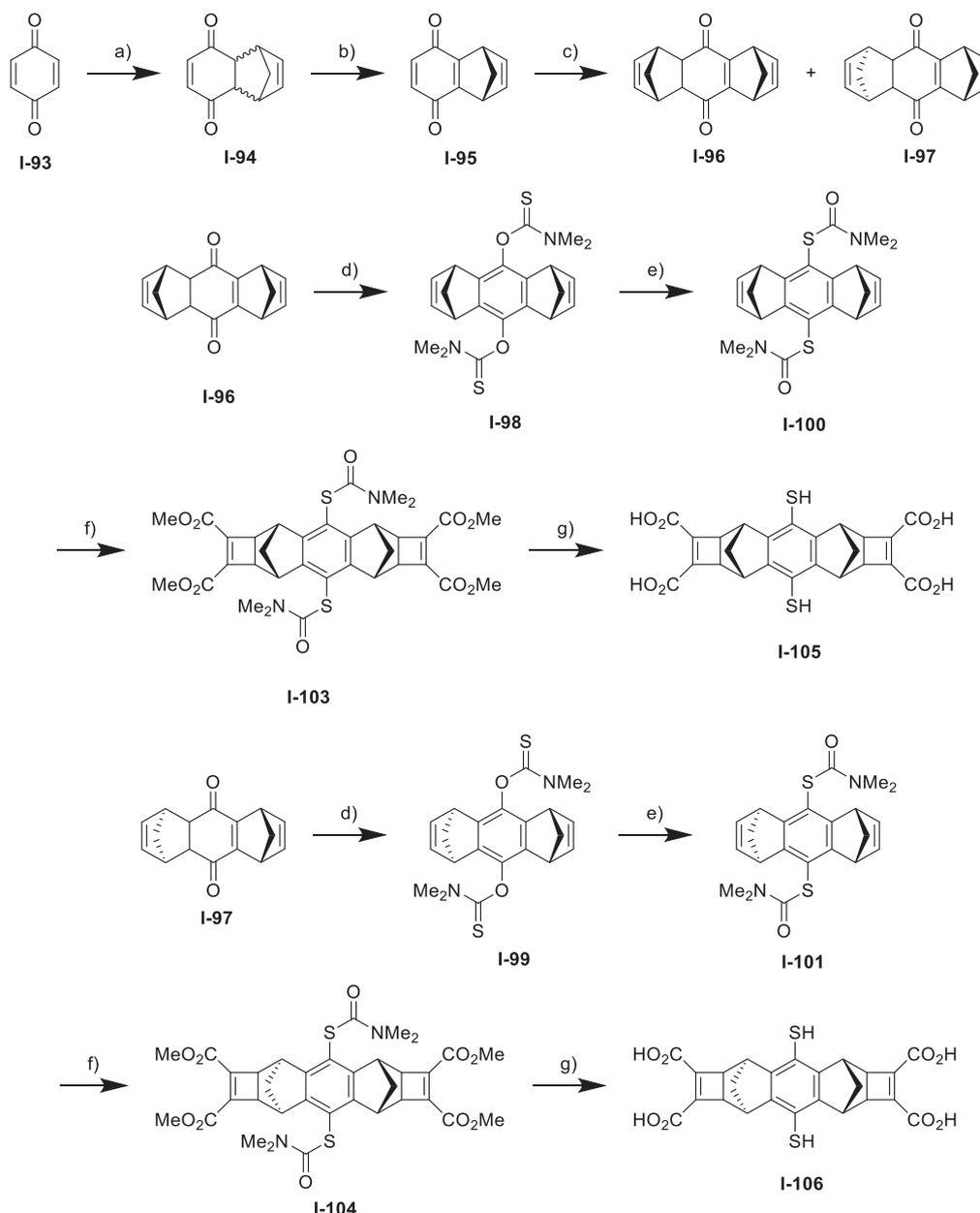
The London dispersion play a part in many areas in chemistry. The balance between attractive London forces and Pauli repulsion is difficult to predict but can be responsible for stabilisation and selectivity.

b. Synthesis of a pair of bulk diastereoisomeric **I-105** and **I-106** building blocks

The aim of this subpart was to synthesise rigid and sterically crowded building blocks in order to obtain bulk groups which can develop strong steric effects. The building blocks **I-105** and **I-106** were synthesised following a 7-steps procedure adapted from the work of a previous PhD student¹⁸⁷ (Scheme 62). Firstly, commercially available benzoquinone and freshly distilled cyclopentadiene obtained by a retro Diels-Alder reaction¹⁸⁸ formed quantitatively **I-94**. Then, the corresponding benzoquinone form **I-95** was restored thanks to a rearomatisation with Et₃N and oxidation with benzoquinone¹⁸⁹ with a satisfying 87 % yield. A second [4 + 2] cycloaddition with cyclopentadiene¹⁸⁹ **I-92** gave **I-96** (22 %) and **I-97** (22 %). The diastereoisomers were separated through a chiral semi-preparative HPLC separation¹⁴. The summed yield (44 %) was poor compared to the reported 97 %¹⁸⁷. Toluene was replaced by methanol for this step. Methanol facilitated the work-up but could be responsible for the lower yield. Regarding the stereochemistry, resulting from the orientation of the methylene bridges, **I-96** was named the *cis* product because its 2 methylene bridges pointed in the same direction regarding the aromatic plan. At the opposite, **I-97** was named the *trans* product. Then, reaction with sodium hydride enabled the rearomatisation of the cores while dimethyl carbamoyl

¹⁴Separation done by Dr. Nicolas Vanthuyne at iSm2, AMU university.

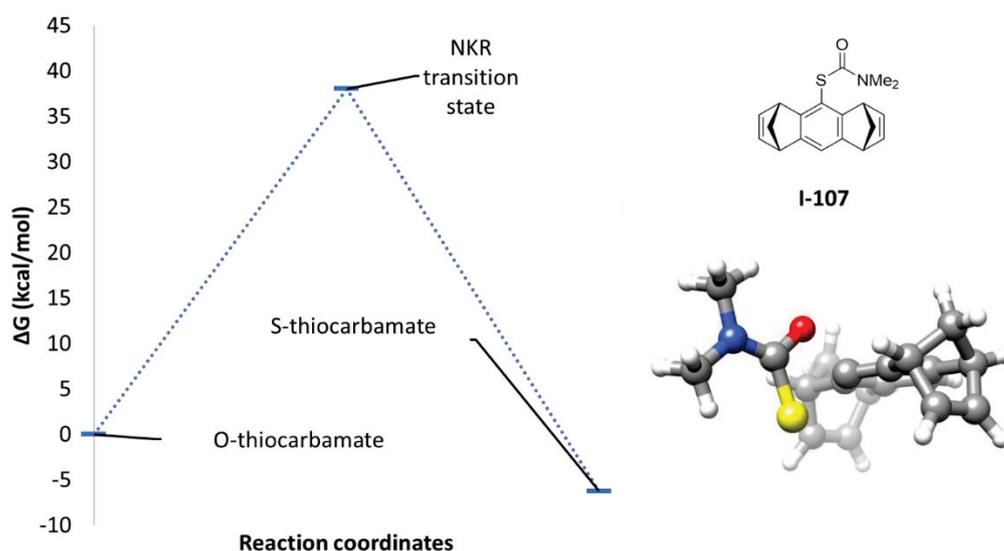
chloride trapped the corresponding phenolate ions into O-thiocarbamate **I-98** (99 %) and **I-99** (98 %) ^{79b,190}. The yields were satisfying.



*Scheme 62: Synthesis of bulk diastereoisomeric building blocks **I-105** and **I-106**, a) cyclopentadiene, MeOH, -78°C to r.t., 16 h., 98 %; b) 1. Et₃N, MeOH, r.t., 16 h., 2. benzoquinone **I-93**, CHCl₃, 50°C, 5 h., 87 %; c) cyclopentadiene, MeOH, -78°C to r.t., 16 h., 44 % (a) 22 % (b) 22 %; d) dimethyl thiocarbamoyl chloride, NaH, DMF, r.t., 16 h., 99 % (a), 98 % (b); g) diphenyl-ether, 230 °C, 3 h., (a) 49 % (b) 11 %; f) RuH₂CO(PH₃)₂, DMAD, THF, reflux, 10 d. (a) 45 %, (b) 41 %; g) NaOH, H₂O/EtOH, reflux, 16 h., (a) 90 %, (b) 37 %.*

At that point, a thermal Newman-Kwart rearrangement^{79b} changed the O-thiocarbamates into S-thiocarbamates in **I-100** (49 %) and **I-101** (11 %). The *cis*-product **I-100** of the rearrangement showed an appreciable yield for a thermal Newman-Kwart rearrangement in diphenyl-ether⁸¹. The yield (49 %) was similar to the microwave condition yield (49 %) ¹⁸⁷ while the *trans* diastereoisomer showed a poorer yield (11 %). The *trans* semi-rearranged product **I-102** was isolated and characterised but was not used. The Newman-Kwart rearrangement was modelled by DFT (computational details in the

experimental section) to get a further insight into its energetic cost (see Scheme 63). As expected, the TS was high in energy, requiring harsh conditions. The S-thiocarbamate was found to be lower in energy by 6 kcal.mol⁻¹, which explained the tendency of going from the O to the S-thiocarbamate.



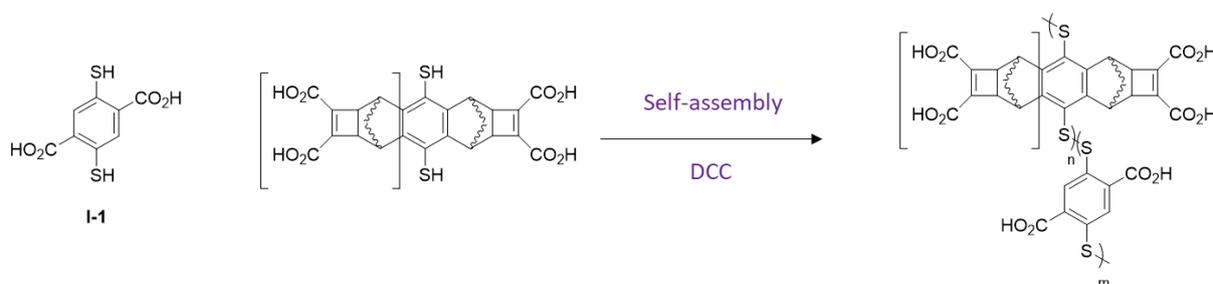
Scheme 63: Energetic profile of the NKR rearrangement of **I-107** analogue and TS representation Yellow: sulfur, red: oxygen, blue: nitrogen.

The next step was a ruthenium catalysed [2 + 2] cycloaddition with dimethylacetylenedicarboxylate¹⁹¹ in order to obtain **I-103** (45 %) and **I-104** (41 %). The yields were moderate, but the cycloaddition was not optimised. For instance, photochemical activation might improve the yields and avoid the high temperature and the long reaction time. Finally, a basic deprotection cleaved both S-thiocarbamates and methyl esters¹⁶ to get **I-105** (90 %) and **I-106** (38 %). The yields for **I-105** was higher than reported (80 %)¹⁸⁷. Finally, the global yields, starting from *p*-benzoquinone, were respectively 7.8 % (a) and 0.6 % (b) for 7 steps.

Two bulk building blocks were synthesised with a yield around 8 and 1 % over 7 steps.

c. Self-assembly through DCC by oxidation with dioxygen

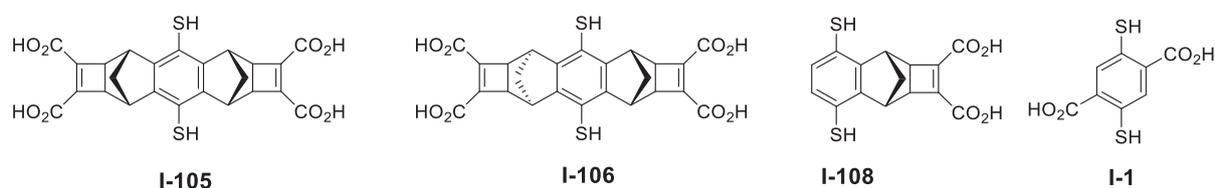
The aim of this subpart was to study the self-assembly of the previous building blocks (Scheme 64). They were individually and simultaneously dissolved in aqueous buffers in DCLs. The mixtures were left open, stirred and analysed by UPLC and LC/MS.



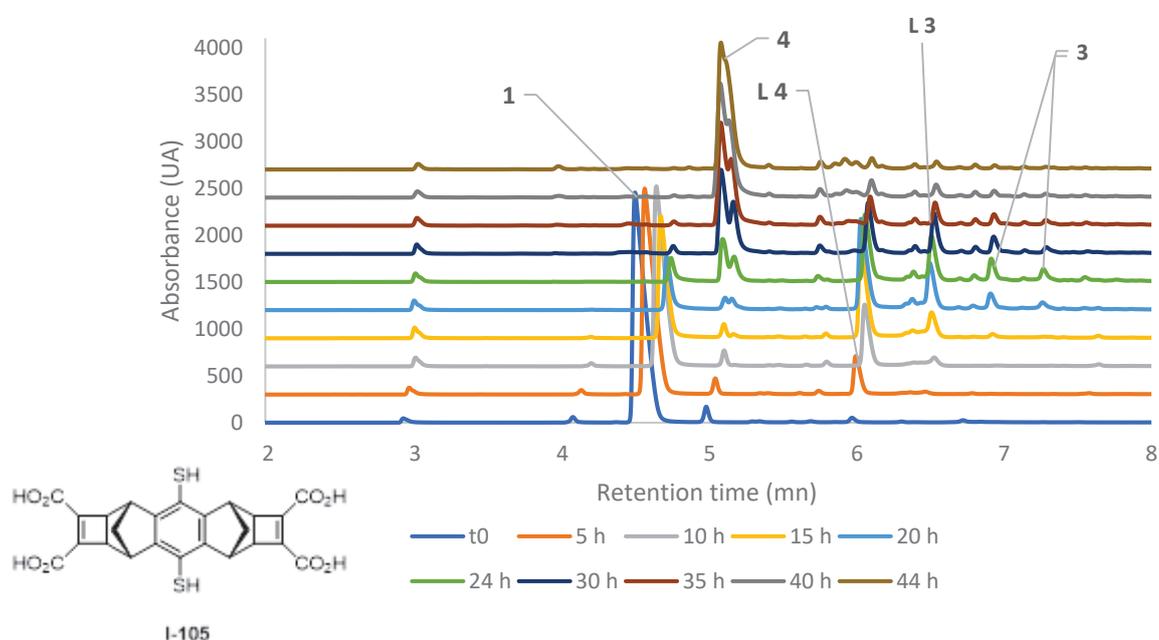
Scheme 64: Combination of sterically crowded building blocks with or without **I-1** in order to obtain disulfide macrocycles by self-assembly upon DCC.

i. DCLs of one building blocks

The building blocks **I-1**, **I-108**¹⁵, **I-105** and **I-106** were individually dissolved at 5 mM in aqueous buffers (Scheme 65). The details are given in appendix (Appendix I-II: Additional conditions of reactions). The HPLC signal of the building block *cis*-**I-105** disappeared in 24 h without the formation of other signals in Tris buffer (200 mM pH 7.4). The same behaviour was observed in ammonium acetate buffer (50 mM pH 7.4). A ESI-MS analysis did not show the presence of sulfenic or sulfinic acid derived from the 1,4-bisthiophenol or its oligomers. Consequently, the main reaction path might be a polymerisation of the building block, explaining the loss of the HPLC signal as the hypothetic polymers were not eluted. This hypothesis was not verified and could be assessed by DLS for instance. Nevertheless, the desired cyclo-oligomerisation occurred in ammonium acetate 50 mM but only without adjustment of the pH (starting pH at 6, final pH at 4.69 during the DCL). Thus, the influence of the pH seemed to be determinant. A lower pH meant a lower rate of disulfide bond formation, which might prevent polymerisation. The DCL showed an unexpected amplification of a cyclo-oligomer. The kinetic data are presented below (Figure 14, Figure 15). 24 hours were required to completely oxidised the building block. Simultaneously, a linear dimer and trimer, and a cyclic trimer were formed. All the materials were converted into a unique signal after 1.5 days, corresponding to a cyclic tetramer **I-105**₄, which was stable over 2 weeks.



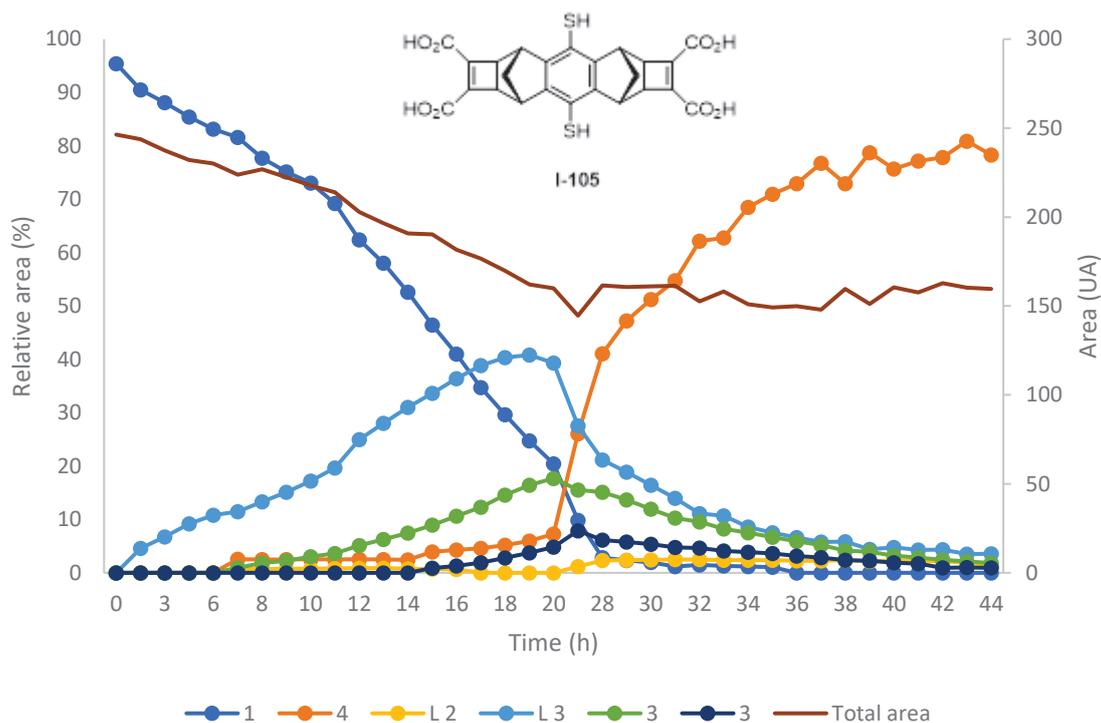
Scheme 65 : Building blocks used in the DCLs



Scheme 66: Evolution of **I-105** in ammonium acetate (50 mM, pH without adjustment), method K¹⁶. 1 building block, n: cyclic oligomer including n monomeric units, L linear oligomer

¹⁵ Batch from Dr Zhang.

¹⁶ The form of the tetramer signal is due to a saturation of the column by too high concentrations.



Scheme 67: Building block and macrocyclic population evolution of **I-105** in ammonium acetate (50 mM, pH without adjustment), method K, 1 building block, n : cyclic oligomer including n monomeric units, L linear oligomer

The cyclo-oligomerisation was also followed by ^1H NMR (5 mM, $\text{ND}_3\text{CD}_3\text{COOD}$ 50 mM without pH adjustment) with an internal standard (dioxane). Thanks to the calibration with dioxane, the conversion into the cyclic tetramer was 79 % at 44 hours, which was coherent with the UPLC (78 %) (Appendix I-IX: NMR experiments of **I-105** / **I-105₄**).

Interestingly, this amplification was reproduced at the same concentration of **I-105** dissolved in aqueous solutions of ammonium salts and in pure water. Hence, the selective formation of the tetramer was not driven by a template effect of an ion.

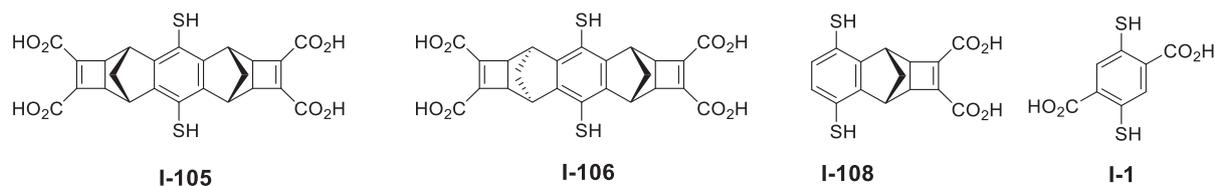
Surprisingly, the *trans* diastereoisomer **I-106** did not form oligomers in ammonium acetate 50 mM without pH adjustment. Thus, the particular shape of the *cis* diastereoisomer **I-105** was determinant to form short oligomers. At this point, the possible role of stabilising London forces inside the cavity of the tetramer was envisioned. Finally, iodine¹¹² oxidation of thiols of **I-105** did not lead to discrete cyclo-oligomers.

Furthermore, the DCL of **I-108** building block dissolved at 5 mM in ammonium acetate 50 mM without adjustment of pH surprisingly did not lead to a cyclo-oligomerisation despite a structure similar with **I-105**. The HPLC signal disappeared in 24 hours. So, the two bulk units in **I-105** seemed determinant for the tetramer formation. A preparative DCL was performed to isolate the macrocycle **I-105₄** in ammonium acetate 50 mM without pH adjustment, with a satisfying yield of 71 %, comparable to the **I-1₄** synthesis.

The oligomerisation occurred for I-105, leading to a cyclic tetramer. This DCL was influenced by the pH but not by the nature of the aqueous solution.

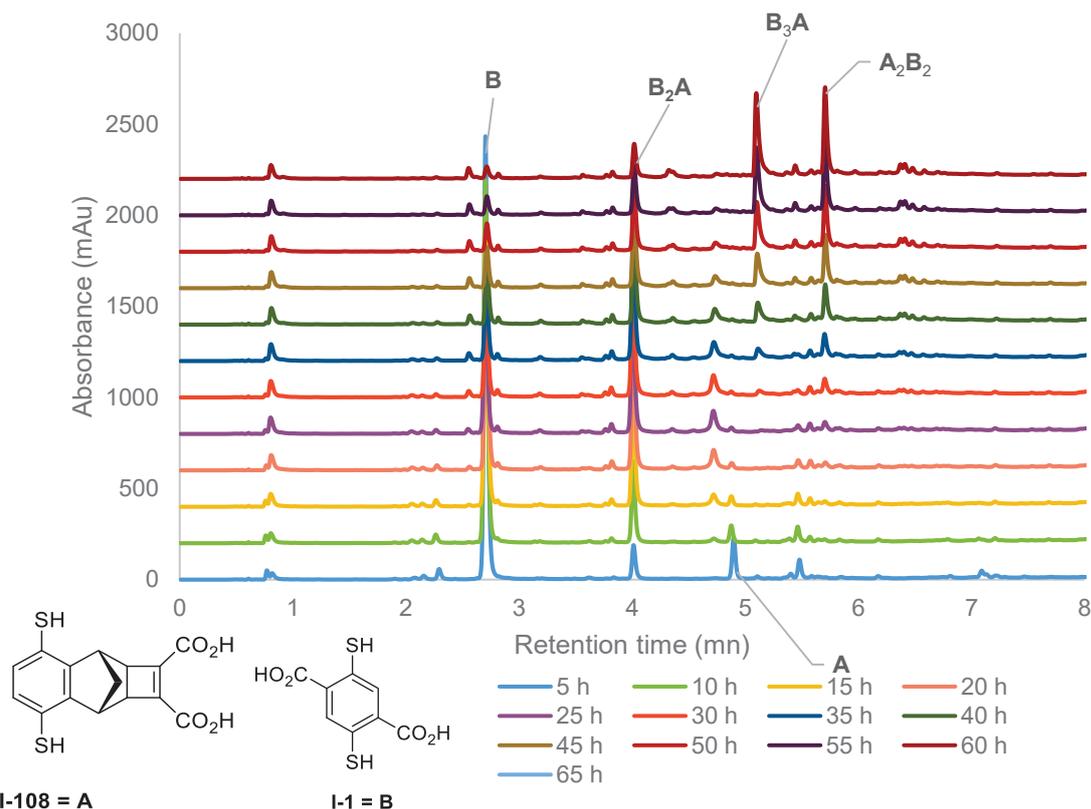
ii. DCLs of two building blocks simultaneously

The building blocks **I-1** and alternatively **I-105**, **1-106** and **1-108** were dissolved at 2.5 mM, mainly in ammonium acetate buffer 50 mM without pH adjustment (Scheme 68). The details are given in appendix (Appendix I-II: Additional conditions of reactions). Firstly, both DCLs with **I-105** and **I-106** did not contain hetero-oligomers with **I-1** or homo-oligomers. The HPLC signal of the building blocks disappeared in few hours. This could correlate with a possible polymerisation, as ESI-MS analysis did not show any evident clues of sulfenic or sulfenic acid formation. Other buffers (Tris 200 mM pH 7.4), oxidation of thiols with iodine¹¹² and the use of cationic amines as templates did not help to form heterocyclo-oligomers.

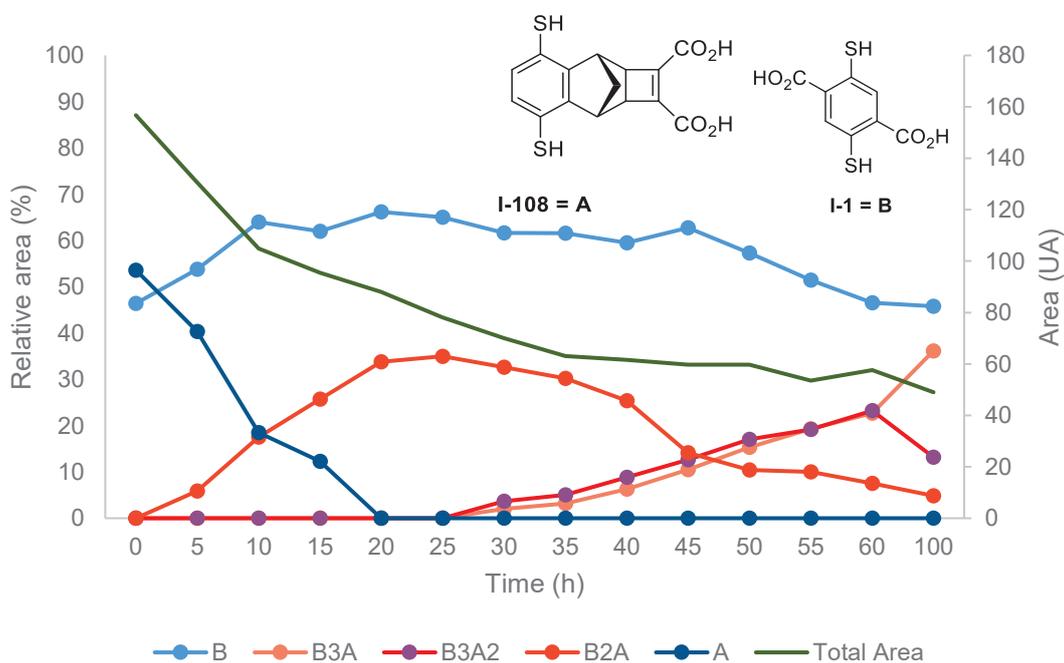


Scheme 68 : Building blocks used in the DCLs

Heterocyclo-oligomers between DCLs of the building blocks **I-108** and **I-1** dissolved at 2.5 mM in ammonium acetate buffer (50 mM without pH adjustment) were detected by LC-MS analysis. The kinetic data are presented below. The concentrations of the macrocyclic products formed in the DCL were obtained from their corresponding HPLC-UV areas after determining the effective absorption coefficients of the two starting building blocks thanks to the Beer-Lambert law. The absorption coefficient of the macrocyclic product was the relative sum of the coefficient of its components. The building block **I-108** was quickly consumed, while the conversion of **I-1** took nearly 48 h. Firstly, a cyclic heterotrimer **I-1₂I-108** was formed, followed by its disappearance into tetrameric forms. The DCL indeed evolved through the formation of cyclic **I-1₂I-108₂** and **I-1₃I-108** as main components. The decrease of the total area might be related to the formation of polymer as well as the formation of aggregate species, which was not verified. At longer times (one week), the HPLC signal was weak but was mainly composed of **I-1₃I-108**. No amplification of macrocyclic products was observed with cationic polyamines. The heterocyclo-oligomerisation did not occur in pure water in similar concentrations of building blocks.



Scheme 69: Evolution of **I-108 = A** and **I-1 = B** in ammonium acetate (50 mM, pH without adjustment), method L^{17} . X_nY_m disulfide macrocycles including n X units and m Y units



Scheme 70: Building block and macrocyclic population evolution of **I-108 = A** and **I-1 = B** in ammonium acetate (50 mM, pH without adjustment), method K^{18} , X_nY_m disulfide macrocycles including n X units and m Y units

¹⁷ The absorption coefficients for each building blocks are different.

¹⁸ The absorption coefficients for each building blocks were determined and used as a combination to obtain those of the corresponding macrocycles.

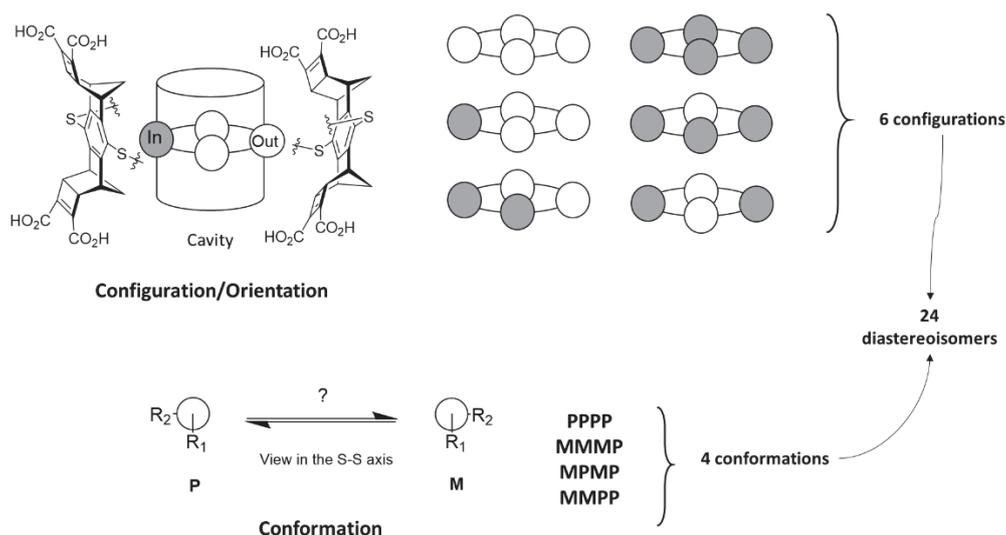
To conclude, efficient conditions of oligomerisation were found between **I-1**, **I-106**, **I-105** and **I-108**. The architecture of the building blocks appeared to be central. The DCLs of the diastereoisomers **I-105** and **I-106** had different compositions in similar conditions, as well as the DCL of the nearly identical **I-108**. Heterocyclo-oligomerisation occurred only between **I-1** and **I-108** in the tested conditions. The steric effects were suspected of being a major actor of the formation of **I-105₄**.

The combination of I-1 with the pair of bulk diastereoisomers into heterocyclo-oligomers was not possible. At the opposite, the I-108 version formed cyclic hetero-tetramers with I-1.

d. Structural properties of **I-105₄**

The aim of this subpart was to precisely assign the conformation and the configuration of the **I-105₄** cyclic tetramer by means of analytical techniques. The HPLC signals could account for several superposed traces of tetramers. In fact, up to 24 different cyclic tetramers of **I-105₄** could be formed during the self-assembly. The building block could orient itself inside or outside the cavity. Thus, a combination of 6 different orientations was possible regarding the position of the methylene bridge in relation with the macrocyclic cavity (Scheme 71). The configuration was named « In » if the methylene bridge pointed inside the cavity, and « out » if the methylene bridge pointed outside the cavity.

In addition, each disulfide bonds could adopt a specific conformation⁸³, P or M referring to the nomenclature⁸⁴. Consequently, 4 different conformations were possible for the disulfide bridges, PPPP, MMMP, MMPP and MPMP. Finally, there was no prochiral R/S configurational contribution from the disulfide bridges because the building block unit was symmetric⁸⁶. That made 24 possible diastereoisomers combined with the building block orientation, illustrating the molecular diversity of the disulfide bonds. The crowded architecture of the building block led to the hypothesis of rigid disulfide bonds, which could not change their conformation during the DCLs. This hypothesis could be eventually verified thanks to ¹H NMR experiments with different temperatures. ¹H NMR, ¹³C NMR, 2D NMR, crystallography and potentiometric titrations, supported by DFT calculations were combined in order to leave the veil on the stereochemistry of **I-105₄**.

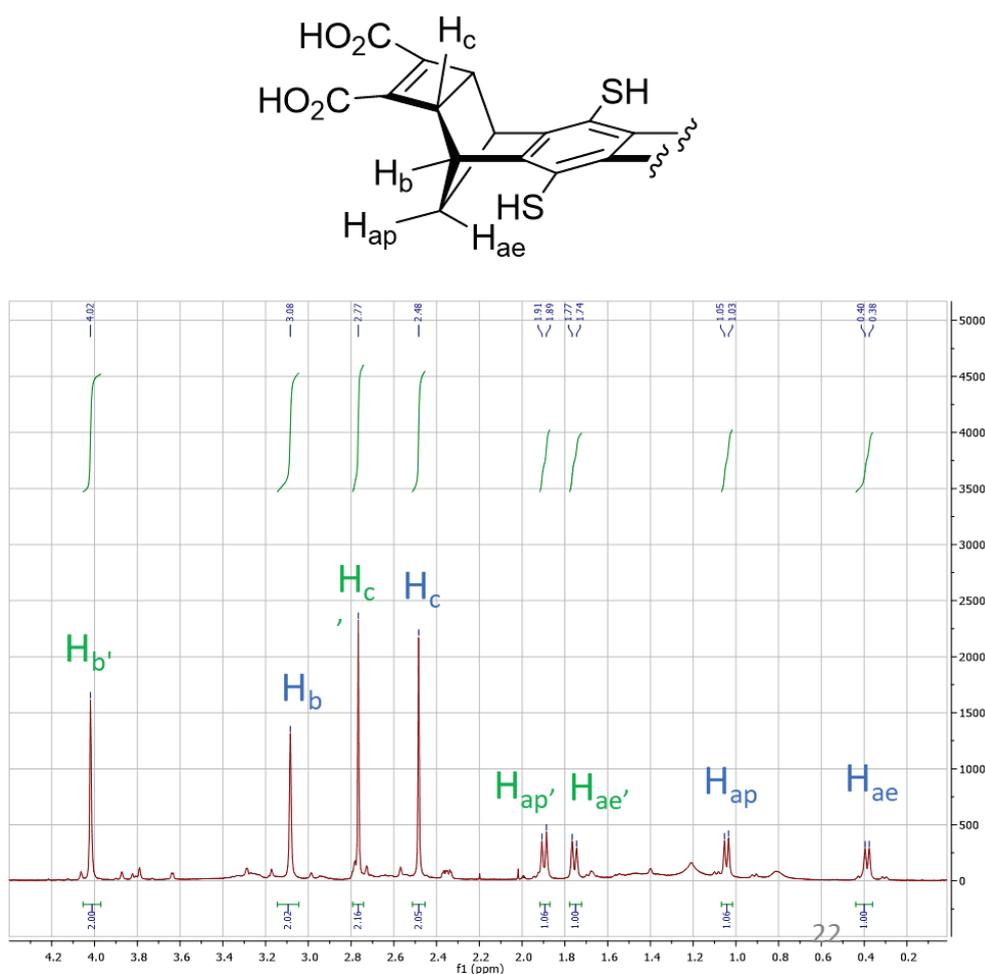


*Scheme 71 : Representation of the molecular diversity source in the case of the cyclic tetramer **I-105₄** obtained from spontaneous self-assembly. No configurational pro R/S contribution because of the building block unit symmetry, but contribution from the building block orientation inside the cavity and from the conformation of the disulfide bonds.*

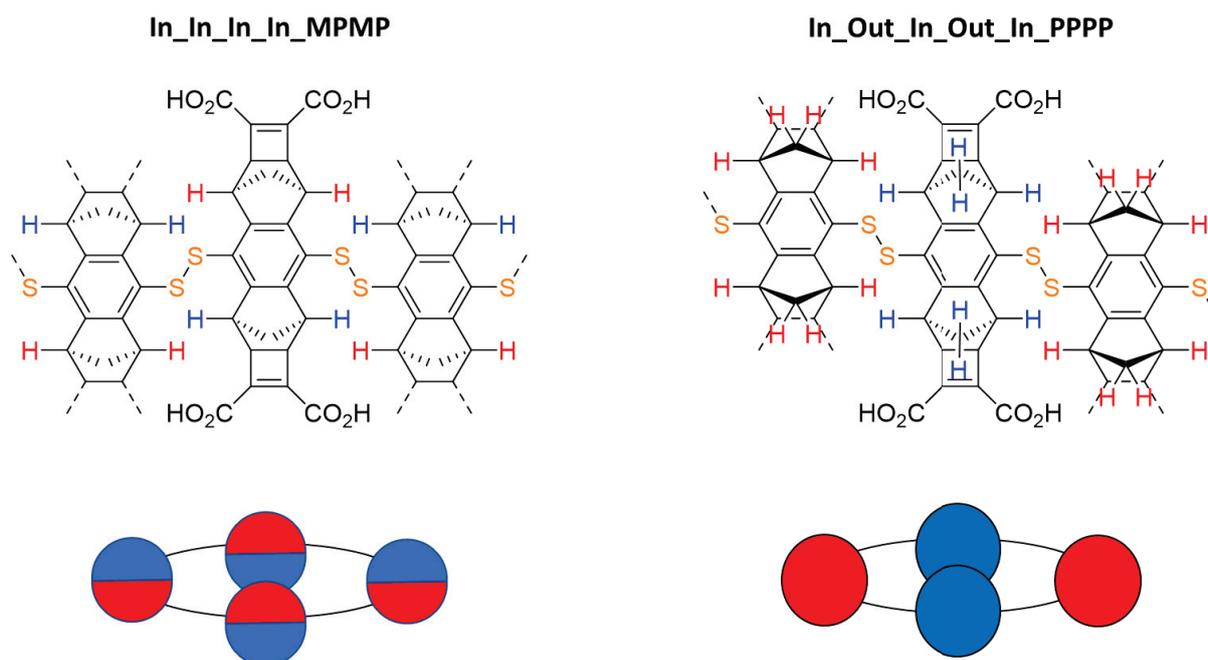
i. NMR analysis

^1H NMR spectrum of the precipitate revealed two similar sets of signals, each accounting for the same number of protons (Scheme 72). Consequently, the precipitate was composed of 1) one species with two sets of magnetically different protons, or 2) two symmetric species with only one set of protons, in an equal proportion. However, the last option was considered improbable given that there were no evident reasons to obtain two different macrocycles in the same proportions during the DCL.

The configuration and the conformation of the 24 possible diastereoisomers were classified following their symmetry. There were only two diastereoisomers that contained two magnetically different sets of protons. The two NMR-compatible diastereoisomer candidates are "In_In_In_In_MPMP" and "In_Out_In_Out_PPPP" (Scheme 73). For the first one, the two sets came from a simple symmetry breaking due to the configuration of the disulfide bond. Regarding the second candidate, the disulfide bonds had the same configuration but the alternate orientation of the faces was responsible for the two sets of signals. That led to the hypothesis of a frozen exchange of the disulfide bond configuration. That was coherent with the crowded cylindrical cavity of the macrocycle. 2D NMR experiments (HSQC, NOESY and HMBC) contributed to the attribution of each set of signals in ^1H and ^{13}C NMR. The detailed spectra with the analysis of the coupling was given in appendix (Appendix I-IX: NMR experiments of **I-105** / **I-105₄**).



Scheme 72 : ^1H NMR spectra of **I-105₄**, with the unit of interest. 2 identical sets of 4 different protons are visible. The macrocycle contains two magnetically different units.



Scheme 73 : Representation of the two configurations coherent with 2 magnetically different sets of protons.

The ^1H NMR control of the DCL leading to **I-105₄** indicated a conversion of 79 % of **I-105** into the unique diastereoisomer **I-105₄** in deuterated ammonium acetate at 50 mM (5 mM of **I-105**, no pH adjustment), which was an impressive selection. DOSY experiments were simultaneously performed and gave the hydrodynamic radius of the members of the DCL (Appendix I-IX: NMR experiments of **I-105** / **I-105₄**). The hydrodynamic radius of **I-105** was 5.71 ± 0.07 Å while the hydrodynamic radius of **I-105₄** was 9.08 ± 0.75 Å. This last value was compared to the external and hydrodynamic radius of the optimised structures of the diastereoisomer **In_In_In_In_MPMP**. The modelled external radius was 8.30 Å and the hydrodynamic radius was 10.40 Å. The computed and experimental hydrodynamic radius were in good agreement. So far, we have not obtained a crystallographic structure of **I-105₄** which can be compared to the modelling and this experimental data.

The precipitate was composed of only one macrocycle. Regarding its configuration, two diastereoisomers are NMR compatible over the 24 possibilities.

ii. Crystallographic study

Samples of the macrocycle were sent to the Ivan Huc's group at the University of Munich. The team has a knowledge in crystallisation of macrocyclic compounds. Different sets of crystals were obtained by our partner¹⁹. The first set were grown in the following media: 1) KH_2PO_4 , NaH_2PO_4 , HEPES (pH 7.5). 2) $\text{NH}_4\text{H}_2\text{PO}_4$, tri-Sod. citrate (pH 5.6) ; 3) $\text{NH}_4\text{H}_2\text{PO}_4$, TRIS (pH 8.5), 4) K_2HPO_4 , NaH_2PO_4 , HEPES (pH 7.5). However, the resolution of the structure was not possible. A second set of crystals were grown in $(\text{NH}_4)_2\text{SO}_4$, Glycerol, TRIS (pH 8.0), Anderson–Evans polyoxotungstate $[\text{TeW}_6\text{O}_{24}]^{6-}$ and were resolved on the Synchrotron of Geneva. Surprisingly, the structure elucidation and preliminary refinement indicated an octamer instead of tetramer (Figure 13). We hypothesised that under slow evaporation of the media, two tetramers reacted by metathesis of disulfide bond, leading to an octamer which

¹⁹ Experimental work by Dr. Pradeep Mandal

crystallised. The structure could not give the orientation of the faces, but the disulfide configuration could be observed. It was MPPPMPPP. That was coherent with the hypothesis of a In_In_In_In_MPMP. In fact, the disulfide metathesis proceeds by a S_N2 -like mechanism, inverting the configuration of the disulfide bonds reacting. Regarding our supposed tetramer, the configuration MPMP + MPMP became MPPPMPPP (Scheme 74). A modelling by DFT of the process was not obtained but could be optimised in the future in order to observe the inversion of the configuration and the energetic cost of the metathesis.

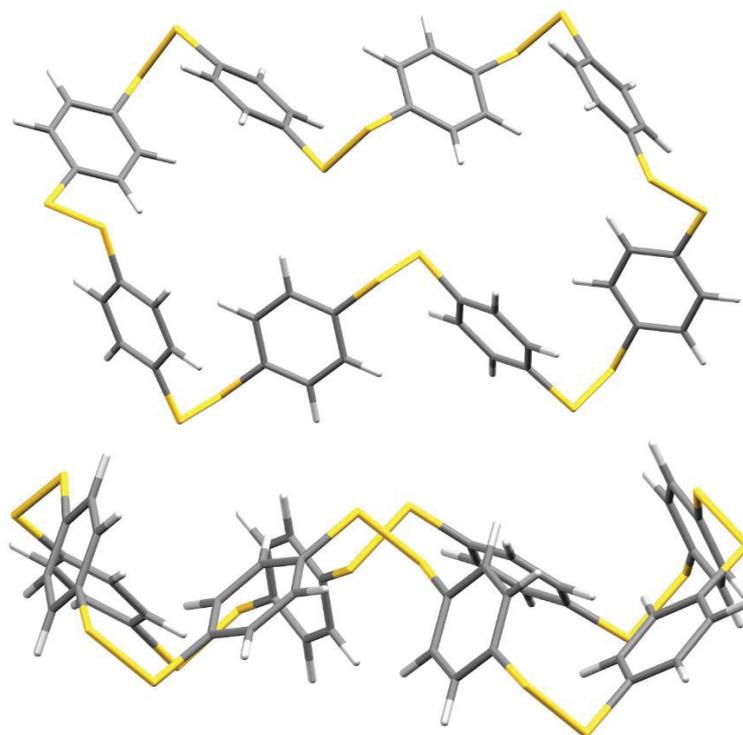
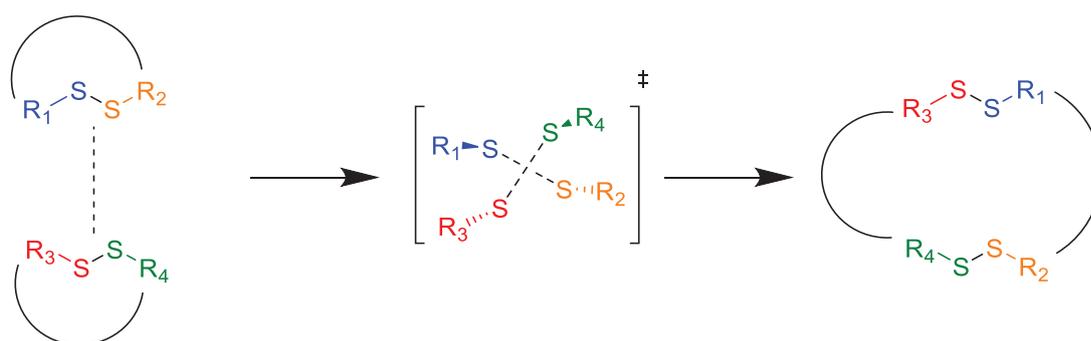


Figure 13: View (Mercury¹³²) of the octameric crystallographic structure, with oxygen on the aromatic rings for clarity. The two views are centred on the two disulfide bonds resulting from the metathesis.



Scheme 74 : Dimerisation of two **I-105₄** upon crystallisation conditions by a metathesis of disulfide bonds with inversion of disulfide configuration.

A third set of crystals were obtained by a stoichiometric dissolution of **I-105₄** and spermine **I-8**. The amine should help to obtain better resolution of the crystals and an efficient crystallisation of the tetramer. They are currently resolved thanks to the Synchrotron of Geneva. Unfortunately, we had not

solved structures now. The crystallographic structure will be determinant to assess definitely the configuration in the solid state, combined with NMR and modelling.

iii. pH titration

The macrocycle was titrated with an excess of HCl. The titration curves are given in appendix. Two pK_a were determined at 3.61 and 6.71, suggesting two distinct acidities. That was coherent with the symmetric structure involving a stabilisation of one carboxylate by an internal hydrogen bond. The pK_a of the similar moiety maleic acid was coherent¹⁹². The pH was 4.69 at the end of the DCL giving the macrocycle. Therefore, the protonation state was 50 %, i.e. – 8. This was an important parameter for DFT modelling.

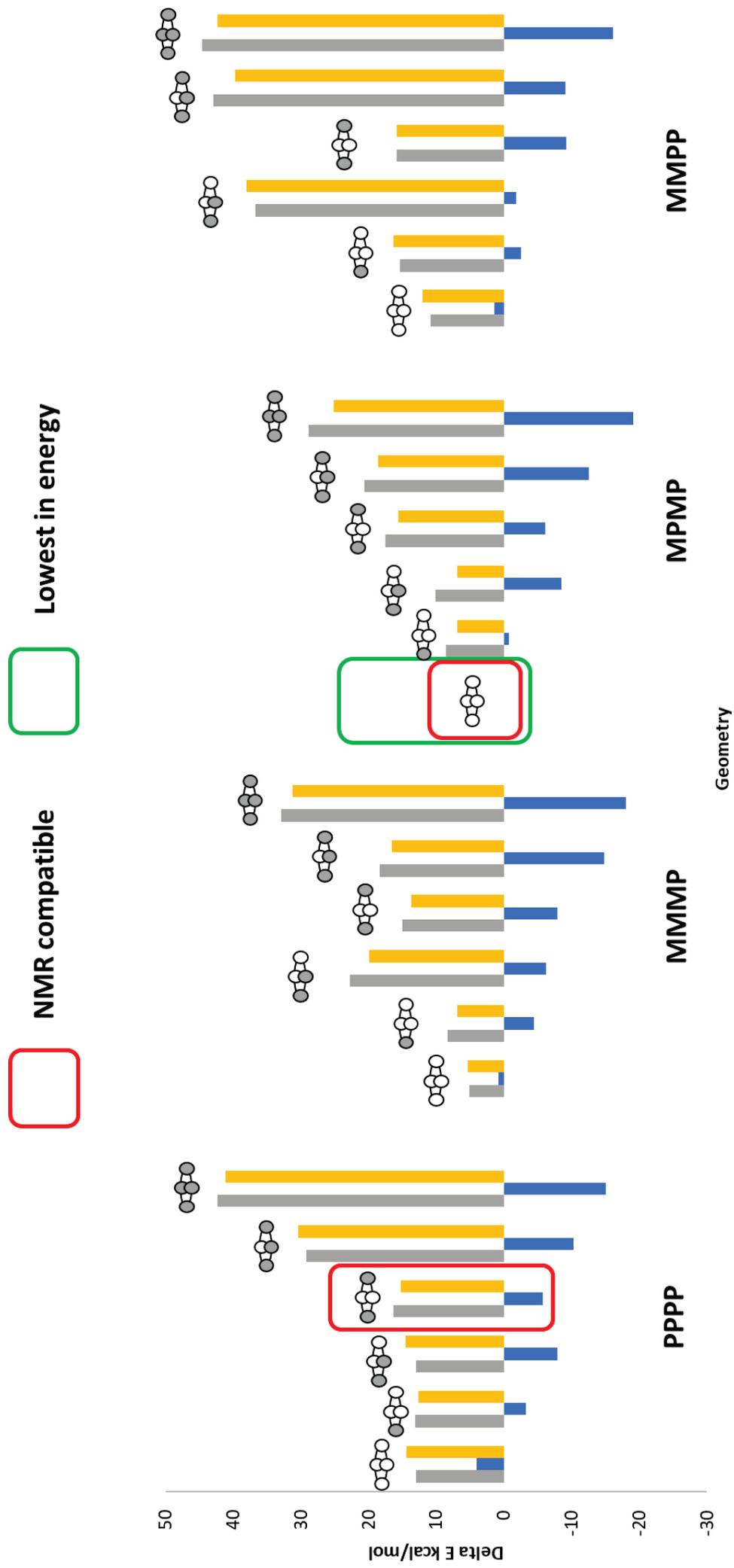
The macrocycle had two pK_a and was half deprotonated (-8) at the end of its DCL.

e. DFT optimisation of the diastereoisomers

The aim of this subpart was to confront the previous experimental assignment of the stereochemistry of **I-105₄** and to study the contribution of London forces to the self-assembly by a computational modelling. A DFT study was conducted in order to predict the most stable diastereoisomer of **I-105₄** including the contribution of dispersion. In fact, DFT could provide information on the possible driving forces of the self-assembly, including the dispersion, as previously discussed (Computational description). The 24 diastereoisomers were built thanks to the GaussView5 software¹⁹³ and optimised with the B3LYP functional⁷². Then, a single point (SP) energy calculation gave the energy of the system. The B3LYP functional did not contain a dispersion term. Thus, the optimised geometries were further submitted to a SP calculation with two additional functionals. The B3LYP was corrected with the D3 version of Grimme's dispersion with Becke-Johnson damping¹⁹⁴ (level 2 Computational description). That enabled to extract the dispersion contribution by difference with the B3LYP energy. Also, a recent parametric functional, M062X^{147a} (level 1 Computational description), was used, containing a description of the dispersion and not an empirical extra term.

A first optimisation with B3LYP-D3BJ was tested. However, the structures were highly deformed and were not acceptable. The dispersion forces were probably overestimated with this functional given that the systems contained a large part of alkyl groups. That resulted in a tendency to shorter the distance between the units of the macrocycle, inducing an important strain on the disulfide bonds. This phenomenon was observed in other system, adopting a particular folding following the B3LYP or B3LYP-D3 optimisation¹⁸⁵. That is why the structures were first optimised with B3LYP which gave coherent structures in terms of symmetry and strain. The computational details are given in the experimental section (DFT optimisation of the diastereoisomers).

The full data are gathered in the appendix (Appendix I-XII: DFT data for **I-105₄**). The summary of the optimisation and energy calculations are presented below (Scheme 75). The energy E was recorded, not the Gibbs free energy G. The calculation of G was costly given the number of atoms and the level of theory and was not performed. The study could be based on E, even though G would be more precise. The diastereoisomer In_In_In_In_MPMP was the lowest in energy with the three functionals. The difference of energy for each diastereoisomer compared to the In_In_In_In_MPMP is presented. The dispersion difference is referenced with In_In_In_In_MPMP for clarity. The different functionals gave coherent results.



Scheme 75: Energies and dispersion contribution of the 24 diastereoisomers of I-105₄, relatively to the lowest in energy: In_In_In_MPMP. The structures were optimised with B3LYP and then single point calculations with B3LYP-D3BJ, B3LYP and M062X gave the energies. Grey: out orientation, white: in orientation

For each disulfide bond geometry, the same tendency could be observed. The all « in » configuration was the lowest in energy over the 6 possibilities. Then, the energy increased for each additional « out » unit in the macrocycle until the all « out ». Regarding the alternate In_Out_In_Out, it broke the tendency and was higher in energy than the In_In_In_Out, In_In_Out_Out and the In_Out_Out_Out geometries. This singularity was observed for all disulfide bonds geometries apart from MPMP. Consequently, a methylene bridge pointing inside the cavity increased the system energy systematically. Thus, the alkyl groups tended to interact unfavourably inside the cavity with the neighbouring other bridges and 4-membered rings. Thus, the steric effects were mainly repulsive. The dispersion attraction followed the opposite trend than the energy. The contribution was more and more stabilising following the addition of « in » units inside the cavity. That could be explained by the increasing steric effects inside the cavity compared to fewer steric effects outside the cavity. Even though these effects increased for each methylene bridges pointing inside, the attractive dispersion did not take the top on the Pauli repulsion. To sum up, the predicted most stable geometry was In_In_In_In_MPMP, which did not show the most stabilising London dispersion contribution. That disqualified unambiguously the hypothesis of the dispersion as driving force for that self-assembly, upon the present calculation level.

The upper level of dispersion description could be tested, such as *VdW-DFT* or *DHDF*. However, the systems contain a high number of atoms, so the calculation cost will be too expensive. Several computational methods could give more information, such as a recently developed universal quantitative descriptor of the dispersion interaction potential, producing a London surface map¹⁹⁵. However, the method cannot be applied for systems with more than 100 atoms. A decomposition of the energy will provide useful data, for instance with the recent local energy decomposition at the domain-based local pair natural orbital level of theory¹⁹⁶. Finally, the NCI plot could also be relevant to explain the selectivity for this self-assembly¹⁹⁷. Those data are coherent with the NMR analysis. As a matter of fact, the most stable diastereoisomer corresponded to one of the two possibilities of configuration dictated by the symmetry of the signals in NMR. Based on the computational study and the NMR analysis, the precipitated species must be the In_In_In_In_MPMP (Figure 14). Nevertheless, the crystallographic structure will be last element to validate this structure.

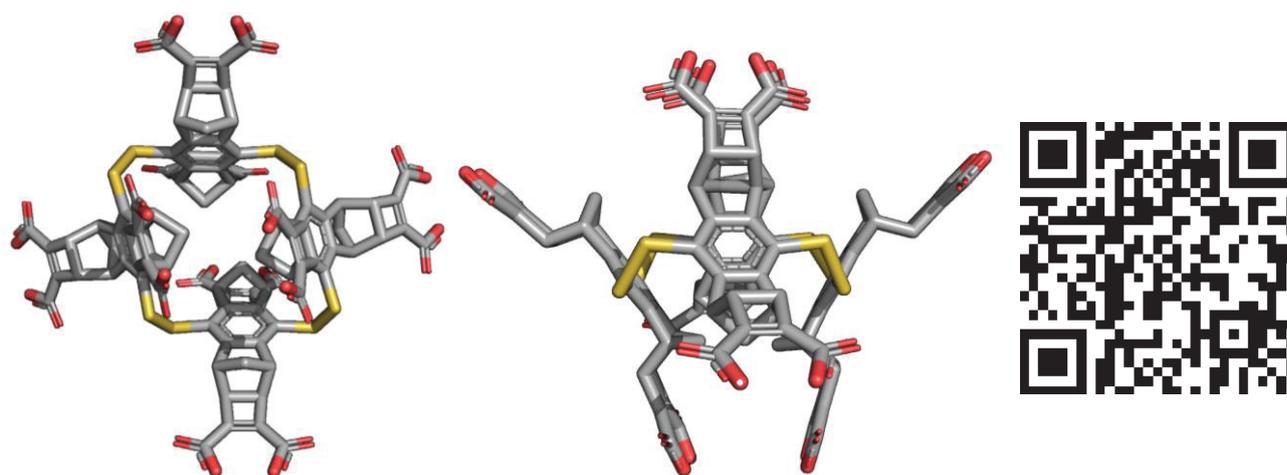
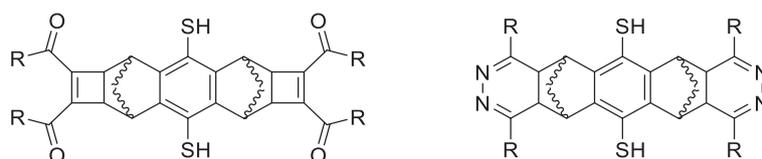


Figure 14 : Pictures of the the diastereoisomer In_In_In_In_MPMP of I-1054 from DFT optimisation (Pymol Software¹⁹⁸)

The most stable geometry was predicted to be the *In_In_In_In_MPMP*. The energy increased from a all “in” to a all “out” geometry. The steric effects were dominated by steric repulsion rather than London attraction.

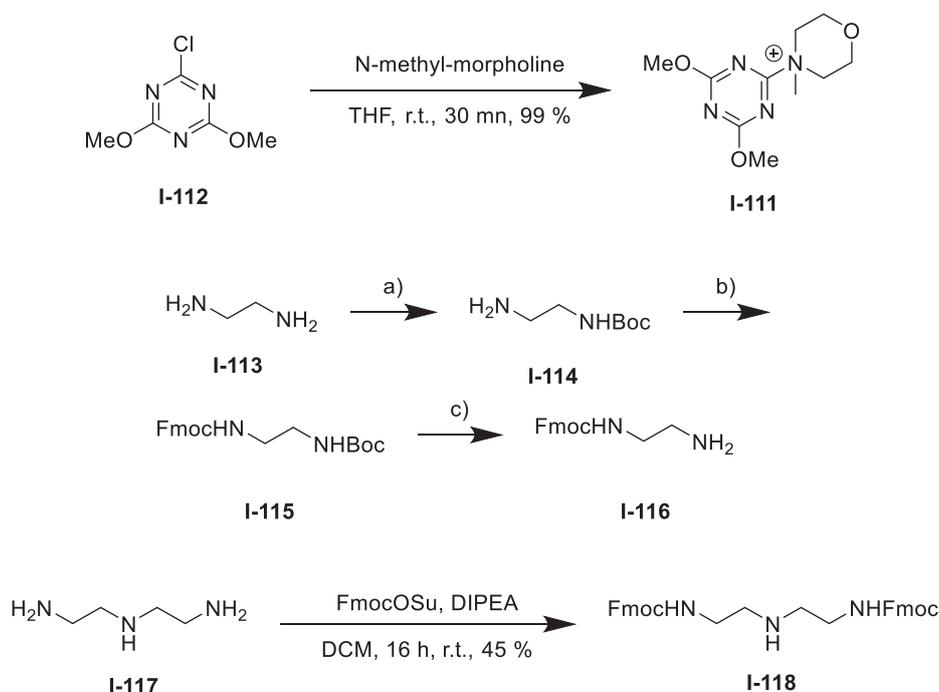
- f. Modification of the external part and of the functionalisation of the sterically crowded 1,4-bisthiophenol

The aim of this subpart was to find a synthetic approach in order to obtain variations on the chemical functions of the sterically crowded building blocks (Scheme 76).



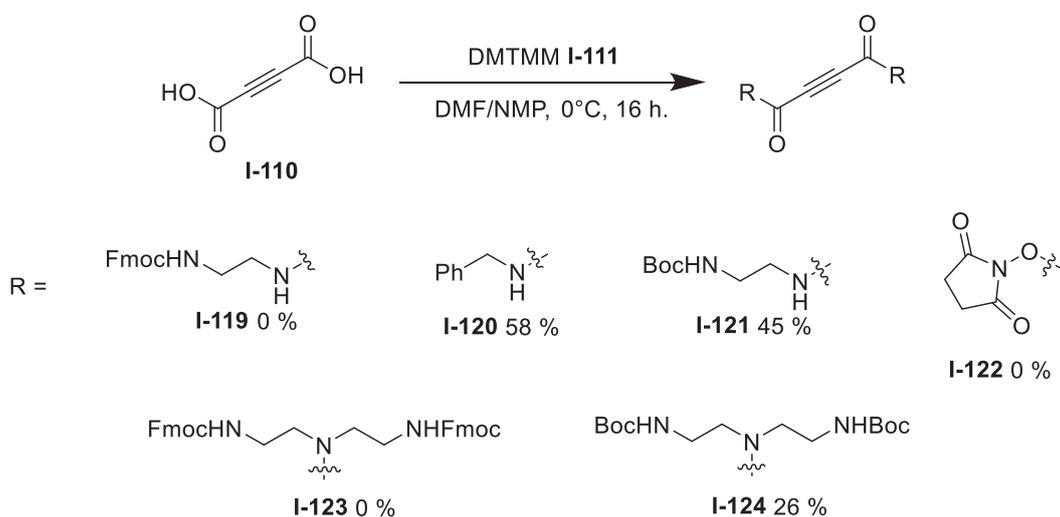
Scheme 76: Desired modifications on the bulk groups and the chemical functions. *R* any chemical functions.

Thus, the synthesis of **I-105** analogues was proposed in order to vary the chemical functions bore by the building block. The first approach consisted of the use of a derivative of acetylenedicarboxylic acid **I-110** instead of dimethylacetylenedicarboxylate for the [2 + 2] cycloaddition (Scheme 62). Protected amines were chosen to be added to acetylenedicarboxylic acid **I-110** thanks to an amide bond. Indeed, a scope of different acetylenedicarboxamides was described in the literature¹⁹⁹. The DMTMM coupling agent **I-111** was freshly prepared as recommended²⁰⁰, with a satisfying yield of 99 % (Scheme 77). Fmoc-protected amine hydrochloride **I-116** was synthesised in 3 steps with a satisfying global yield 75 % from diethylene amine²⁰¹ (Scheme 77). **I-118** was synthesised from **I-117**²⁰² with a satisfying yield of 45 % (Scheme 77).



Scheme 77 : Preparation of the DMTMM coupling agent **I-111** and protected amines. a) Boc_2O , CHCl_3 , 0°C , 16 h., 99 %; b) FmocO-Su , NaHCO_3 , H_2O / MeCN , r.t., 16 h., 96 %; c) HCl , H_2O / AcOEt , r.t., 4 h, 79 %.

The amide bond formation with acetylenedicarboxylic acid **I-110** was then explored. Few products of reaction between the diacid and the amines were obtained (Scheme 78). The reaction between benzylamine and **I-110** gave the corresponding derivate **I-120** with acceptable yield of 58 (reported 77 %¹⁹⁹). The amide bond formation with benzylamine, HOBt and EDCI²⁰⁰ was not efficient, highlighting the importance of DMTMM **I-111**. The product **I-122** from the reaction with 1-hydroxypyrrolidine-2,5-dione was not observed. Interestingly, the product **I-123** of the reaction between **I-110** and **I-116** was not detected while the Boc analogue **I-114** led to **I-121** with a yield of 45 %. The product of the reaction between the **I-118** and **I-110** was not observed, while the product between **I-110** and di-tert-butyl (azanediylbis(ethane-2,1-diyl))dicarbamate was obtained with a yield of 26 %. For all the substrates, the conversion was total, as attested by the disappearance of the ¹³C NMR alkyne signal of **I-110** and a shift of this signal in the region of C=C double bond, suggesting a Michael's addition instead of an amide bond formation. Other coupling agents were employed as well as other solvents (HCTU, DMF²⁰³; EDCI, MeOH²⁰⁰; HOBt, EDCI, MeOH²⁰⁰/DMF²⁰⁰, PyBrop²⁰⁴) without obtaining more desired products. The use of a reactive acyl chloride²⁰⁵ or succinimide²⁰⁴ intermediate also failed to enlarge the scope of acetylenedicarboxamides.

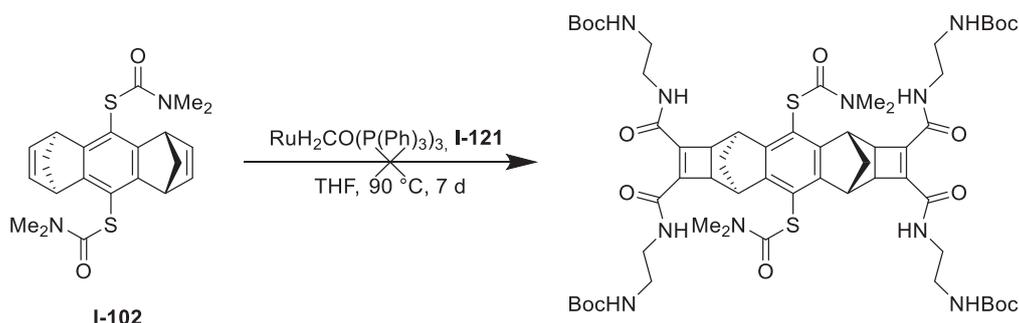


*Scheme 78: Derivatisation of acetylenedicarboxylic acid by amide bond formation with various nucleophiles and DMTMM agent **I-111***

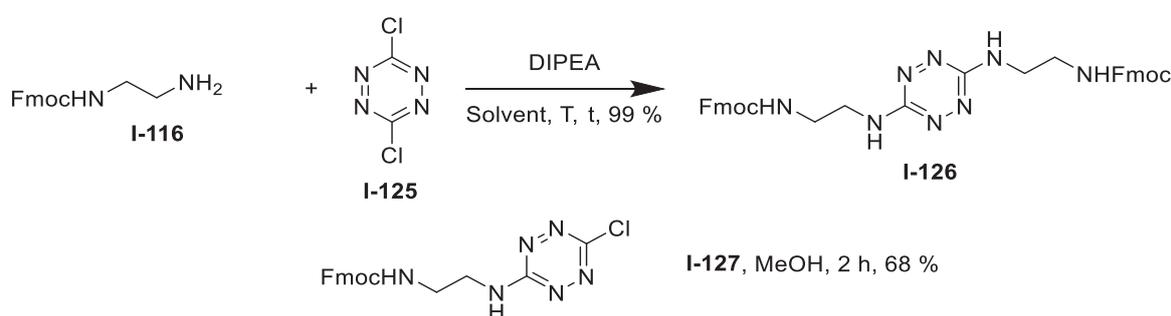
The product of the ruthenium catalysed cycloaddition¹⁹¹ between **I-102** and **I-121** was not observed (Scheme 79). Indeed, no conversion was observed for both starting materials after 7 days at reflux, and higher temperature degraded the acetylene derivative. A microwave activation or a photocatalysis could be an alternative solution.

Another derivatisation plan was explored, involving tetrazines. In fact, they are efficient partners for [2 + 2] cycloadditions²⁰⁶. The synthesis of substituted 1,2,4,5-tetrazines from **I-125** (synthesis²⁰⁷ in chapter 2) was tested by S_{Nar} between amine **I-116** and **I-125**²⁰⁸ (Scheme 80). Different conditions were employed (acetonitrile, DMF, methanol, r.t., reflux, 16 h., 2 h.). A mono-substituted product **I-127** was isolated in 2 h at r.t. in methanol with an excellent yield of 68 %. At longer times or higher temperature (reflux, 16 h.), the Fmoc protecting group was cleaved partially (ESI-MS analysis) and the tetrazine moiety was degraded by the reaction conditions (¹³C NMR), even though there are many examples in the literature for similar substitutions²⁰⁸. Other conditions and a careful control of the reaction

composition could be a solution. The direct synthesis of substituted tetrazines from hydrazine could be explored alternatively²⁰⁹.



Scheme 79 : Ruthenium catalysed [2+2] cycloaddition of **I-121** with **I-102**



Scheme 80: Substitution of **I-125** with **I-116**

The modification of the chemical functions of **I-105** encountered synthetic issues.

g. Conclusion

The synthesis of building blocks with distant carboxylates and thiols moieties including a bulk separation was achieved. The building blocks did not self-assemble through DCC similarly with **I-1** in hetero-oligomers. In addition, two diastereoisomers showed opposite self-assemblies through DCC in aqueous solutions. Especially, **I-105** formed a single cyclic tetramer over all the oligomers and over the 24 accessible macrocyclic configurations. The assignment In_In_In_In_MPMP resulted from a combination of NMR and DFT study and will be confirmed by a crystallographic structure, currently under resolution. Under some conditions of crystallisation, an unexpected metathesis of disulfide bonds converted two **I-105**₄ into an octamer with a disulfide configuration MPPPMPPP. That impressive selection during the self-assembling was not directed by attractive London forces but by repulsive steric effects. The modification of the chemical functions of **I-105** met synthetic issues but optimisation of the Ru-catalysed [2+2] could address the problem.

6. Conclusion

The exploration of different polycarboxylate disulfide bridge cyclophanes led to fundamental and applied study, validating their ability to develop theoretical concept but also various applications. The polyanionic macrocycle **I-1**₄ had high affinity for polycationic amines in solutions with K_a up to 10^{11} M^{-1} and biological applications currently under development. The selectivity of the receptor for α,ω -alkyldiammoniums was governed by the solvation of the complexes and by the desolvation of the

guests. Regarding the binding of **I-14** with spermine **I-8** and spermidine **I-14**, the solvation effects are also discriminant in the selectivity for spermidine **I-14**, even though the exact picture is not yet depicted. Spermidine **I-14** is less charged but was more mobile in the complex, which contributed to the selectivity.

The formation of heterocyclo-oligomers that are mutants of **I-14**, for both fundamental and applied studies partially failed for the modification of the rims, notably their chemical functions. Nevertheless, a successful expansion of the diameter of the cavity was obtained by the self-assembly through DCC of a benzylic analogue of **I-1**. The building block revealed different schemes of self-assembly following the oxidant (O_2 vs I_2). Discrete macrocycles were observed, and a strained dimer was isolated and crystallised thanks to iodine. The distribution of macrocycles was influenced by templating. The cyclic dimer **I-15₂** was too small to accommodate a guest but was used as a model of post-functionalisation of the carboxylic acids. A further modification of the link between the aromatic units, turning disulfide bridges into thioether was observed but could not be managed.

Finally, the synthesis of building blocks with distant carboxylates and thiols moieties including a bulk separation was achieved. That resulted in a possible expansion of the height of the macrocyclic cavity. Two diastereoisomers showed opposite self-assemblies in aqueous solutions. Especially, a single cyclic tetramer over all the oligomers and over the 24 accessible macrocyclic configurations was obtained by a spectacular selective self-assembly of **I-105**. The assignment In_In_In_In_MPMP resulted from a combination of experimental and computational study and will be confirmed by a crystallographic structure. Under conditions of crystallisation, an unexpected metathesis of disulfide bonds converted two **I-105₄** into an octamer with a disulfide configuration MPPPMPPP. The driving force of the self-assembly was first supposed to be the London forces, which was invalidated by DFT. The assembly was probably governed by repulsive steric effects. The modification of the chemical functions of this macrocycle met synthetic issues. During the exploration of all those backbones, the iodine gave contrasted results, between efficient oligomerisation and polymerisation promoters. Water-soluble cyclophanes will continue to play a major role in molecular recognition and fundamental studies, as highlighted by the investigation of solvation and steric effects. Self-assembly provided efficiently that family of receptors, which are promising platform for different applications.

Experimental section

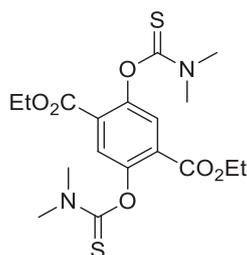
a. Synthetic procedures and characterisations

Only new compounds are fully characterised. Otherwise, only ^1H and ^{13}C NMR were recorded. "Match with literature" means a reference for the compound spectra. The protocol references are given in the main text. Adaptations and optimisations were done without specific mention. The new compounds were fully characterised by HRMS, 1D, 2D NMR, and mp. The spectral data are described after their synthetic procedure. The corresponding spectra are not given in appendix for conciseness. They were compiled on previous reports and are electronically available. If a molecule was not fully characterised, for instance with macrocycles, spectral data could be delivered in appendix to support the discussion.

i. Materials and methods

All commercial reagents were used as received. Solvents were dried using conventional methods. All reactions were carried out under nitrogen or argon. Column chromatography was performed using silica gel (0.040–0.063 nm). Reactions were monitored via TLC on a silica gel plate and visualised under UV light. ^1H NMR and ^{13}C NMR spectra were recorded on a Bruker DRX at 400 MHz or on a Bruker ALS at 300 MHz. J values are given in Hz. Mass spectra were acquired on an LCQ Advantage ion trap instrument, detecting positive or negative ions in the ESI mode. Melting points were determined by a Büchi Melting point B-540.

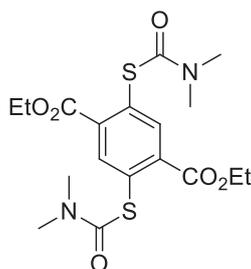
ii. Preparative procedures and characterisations



I-6 diethyl 2,5-bis((dimethylcarbamothioyl)oxy)terephthalate

To a solution of **I-5** diethyl 2,5-dihydroxyterephthalate (10 g, 39 mmol, 1 eq) in dry DMA (100 mL) was added dropwise at 0°C a solution of dimethylcarbamothioic chloride (19 g, 157 mmol, 4 eq) and DABCO (18 g, 157 mmol, 4 eq) in dry DMA (50 mL). The mixture was stirred under nitrogen at room temperature for 24 h. The precipitate was filtrated and fwashed extensively with water (4* 250 mL). The solid was dried under vacuum to give a white powder (17 g, 99 %).

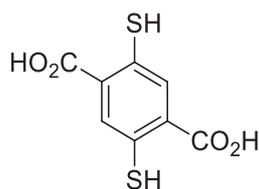
^1H NMR (300 MHz, CDCl_3) δ ppm = 7.74 (s, 2 H, Ar), 4.31 (q, J = 7.2 Hz, 4 H, CH_2), 3.47 (s, 6 H, NCH_3), 3.41 (s, 6 H, NCH_3), 1.34 (t, J = 7.2 Hz, 6 H, CH_3) data matched with literature reference¹⁶



I-7 diethyl 2,5-bis((dimethylcarbamoyl)thio)terephthalate

I-6 diethyl 2,5-bis((dimethylcarbamothioyl)oxy)terephthalate (3 g, 6.88 mmol) was heated under neat conditions at 210°C for 2 h. The solid was cooled at 90°C. Then, ethanol was added (90 mL) and the mixture was cooled at 0°C overnight. The precipitate was washed with cooled ethanol (2*45 mL) and dried under vacuum to give a beige powder (2.46 g, 82 %).

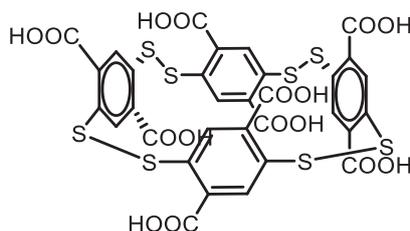
¹H NMR (300 MHz, CDCl₃) δ ppm = 8.09 (s, 2 H, Ar), 4.32 (q, J = 7.2 Hz, 4 H, CH₂), 3.10 (br. s., 6 H, NCH₃), 2.99 (br. s., 6 H, NCH₃), 1.34 (t, J = 7.2 Hz, 6 H, CH₃) data matched with literature reference¹⁶



I-1 2,5-dimercaptoterephthalic acid

A solution of **I-1** diethyl 2,5-bis((dimethylcarbamoyl)thio)terephthalate (2.6 g, 6.06 mmol, 1 eq) in degassed 1.3 M KOH (5.8 g, 104 mmol, 17 eq) in EtOH/H₂O (1:1, 80 mL) was refluxed under an inert atmosphere for 3 h. The reaction mixture was cooled in ice, and concentrated HCl (15 mL) was added until pH 1. A bright yellow precipitate was formed, filtered, and washed extensively with water, yielding compound **A** as a yellow solid (1.215 g, 95 %).

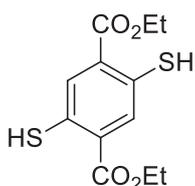
¹H NMR (300 MHz, MeOD) δ = 8.06 (s, 2 H, Ar) data matched with literature reference¹⁶



I-14 (2,3,8,9,14,15,20,21-Octathiapentacyclo[20.2.2.24,7.210,13.216,19]dotriaconta-4,6,10,12,16,18,22,24,25,27,29,31-dodecaene-5,11,17,23,25,27,29,31-octacarboxylic acid

A solution of **I-1** 2,5-dimercaptoterephthalic acid (230 mg, 1 mmol, 1 eq) and spermine (50.6 mg, 0.25 mmol, 0.25 eq) in Tris [200 mM] buffer pH 7.4 (3.05 g in 100 mL of milliQ water) was stirred under air for 24 h. Then, TFA was added until pH 1. The resulting precipitate was filtrated, and dissolved in a minimum amount of borate buffer [200 mM] pH 8.3 and stirred for 5 mn. Then, TFA was added until pH 1. The sequence was done three times. Then, the solid was washed with milliQ water three times, and dried under vacuum to yield yellow powder (185 mg, 80 %).

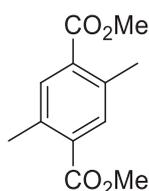
¹H NMR (300 MHz, D₂O, NaOD, pH 7.4) δ = 7.98 (s, 8 H, Ar) data matched with literature reference¹⁶



I-11 diethyl 2,5-dimercaptoterephthalate

A 0.75 M solution of freshly prepared EtONa in distilled EtOH (642 mg of Na, 27.96 mmol, 12 eq, in 37 mL of EtOH) was added to diethyl 2,5-bis((dimethylcarbamoyl)thio)terephthalate (1 g, 2.33 mmol, 1 eq). The mixture was heated under reflux for 1 h 30) under N₂, and then cooled to 0°C. A 1 M solution of HCl was added until pH 1. The resulting precipitate was filtrated and washed with chilled water. The solid was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 7:3) to give a yellow solid (147 mg, 22 %).

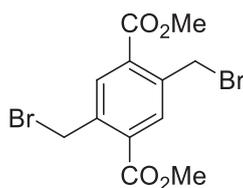
Yellow solid; mp: 132.7-134.8 °C; HRMS (ESI) [M + Na]⁺ found 309.0222, calculated 309.0226 for [C₁₂H₁₄NaO₄S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.96 (s, 2 H, Ar), 4.69 (s, 2 H, SH), 4.42 (q, J = 7.15 Hz, 4 H, CH₂), 1.44 (t, J = 7.15 Hz, 6 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.5 (CO), 133.5 (C_{Ar}-CO), 133.3 (C_{Ar}), 129.4 (C_{Ar}-S), 61.9 (CH₂), 14.2 (CH₃)



I-32 dimethyl 2,5-dimethylterephthalate

To a suspension of **I-31** 2,5-dimethylterephthalic acid (800 mg, 4.12 mmol, 1 eq) in dry methanol (12.5 mL) was added dropwise at 0°C thionyl chloride (1.25 mL, 17.1 mmol, 3.5 eq). The mixture was heated at reflux and stirred overnight under inert atmosphere. After cooling, the resulting solid was washed with methanol (10 mL) and twice with hexane (2*10 mL).The solid was dried under vacuum to give a white powder (915 mg, 99 %).

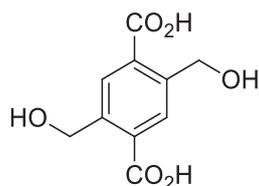
White solid; mp: 113.7-115.1 °C; HRMS (ESI) [M + H]⁺ found 223.0954, calculated 223.0965 for [C₁₂H₁₅O₄]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.76 (s, 2 H, Ar), 3.91 (s, 6 H, COOCH₃), 2.57 (s, 6 H, Ar-CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 167.5 (CO), 137.0 (C_{Ar}-CH₃), 133.5 (C_{Ar}-CO), 132.4 (C_{Ar}H), 52.0 (OCH₃), 20.9 (Ar-CH₃)



I-33 dimethyl 2,5-bis(bromomethyl)terephthalate

A solution of **I-32** dimethyl 2,5-dimethylterephthalate (600 mg, 2.70 mmol, 1 eq), NBS (1.05 g, 6 mmol, 2.2 eq), AIBN (36 mg, 0.220 mmol, 0.08 eq) in dry DCM (7.5 mL) was heated at reflux until conversion of starting material. The reaction is monitored by NMR control (approximately 16 h). After cooling, the organic phase was washed with a solution of NaHCO₃ (2*10 mL) and brine (2*10 mL). The organic phase was dried with MgSO₄ and evaporated to dryness. The crude mixture was purified by column chromatography (SiO₂, cyclohexane : DCM = 1:1) to give a white powder (591 mg, 86 %).

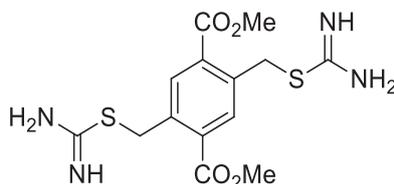
White solid; mp: 109.3-111.7 °C; HRMS (ESI) [M + Na]⁺ found 400.8989, calculated 400.8995 for [C₁₂H₁₂Br₂NaO₄]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 8.06 (s, 2 H, Ar), 4.93 (s, 4 H, CH₂), 3.98 (s, 6 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.7 (CO), 139.4 (C_{Ar}-CH₂), 134.4 (C_{Ar}H), 132.3 (C_{Ar}-CO), 52.7 (CH₂), 30.0 (CH₃)



I-40 2,5-bis(hydroxymethyl)terephthalic acid

To a solution of **I-33** dimethyl 2,5-bis(bromomethyl)terephthalate (500 mg, 1.32 mmol, 1 eq) in THF (5 mL) was added 5 mL of an aqueous solution of potassium hydroxide (390 mg, 6.96 mmol, 5 eq). The biphasic mixture was vigorously stirred at room temperature for 24 h, and then acidified with concentrated HCl until pH 1. The resulting precipitate was filtrated and then dried under vacuum to yield a white solid (230 mg, 77 %).

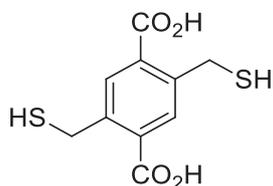
White solid; mp: degradation 308.3 °C; HRMS (ESI) $[M - H]^-$ found 225.0401, calculated 225.0405 for $[C_{10}H_9O_6]$; 1H NMR (400 MHz, DMSO- d_6) δ ppm = 8.16 (s, 2 H, Ar), 4.82 (s, 4 H, CH₂); ^{13}C NMR (400 MHz, DMSO- d_6) δ ppm = 168.1 (CO), 142.1 (C_{Ar}-CH₂), 131.1 (C_{Ar}H), 128.5 (C_{Ar}-CO), 60.7 (CH₂)



I-34 dimethyl 2,5-bis((carbamimidoylthio)methyl)terephthalate

A suspension of **I-33** dimethyl 2,5-bis(bromomethyl)terephthalate (275 mg, 0.715 mmol, 1 eq) and thiourea (115 mg, 1.66 mmol, 2.33 eq) in MeOH (11 mL) was stirred at room temperature for 16 h. The solvent was evaporated, and the crude mixture was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 8:2). Evaporation of the solvent gave a white powder (263 mg, 99 %).

White solid; mp: degradation 250.3 °C; HRMS (ESI) $[M + H]^+$ found 371.0836, calculated 371.0842 for $[C_{14}H_{19}N_4O_4S_2]^+$; 1H NMR (400 MHz, DMSO- d_6) δ ppm = 9.40 (br. d., 6 H, NH₂), 8.16 (s, 2 H, Ar), 4.76 (s, 4 H, CH₂), 3.91 (s, 6 H, CH₃); ^{13}C NMR (400 MHz, DMSO- d_6) δ ppm = 169.0 (CO), 165.6 (C=NH), 136.9 (C_{Ar}-CH₂), 133.4 (C_{Ar}H), 132.3 (C_{Ar}-CO₂), 52.9 (CH₃), 32.4 (CH₂-S)

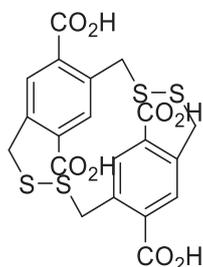


I-15 2,5-bis(mercaptomethyl)terephthalic acid

A solution of **I-34** dimethyl 2,5-bis((carbamimidoylthio)methyl)terephthalate (100 mg, 0.27 mmol, 1 eq) in degassed 2 M KOH (896 mg, 16 mmol, 60 eq) in MeOH/H₂O (1:1, 8 mL) was refluxed under an inert atmosphere for 2 h. The reaction mixture was cooled in ice, and concentrated TFA (1 mL) was added until pH 1. The methanol was evaporated, and the water phase was cooled. A white yellow

precipitate was formed, filtered, and washed with water, yielding compound **4-8** as a white solid (40 mg, 57 %).

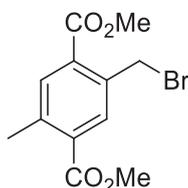
White solid; mp: degradation 314.1 °C; HRMS (ESI) $[M + Na]^+$ found 280.9902, calculated 280.9913 for $[C_{10}H_{10}NaO_4S_2]^+$; 1H NMR (400 MHz, MeOD) δ ppm = 7.93 (s, 2 H, Ar), 4.06 (s, 4 H, CH₂); ^{13}C NMR (400 MHz, MeOD) δ ppm = 169.5 (CO), 144.0 (C_{Ar}-CH₂), 134.4 (C_{Ar}H), 133.7 (C_{Ar}-CO₂H), 27.4 (CH₂)



1-15₂

To a stirred solution of **I-15** 2,5-bis(mercaptomethyl)terephthalic acid (20 mg, 0.04 mmol, 1 eq) in methanol (10 mL) was added dropwise over 5 mn [50 mM] iodine solution in MeOH until a yellow color was persistent. A white precipitate was formed, filtered, and washed with water, yielding compound **I-15₂** as a white solid (12 mg, 61 %).

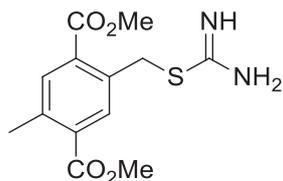
White solid; mp: degradation 270.1 °C; HRMS (ESI) $[M - H]^-$ found 510.9630, calculated 510.9655 for $[C_{20}H_{15}O_8S_4]^-$; 1H NMR (500 MHz, D₂O + NaOD) δ ppm = 6.96 (s, 2 H, Ar), 4.21 (d, J = 15.6 Hz, 2 H, CH₂, AB system), 3.27 (d, J = 15.0 Hz, 2 H, CH₂, AB system); ^{13}C NMR (400 MHz, DMSO-d₆) δ ppm = 166.2 (COOH), 137.5 (C_{Ar}-CH₂), 133.9 (C_{Ar}H), 130.9 (C_{Ar}-COOH), 40.8 (CH₂)



I-35 dimethyl 2-(bromomethyl)-5-methylterephthalate

A solution of **I-32** dimethyl 2,5-dimethylterephthalate (400 mg, 1.80 mmol, 1 eq), NBS (320 mg, 1.80 mmol, 1 eq), AIBN (23 mg, 0.144 mmol, 0.08 eq) in dry DCM (6 mL) was heated at reflux until conversion of starting material. The reaction is monitored by NMR control (approximately 16 h). After cooling, the organic phase was washed with a solution of NaHCO₃ (2*10 mL) and brine (2*10 mL). The organic phase was dried with MgSO₄ and evaporated to dryness. The crude mixture was purified by column chromatography (SiO₂, cyclohexane : ethyl acetate = 9:1) to give a white powder (542 mg, 92 %).

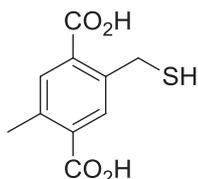
White solid; mp: 52.9 – 55.4 °C; HRMS (ESI) $[M + Na]^+$ found 322.9876, calculated 322.9889 for $[C_{12}H_{13}BrNaO_4]^+$; 1H NMR (400 MHz, CDCl₃) δ ppm = 8.06 - 7.97 (m, 1 H, Ar), 7.84 - 7.74 (m, 1 H, Ar), 4.95 - 4.89 (m, 2 H, CH₂-Br), 3.99 - 3.89 (m, 6 H, OCH₃), 2.63 - 2.55 (m, 3 H, CH₃); ^{13}C NMR (400 MHz, CDCl₃) δ ppm = 13C NMR (101MHz, CHLOROFORM-d) Shift = 166.6 (CO), 166.4 (CO), 140.5 (C_{Ar}-CH₃), 136.4 (C_{Ar}-CH₂), 134.4 (C_{Ar}-CO), 133.6 (C_{Ar}H), 52.5 (OCH₃), 52.2 (OCH₃), 30.7 (CH₂), 21.2 (CH₃)



I-36 dimethyl 2-((carbamimidoylthio)methyl)-5-methylterephthalate

A suspension of **I-35** dimethyl 2-(bromomethyl)-5-methylterephthalate (510 mg, 1.69 mmol, 1 eq) and thiourea (155 mg, 2.03 mmol, 1.20 eq) in MeOH (24 mL) was stirred at room temperature for 16 h. The solvent was evaporated, and the crude mixture was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 8:2). Evaporation of the solvent gave a white powder (343 mg, 68 %).

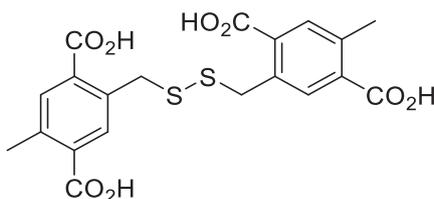
White solid; mp: 150.0 – 152.6 °C; HRMS (ESI) [M + H]⁺ found 297.0890, calculated 397.0904 for [C₁₃H₁₇N₂O₄S₂]⁺; ¹H NMR (400 MHz, DMSO-d₆) δ ppm = 9.54 (br. s., 1 H, NH), 9.22 (br. s., 1 H, NH), 8.01 (s, 1 H, Ar), 7.87 (s, 1 H, Ar), 4.72 (s, 2 H, CH₂), 3.88 (s, 3 H, OCH₃), 3.87 - 3.85 (m, 3 H, OCH₃), 2.53 (s, 3 H, CH₃); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 169.2 (C=NH), 166.3 (CO), 166.1 (CO), 139.4 (C_{Ar}-CH₂), 134.0 (C_{Ar}-CO), 134.0 (C_{Ar}-CO), 133.0 (C_{Ar}H), 132.5 (C_{Ar}H), 131.6 (C_{Ar}-CH₃), 52.7 (OCH₃), 52.4 (OCH₃), 32.5 (CH₂-S), 20.5 (CH₃)



I-37 2-(mercaptomethyl)-5-methylterephthalic acid

A solution of **I-36** dimethyl 2-((carbamimidoylthio)methyl)-5-methylterephthalate (343 mg, 0.866 mmol, 1 eq) in degassed 2 M KOH (2.86 g, 51 mmol, 60 eq) in MeOH/H₂O (1:1, 24 mL) was refluxed under an inert atmosphere for 2 h. The reaction mixture was cooled in ice, and concentrated TFA was added until pH 1. The methanol was evaporated, and the water phase was cooled. A white yellow precipitate was formed, filtered, and washed with water, yielding compound **I-37** as a white solid (132 mg, 67 %).

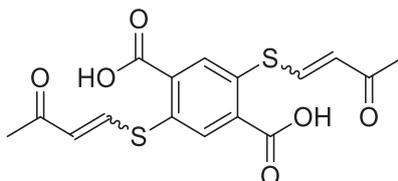
White solid; mp: degradation 196.1 °C; HRMS (ESI) [M + Na]⁺ found 249.0184, calculated 249.0192 for [C₁₀H₁₀NaO₄S]⁺; ¹H NMR (400 MHz, MeOD) δ ppm = 7.90 (s, 1 H, Ar), 7.82 (s, 1 H, Ar), 4.06 (s, 2 H, CH₂), 2.59 (s, 3 H, CH₃); ¹³C NMR (400 MHz, MeOD) δ ppm = 170.2 (CO), 169.9 (CO), 142.7 (C_{Ar}-CH₂), 139.5 (C_{Ar}-CH₃), 135.3 (C_{Ar}H), 135.0 (C_{Ar}H), 133.7 (C_{Ar}-CO₂H), 133.1 (C_{Ar}-CO₂H), 27.4 (CH₂), 21.4 (CH₃)



I-37₂ 5,5'-(disulfanediy)bis(methylene))bis(2-methylterephthalic acid)

To a stirred solution of **I-37** 2-(mercaptomethyl)-5-methylterephthalic acid (20 mg, 0.088 mmol, 1 eq) in methanol (2 mL) was added dropwise over 5 mn [50 mM] iodine solution in MeOH until a yellow color was persistent. A white precipitate was formed, filtered, and washed with water, yielding compound **I-37₂** as a white solid (14 mg, 71 %).

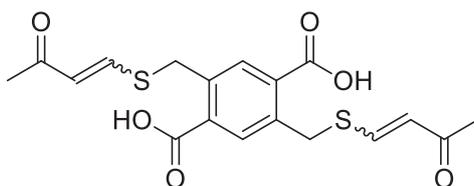
White solid; mp: degradation 327.1 °C; HRMS (ESI) [M + Na]⁺ found 473.0324, calculated 473.0335 for [C₂₀H₁₈NaO₈S₂]⁺; ¹H NMR (400 MHz, D₂O + NaOD) δ ppm = 7.27 (s, 2 H, Ar), 7.20 (s, 2 H, Ar), 3.98 (s, 4 H, CH₂), 2.31 (s, 6 H, CH₃); ¹³C NMR (400 MHz, D₂O + NaOD) δ ppm = 178.3 (COOH), 177.2 (COOH), 140.0 (C_{Ar}), 138.7 (C_{Ar}), 133.3 (C_{Ar}), 132.0 (C_{Ar}), 129.7 (C_{Ar}H), 128.4 (C_{Ar}H), 39.8 (CH₂), 18.8 (CH₃)



I-78 2,5-bis((3-oxobut-1-en-1-yl)thio)terephthalic acid

A mixture of thiol **I-1** (100 mg, 0.435 mmol, 1 eq.), butynone (34 μL, 0.435 mmol, 1 eq.) and sodium acetate (534 mg, 6.52 mmol, 15 eq.) in degassed THF/water (5 mL, 1:1) was stirred at room temperature for 60 min. The reaction mixture was concentrated in vacuum. Water was added. The aqueous mixture was acidified with concentrated HCl until pH 1. Then, the resulting solid and the aqueous layer were extracted. The combined organic layers were washed with brine and dried over Na₂SO₄, concentrated in vacuum to yield a yellow powder (124 mg, 76 %)

Yellow powder, no Mp (mixture of diastereoisomers); HRMS (ESI) [M + H]⁺ found 367.0301, calculated 367.0305 for [C₁₆H₁₅O₆S₂]⁺; major diastereoisomer ¹H NMR (400 MHz, DMSO-d₆) δ ppm = 7.96 (s, 2 H, Ar), 7.51 (d, J = 9.8 Hz, 2 H, H_{C=C}), 6.61 (d, J = 9.8 Hz, 2 H, H_{C=C}), 2.25 (s, 6 H, CH₃), 3.73 - 3.66 (m, 12 H, CH₂-O, CH₂-CH₂-COO), 3.38 (t, J = 4.9 Hz, 4 H, CH₂-N₃); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 195.1 (CO), 166.2 (COOH), 143.4 (C=C-S), 133.3 (C_{Ar}-COO), 131.9 (C_{Ar}-S), 129.0 (C_{Ar}H), 122.7 (C=C-CO), 27.3 (CH₃)

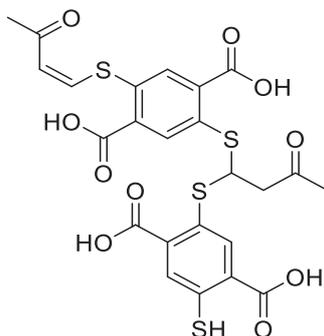


I-79 2,5-bis(((3-oxobut-1-en-1-yl)thio)methyl)terephthalic acid

A mixture of thiol **I-15** (112 mg, 0.435 mmol, 1 eq.), butynone (34 μL, 0.435 mmol, 1 eq.) and sodium acetate (534 mg, 6.52 mmol, 15 eq.) in degassed THF/water (5 mL, 1:1) was stirred at room temperature for 60 min. The reaction mixture was concentrated in vacuum. Water was added. The aqueous mixture was acidified with concentrated HCl until pH 1. Then, the resulting solid and the aqueous layer were extracted. The combined organic layers were washed with brine and dried over Na₂SO₄, concentrated in vacuum to yield a yellow powder (64 mg, 40 %)

White powder, no Mp (mixture of diastereoisomers 81 / 19 %); HRMS (ESI) [M + Na]⁺ found 417.0436, calculated 417.0437 for [C₁₈H₁₈NaO₆S₂]⁺; major diastereoisomer ¹H NMR (400 MHz, DMSO-d₆) δ ppm = 8.00 (s, 2 H, Ar), 7.48 (d, J = 9.8 Hz, 2 H, H_{C=C}), 6.34 (d, J = 9.8 Hz, 2 H, H_{C=C}), 4.37 (s, 4 H, CH₂), 2.08 (s, 6

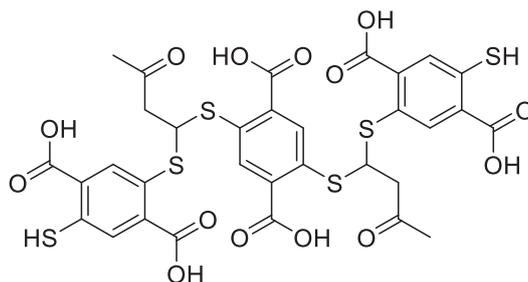
H, CH₃); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 196.0 (CO), 167.5 (COOH), 149.2 (C_{Ar}-C), 139.3 (C_{Ar}-COOH), 133.0 (C=C-CO), 132.7 (C_{Ar}H), 120.4 (C=C-CO), 37.6 (CH₂), 29.8 (CH₃)



2-(((Z)-1-((2,5-dicarboxy-4-(((Z)-3-oxobut-1-en-1-yl)thio)phenyl)thio)-3-oxobut-1-en-1-yl)thio)-5-mercaptoterephthalic acid

A mixture of **I-78** 2,5-bis((3-oxobut-1-en-1-yl)thio)terephthalic acid (10 mg, 0.027 mmol, 1 eq), **I-1** (7.8 mg, 0.055 mmol, 2 eq), FeCl₃ (0.89 mg, 0.055 mmol, 2 eq) in degassed DMSO (1 mL) was stirred at 60 °C during 48 h. Then, 0.1 M HCl (10 mL) was added. The resulting precipitate was filtrated and evaporated under vacuum. The crude product was analysed by ¹H NMR and mass spectrometry.

HRMS (ESI) [M - H]⁻ found 592.9711, calculated 592.9710 for [C₂₄H₁₇NO₁₀S₄]⁻



5,5'-((((1Z,1'Z)-((2,5-dicarboxy-1,4-phenylene)bis(sulfanediy))bis(3-oxobut-1-ene-1,1,1-triyl))bis(sulfanediy))bis(2-mercaptoterephthalic acid)

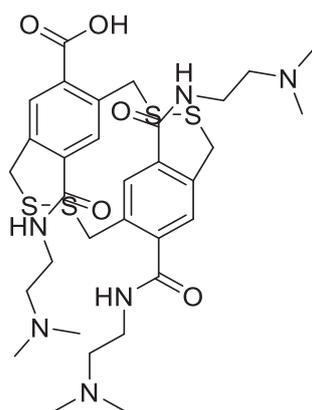
A mixture of **I-78** 2,5-bis((3-oxobut-1-en-1-yl)thio)terephthalic acid (10 mg, 0.027 mmol, 1 eq), **I-1** (7.8 mg, 0.055 mmol, 2 eq), FeCl₃ (0.89 mg, 0.055 mmol, 2 eq) in degassed DMSO (1 mL) was stirred at 60 °C during 48 h. Then, 0.1 M HCl (10 mL) was added. The resulting precipitate was filtrated and evaporated under vacuum. The crude product was analysed by ¹H NMR and mass spectrometry.

HRMS (ESI) [M - H]⁻ found 820.9235, calculated 920.9261 for [C₃₂H₂₁O₁₄S₆]⁻



Cyclopentadiene

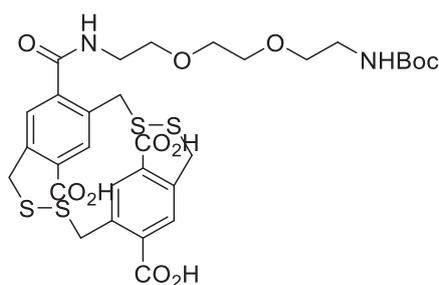
Cyclopentadiene was obtained by distillation of dicyclopentadiene. A boiler was charged with dicyclopentadiene and fitted with a condenser tube. The product was stirred and heated to 165°C. Cyclopentadiene (T_{eb} = 40°C) was obtained in a refrigerated Erlenmeyer and then directly used (inflammable product, avoid storage and heat source).



I-84

To a solution of **I-15₂** (10 mg, 0.0039 mmol, 1 eq) with HCTU (64 mg, 0.016 mmol, 6 eq) in DMF (1 mL) was added *N,N*-dimethylethylenediamine **I-89** (46 μ L, 0.039 mmol, 10 eq) and the mixture was stirred for 24 hours at room temperature under N_2 . The coupling products were purified by semi-preparative HPLC (method PREP₁ 44 %) to give a white solid.

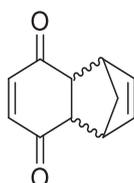
HRMS (ESI) $[M + H]^+$ found 723.2425, calculated 723.24851590 for $[C_{32}H_{47}N_6O_5S_4]^+$



I-85

To a solution of **1-15₂** (1.0 mg, 0.00195 mmol, 1 eq) with HCTU (8 mg, 0.0195 mmol, 10 eq) in DMF (0.8 mL) was added tert-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate (5 mg, 0.0195 mmol, 10 eq) and the mixture was stirred for 24 hours at room temperature under N_2 . The crude product was analysed by mass spectrometry.

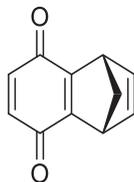
HRMS (ESI) $[M + H]^+$ found 743.1431, calculated 743.1431 for $[C_{31}H_{39}N_2O_{11}S_4]^+$



I-94 1,4,4a,8a-tetrahydro-1,4-methanonaphthalene-5,8-dione

p-benzoquinone **I-93** (36 g, 0.33 mol) was dissolved in 500 mL of MeOH. The mixture was cooled down to -78 °C and freshly distilled cyclopentadiene (28 mL, 0.33 mol) was added. The reaction mixture was stirred at room temperature for 2h. The solvent was removed under vacuum and the resulting solid was recrystallised in hexane if necessary, to give a brown powder (57,8 g, 99 %).

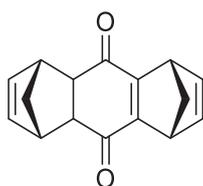
Brown powder, mp: 63,4 – 65,8 °C; HRMS (ESI) $[M + Na]^+$ found 197.0572, calculated 197.0573 for $[C_{11}H_{11}NaO_2]^+$; 1H NMR (400 MHz, $CDCl_3$) δ ppm = 6.46 (s, 2 H, =CH-CO), 5.97 - 5.92 (m, 2 H, CH=CH), 3.44 - 3.40 (m, 2 H, CH-CO), 3.14 - 3.11 (m, 2 H, CH-CH=CH), 1.45 - 1.39 (m, 1 H, CH_2), 1.36 - 1.31 (m, 1 H, CH_2); ^{13}C NMR (400 MHz, $CDCl_3$) δ ppm = 199.1 (CO), 141.7(=CH-CO), 135.0 (CH-CH=CH), 48.4 (CH-CO), 48.4 (CH-CH=CH), 48.0 (CH_2)



I-95 (1R,4S)-1,4-dihydro-1,4-methanonaphthalene-5,8-dione

To 240 mL of methanol was added **I-94** 1,4,4a,8a-tetrahydro-1,4-methanonaphthalene-5,8-dione (38 g, 216 mmol, 1 eq). The resulting solution was cooled down to 10°C and triethylamine (0.38 mL, 2.72 mmol, 0.013 eq) was added. The reaction mixture was stirred for 5 h at 10°C then was enabled to warm up to room temperature and was stirred for further 15 h. The solvent was removed under vacuum to give the corresponding hydroquinone form, which was suspended in p-benzoquinone (25.9 g, 238 mmol, 1.1 eq) in 590 mL of chloroform. The stirred suspension was heated firstly at 50°C for 4h, then at 40°C for one more hour. The reaction mixture was cooled down to room temperature. The hydroquinone precipitate was filtered off and washed with 30 mL of chloroform. The combined chloroform layers were washed with a 1% aqueous solution of NaOH, and then dried over $MgSO_4$. The solvent was removed under vacuum to give a yellow oil (36 g, 87 %).

Yellow powder, mp : 62.6 – 64.1; HRMS (ESI) $[M + H]^+$ found 173.0599, calculated 173.0597 for $[C_{11}H_9O_2]^+$; 1H NMR (400 MHz, $CDCl_3$) δ ppm = 6.80 (t, $J = 1.8$ Hz, 2 H, =CH-CH), 6.52 (s, 2 H, =CH-CO), 4.04 (t, $J = 1.3$ Hz, 2 H, CH), 2.30 - 2.24 (m, $J = 1.5, 1.5, 7.2$ Hz, 1 H, CH_2), 2.23 - 2.19 (m, $J = 1.5, 1.5, 7.2$ Hz, 1 H, CH_2); ^{13}C NMR (400 MHz, $CDCl_3$) δ ppm = 183.5 (CO), 160.2 ($C_{IV}=C_{IV}$), 142.1 (CH- $C_{III}=C_{III}$ -CH), 135.3 (CO- $C_{III}=C_{III}$ -CO), 73.4 (CH_2), 47.9 (CH)

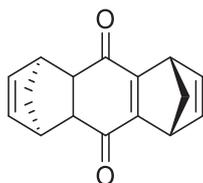


I-96 (1R,4S,5R,8S)-1,4,4a,5,8,9a-hexahydro-1,4:5,8-dimethanoanthracene-9,10-dione

A solution of **I-95** (1R,4S)-1,4-dihydro-1,4-methanonaphthalene-5,8-dione (37 g, 189 mmol, 1 eq) in MeOH (280 mL) was maintained at -78°C and freshly distilled cyclopentadiene (16 mL, 189 mmol, 1 eq) was added at once. The stirred orange solution was enabled to warm up overnight. The solvent was removed under vacuum and the resulting yellow solid was recrystallised in ethanol, and then purified by column chromatography (SiO_2 , pentane : ethyl acetate = 8:2) to give a yellow powder (21,109 g, 44 %). The diastereoisomeric mixture was separated into pure diastereoisomers through semi-preparative chiral HPLC²⁰ (Lux-Amylose-2 (250 x 10 mm), hexane : EtOH 80 : 20) (10,55 g * 2, 2 * 22 %).

²⁰ Dr. Nicolas Vanthuyne, AMU, ISM² Plateforme de Chromatographie Chirale et de Stéréochimie Dynamique

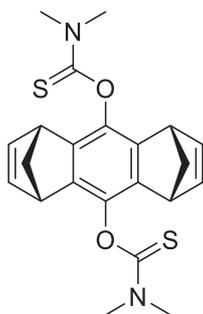
Yellow powder, mp : 152.8 – 155.3; HRMS (ESI) $[M + H]^+$ found 261.0890, calculated 261.0886 for $[C_{16}H_{14}NaO_2]^+$; 1H NMR (400 MHz, $CDCl_3$) δ ppm = 6.77 (t, $J = 2.0$ Hz, 2 H, HC=CH, opposite side of the central double bond), 6.04 - 5.98 (m, 2 H, HC=CH, side of the central double bond), 3.92 (t, $J = 2.0$ Hz, 2 H, =CH-CH, opposite side of the central double bond), 3.51 - 3.31 (m, 2 H, =CH-CH, side of the central double bond), 3.15 (t, $J = 2.2$ Hz, 2 H, CH-CH), 2.18 (td, $J = 1.5, 7.1$ Hz, 1 H, CH_2 , side of the central double bond), 2.03 (td, $J = 1.5, 7.1$ Hz, 1 H, CH_2 , side of the central double bond), 1.53 (td, $J = 1.7, 8.6$ Hz, 1 H, CH_2 , opposite side to the central double bond), 1.41 (td, $J = 1.7, 8.6$ Hz, 1 H, CH_2 , opposite side of the central double bond); ^{13}C NMR (400 MHz, $CDCl_3$) δ ppm = 195.2 (CO), 166.7 ($C_{IV}=C_{IV}$), 142.2 (CH- $C_{III}=C_{III}$ -CH), 134.6 (CO- $C_{III}=C_{III}$ -CO), 73.5 (CH_2), 51.0 (CH-CO), 49.3 (CH-CO), 48.6 (CH), 48.2 (CH)



I-97 (1R,4S,5S,8R)-1,4,4a,5,8,9a-hexahydro-1,4:5,8-dimethanoanthracene-9,10-dione

A solution of **I-95** (1R,4S)-1,4-dihydro-1,4-methanonaphthalene-5,8-dione (37 g, 0, 189 mmol, 1 eq) in MeOH (280 mL) was maintained at $-78^\circ C$ and freshly distilled cyclopentadiene (16 mL, 189 mmol, 1 eq) was added at once. The stirred orange solution was enabled to warm up overnight. The solvent was removed under vacuum and the resulting yellow solid was purified by column chromatography (SiO_2 , pentane : ethyl acetate = 8:2) to give a yellow powder (21,109 g, 44 %). The diastereoisomeric mixture was separated into pure diastereoisomers through semi-preparative chiral HPLC²¹ (Lux-Amylose-2 (250 x 10 mm), hexane : ethanol 80 : 20) (10,55 g * 2, 2 * 22 %).

Yellow powder, mp : 148.3 – 150.1; HRMS (ESI) $[M + H]^+$ found 261.0890, calculated 261.0886 for $[C_{16}H_{14}NaO_2]^+$; 1H NMR (400 MHz, $CDCl_3$) δ ppm = 6.76 (t, $J = 1.7$ Hz, 2 H, HC=CH, opposite side of the central double bond), 5.77 (t, $J = 1.8$ Hz, 2 H, HC=CH, side of the central double bond), 3.96 (t, $J = 1.5$ Hz, 2 H, =CH-CH, opposite side of the central double bond), 3.51 - 3.38 (m, 2 H, =CH-CH, side of the central double bond), 3.29 - 3.16 (m, 2 H, CH-CH), 2.20 (td, $J = 1.5, 7.3$ Hz, 1 H), 2.12 (td, $J = 1.5, 7.3$ Hz, 1 H), 1.47 (td, $J = 1.8, 8.3$ Hz, 1 H, CH_2 , opposite side to the central double bond), 1.39 (td, $J = 1.8, 8.6$ Hz, 1 H, CH_2 , opposite side of the central double bond); ^{13}C NMR (400 MHz, $CDCl_3$) δ ppm = 195.6 (CO), 166.9 ($C_{IV}=C_{IV}$), 142.7 (CH- $C_{III}=C_{III}$ -CH), 134.5 (CO- $C_{III}=C_{III}$ -CO), 72.4 (CH_2), 50.6 (CH-CO), 48.8 (CH-CO), 48.4 (CH), 48.3 (CH)

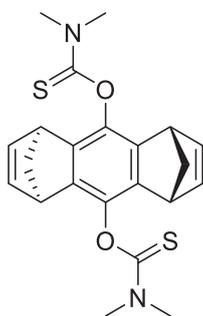


²¹ Dr. Nicolas Vanthuyne, AMU, ISM² Plateforme de Chromatographie Chirale et de Stéréochimie Dynamique

I-98 O,O'-((1R,4S,5R,8S)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl)
bis(dimethylcarbamothioate)

To a solution of **I-96** (1R,4S,5R,8S)-1,4,4a,5,8,9a-hexahydro-1,4:5,8-dimethanoanthracene-9,10-dione (3,400 g, 14.3 mmol, 1 eq) and NaH (1,06 g, 42.1 mmol, 4 eq) in DMF (30 mL) was added at 0°C dimethylthiocarbamoyl chloride (7.023 g, 42.1 mmol, 4 eq) in 13 mL of dry DMA. The mixture was stirred at room temperature for 24 h. The precipitate was then filtrated and washed extensively with water. Drying of the resulting solid gave a brown powder (5.843 g, 99 %).

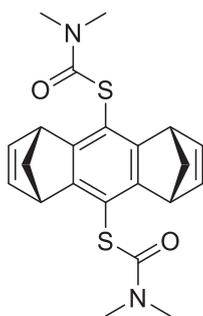
Brown powder, mp: degradation, 260.1 °C; HRMS (ESI) [M + H]⁺ found 413.1351, calculated 413.1352 for [C₂₂H₂₅N₂O₂S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 6.78 (t, J = 1.5 Hz, 4 H, HC=CH), 3.82 (t, J = 1.7 Hz, 4 H, CH), 3.49 (s, 6 H, CH₃), 3.38 (s, 6 H, CH₃), 2.33 (td, J = 1.6, 6.8 Hz, 2 H, CH₂), 2.21 (td, J = 1.5, 7.1 Hz, 2 H, CH₂); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 187.6 (CS), 142.7 (HC=CH), 140.6 (C_{Ar}-O), 69.5 (CH₂), 47.5 (CH), 43.2 (CH₃), 38.5 (CH₃)



I-99 O,O'-((1R,4S,5S,8R)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl)
bis(dimethylcarbamothioate)

To a solution of **I-97** (1R,4S,5S,8R)-1,4,4a,5,8,9a-hexahydro-1,4:5,8-dimethanoanthracene-9,10-dione (4,50 g, 18.9 mmol, 1 eq) and NaH (1,822 g, 75.9 mmol, 4 eq) in DMF (39 mL) was added at 0°C dimethylthiocarbamoyl chloride (9.34 g, 75.9 mmol, 4 eq) in 16 mL of dry DMA. The mixture was stirred at room temperature for 24 h. The precipitate was then filtrated and washed extensively with water. Drying of the resulting solid gave a brown powder (7.78 g, 98 %).

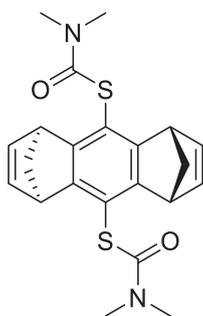
Brown powder, mp: degradation, 273.4 °C; HRMS (ESI) [M + H]⁺ found 413.1351, calculated 413.1352 for [C₂₂H₂₅N₂O₂S₂]⁺; ¹H NMR (400 MHz, CHLOROFORM-d) Shift = 6.94 - 6.76 (m, 4 H, HC=CH), 3.88 - 3.77 (m, 4 H, CH), 3.52 (s, 6 H, CH₃), 3.38 (s, 6 H, CH₃), 2.29 (td, J = 1.6, 6.8 Hz, 2 H, CH₂), 2.19 (td, J = 1.7, 7.1 Hz, 2 H, CH₂); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 187.7 (CS), 140.6 (HC=CH), 69.4 (CH₂), 47.5 (CH), 43.3 (CH₃), 38.5 (CH₃)



I-100 S,S'-((1R,4S,5R,8S)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl) bis(dimethylcarbamothioate)

A solution of **I-98** O,O'-((1R,4S,5R,8S)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl) bis(dimethylcarbamothioate) (357 mg, 0.87 mmol, 1 eq) in diphenyl ether (3 mL) was heated at 230 °C for 4 h. The mixture was then cooled to room temperature. The resulting mixture was purified by column chromatography (SiO₂, pentane : ethyl acetate = 1:0 -> 1:1 -> 0:1) to give a yellow powder (175 mg, 49 %).

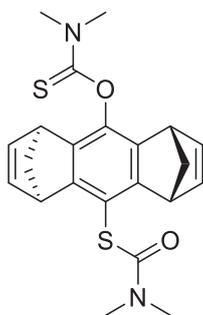
Yellow powder, mp: 226.3 - 228.1 °C; HRMS (ESI) [M + H]⁺ found 413.1351, calculated 413.1352 for [C₂₂H₂₅N₂O₂S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 6.78 (t, J = 1.7 Hz, 4 H, HC=CH), 4.02 (t, J = 1.5 Hz, 4 H, CH), 3.19 - 2.99 (m, 12 H, CH₃), 2.35 (td, J = 1.5, 7.1 Hz, 2 H, CH₂), 2.24 (td, J = 1.5, 7.1 Hz, 2 H, CH₂); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.7 (CO), 154.1 (HC=CH), 143.2 (C_{Ar,IV}), 118.0 (C_{Ar-S}), 69.6 (CH₂), 50.5 (CH), 36.9 (CH₃)



I-101 S,S'-((1R,4S,5S,8R)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl) bis(dimethylcarbamothioate)

A solution of **I-99** O,O'-((1R,4S,5S,8R)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl) bis(dimethylcarbamothioate) (357 mg, 0.87 mmol, 1 eq) in diphenylether (3 mL) was heated at 230 °C for 4 h. The mixture was then cooled to room temperature. The resulting mixture was purified by column chromatography (SiO₂, pentane : ethyl acetate = 1:0 -> 1:1 -> 0:1) to give a yellow powder (41 mg, 11 %).

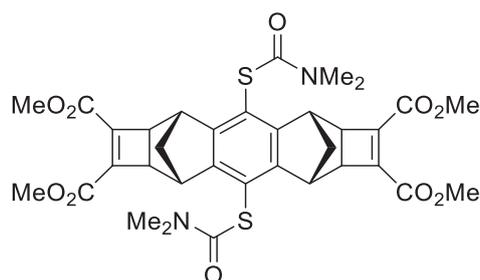
Yellow powder, mp: degradation, 245.6 – 247.9 °C; HRMS (ESI) [M + Na]⁺ found 435.1173, calculated 435.1171 for [C₂₂H₂₅NaN₂O₂S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 6.84 (t, J = 1.2 Hz, 2 H, HC=CH), 4.01 (t, J = 1.7 Hz, 2 H, CH), 3.21 - 2.99 (m, 6 H, CH₃), 2.24 (dd, J = 6.9, 28.0 Hz, 2 H, CH₂); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.7 (CO), 154.0 (HC=CH), 143.2 (C_{Ar,IV}), 118.0 (C_{Ar-S}), 69.5 (CH₂), 50.6 (CH), 37.0 (CH₃)



I-102 S,S'-((1R,4S,5S,8R)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl) bis(dimethylcarbamothioate)

A solution of **I-99** O,O'-((1R,4S,5S,8R)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl) bis(dimethylcarbamothioate) (357 mg, 0.87 mmol, 1 eq) in diphenylether (3 mL) was heated at 230 °C for 4 h. The mixture was then cooled to room temperature. The resulting mixture was purified by column chromatography (SiO₂, pentane : ethyl acetate = 1:0 -> 1:1 -> 0:1) to give a yellow powder (11 mg, 3 %).

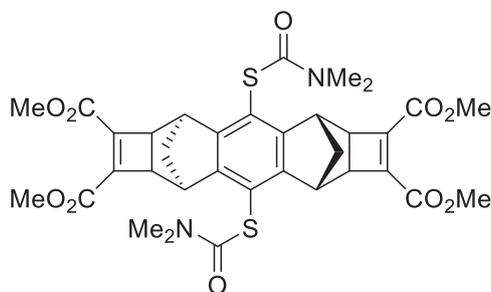
Yellow powder, mp: degradation, 238.6 – 241.0 °C; HRMS (ESI) [M + H]⁺ found 413.1349, calculated 413.1351 for [C₂₂H₂₅N₂O₂S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 6.94 - 6.73 (m, 2 H, HC=CH), 4.15 - 3.92 (m, 1 H, CH), 3.89 - 3.67 (m, 1 H, CH), 3.59 - 3.38 (m, 3 H, CH₃), 3.24 - 2.99 (m, 3 H, CH₃), 2.28 (td, J = 1.2, 6.6 Hz, 1 H, CH₂), 2.20 (td, J = 1.5, 7.1 Hz, 1 H, CH₂); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 187.3 (CS), 166.1 (CO), 155.9 (C_{Ar}-O), 144.0 (HC=CH), 140.6 (C_{Ar,IV}), 129.7 (C_{Ar,IV}), 114.0 (C_{Ar}-S), 69.5 (CH₂), 50.2 (CH), 47.8 (CH), 47.6 (CH₃), 43.2 (CH₃), 38.6 (CH₃), 37.0 (CH₃)



I-103 tetramethyl (3R,5S,8R,10S)-4,9-bis((dimethylcarbamoyl)thio)-2a,3,5,5a,7a,8,10,10a-octahydro-3,10:5,8-dimethanodicyclobuta[b,i]anthracene-1,2,6,7-tetracarboxylate

DMAD (2.84 mL, 23.1 mmol, 2.6 eq) was added dropwise to a solution of **I-100** S,S'-((1R,4S,5R,8S)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl) bis(dimethylcarbamothioate) (3.76 g, 8.89 mmol, 1 eq) with RuH₂CO(PPh₃)₂ (816 mg, 0.89 mmol, 0.1 eq) in anhydrous and degassed THF (40 mL). The mixture was refluxed for 10 d. under inert atmosphere. The solvent was removed under vacuum and the resulting residue was purified by SiO₂ column chromatography (dichloromethane : ethylacetate 1 : 1) to give a yellow solid (2.9 g, 47 %).

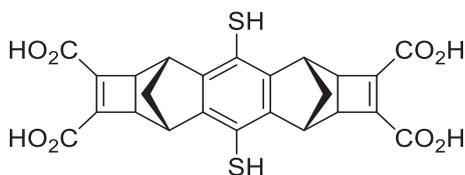
Grey powder, mp: degradation, 301.3 °C; HRMS (ESI) [M + Na]⁺ found 719.1704, calculated 719.1703 for [C₃₄H₃₆NaN₂O₁₀S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 3.81 (m, 12 H, CH₃ ester), 3.44 (m, 4 H, CH₆-membered ring), 3.27 - 2.99 (m, 12 H, CH₃ carbamate), 2.87 (m, 4 H, CH₄-membered ring), 1.92 - 1.79 (m, 2 H, CH₂), 1.75 - 1.69 (m, 2 H, CH₂); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.6 (CO_{ester}), 161.8 (CO_{carbamate}), 150.0 (C=C), 143.2 (C_{Ar,IV}), 118.7 (C_{Ar}-S), 51.9 (CH₃), 45.3 (CH₂), 40.8 (CH), 40.3 (CH₃), 37.1 (C_{IV})



I-104 tetramethyl (3R,5R,8S,10S)-4,9-bis((dimethylcarbamoyl)thio)-2a,3,5,5a,7a,8,10,10a-octahydro-3,10:5,8-dimethanodicyclobuta[b,i]anthracene-1,2,6,7-tetracarboxylate

DMAD (0,66 mL, 5.28 mmol, 2.6 eq) was added dropwise to a solution of **I-101** *S,S'*-((1*R,4S,5R,8S*)-1,4,5,8-tetrahydro-1,4:5,8-dimethanoanthracene-9,10-diyl) bis(dimethylcarbamothioate) (855 mg, 2.07 mmol, 1 eq) with RuH₂CO(PPh₃)₂ (190 mg, 0,207 mmol, 0.1 eq) in anhydrous and degassed THF (9.3 mL). The mixture was refluxed for 10 d. under inert atmosphere. The solvent was removed under vacuum and the resulting residue was purified by SiO₂ column chromatography (dichloromethane : ethylacetate 1 : 1) to give a grey solid (587 mg, 41 %).

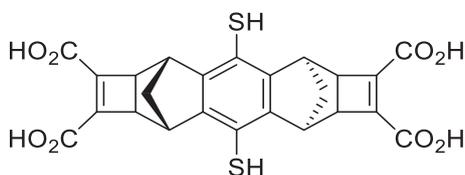
Grey powder, mp: 261.2 – 263.3 °C; HRMS (ESI) [M + Na]⁺ found 719.1707, calculated 719.1703 for [C₃₄H₃₆NaN₂O₁₀S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 3.80 (m, 12 H, CH₃ ester), 3.44 (m, 4 H, CH₆-membered ring), 3.25 - 2.96 (m, 12 H, CH₃ carbamate), 2.81 (m, 4 H, CH₄-membered ring), 1.85 (m, 2 H, CH₂), 1.74 - 1.65 (m, 2 H, CH₂); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.4 (CO_{ester}), 161.7 (CO_{carbamate}), 149.8 (C=C), 143.2 (C_{Ar,IV}), 118.7 (C_{Ar-S}), 51.8 (CH₃), 45.3 (CH₂), 40.8 (CH), 40.0 (CH₃), 37.0 (C_{IV})



I-105 (3*R,5S,8R,10S*)-4,9-dimercapto-2*a,3,5,5a,7a,8,10,10a*-octahydro-3,10:5,8-dimethanodicyclobuta[*b,i*]anthracene-1,2,6,7-tetracarboxylic acid

A solution of **I-103** (3*R,5S,8R,10S*)-tetramethyl 4,9-bis((dimethylcarbamoyl)thio)-3,5,8,10-tetrahydro-3,10:5,8-dimethanodicyclobuta[*b,i*]anthracene-1,2,6,7-tetracarboxylate (1.729 g, 2.52 mmol, 1 eq) in degassed 1.1 M NaOH (5.0 g, 125.7 mmol, 50 eq) in a degassed mixture of EtOH 1 / 1 H₂O (117 mL) was refluxed under an inert atmosphere for 16 h (monitored by NMR). The reaction mixture was cooled in ice, and concentrated HCl was added until pH 1. A precipitate was formed, filtered, and washed with water, yielding compound **I-105** as a brown solid (1.126 g, 90 %).

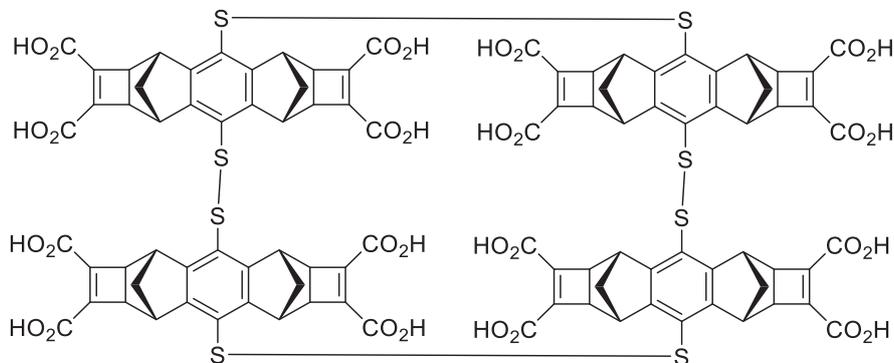
Grey solid; mp: decomposition 306.1 °C; HRMS (ESI) [M - H]⁺ found 497.0369, calculated 497.0370 for [C₂₄H₁₇O₁₈S₂]⁻; ¹H NMR (400 MHz, D₂O, NaOD) δ ppm = 3.34 - 3.31 (m, 4 H, CH₆-membered ring), 2.58 - 2.55 (m, 4 H, CH₄-membered ring), 1.61 - 1.55 (m, 2 H, CH₂), 1.48 - 1.43 (m, 2 H, CH₂); ¹³C NMR (400 MHz, D₂O, NaOD) δ ppm = 168.1 (CO), 147.2 (C=C), 144.6 (C_{IV Ar}), 119.9 (C_{Ar-S}), 44.0 (CH₆-membered ring), 39.8 (CH₄-membered ring), 38.7 (CH₂)



I-106 (3*R,5R,8S,10S*)-4,9-dimercapto-2*a,3,5,5a,7a,8,10,10a*-octahydro-3,10:5,8-dimethanodicyclobuta[*b,i*]anthracene-1,2,6,7-tetracarboxylic acid

A solution of **I-104** (3*R,5R,8S,10S*)-tetramethyl 4,9-bis((dimethylcarbamoyl)thio)-3,5,8,10-tetrahydro-3,10:5,8-dimethanodicyclobuta[*b,i*]anthracene-1,2,6,7-tetracarboxylate (587 mg, 0.86 mmol, 1 eq) in degassed 1.1 M NaOH (1.720 mg, 42.9 mmol, 50 eq) in a degassed mixture of EtOH 1 / 1 H₂O (40 mL) was refluxed under an inert atmosphere for 16 h (monitored by NMR). The reaction mixture was cooled in ice, and concentrated HCl was added until pH 1. A precipitate was formed, filtered, and washed with water, yielding compound **I-106** as a brown solid (161 mg, 38 %).

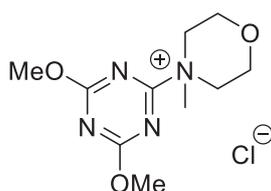
Grey solid; mp: decomposition 372.6 °C; HRMS (ESI) $[M - H]^+$ found 497.0368, calculated 497.0370 for $[C_{24}H_{17}O_{18}S_2]$; 1H NMR (400 MHz, D_2O , NaOD) δ ppm = 3.47 - 3.41 (m, 4 H, CH_{6-membered ring}), 2.51 - 2.46 (m, 4 H, CH_{4-membered ring}), 1.80 - 1.72 (m, 2 H, CH₂), 1.55 - 1.48 (m, 2 H, CH₂); ^{13}C NMR (400 MHz, D_2O , NaOD) δ ppm = 172.1 (CO), 145.5 (C=C), 143.6 (C_{IV Ar}), 124.8 (C_{Ar-S}), 44.0 (CH_{6-membered ring}), 40.9 (CH_{4-membered ring}), 38.2 (CH₂)



I-105 a name could not be generated for this structure

A solution of **I-105** (3R,5S,8R,10S)-4,9-dimercapto-2a,3,5,5a,7a,8,10,10a-octahydro-3,10:5,8-dimethanodicyclobuta[b,i]anthracene-1,2,6,7-tetracarboxylic acid (123 mg, 0.25 mmol, 1 eq) in ammonium acetate buffer (50 mL, 50 mM, pH no-adjusted) was stirred for 36 h. under air. Then, TFA was added until pH 1. The resulting precipitate was filtrated and washed with water. Evaporation of the solvent gave a brown powder (71 mg, 71 %).

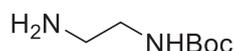
Brown solid; mp: degradation, 321.7°C; MS (ESI) $[M - 2 H]^{2-}$ found 991.047499, calculated 991.050042 for $[C_{96}H_{62}O_{32}S_8]^{2-}$; 1H NMR (500 MHz, $D_2O + 50$ mM $ND_4 CD_3COOD$) δ ppm = 4.04 - 3.99 (m, 8 H), 3.11 - 3.06 (m, 8 H), 2.78 - 2.74 (m, 8 H), 2.50 - 2.46 (m, 8 H), 1.90 (d, J = 10.5 Hz, 4 H), 1.76 (d, J = 10.1 Hz, 4 H), 1.04 (d, J = 9.6 Hz, 4 H), 0.39 (d, J = 9.2 Hz, 4 H) (full attribution is described in the main text); ^{13}C NMR (500 MHz, $D_2O + 50$ mM $ND_4 CD_3COOD$) δ ppm = 177.9, 167.1, 166.5, 149.5, 148.2, 147.8, 124.0, 123.9, 44.4, 43.8, 40.0, 39.4, 39.1, 39.0 (full attribution is described in the main text).



I-111 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholin-4-ium chloride

N-methylmorpholine (2.22 g, 22 mmol, 1 eq) was added to a solution of 2-Chloro-4,6-dimethoxy-1,3,5-triazine (3.86 g, 22 mmol, 1 eq) in THF (60 mL) at room temperature. A white solid appeared within several minutes. After stirring for 30 min at room temperature., the solid was collected by filtration and washed with THF and dried to give a white solid (5.50 g, 99 %).

White solid; 1H NMR (300 MHz, D_2O) δ ppm = 4.11 (s, 6 H, O-CH₃), 4.01 - 3.74 (m, 4 H, CH₂), 3.51 (s, 3 H, N-CH₃), data matched with literature²⁰⁰.



I-114 tert-butyl (2-aminoethyl)carbamate

To a solution of 1,2-diaminoethane (1.4 mL, 20 mmol) in chloroform (20 mL) was added dropwise at 0°C a solution of di-tertbutyldicarbonate (4.37 g, 20 mmol) in chloroform (10 mL) over a period of 3 h. After stirring at room temperature for 16 h, the mixture was washed with brine (3 * 20 mL) and water (20 mL), dried over MgSO₄ and concentrated in vacuo to afford a colourless oil (3.2 g, 99 %).

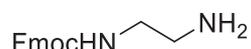
¹H NMR (300 MHz, CDCl₃) δ ppm = 4.89 (br. s., 1 H, NH), 3.17 (q, J = 5.9 Hz, 2 H, CH₂-NH), 2.79 (t, J = 6.0 Hz, 2 H, CH₂-NH₂), 1.48 - 1.42 (m, 9 H, CH₃), 1.36 (br. s., 2 H, NH₂); data matched with literature reference^{201b}.



I-115 (9H-fluoren-9-yl)methyl tert-butyl ethane-1,2-diyldicarbamate

To a solution of I-114 tert-butyl (2-aminoethyl)carbamate (2.699 g, 16.9 mmol, 1 eq) and NaHCO₃ (2.12 g, 25.3 mmol, 1.5 eq) in water (96 mL) and MeCN (96 mL) was added dropwise a solution of Fmoc-OSu (5.70 g, 16.9 mmol, 1 eq). The solution was stirred overnight. Then, DCM (200 mL) was added. The aqueous layer was extracted twice with DCM (2 * 100 mL). The organic layer was then dried under MgSO₄ and evaporated, giving a white solid (6,174 mg, 96 %).

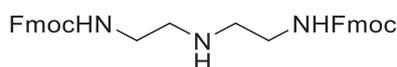
White solid; ¹H NMR (300 MHz, CDCl₃) δ ppm = 7.77 (d, J = 7.3 Hz, 2 H, H_{Ar Fmoc}), 7.60 (d, J = 7.3 Hz, 2 H, H_{Ar Fmoc}), 7.41 (t, J = 7.0 Hz, 2 H, H_{Ar Fmoc}), 7.31 (t, J = 7.7 Hz, 2 H, H_{Ar Fmoc}), 5.30 (br. s., 1 H, NH), 4.90 (br. s., 1 H, NH), 4.41 (d, J = 7.0 Hz, 2 H, CH_{2 Fmoc}), 4.21 (t, J = 6.2 Hz, 1 H, CH_{Fmoc}), 3.35 - 3.21 (m, 4 H, CH_{2-N}), 1.46 (s, 9 H, CH₃), data matched with literature^{201a}.



I-116 (9H-fluoren-9-yl)methyl (2-aminoethyl)carbamate

To a solution of I-115 (9H-fluoren-9-yl)methyl tert-butyl ethane-1,2-diyldicarbamate (6.464 g, 16.9 mmol, 1 eq) in ethyl acetate (550 mL) was added dropwise 18 M HCl (9.41 mL, 169 mmol, 10 eq). The solution was stirred for 4 h. Then, the resulting solid was filtrated, washed with AcOEt and dried under vacuum to yield a white hydrochloride solid (4.251 g, 79 %).

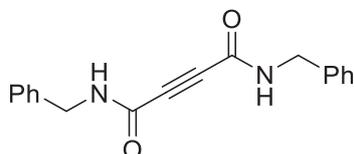
White solid; ¹H NMR (300 MHz, DMSO-d₆) δ ppm = 8.20 (br. s., 3 H, NH₃⁺), 7.89 (d, J = 7.3 Hz, 2 H, H_{Ar Fmoc}), 7.70 (d, J = 7.3 Hz, 2 H, H_{Ar Fmoc}), 7.55 (t, J = 5.4 Hz, 1 H, NH), 7.41 (t, J = 7.2 Hz, 2 H, H_{Ar Fmoc}), 7.33 (t, J = 7.3 Hz, 2 H, H_{Ar Fmoc}), 4.35 - 4.30 (m, 2 H, CH_{2 Fmoc}), 4.26 - 4.18 (m, 1 H, CH_{Fmoc}), 3.27 (q, J = 6.2 Hz, 2 H, CH_{2-NH₂}), 2.84 (q, J = 5.5 Hz, 2 H, CH_{2-NHFmoc}), data matched with literature^{201a}.



I-118 bis((9H-fluoren-9-yl)methyl) (azanediylbis(ethane-2,1-diyl))dicarbamate

To a solution of DETA (0,836 mL, 7,65 mmol, 1 eq) and DIPEA (2.65 mL, 15.2 mmol, 2 eq) in DCM (50 mL) was added dropwise at 0°C a solution of (9-fluorenylmethoxycarbonyl)succinimide (5.00 g, 14.8 mmol, 2 eq) in DCM (33 mL). The solution was stirred overnight. The mixture was then shake with 50 mL of HCl 1 M. The resulting precipitate was filtrated and dried under vacuum, giving a white solid (2.347 g, 56 %)

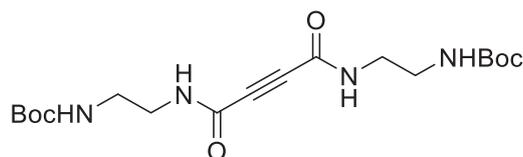
White solid; ^1H NMR (300 MHz, MeOD) δ ppm = 7.76 (d, J = 7.3 Hz, 4 H, $\text{C}_{\text{Ar Fmoc}}$), 7.59 (d, J = 7.3 Hz, 4 H, $\text{C}_{\text{Ar Fmoc}}$), 7.36 (t, J = 7.3 Hz, 4 H, $\text{C}_{\text{Ar Fmoc}}$), 7.27 (t, J = 7.3 Hz, 4 H, $\text{C}_{\text{Ar Fmoc}}$), 4.35 (d, J = 6.4 Hz, 4 H, $\text{CH}_2_{\text{Fmoc}}$), 2.00 (t, J = 6.1 Hz, 2 H, CH_{Fmoc}), 3.45 - 3.36 (m, 4 H, $\text{CH}_2\text{-NCOO}$), 3.22 - 2.88 (m, 4 H, $\text{CH}_2\text{-NH}$), data matched with literature²⁰².



I-120 N^1,N^4 -dibenzylbut-2-ynediamide

To a solution of acetylenedicarboxylic acid (171 mg, 1.50 mmol, 1 eq) with HOBt (486 mg, 3.60 mmol, 2.4 eq) in dry DMF (10 mL) was added at 0°C EDC (636 μL , 3.60 mmol, 2.4 eq) and the mixture was stirred for 1 hour under N_2 . Benzylamine (393 μL , 3.60 mmol, 2.4 eq) was then added at 0°C. The mixture was stirred at room temperature for 16 h. Ethyl acetate and water was then added. Organic layer was washed with brine and evaporated. The remaining residue was purified by column chromatography (SiO_2 , pentane : ethyl acetate = 50 : 50) to give a yellow solid (218 mg, 58 %).

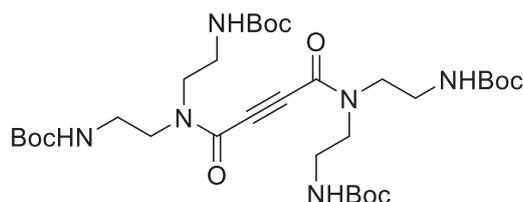
Yellow solid; ^1H NMR (300 MHz, DMSO-d_6) δ ppm = 7.39 - 7.18 (m, 10 H, H_{Ar}), 6.47 (t, J = 5.9 Hz, 2 H, NH), 4.26 (d, J = 5.9 Hz, 4H, CH_2), data matched with literature¹⁹⁹.



I-121 di-tert-butyl (((but-2-ynedioyl)bis(azanediyl))bis(ethane-2,1-diyl))dicarbamate

To a solution of acetylenedicarboxylic acid (66 mg, 0.576 mmol, 1 eq) and **I-114** tert-butyl (2-aminoethyl)carbamate (207 mg, 1.38 mmol, 2.4 eq) in dry DMF (5 mL) was added at 0°C **I-111** 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholin-4-ium chloride (382 mg, 1.38 mmol, 2.4 eq) and the mixture was stirred for 16 hour under N_2 . Ethyl acetate and water was then added. Organic layer was washed with brine and evaporated. The remaining residue was purified by column chromatography (SiO_2 , pentane : ethyl acetate = 50 : 50) to give a white solid (113 mg, 45 %).

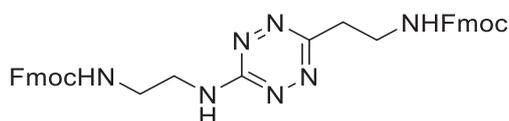
White solid; mp: 186.2 – 188.9 °C; HRMS (ESI) $[\text{M} + \text{Na}]^+$ found 421.2054, calculated 421.2058 for $[\text{C}_{18}\text{H}_{30}\text{N}_4\text{NaO}_6]^+$; ^1H NMR (400 MHz, DMSO-d_6) δ ppm = 8.98 (t, J = 5.2 Hz, 2 H, NH), 6.85 (t, J = 5.4 Hz, 2 H, NH), 3.13 (q, J = 6.0 Hz, 4 H, $\text{CH}_2\text{-NH-Boc}$), 2.99 (q, J = 6.0 Hz, 4 H, $\text{CH}_2\text{-NH-AD}$), 1.37 (s, 18 H, CH_3), ^{13}C NMR (400 MHz, DMSO-d_6) δ ppm = 155.6 (CO_{Boc}), 151.0 ($\text{CO}_{\text{acetylene}}$), 77.8 ($\text{C}_{\text{IV alkyne}}$), 76.6 ($\text{C}_{\text{IV Boc}}$), 39.0 (2^*CH_2), 28.2 (CH_3)



I-124 tetra-tert-butyl (((but-2-ynedioyl)bis(azanetriyl))tetrakis(ethane-2,1-diyl))tetracarbamate

To a solution of **I-128** di-*tert*-butyl (azanediylbis(ethane-2,1-diyl))dicarbamate (1 g, 3.3 mmol, 2.2 eq) and acetylenedicarboxylic acid (171 mg, 1.5 mmol, 1 eq) in DMF (5 mL) was added HCTU (1.37 g, 3.3 mmol, 2.2 eq). The solution was stirred overnight. Then, DCM and water were added. The organic layer was washed twice with brine, dried under MgSO₄ and evaporated. The resulting oil was purified by column chromatography (SiO₂, pentane : ethyl acetate = 1:1 -> 0:1) to give a yellow oil (267 mg, 26 %).

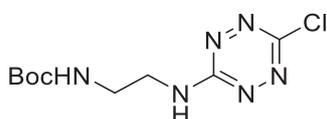
Yellow oil; HRMS (ESI) [M + H]⁺ found 685.4128, calculated 685.4131 for [C₂₂H₂₅N₂O₂S₂]⁺.



I-126 (9H-fluoren-9-yl)methyl (2-((6-chloro-1,2,4,5-tetrazin-3-yl)amino)ethyl)carbamate

To a solution of **I-116** (9H-fluoren-9-yl)methyl (2-aminoethyl)carbamate hydrochloride (200 mg, 0.62 mmol, 2 eq) with DIPEA (108 μL, 0.62 mmol, 2 eq) in methanol (10 mL) was added 3,6-dichlorotetrazine (50 mg, 0.31 mmol, 1 eq). The mixture was stirred for 2 hours at reflux. Then, solvent was evaporated. The remaining residue was purified by column chromatography (SiO₂, dichloromethane : methanol = 95 : 5) to give an orange solid (122 mg, 99 %).

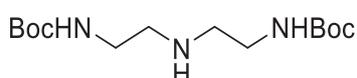
Orange solid, no melting point; MS (ESI) [M + Na]⁺ found 419.1, calculated 419.1 for [C₁₉H₁₇ClN₆NaO₂]⁺; ¹H NMR (300 MHz, DMSO-d₆) δ ppm = 8.93 - 8.81 (m, 1 H, NH), 7.87 (d, J = 7.2 Hz, 2 H, H_{Ar}), 7.66 (d, J = 6.8 Hz, 2 H, H_{Ar}), 7.48 - 7.25 (m, 4 H, H_{Ar}), 4.31 (d, J = 6.2 Hz, 2 H, CH₂Fmoc), 4.21 (d, J = 5.7 Hz, 1 H, CH), 3.49 (d, J = 5.6 Hz, 2 H, CH₂-NH-C_{tetrazine}), 3.27 (d, J = 5.9 Hz, 2 H, CH₂-NHFmoc); ¹³C NMR (300 MHz, DMSO-d₆) δ ppm = 161.5 (C_{IV} tetrazine), 158.6 (CO), 156.2 (C_{Ar}-Cl), 143.9 (C_{Ar}), 140.7 (C_{Ar}), 127.6 (C_{Ar}), 127.0 (C_{Ar}), 125.1 (C_{Ar}), 120.1 (C_{Ar}), 65.4 (CH₂Fmoc), 46.7 (CH_{Fmoc}), 40.5 (CH₂-NH_{tetrazine}), 38.8 (CH₂-NHFmoc).



I-127 di-*tert*-butyl (((1,2,4,5-tetrazine-3,6-diyl)bis(azanediyl))bis(ethane-2,1-diyl))dicarbamate

To a solution of **I-114** *tert*-butyl (2-aminoethyl)carbamate (422 mg, 2.64 mmol, 2 eq) with DIPEA (460 μL, 2.64 mmol, 2 eq) in acetonitrile (20 mL) was added 3,6-dichlorotetrazine (200 mg, 1.32 mmol, 1 eq). The mixture was stirred for 2 hours at room temperature. Then, solvent was evaporated. The remaining residue was purified by column chromatography (SiO₂, pentane : ethyl acetate = 70 : 30) to give an orange oil (359 mg, 68 %).

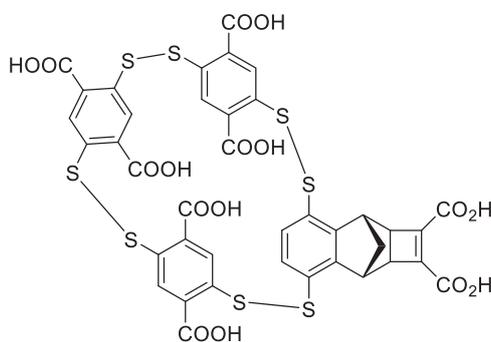
Red solid; mp: 106.1 – 109.0 °C; HRMS (ESI) [M + Na]⁺ found 421.2289, calculated 421.2282 for [C₁₆H₃₀N₈NaO₄]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.43 (t, J = 5.1 Hz, 2 H, NH_{tetrazine}), 5.40 (t, J = 5.5 Hz, 2 H, NH_{Boc}), 3.65 (q, J = 5.7 Hz, 4 H, CH₂-NH_{tetrazine}), 3.42 (q, J = 5.7 Hz, 4 H, CH₂-NHBoc), 1.34 (s, 18 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 161.5 (C_{IV} tetrazine), 156.6 (CO), 79.7 (C_{IV} Boc), 42.2 (CH₂-N_{tetrazine}), 39.3 (CH₂-N_{Boc}), 28.2 (CH₃)



I-128 di-*tert*-butyl (azanediylbis(ethane-2,1-diyl))dicarbamate

To a solution of imidazole (1.566 g, 23 mmol, 1 eq) in DCM (10 mL) was added di-tert-butyl dicarbonate (5.24 g, 24 mmol, 1.05 eq) portion wise. The reaction mixture was stirred for one hour at room temperature. The reaction mixture was washed with 3*20 mL water, dried over Na₂SO₄, filtered and the volatiles were removed under reduced pressure. The residue was dissolved in 15 mL toluene and diethylenetriamine (1.2 mL, 11 mmol, 0.5 eq) was added. The reaction mixture was stirred for two hours at 60°C. Then, 20 mL DCM was added, and the organic phase was washed with 2*20 mL water. The organic phase was dried over Na₂SO₄, filtered and reduced under reduced pressure. The remaining residue was purified by column chromatography (SiO₂, DCM : MeOH : NH₃ = 10:1:0.1) to give a colourless oil (2.889 g, 86 %).

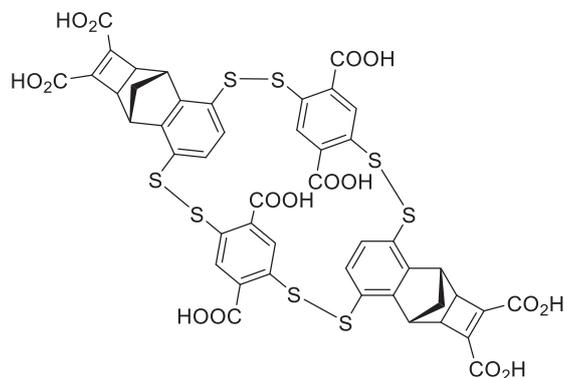
¹H NMR (300 MHz, CDCl₃) δ ppm = 5.04 (br. s., 2 H, (CO)NH), 3.24 - 3.09 (m, 4 H, CH₂-NH(CO)), 2.69 (t, J = 5.9 Hz, 4 H, CH₂-NH), 1.41 (s, 18 H, CH₃); ¹³C NMR (300 MHz, CDCl₃) δ ppm = 156.1 (CO), 79.0 (C_{tBu}), 48.7 (CH₂-N), 40.3 (CH₂-NH(CO)), 28.3 (CH₃), data matched with literature reference²¹⁰



1-1₃-108

To a stirred solution of **I-1** 2,5-dimercaptoterephthalic acid (1.15 mg, 0.005 mmol, 1 eq) and **I-108** (3R,8S)-2-(hydroperoxy-λ²-methyl)-4,7-dimercapto-2a,3,8,8a-tetrahydro-3,8-methanocyclobuta[b]naphthalene-1-carboxylic acid (1.65 mg, 0.005 mmol, 1 eq) in ammonium acetate buffer (50 mM, pH without adjustment) (1 mL). The mixture was analysed by mass spectrometry after 72 h of stirring at room temperature.

HRMS (ESI) [M - H]⁺ found 1000.8621, calculated 1000.8601 for [C₃₉H₂₁O₁₆S₈]⁻



1-1₂-108₂

To a stirred solution of **1-1** 2,5-dimercaptoterephthalic acid (1.15 mg, 0.005 mmol, 1 eq) and **1-108** (3R,8S)-2-(hydroperoxy-λ²-methyl)-4,7-dimercapto-2a,3,8,8a-tetrahydro-3,8-methanocyclobuta[b]naphthalene-1-carboxylic acid (1.65 mg, 0.005 mmol, 1 eq) in ammonium acetate

buffer (50 mM, pH without adjustment) (1 mL). The mixture was analysed by mass spectrometry after 72 h of stirring at room temperature.

HRMS (ESI) $[M - H]^+$ found 1090.9116, calculated 1090.9070 for $[C_{46}H_{27}O_{16}S_8]^-$

b. DOSY NMR experiments

DOSY experiments were driven on a 500 MHz NMR Bruker Advance III spectrometer with BBI probe with standard 50 G.cm^{-1} Z gradient. To avoid convection due to inhomogeneous temperature in the sample, the height of liquid was limited to approximately 37 mm. Pulses were calibrated prior acquisition. The standard ledbpgp1s were generally used. The gradient pulses δ and the delay Z were adjusted to reach an attenuation of approximately 95% at maximum gradient. Processing used the Topspin package and Dynamics center software. Dioxane as internal reference was introduced for each experiment. Stokes-Einstein equation (1) was used to evaluate hydrodynamic radii in the solvent of analysis from diffusion coefficients:

$$R_H = \frac{k_b T}{6\pi\eta D} \quad (1)$$

For each analysis, either the reported values of the viscosity η of the medium in the same experimental conditions (temperature, concentration), following reported DOSY analyses or reported hydrodynamic radii value $R_H = (2.12 \text{ \AA})$ of dioxane in the same experimental conditions (solvent, concentration and temperature) was used through equation (2) derived from (1) :

$$R_H = \frac{R_H(\text{ref})D(\text{ref})}{D} \quad (2)$$

Where $R_H(\text{ref})$ and $D(\text{ref})$ are reported values of the reference's hydrodynamic radius and $D(\text{ref})$ the value of the diffusion coefficient obtained experimentally in this study.

The modelled external and hydrodynamic radii were measured from the optimised structures in DFT of the diastereoisomer In_In_In_In_MPMP (Figure 15). The centroid of all atoms was calculated. The external radius was defined as the distance between the centroid and the more distant C_{sp^2} atoms of the 4-membered ring, while the hydrodynamic radius was defined as the distance between the centroid and the more distant oxygen of the carboxylate moieties.

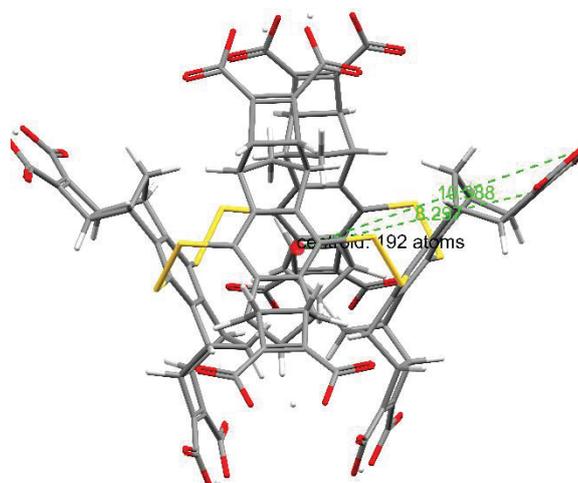


Figure 15 : Measurement of the modelled external and hydrodynamic radii of the In_In_In_In_MPMP diastereoisomer of I-105₄ from the centroid of all atoms.

c. DCC, HPLC and UPLC/HRMS analysis

i. Library preparation

In a typical experiment, the building blocks were individually dissolved in an aqueous buffer (Tris pH 7.4 200 mM for example) at a given pH or in organic media to obtain building block stock solutions at [2-10] mM. The stock solutions were freshly prepared and quickly engaged in the DCLs: in 1.5 mL vials equipped with a turbulent were placed 500 μ L of the building block 1 and 500 μ L of building block 2 solution stocks. The vials were left open to the air and stirred at room temperature for 48 hours or more. Templates were freshly prepared at 100 mM solution stocks in the corresponding media. If required, 10 μ L were added in the library.

ii. Antiparralel experiments

The 50 mM borate buffer was prepared from boric acid dissolved in distilled water, and adjusted with 1.0 M KOH to pH 8.2. Afterward, it was degassed by argon purging for 60 min. Libraries were prepared in clear vials under inert atmosphere. Library solutions were stirred using Teflon-coated microstirrer bars on a magnetic stirrer at 1100 rpm. Stock solutions at 5.0 mM of vinyl sulfide, 10 mM of bisthiol, and a 10 mM stock solution of NaBO₃ was prepared. Equal volumes of bisthiol and NaBO₃ solutions were mixed to obtain 5.0 mM oxidised bisthiol and left stirring for 3 h before further use. Simultaneously, bisthiol was diluted twice with buffer solution to obtain 5.0 mM unoxidized A. Libraries were prepared by mixing adequate volumes of vinyl sulfide, reduced bisthiol, and oxidised bisthiols. The volume of each library was 1 mL.

iii. HPLC analysis

Analysis are performed using a 1260 infinity Agilent Technologies system at 25 °C.

Method A: column Agilent Eclipse Plus C₈, 3.5 μ m, 4.6 x 150 mm, V_{inj} = 3 μ L, flow 1 mL/mn. Gradient water + 0.05 % TFA / MeCN + 0.05 % TFA in 25 mn method (t₀ 80/20, t₅ 60/40, t₈ 0/100).

Method B: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 μ L, 2.7 μ m, 2.1 x 50 mm flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 22 mn method (t₀ 100/0, t₁₅ 70/30, t₂₂ 0/100).

Method C: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 μ L, 2.7 μ m, 2.1 x 50 mm flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 28 mn method (t₀ 100/0, t₂ 100/00, t₂₀ 80/20, t₂₅ 0/100, t₂₈ 0/100)

Method D: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 μ L, 2.7 μ m, 2.1 x 50 mm flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 25 mn method (t₀ 100/0, t_{1.5} 100/00, t₁₅ 85/15, t₂₄ 0/100, t₂₅ 0/100)

Method E: Agilent Nucleodur HILIC EC 250 4.6 V_{inj} = 5 μ L, 5 μ m, 4.6 x 250 mm flow 0.8 mL/mn. Gradient AcONH₄ 50 mM / MeCN in 25 mn method (t₀ 10/90, t₂₅ 60/40)

Method F: Restek Raptor Biphenyl, 2.7 μ m, 3.00 x 50 mm, V_{inj} = 3 μ L, flow 0.6 mL/mn. Gradient water + 0.1 % formic acid / MeOH + 0.1 % formic acid in 17 mn method (t₀ 99/1, t_{0.5} 99/1, t₁₅ 50/50, t₁₆ 1/99, t₁₇ 1/99)

Method G: Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 10 μL, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t₀ 90/10, t₁₀ 50/50, t₁₅ 40/60, t₁₇ 0/100, t₂₀ 0/100)

Method H: Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 10 μL, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t₀ 100/0, t₁ 100/0, t₁₅ 40/60, t₁₇ 0/100, t₁₉ 0/100)

Method I: Restek Raptor Biphenyl, 2.7 μm, 3.00 x 50 mm, V_{inj} = 5 μL, flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeOH + 0.1 % formic acid in 17 mn method (t₀ 70/30, t₁₅ 0/100, t₁₇ 0/100)

Method PREP₁: Zorbax Eclipse XDB C₈ 5.0 μm 250 x 9.4 mm V_{inj} = 100 μL, flow 5.0 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t₀ 92/8, t₂ 92/8, t₁₄ 50/50)

iv. UPLC analysis

Analysis are performed using a DIONEX ultimate 3000 Thermo Scientific system.

Method J: 30 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 4 μL, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 20 mn method (t₀ 90/10, t₂₀ 50/50)

Method K: 45 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 4 μL, flow 0.4 mL/mn. Gradient water + 0.1 % TFA / MeCN + 0.1 % TFA in 8 mn method (t₀ 85/15, t₈ 58/42)

Method L: 45 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 4 μL, flow 0.4 mL/mn. Gradient water + 0.1 % TFA / MeCN + 0.1 % TFA in 8 mn method (t₀ 85/15, t₈ 40/60)

v. UPLC/HRMS analysis

Analysis are performed using a U3000 Thermo Fisher Scientific / QTOF Impact II Bruker system. The column used was an Agilent Poroshell 120 EC C₈, 2.7 μm, 5 x 2.1 mn, V_{inj} = 3 μL, flow 0.8 mL/mn, Gradient water + 0.1 % formic acid / MeCN:MeOH (50/50) (or MeCN) + 0.1 % formic acid for typical run, other methods have been used. Mass spectrometry parameters depend on the analysis, UV detection performed at 250 and 350 nm.

vi. Semi-preparative chiral HPLC

Sample preparation: About 9 g of compound **I-98** + **I-99** are dissolved in 380 mL of ethanol. Chromatographic conditions: Lux-Amylose-2 (250 x 10 mm), Hexane / EtOH (80/20) as mobile phase, flow-rate = 5 mL / min, UV detection at 290 nm. Injections (stacked): 475 times 800 μL, every 4 minutes

d. Crystallography

Crystals were obtained by slow diffusion of hexane into a solution of the compound of interest in DCM. Suitable crystals from **1-15₂** (not yet deposited on the CCDC) were selected and mounted on a Gemini kappa-geometry diffractometer (Agilent Technologies UK Ltd) equipped with an Atlas CCD detector using Mo radiation (λ = 0.71073 Å). Intensities were collected at 150 K by means of the CrysAlisPro software²². Reflection indexing, unit-cell parameter refinement, Lorentz-polarisation correction, peak integration and background determination were carried out using CrysAlisPro software. An analytical

²²CrysAlisPro, Agilent Technologies, Version, 1.171.34.49, (release 20–01–2011 CrysAlis171.NET), compiled Jan 20 2011, 15:58:25.

absorption correction was applied using the modelled faces of the crystal²¹¹. The resulting sets of hkl were used for structure solution and refinement. The structures were solved using direct methods with SIR97²¹² and the least-squares refinement on F_2 was achieved using CRYSTALS software²¹³. OLEX2²¹⁴ and SHELXT²¹⁵ software were also used for this purpose. All non-hydrogen atoms were refined anisotropically. The hydrogen atoms were initially positioned geometrically and initially refined with soft restraints on the bond lengths and angles to regularise their geometry (C–H in the range 0.93–0.98 Å) and $U_{iso}(H)$ (in the range 1.2–1.5 times U_{eq} of the parent atom), after which the positions were refined with riding constraints.

Regarding **I-105₄**, high concentrate aqueous solution and use of suitable commercial crystallization screen solutions was employed to try different precipitants, buffers, additives etc. The principle is vapor diffusion method where the tiny droplets are either hanging or sitting are enabled to equilibrate with the crystal solution in reservoir a in sealed tray set ups. Droplets usually comprise of equal volumes of sample and crystal solution (ex. 1 μ L of sample + 1 μ L of crystal sol.) and in the reservoir or well was added a higher volume of crystal solution or precipitant. After sealing the setup, was driven vapor diffusion so that the drop with sample slowly super-saturates to give crystals. A typical screen contains 96-conditions. X-ray was a D8 Venture Bruker with liquid MetalJet.

e. pH titrations

Experiments were performed on a Metrohm 848 Titrino plus at 25°C. The machine was calibrated prior use with adequate buffer solution (7.00; 10.01). NaOH solution was calibrated prior use with analytical grade potassium hydrogen phthalate. 11.0 mg of **I-105₄** were dissolved in 1 mL HCl (0.0958 mM) + 9 mL milliQ water. The solution was then titrated with a solution of NaOH solution (49.0 mM). The same experiment was done without the macrocycle

f. ITC^{136,216}

i. Standard procedure

Typical experiments were performed at 298 K using a Malvern ITC200 instrument. The host solution in Tris buffer pH 7.4 (200mM) was placed into the calorimeter cell (200 μ l) and 20 successive aliquots (2 μ l) of guest solution (10 times more concentrated) were added via a computer-automated injector at 2 mn intervals (Reference power 5, Initial delay 60, Stirring speed 1000, t_{inj} : 4.0 s, filter period 5 s, feedback mode/gain high, edit mode unique). Heat changes were recorded after each addition. Heats of dilution were subtracted from the titration data prior to curve fitting. The first injection was discarded from each dataset to remove the effect of guest diffusion across the syringe tip during the equilibration process. Titrations curves were fitted with the one binding site model using Origin v.5.0 software supplied by MicroCal. The ITC apparatus was provided by the IBCP, Lyon 1, apart from GFPs / **I-1₄**, performed at the CIP of the University of Liege.

i. GFPs / **I-1₄**

ITC titrations were performed on a MicroCal ITC200 (GE-Malvern) equipped with a 200 μ l Hastelloy sample cell and an automated 40 μ l glass syringe rotating at 1000 rpm. GFP variants were dialysed overnight at 4 °C. The receptor was solubilised in the dialysis buffer, sonicated and filtered on 0.22 μ m membranes. In a standard experiment, GFP variants (559 μ M) were titrated by one initiating injection

(0.4 μ l) followed by 19 injections (2 μ l) of receptor (5.6 mM) at an interval of 120 s. Dilution heats of receptor injections into buffer were subtracted from raw data, whereas dilution heats of buffer injections into GFP variants were undetectable. The data so obtained were fitted via nonlinear least squares minimisation method to determine binding stoichiometry (n), association constant (K_a), and change in enthalpy of binding (ΔH_b°) using ORIGIN 7 software v.7 (OriginLab). The Gibbs free energy of binding, ΔG_b° , was calculated from K_a : $\Delta G_b^\circ = -RT \ln K_a$ and the entropic term, $T\Delta S_b^\circ$, was derived from the Gibbs-Helmholtz equation using the experimental ΔH_b° value: $\Delta G_b^\circ = \Delta H_b^\circ - T\Delta S_b^\circ$

g. Octet[®]

Experiments were performed on an Octet[®] HTX Pall ForteBio in Tris 100 mM pH 8. The fibre Ni-NTA part n^o18-5101 was conditioned with buffer, then the GFP was grafted. Baseline was evaluated with buffer, then the measurement was conducted after addition of I-1₄ (5.6 mM) in Tris 100 mM pH 8. The receptor was then released by buffer (baseline 300 s, association 600 s, dissociation 900 s, loading 300 s, 300 s equilibration, flow rate 1 mL/mn).

h. μ -FFE

Electrophoresis were performed at room temperature on a customised “chip” in Tris 20 mM at pH 8. GFPs and I-1₄ were prepared at 0.3 and 3 mM.

i. DFT

All the molecules have been fully optimized at B3LYP⁷² or APFD²¹⁷ functional /6-311G** (C, H, O, N) 6-311++G** (S) level of theory with Gaussian 09 Rev B.02¹⁹³. Vibrational frequencies have been computed to characterize transition states and stationary points. IRC were employed to check transition states.

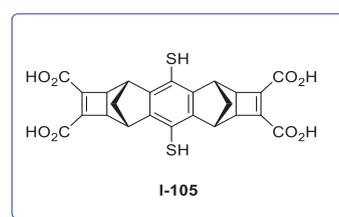
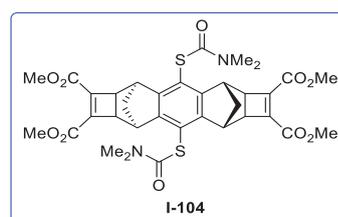
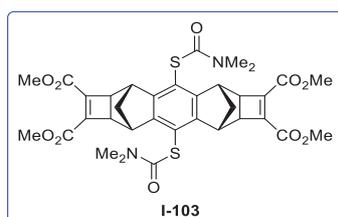
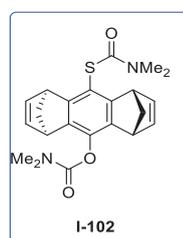
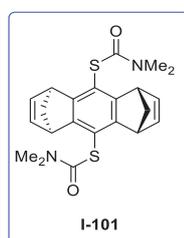
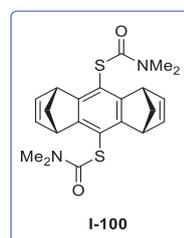
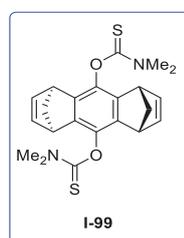
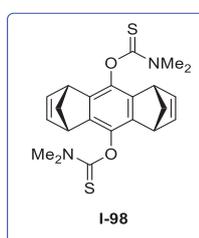
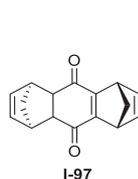
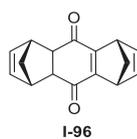
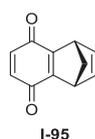
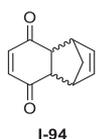
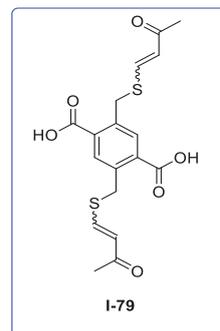
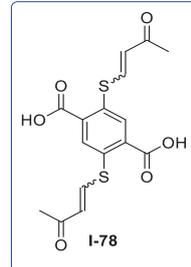
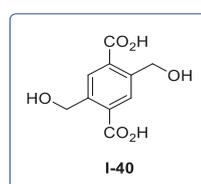
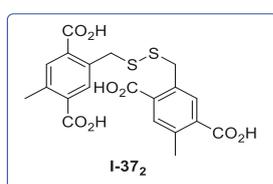
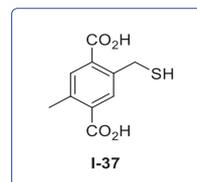
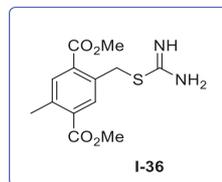
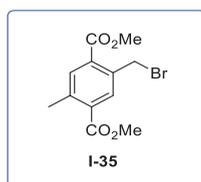
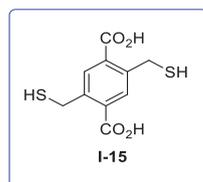
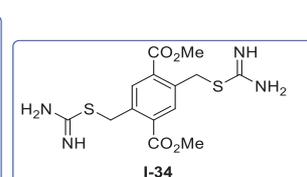
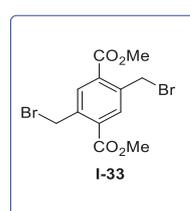
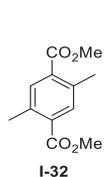
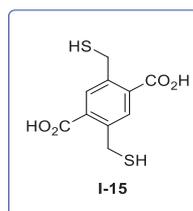
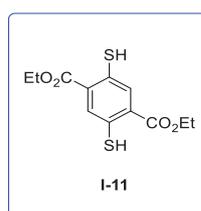
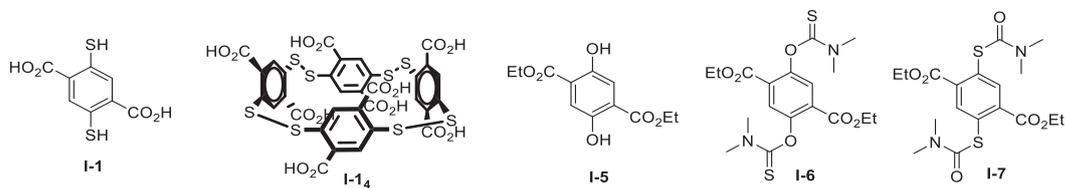
j. Molecular dynamic

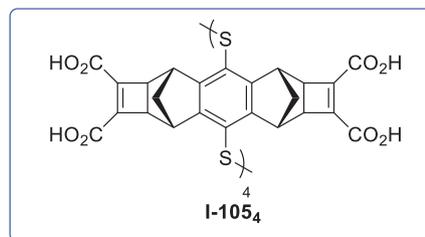
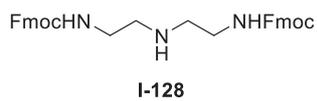
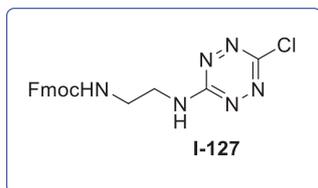
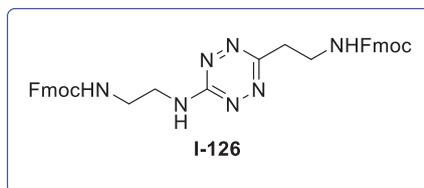
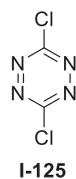
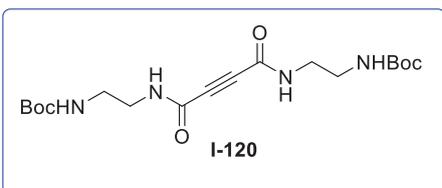
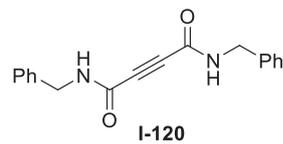
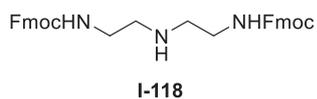
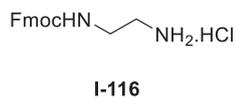
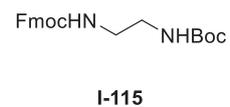
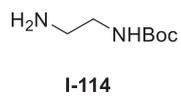
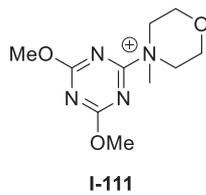
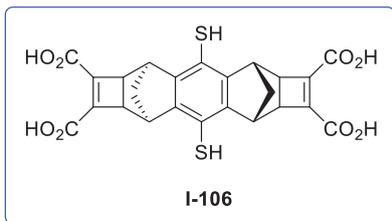
All molecular dynamic simulations and post-processing were performed with the Amber 16 Molecular Dynamics software package^{95b}. The force-field parameters were taken from parm99²¹⁸, while the parameters for the cage were generated using the generalized AMBER force field GAFF²¹⁹. Each compound was previously build using the Gaussview5 software and their geometries were optimized with the Gaussian09 suite of programs¹⁹³ software using density functional theory at the B3LYP level of theory with the 6-31G(d,p) basis set. The different parameters were generated with antechamber and parmcheck subprograms, and atom point charges were computed using the RESP protocol. The guest was inserted inside the host using the xleap module. Potassium cations (K⁺) were added in order to neutralize the systems, which were immersed in a truncated octahedral water TIP3P²²⁰ water box containing around 2500 water molecules. Each system was first minimized in a 5000 steps simulation, including 2500 steps of steepest descent. Then, a thermalization step was performed to heat each system from 0 to 300 K in 30 ps. The temperature was kept constant during the following steps using Langevin thermostat with a collision frequency γ ln of 1 ps⁻¹. A 1030 ps equilibration run was performed in NPT conditions. Finally, a 100 ns production was executed with constant pressure. After all dynamic molecular simulations, analysis were performed using the cpptraj module^{95a} of AMBER, and thermodynamic parameters were extracted using the MM/GBSA method²²¹ with the internal and external dielectric constants set to 4 and 80, respectively. The salt concentration was assigned to 0.150 M. Trajectories were visualized with the chimera software¹¹⁸.

Appendices

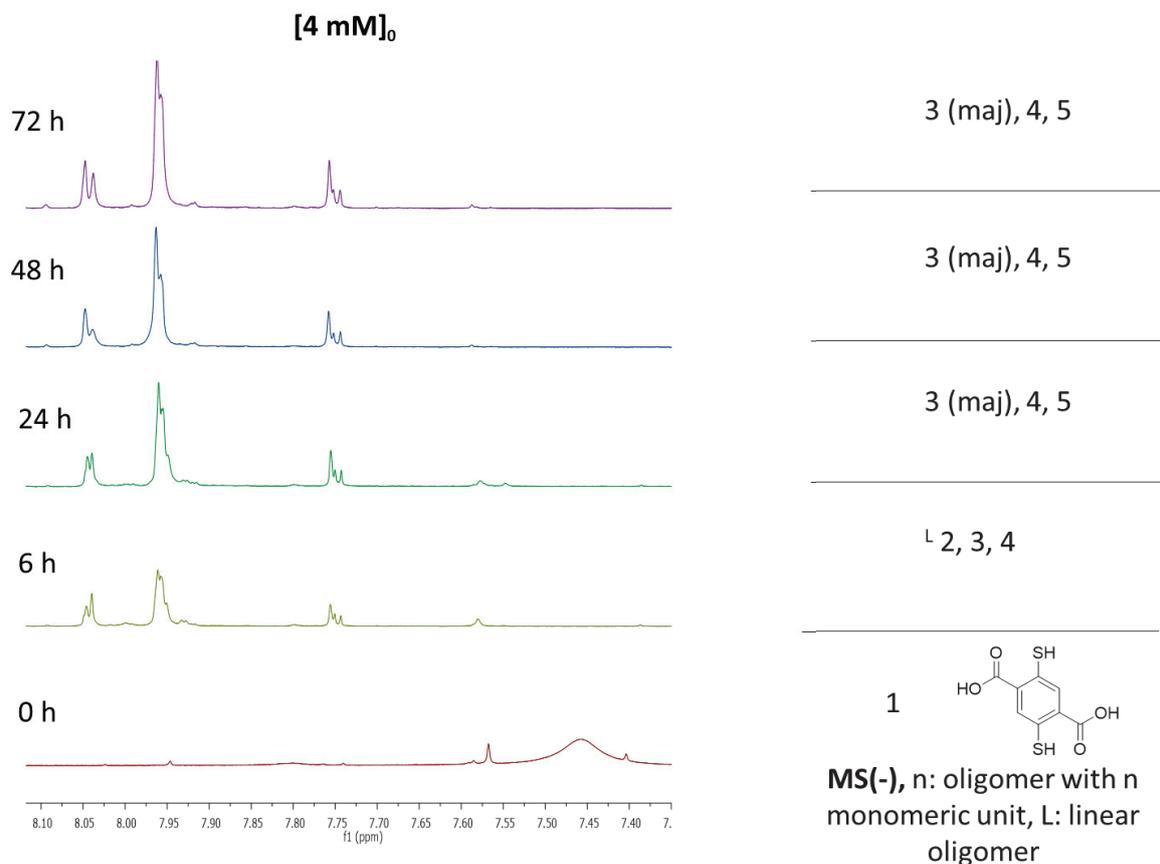
a. Appendix I-I: List of molecules

Synthesised molecules of chapter 1, for isolated and ESI-MS analysis + ^1H NMR identification on pure product. New and fully characterised molecules in a box.





b. Appendix I-II: Additional conditions of reactions



Scheme 81: ¹H NMR and ESI-MS analysis control of the DCL of I-1 alone in D₂O pH 7.4 at 4 mM.

Entry	Organic solvent	I-11 concentration	I-12 concentration	template
1	MeCN	4 mM	/	/
2	DMSO	4 mM		
3	DMSO	1.33 mM		
4	DMSO	/	4 mM	spermine
5	DMSO	1.33 mM	/	
6	DMSO	/	4 mM	

Table 6: DCL conditions in order to obtain a I-1_n mutant, combining the building block I-1, I-11, and I-12. I-1 at 4 mM in aqueous buffer, the other ones in organic solvent before mixing and stirring.

Entry	Conditions
1	NBS, NMP, 60°C, 24 h ²²²
2	NBS, AIBN, DMF, reflux, 24 h ²²³
3	Br ₂ , MnO ₂ , DMF, 50°C, 24 h ²²⁴

Table 7: Screening of the Wohl-Ziegler reaction of I-31

DCL	I-105	I-106	I-108	Conditions
1	5 mM			Tris 200 mM pH 7.4
2	5 mM	/	/	NH ₄ OAc 200 mM pH 7.4
3				NH ₄ OAc 50 mM pH 7.4
4				NH ₄ OAc 50 mM without adjustment
5				NH ₄ F 50 mM pH 6.0
6				NH ₄ Cl 50 mM pH 6.0
7				Ammonium formate 50 mM pH 6.0
8				NaOAc 50 mM pH 6.0
9				MilliQ water, pH 6.0
10				/
11	5 mM	/		MeOH, I ₂
12	/	/	5 mM	NH ₄ OAc 50 mM without adjustment

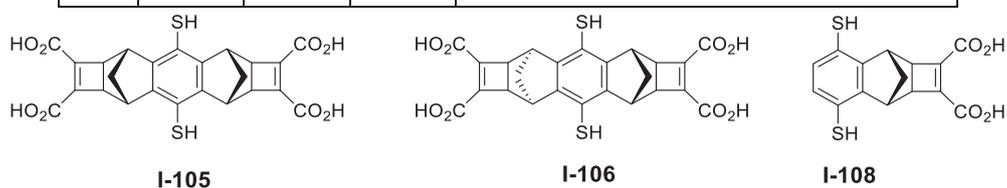
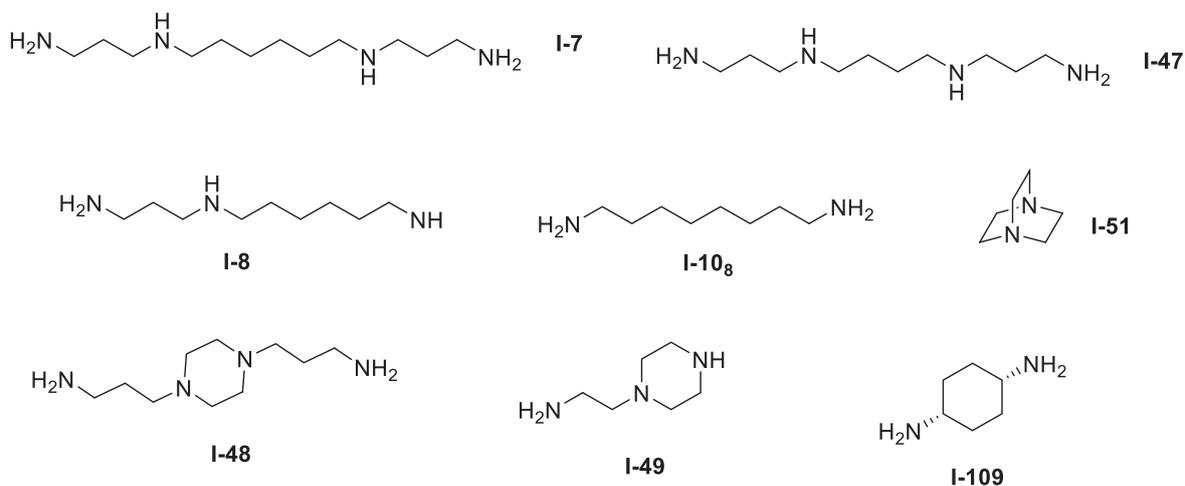
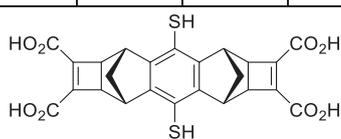


Table 8: DCL conditions involving building blocks I-105, I-106 and I-108

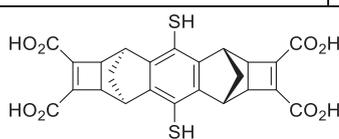


Scheme 82: Cationic templates used for DCLs involving building blocks I-1 and I-108

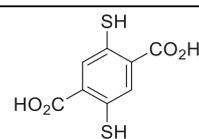
DCL	I-1	I-105	I-106	I-108	Conditions	Template (1.25 mM)			
1	2.5 mM	2.5 mM	/	/	NH ₄ OAc 50 mM without adjustment	/			
2						Spermine I-7			
3						Super spermine I-47			
4						Spermidine I-8			
5						Octadecylamine I-10 ₈			
6						/			
7		/	2.5 mM			/	/	Tris 200 mM pH 7.4	Spermine I-7
8									Super spermine I-47
9									Spermidine I-8
10									Octadecylamine I-10 ₈
11		2.5 mM	/			/	/	Tris 200 mM pH 7.4	/
12		/	2.5 mM						
13		2.5 mM	/						
14		/	/	/	2.5 mM	NH ₄ OAc 50 mM without adjustment	MeOH, I ₂		
15							Water		
16							Spermine I-7		
17							Super spermine I-47		
18							I-54		
19							I-55		
20							I-57		
21							(1s,4s)-cyclohexane-1,4-diamine cis I-109		



I-105



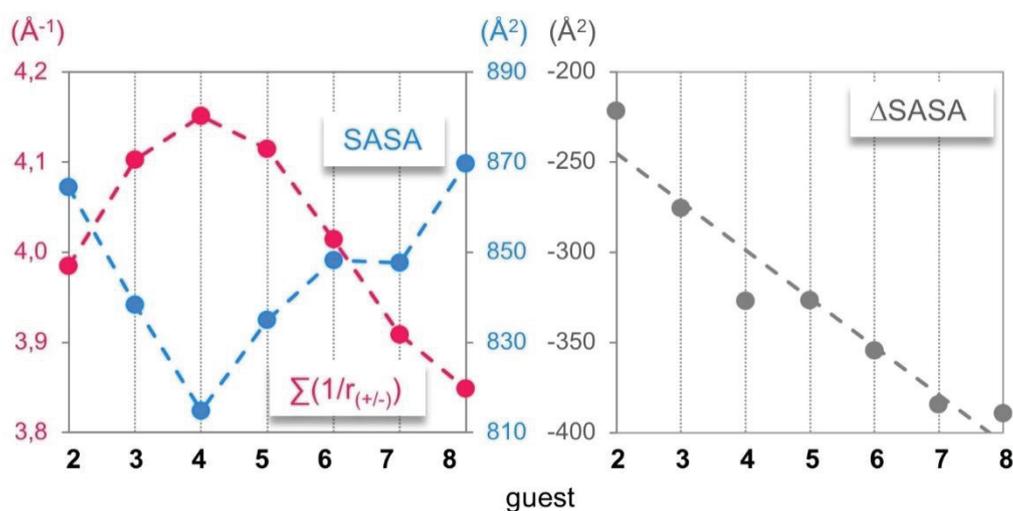
I-106



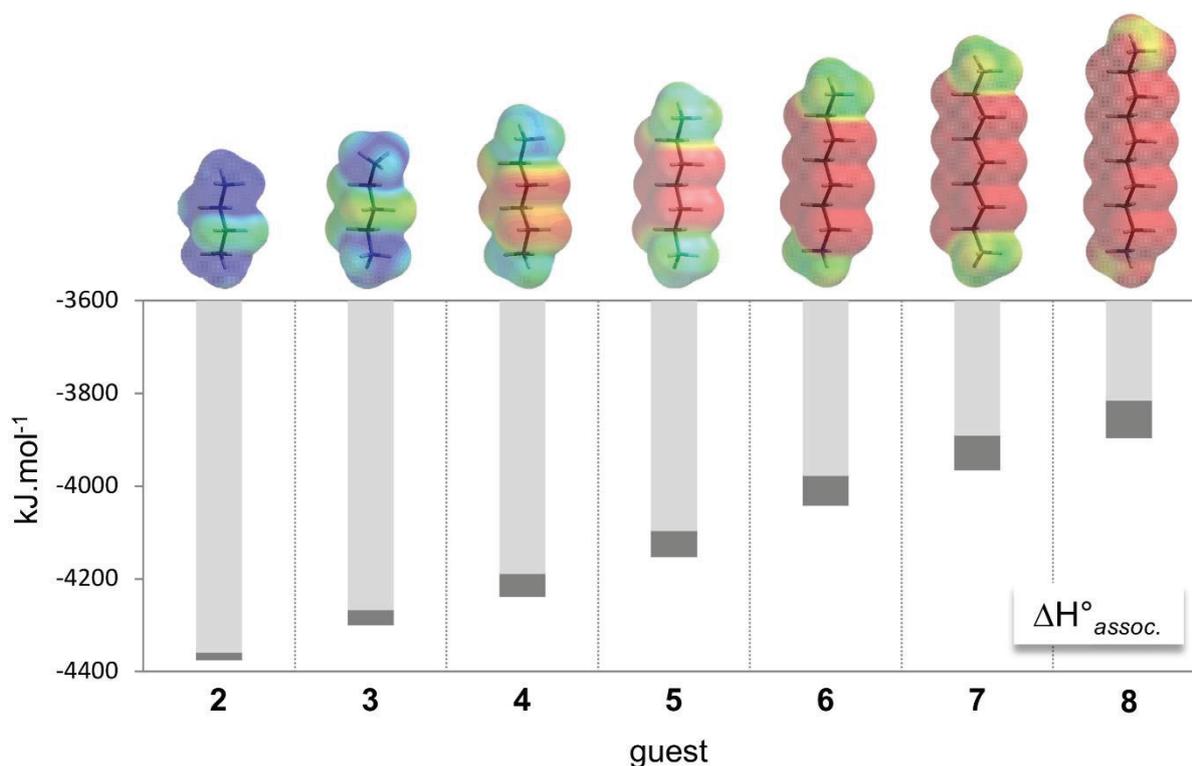
I-1

Table 9: DCL conditions involving building blocks I-1, I-105 and I-106

c. Appendix I-III: Data of the molecular dynamic simulation between I-1 and diamines.



Scheme 83: Electrostatic (+/-), geometric (SASA) complementarity and classical hydrophobic effect (relative SASA) for a lock and key model regarding the association between I-1 and diamines I-10_n with various chain lengths. Extracted from molecular dynamic simulations.



Scheme 84: Evolution of the computed enthalpy of association in the gas phase between host I-1 and guests I-10_n from MD trajectories. The coulombic and van der Waals contributions to $\Delta H^{\circ}_{\text{assoc}}$ are highlighted in light and dark grey, respectively. Each guest is displayed with its corresponding electrostatic potential surface calculated using Spartan 14 software. Color scale from 720 (red) to 980 (blue) kJ mol⁻¹; isovalue $\frac{1}{4}$ 0.002 au.

d. Appendix I-IV: Data of the molecular dynamic simulation between I-1₄ and polyamines I-8 and I-14.

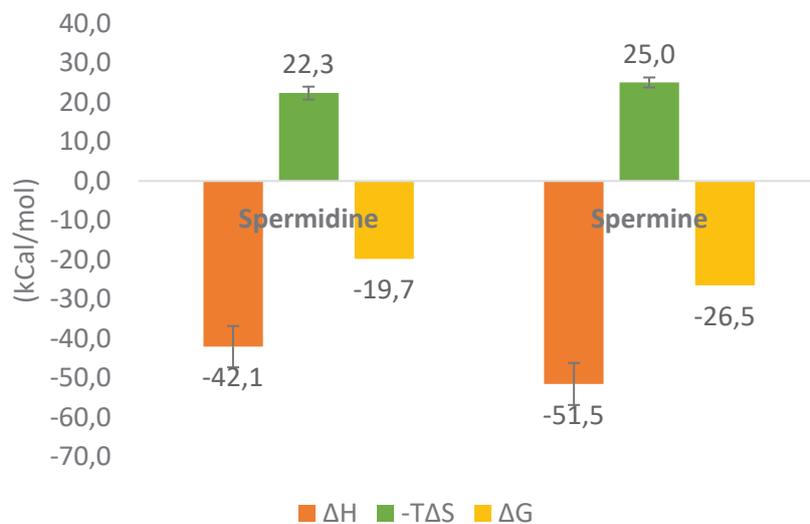
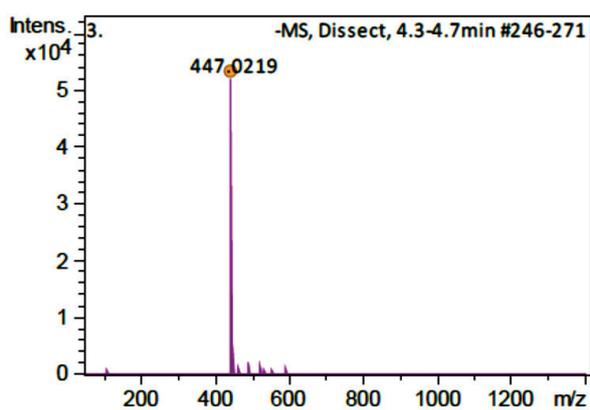
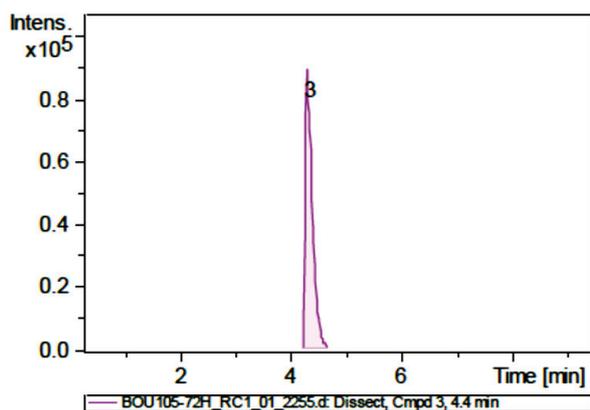
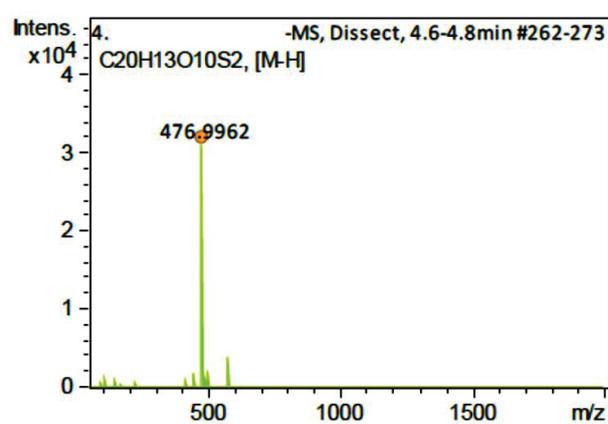
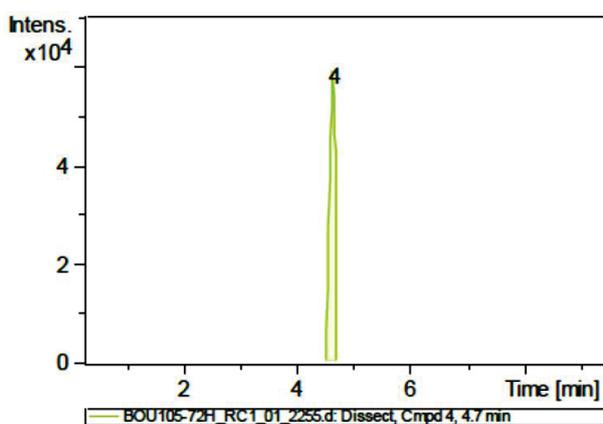


Figure 16: Thermodynamic data of the association between I-1₄ and spermidine/spermine extracted from MM-PBSA calculations.



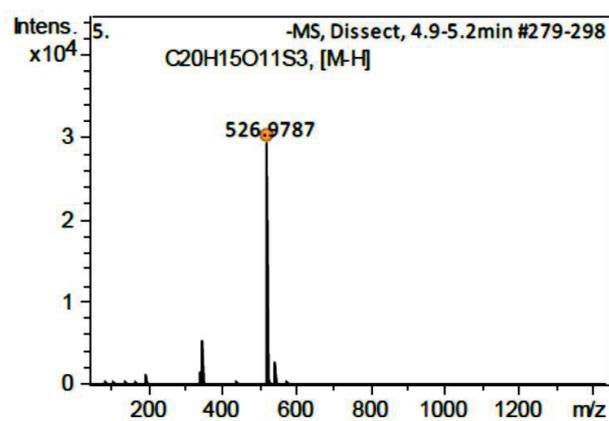
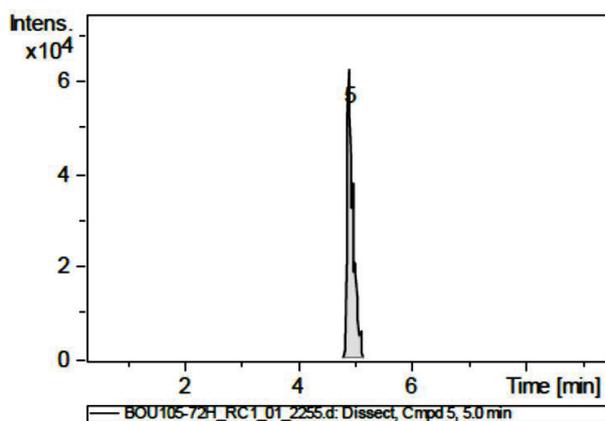
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
447.0219	C ₂₀ H ₁₅ O ₈ S ₂	447.0214	-1.2	15.1	100.00	M-H	1-

Scheme 86: HRMS from LC/MS, corresponding to 2



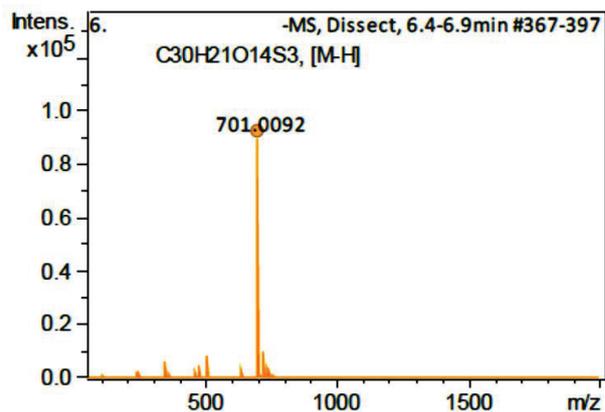
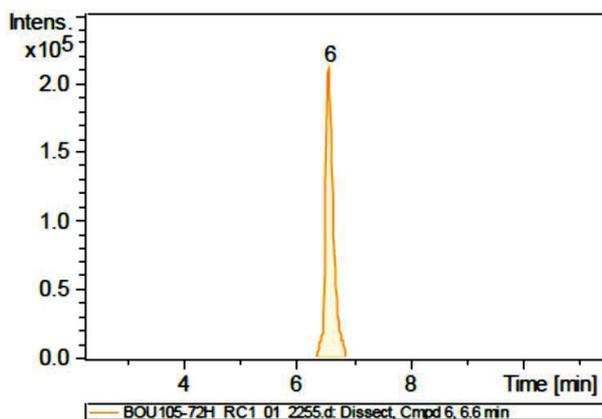
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
476.9962	C ₂₀ H ₁₃ O ₁₀ S ₂	476.9956	-1.4	15.6	100.00	M-H	1-

Scheme 87: HRMS from LC/MS, corresponding to 2 + O²



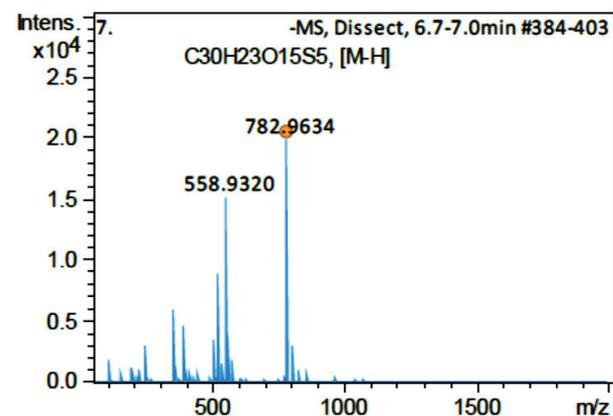
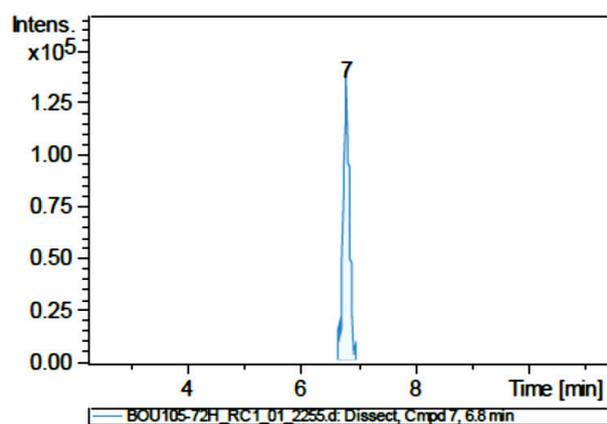
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
526.9787	C ₂₀ H ₁₅ O ₁₁ S ₃	526.9782	-0.9	18.9	100.00	M-H	1-

Scheme 88: HRMS from LC/MS, corresponding to 2 -S + O²



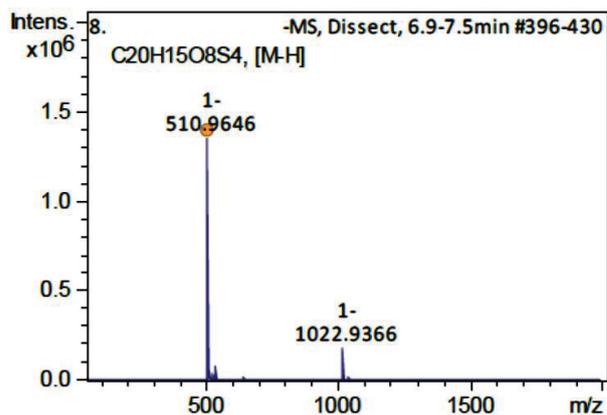
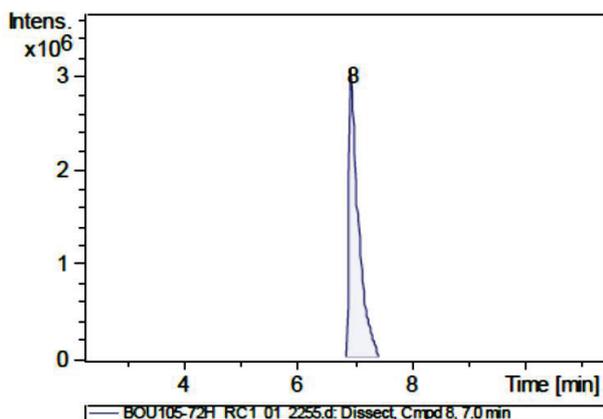
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
701.0092	C30H21O14S3	701.0099	0.9	20.9	85.05	M-H	1-

Scheme 89: HRMS from LC/MS, corresponding to $3-S+O^2-H_2$



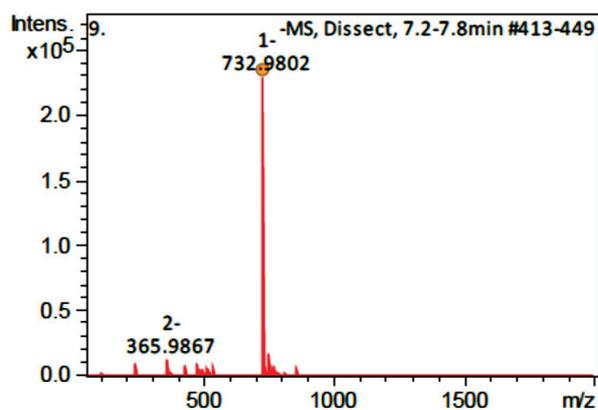
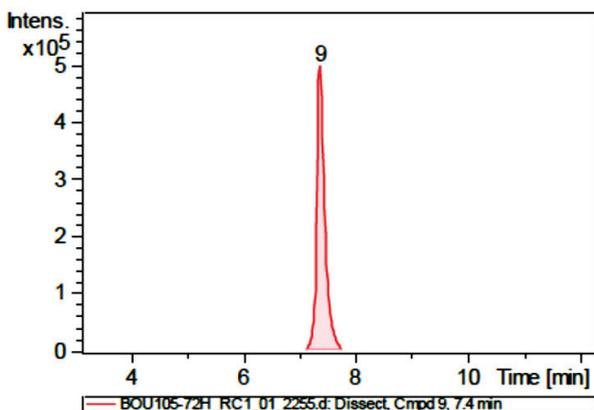
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
782.9634	C30H23O15S5	782.9646	1.6	13.8	100.00	M-H	1-

Scheme 90: HRMS from LC/MS, corresponding to $3-S+O^3$



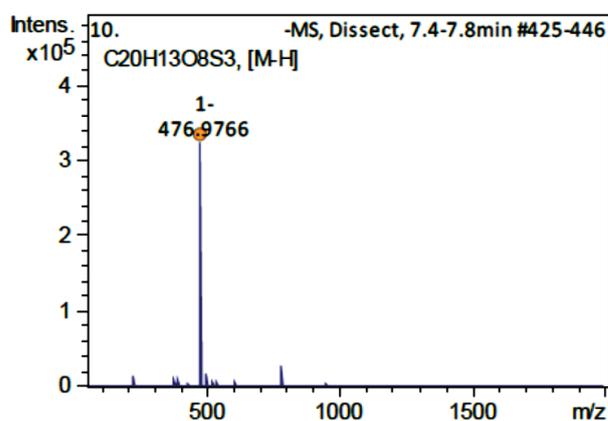
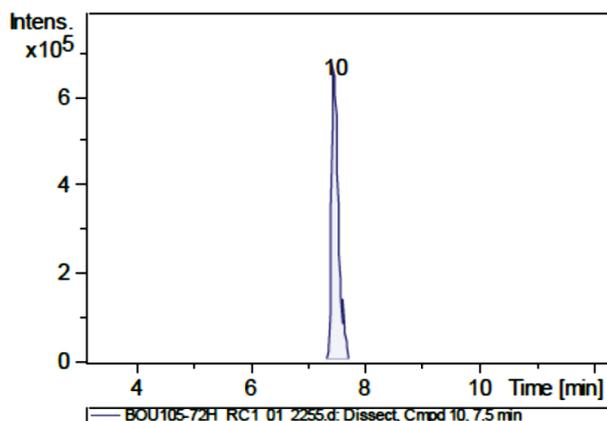
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
510.9646	C20H15O8S4	510.9655	1.8	10.4	100.00	M-H	1-

Scheme 91: HRMS from LC/MS, corresponding to 2



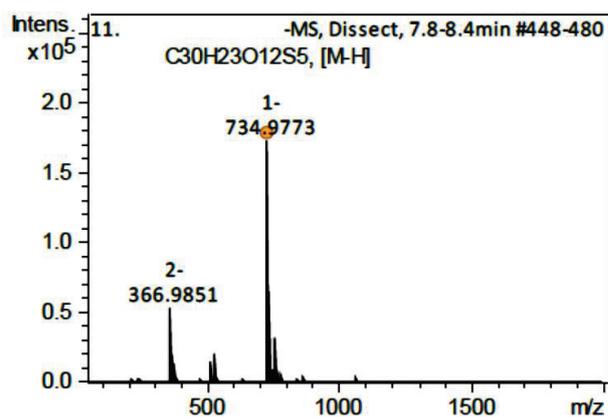
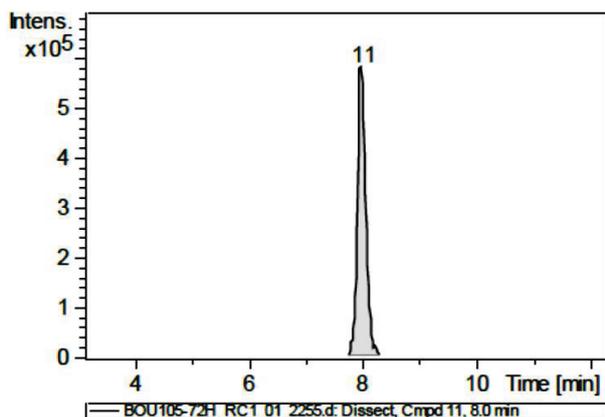
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
732.9802	C30H21O14S4	732.9820	2.4	35.5	41.72	M-H	1-

Scheme 92: HRMS from LC/MS, corresponding to 3 -2S +O²



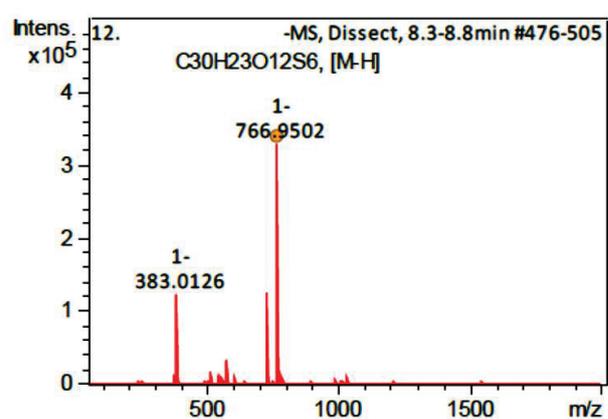
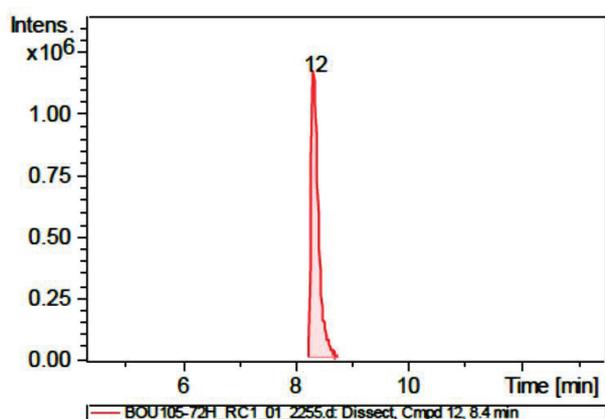
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
476.9766	C20H13O8S3	476.9778	2.6	23.7	76.33	M-H	1-

Scheme 93: HRMS from LC/MS, corresponding to 2 + S -H₂



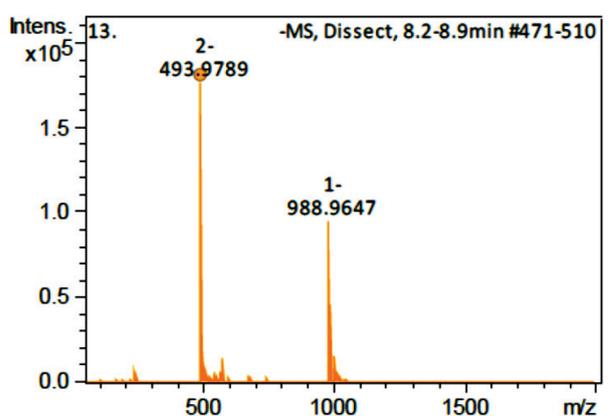
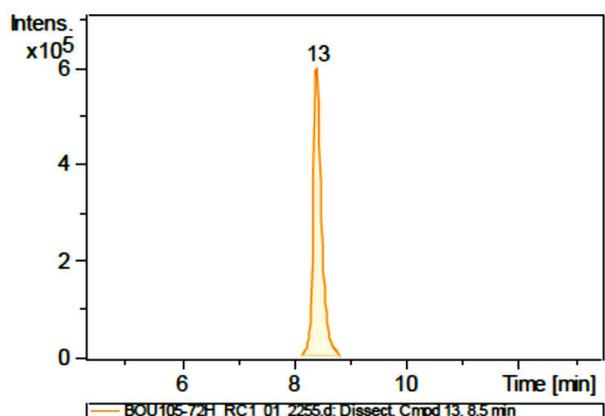
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
734.9773	C30H23O12S5	734.9799	3.5	13.4	25.44	M-H	1-

Scheme 94: HRMS from LC/MS, corresponding to 3 -5



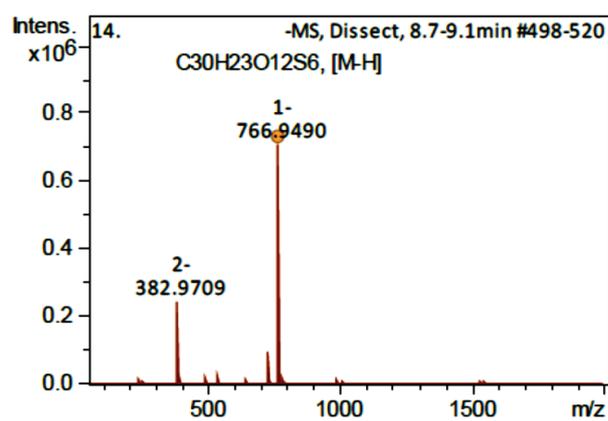
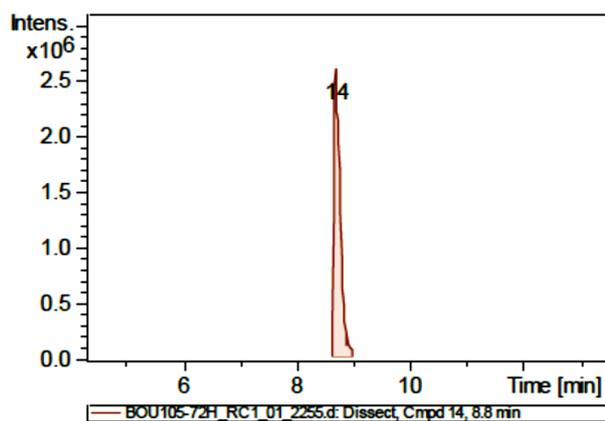
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
766.9502	C30H23O12S6	766.9519	2.3	46.4	67.36	M-H	1-

Scheme 95: HRMS from LC/MS, corresponding to 3



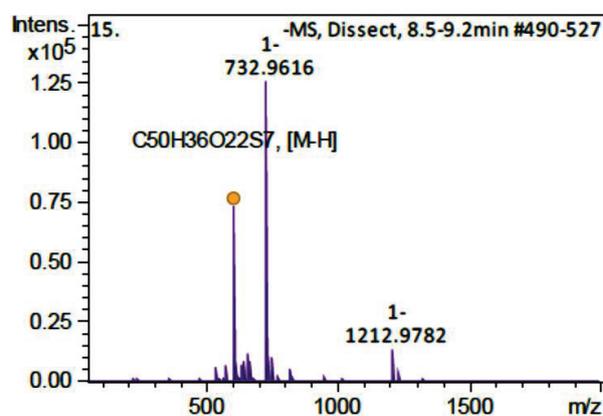
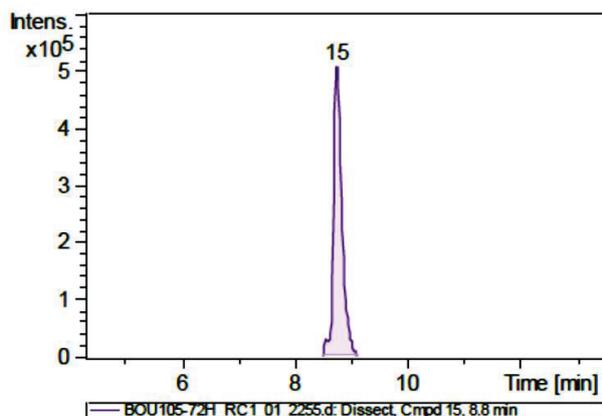
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
493.9789	C40H28O18S6	493.9805	3.4	24.7	35.12	M-H	2-

Scheme 96: HRMS from LC/MS, corresponding to 4 + O² -H₂



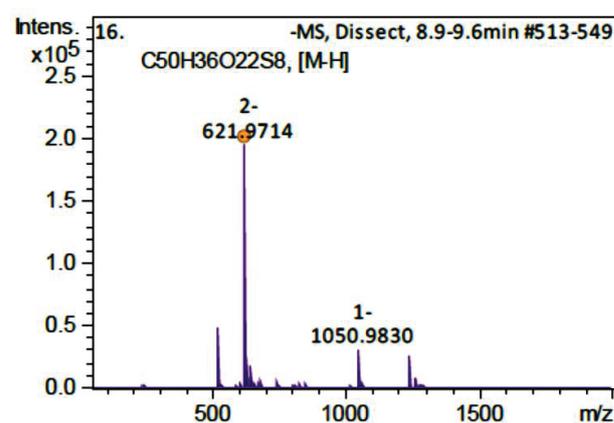
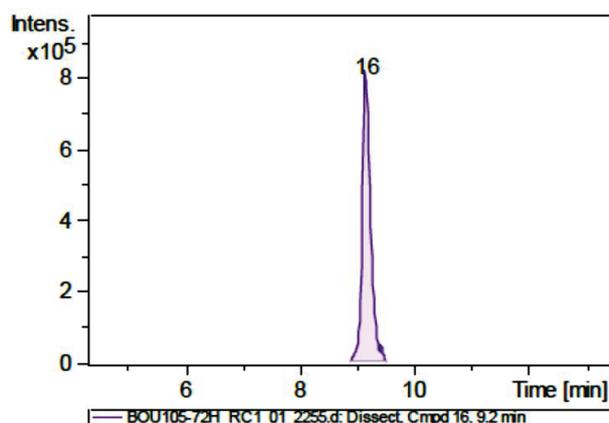
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
766.9490	C30H23O12S6	766.9519	3.8	14.3	16.53	M-H	1-

Scheme 97: HRMS from LC/MS, corresponding to 3



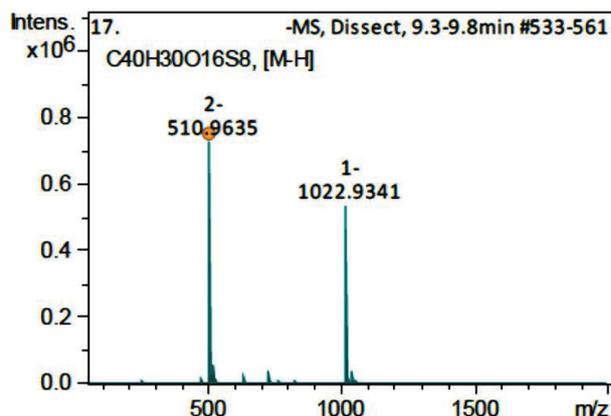
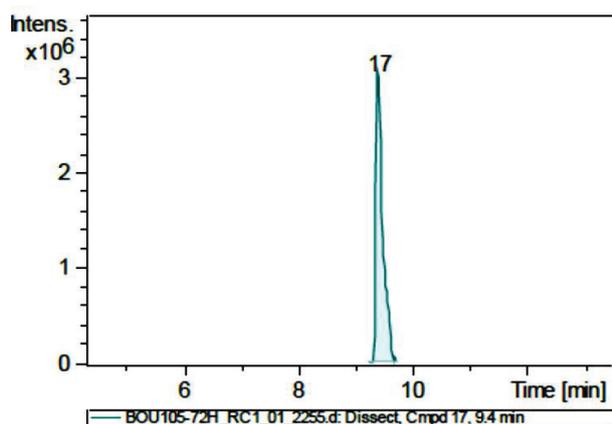
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
605.9856	C50H36O22S7	605.9877	3.5	15.9	26.99	M-H	2-

Scheme 98: HRMS from LC/MS, corresponding to 5 + O² -H₂



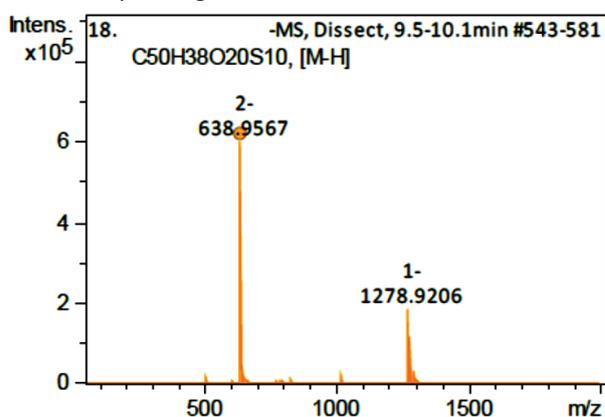
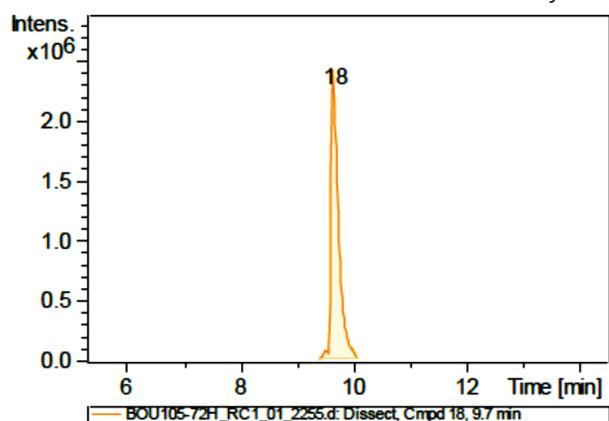
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
621.9714	C50H36O22S8	621.9737	3.7	15.8	23.29	M-H	2-

Scheme 99: HRMS from LC/MS, corresponding to 5 -H₂



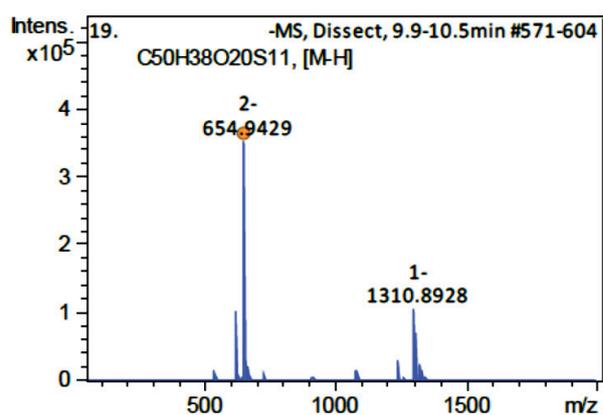
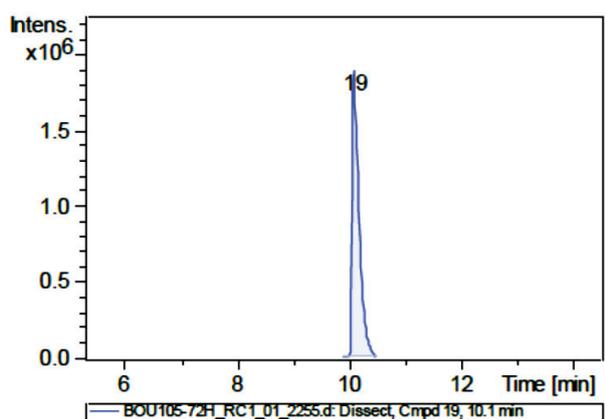
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
510.9635	C40H30O16S8	510.9655	3.9	13.7	33.63	M-H	2-

Scheme 100: HRMS from LC/MS, corresponding to 4



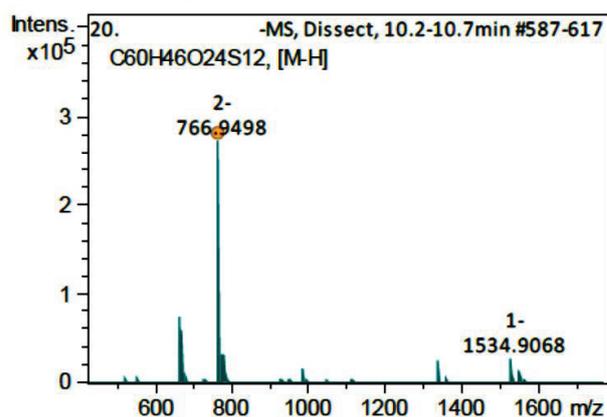
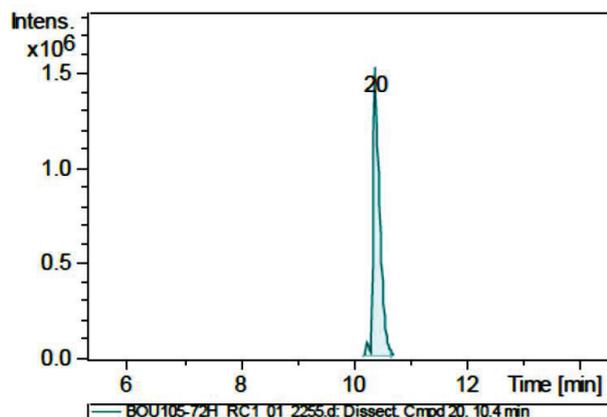
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
638.9567	C50H38O20S10	638.9587	3.2	13.3	26.87	M-H	2-

Scheme 101: HRMS from LC/MS, corresponding to 5



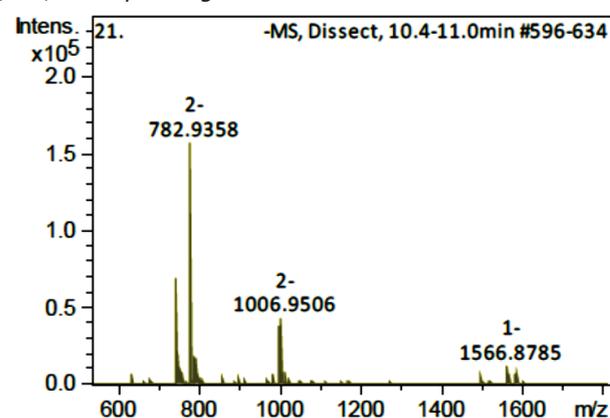
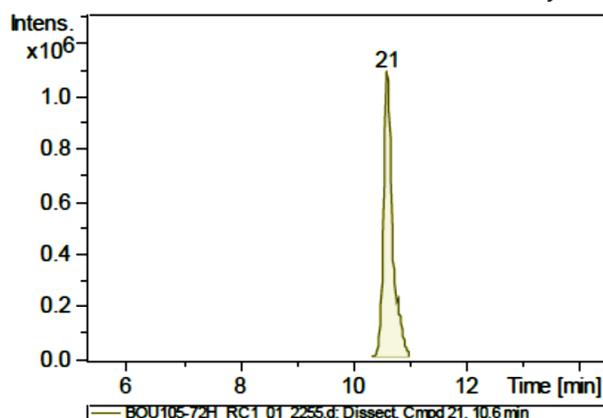
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
654.9429	C50H38O20S11	654.9448	2.9	13.9	28.95	M-H	2-

Scheme 102: HRMS from LC/MS, corresponding to 5 +S



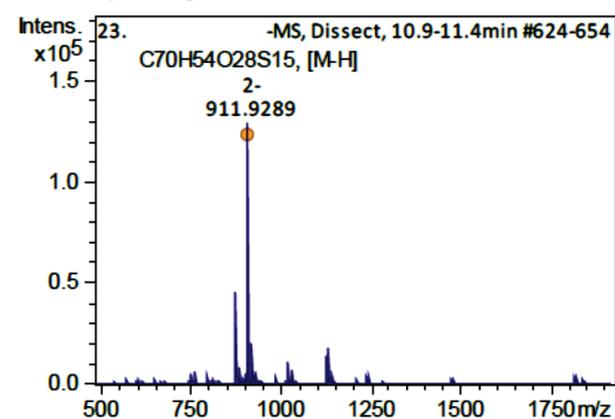
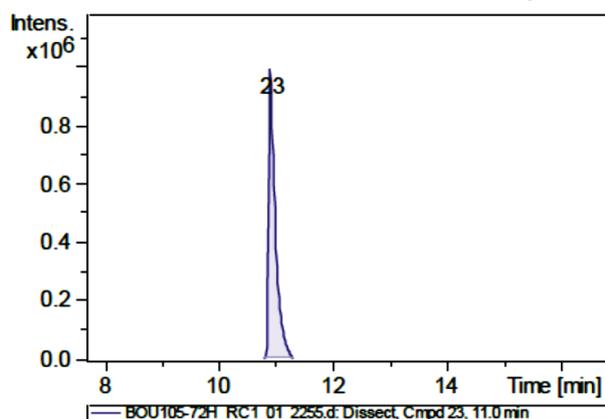
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
766.9498	C60H46O24S12	766.9519	2.8	13.5	42.24	M-H	2-

Scheme 103: HRMS from LC/MS, corresponding to 6



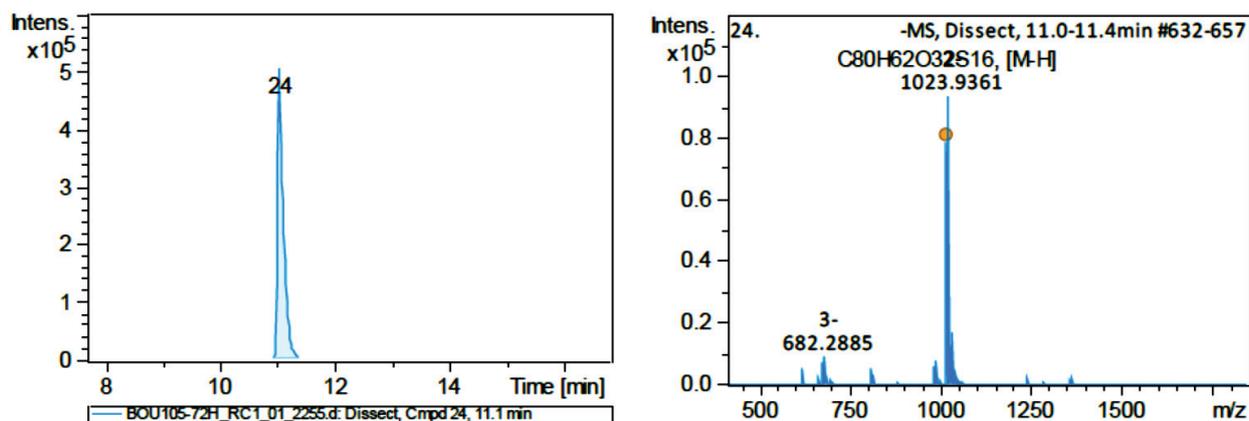
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
782.9358	C60H46O24S13	782.9380	2.8	6.5	41.11	M-H	2-

Scheme 104: HRMS from LC/MS, corresponding to 6 +S



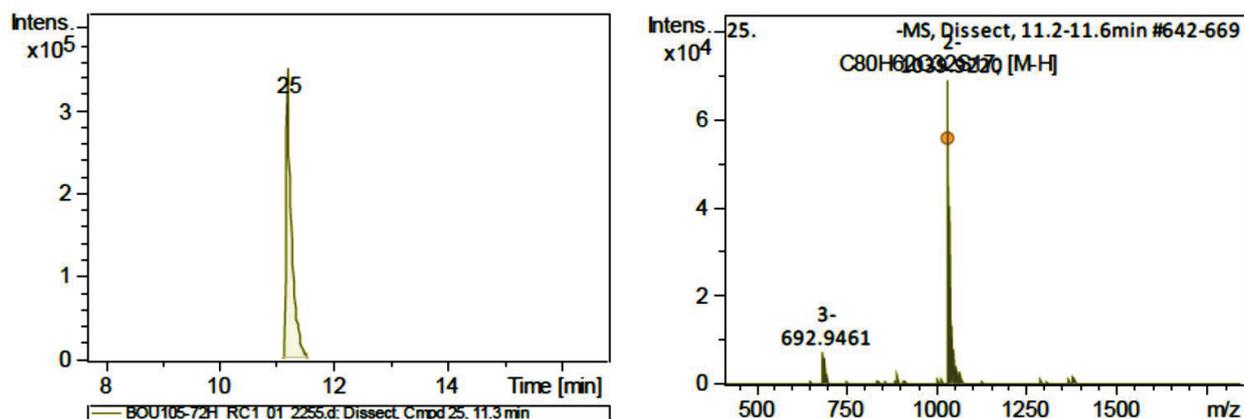
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
910.9290	C70H54O28S15	910.9312	2.4	13.6	26.90	M-H	2-

Scheme 105 : HRMS from LC/MS, corresponding to 7 +S



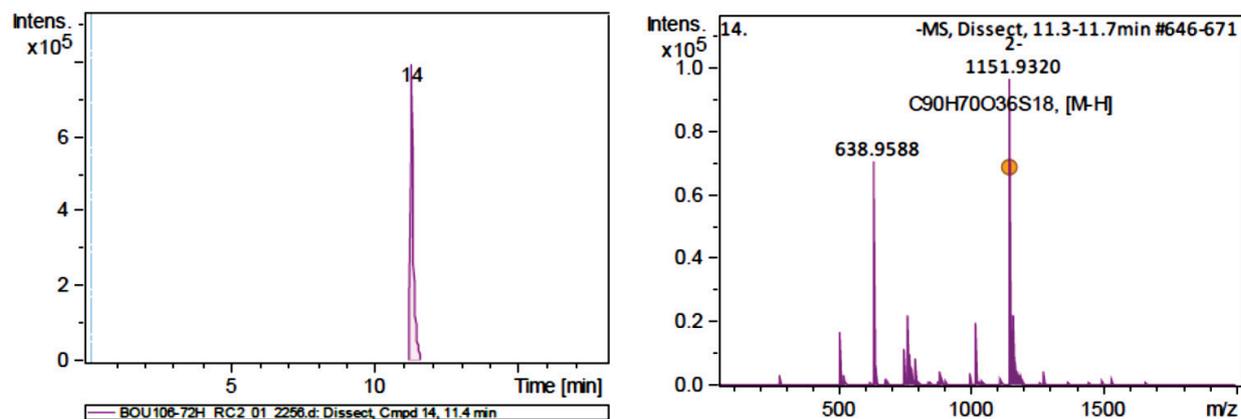
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
1022.9359	C80H62O32S16	1022.9383	2.4	24.2	34.94	M-H	2-

Scheme 106: HRMS from LC/MS, corresponding to 8



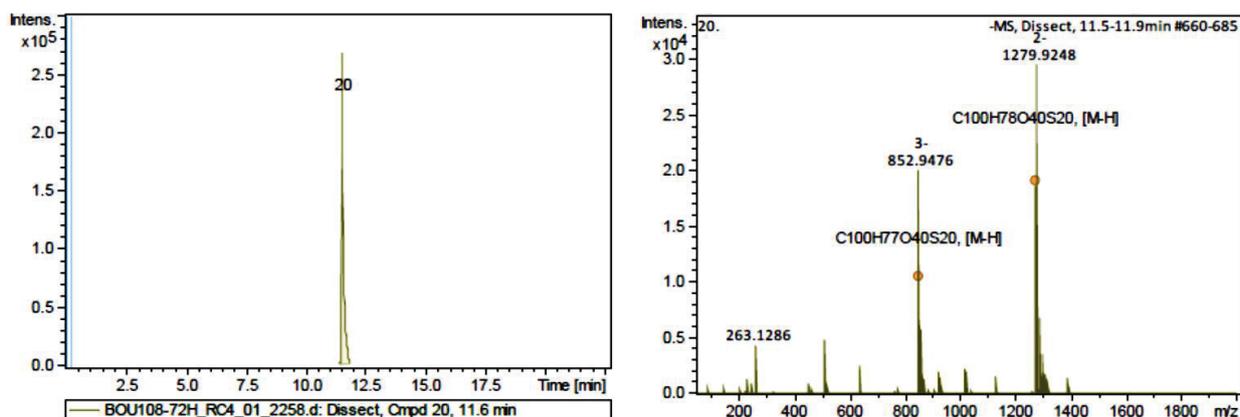
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
1038.9225	C80H62O32S17	1038.9244	1.8	17.4	36.66	M-H	2-

Scheme 107: HRMS from LC/MS, corresponding to 8 +S



m/z	Ion Formula	Meas. m/z	err [ppm]	mSigma	Score	Adduct	z
1150.9315	C90H70O36S18	1150.9319	-0.4	13.1	100.00	M-H	2-

Scheme 108: HRMS from LC/MS, corresponding to 9



Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
852.2807	C100H77O40S20	852.2807	0	35.9	100	M-H	3-
1278.924	C100H78O40S20	1278.9247	0.6	57.6	100	M-H	2-

Scheme 109: HRMS from LC/MS, corresponding to 10

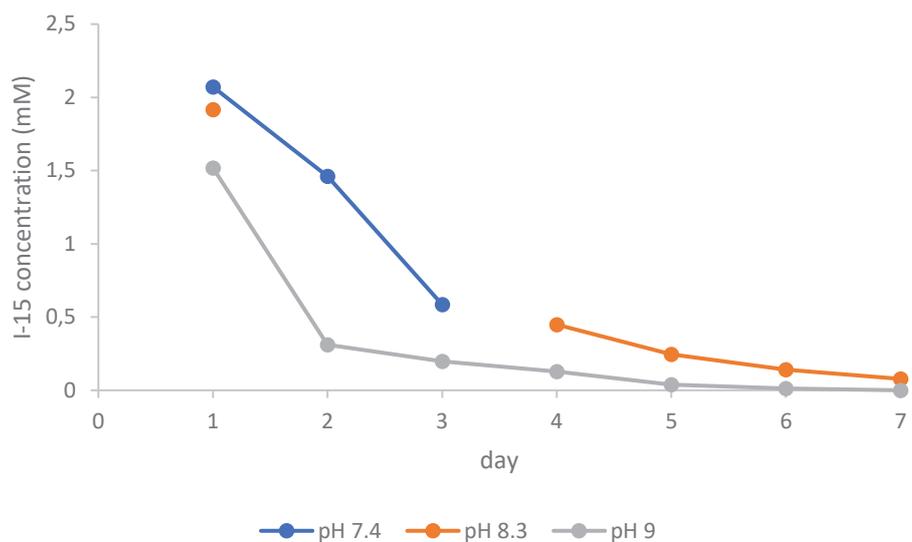
Template	1	2	3	4	5	6	7	8
/	27%	9%	21%	26%	10%	7%	0%	0%
I-105 ₆	22%	8%	17%	30%	11%	9%	3%	0%
I-45	22%	10%	22%	28%	10%	8%	0%	0%
I-7	41%	13%	15%	16%	6%	3%	0%	0%
I-47	33%	17%	15%	20%	6%	7%	0%	0%
I-46	22%	7%	26%	27%	11%	7%	0%	0%
I-48	21%	11%	9%	19%	15%	8%	2%	0%
I-49	0%	0%	29%	9%	16%	4%	5%	5%
I-50	8%	4%	27%	16%	10%	8%	3%	1%
I-51	3%	29%	3%	11%	13%	7%	5%	2%
I-52	3%	9%	24%	17%	13%	12%	6%	3%
I-53	33%	23%	4%	4%	16%	19%	0%	0%
I-54	20%	8%	16%	30%	12%	10%	3%	0%
I-55	0%	14%	38%	4%	2%	14%	8%	4%
I-56	17%	10%	18%	29%	13%	11%	4%	0%
I-57	0%	5%	29%	14%	8%	14%	8%	4%
I-58	3%	6%	29%	13%	10%	11%	6%	3%
I-59	19%	11%	15%	30%	11%	10%	4%	0%
I-60	28%	7%	23%	28%	9%	6%	0%	0%
I-61	66%	22%	5%	3%	0%	0%	0%	0%
I-62	65%	22%	4%	4%	0%	0%	0%	0%

I-63	24%	8%	18%	29%	10%	9%	3%	0%
I-64	36%	16%	0%	0%	0%	0%	0%	0%
I-65	3%	9%	39%	3%	8%	8%	3%	2%
I-66	4%	9%	8%	6%	2%	23%	7%	4%

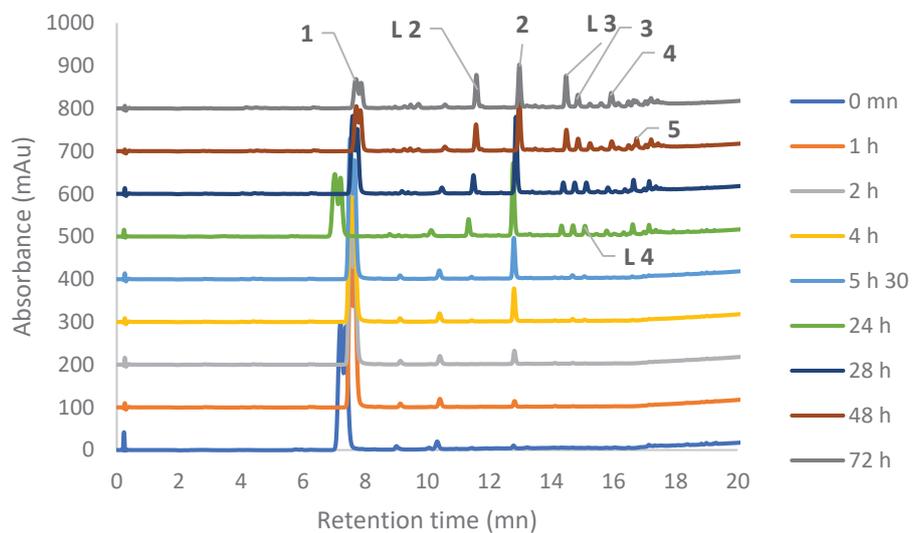
Table 11: Proportions of macrocycles, after 48 h., in Tris 200 mM pH 7.8, I-15 at 4 mM, template at 0.25 eq.

Template	1	2	3	4	5	6	7	8
/	0,8	0,9	0,8	1,2	1,1	1,3	3	0
I-105₆	0,8	1,1	1,0	1,1	1,0	1,1	0	0
I-45	1,5	1,4	0,7	0,6	0,6	0,4	0	0
I-7	1,2	1,9	0,7	0,8	0,6	1,0	0	0
I-47	0,8	0,8	1,2	1,0	1,1	1,0	0	0
I-46	0,8	1,2	0,4	0,7	1,5	1,1	2	0
I-48	0,0	0,0	1,4	0,3	1,6	0,6	5	5
I-49	0,3	0,4	1,3	0,6	1,0	1,1	3	1
I-50	0,1	3,2	0,1	0,4	1,3	1,0	5	2
I-51	0,1	1,0	1,1	0,7	1,3	1,7	6	3
I-52	1,2	2,6	0,2	0,2	1,6	2,7	0	0
I-53	0,7	0,9	0,8	1,2	1,2	1,4	3	0
I-54	0,0	1,6	1,8	0,2	0,2	2,0	8	4
I-55	0,6	1,1	0,9	1,1	1,3	1,6	4	0
I-56	0,0	0,6	1,4	0,5	0,8	2,0	8	4
I-57	0,1	0,7	1,4	0,5	1,0	1,6	6	3
I-58	0,7	1,2	0,7	1,2	1,1	1,4	4	0
I-59	1,0	0,8	1,1	1,1	0,9	0,9	0	0
I-60	2,4	2,4	0,2	0,1	0,0	0,0	0	0
I-61	2,4	2,4	0,2	0,2	0,0	0,0	0	0
I-62	0,9	0,9	0,9	1,1	1,0	1,3	3	0
I-63	1,3	1,8	0,0	0,0	0,0	0,0	0	0
I-64	0,1	1,0	1,9	0,1	0,8	1,1	3	2
I-65	0,1	1,0	0,4	0,2	0,2	3,3	7	4

Table 12: Amplification factors of macrocycles, after 48 h., in Tris 200 mM pH 7.8, I-15 at 4 mM, compared to the DCL without template, template at 0.25 eq



Scheme 110: I-15 concentration at different pH, starting at 4 mM in Tris 200 mM



Scheme 111: Evolution of I-15 (4 mM) in ammonium acetate (200 mM, pH 7.4), method B. n disulfide macrocycle containing n units, 1 starting building block, L linear oligomer

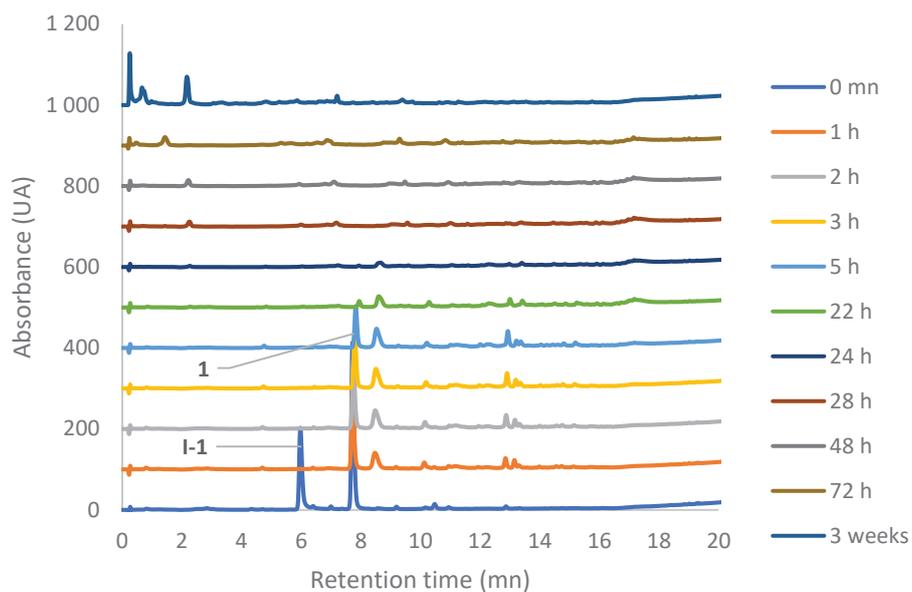


Figure 17 : Evolution of **I-15** (2 mM) and **I-1** (2 mM) in Tris (200 mM, pH 7.4), method B. *n* disulfide macrocycle containing *n* units of **I-15**, 1 starting building block

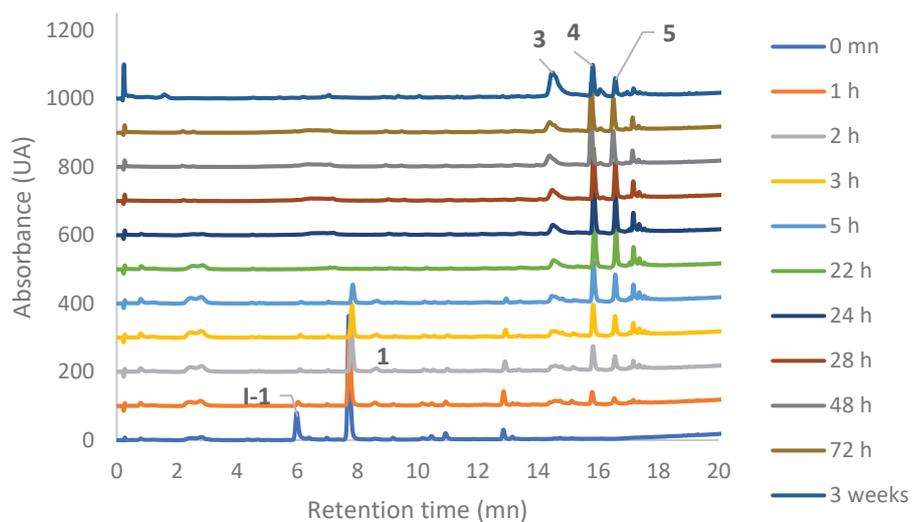
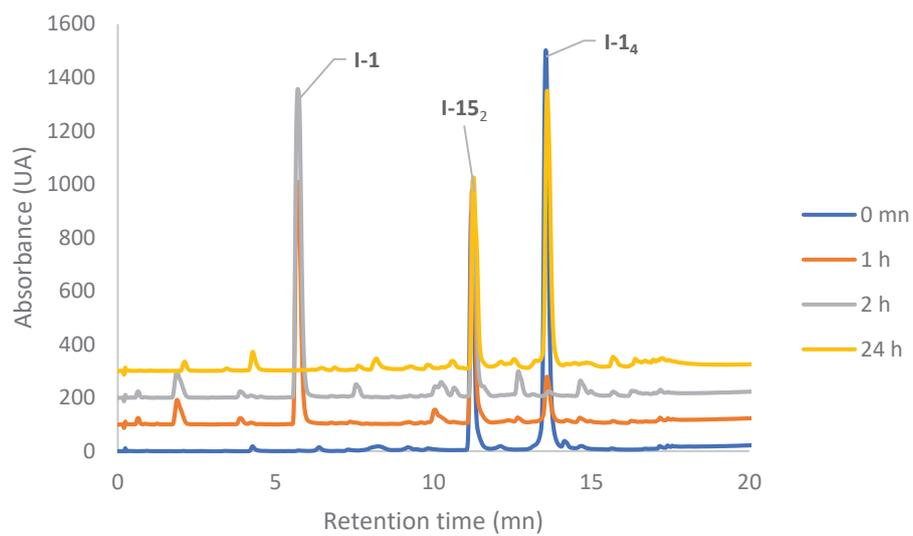


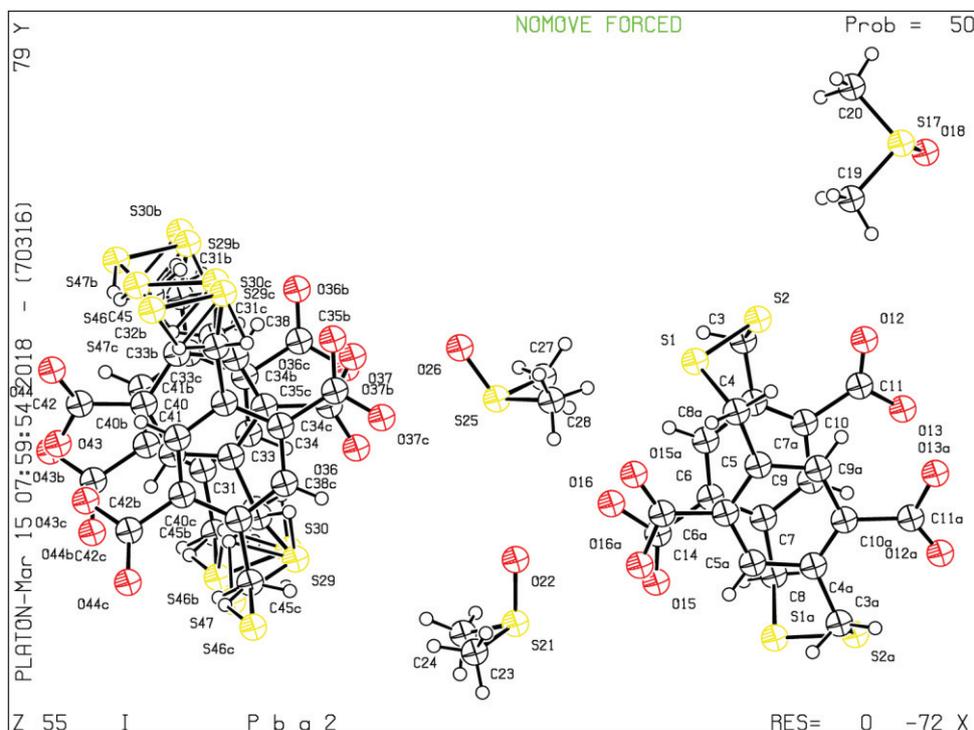
Figure 18: Evolution of **I-15** (2 mM) and **I-1** (2 mM) with spermine (1 mM) in Tris (200 mM, pH 7.4), method B. *n* disulfide macrocycle containing *n* units of **I-15**, 1 starting building block



Scheme 112: Evolution of I-1₄ and I-15₂ (1 mM + 2 mM) in Tris 200 mM (pH 7.4) with spermine (0.25 mM) and DTT (0.05 mM), method E

f. Appendix I-VI: CheckCIFiles and crystallographic data for **I-15₂**

Bond precision:	C-C = 0.0100 Å	Wavelength=0.71073
Cell:	a=7.3353(6) b=20.9956(18) c=23.882(3)	
	alpha=90 beta=90 gamma=90	
Temperature: 150 K		
	Calculated	Reported
Volume	3678.0(6)	3678.0(6)
Space group	P b a 2	P b a 2
Hall group	P 2 -2ab	P 2 -2ab
Moiety formula	C10 H6 O4 S2, 0.5(C20 H12 O8 S4), 3(C2 H6 O S) [+ solvent]	C13 H15 O5.50 S3.50
Sum formula	C26 H30 O11 S7 [+ solvent]	C13 H15 O5.50 S3.50
Mr	742.92	371.49
Dx, g cm-3	1.342	1.342
Z	4	8
Mu (mm-1)	0.478	0.478
F000	1544.0	1544.0
F000'	1548.18	
h,k,lmax	10,29,33	10,27,32
Nref	10364[5299]	8523
Tmin,Tmax	0.942,0.965	0.821,0.936
Tmin'	0.857	
Correction method= #	Reported T Limits: Tmin=0.821 Tmax=0.936	
AbsCorr = ANALYTICAL		
Data completeness= 1.61/0.82	Theta(max)= 29.603	
R(reflections)= 0.0944(7538)	wR2(reflections)= 0.2288(8504)	
S = 0.977	Npar= 427	



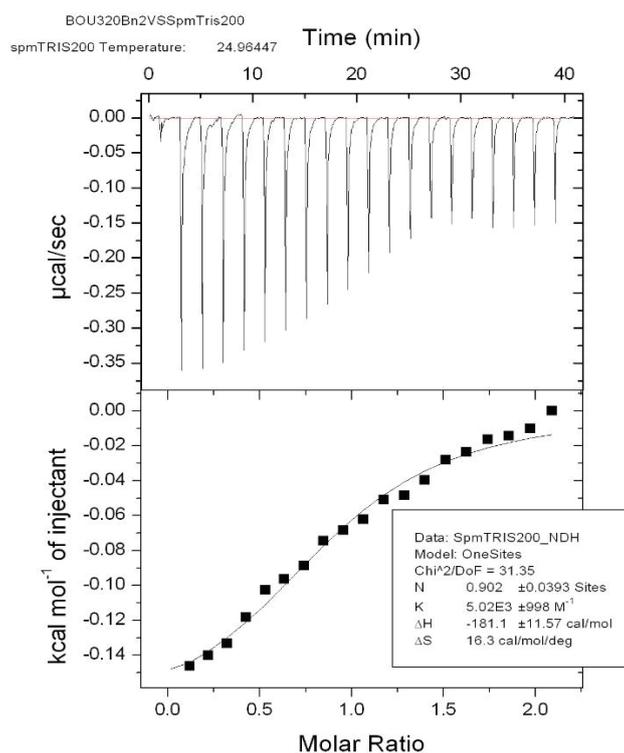
g. Appendix I-VII: DFT data for I-15₂

TS coordinates:				C	8.89515300	9.44115700	6.31086400
S	8.96311800	14.48320300	10.45355300	O	8.74289500	8.20644900	6.10725800
S	6.84846400	14.28126100	10.40810600	O	9.02839600	10.31092000	5.41128100
C	6.43312600	13.34697200	8.88958500	C	9.31921200	11.45758400	11.78598900
C	5.98666100	11.91439500	9.10938300	O	10.45777500	11.90201900	12.07724500
C	5.74338700	11.44587300	10.39605100	O	8.35052400	11.29521800	12.57121200
C	5.33609200	10.14673400	10.68770600	H	8.77666100	7.15766000	9.78236500
C	5.10647400	9.26897000	9.61934700	H	9.42430800	7.18743900	8.13450900
C	4.64130300	7.85804500	9.76013100	H	8.68084600	9.02279400	11.00904800
C	5.29539600	9.75163300	8.32606300	H	10.76840400	13.35570500	9.64035200
C	5.75466500	11.03430700	8.03445600	H	9.67979900	13.70244100	8.29893500
C	5.98592400	11.37169600	6.56885400	H	9.26843400	11.92330300	7.10741100
O	6.36778700	12.53578900	6.27546000				
O	5.77831700	10.45239800	5.73307800				
C	5.16859300	9.78335600	12.15260700				
O	5.35744900	8.58414300	12.48602300				
O	4.86033500	10.72083000	12.93431100				
H	5.63266900	13.91775900	8.41224500				
H	7.27546800	13.40082700	8.20413400				
H	5.89573200	12.11696700	11.23095700				
H	4.12611200	7.68329300	10.70458400				
H	3.97713100	7.59669000	8.93389700				
H	5.09214000	9.09708800	7.48785700				
S	5.97236100	6.57330200	9.80592600				
S	7.09862100	6.91603400	8.09911500				
C	8.65948800	7.59956900	8.79155700				
C	8.78252700	9.08880700	8.87101700				
C	8.82951100	9.66114700	10.14205200				
C	9.12801200	11.00553100	10.35589000				
C	9.30672800	11.83668800	9.24710700				
C	9.72661500	13.28009600	9.30386400				
C	9.16828800	11.28197400	7.97632100				
C	8.93618500	9.92677500	7.75210200				

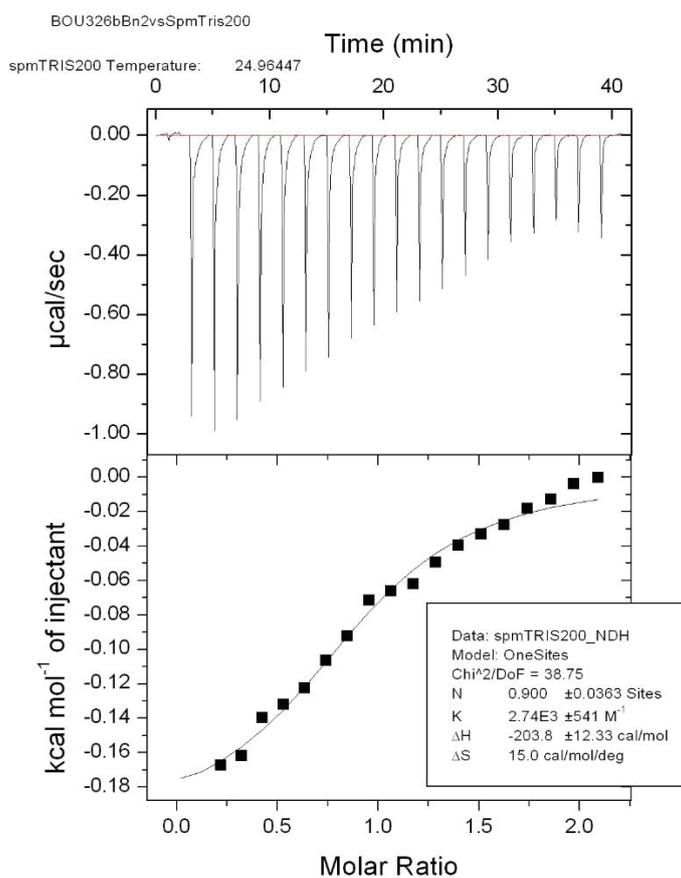
Conformer	ΔG (kcal/mol)
pR2 PP	0
pR2 MM	-0,29
pR2 PM 1	0,50
pR2 PM 2	1,56
pRpS MM 1	0,13
pRpS MM 2	-2,48
pRpS PM 1	0,33
pRpS PM 2	2,64
pRpS PP 1	3,13
pRpS PP 2	2,21
pS2 MM	-1,96
pS2 PM 1	0,45
pS2 PM 2	4,03
pS2 PP	5,36
$TS_{pR2PP / pR2 PM1}$	11.9

Table 13: Energies of **I-15₂** conformers from DFT optimisation

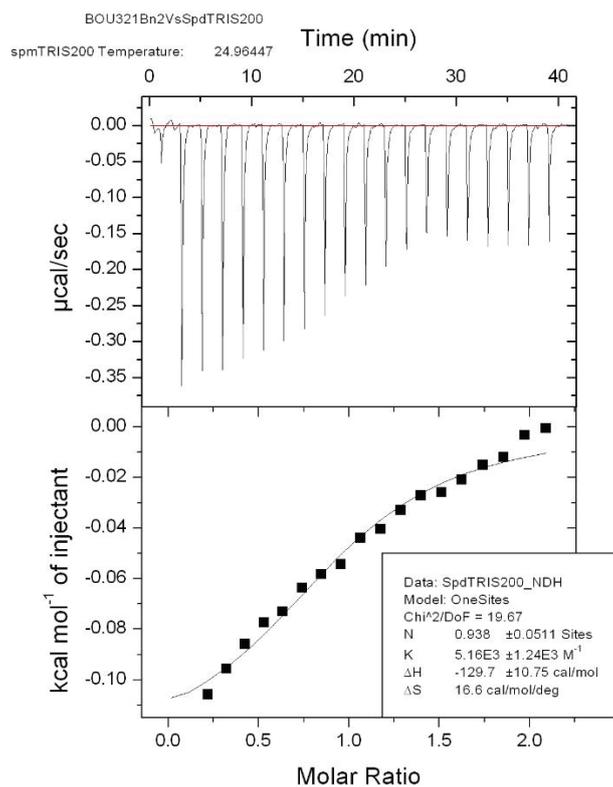
h. Appendix I-V-III: ITC thermograms of I-15₂ vs polyamines and LiCl



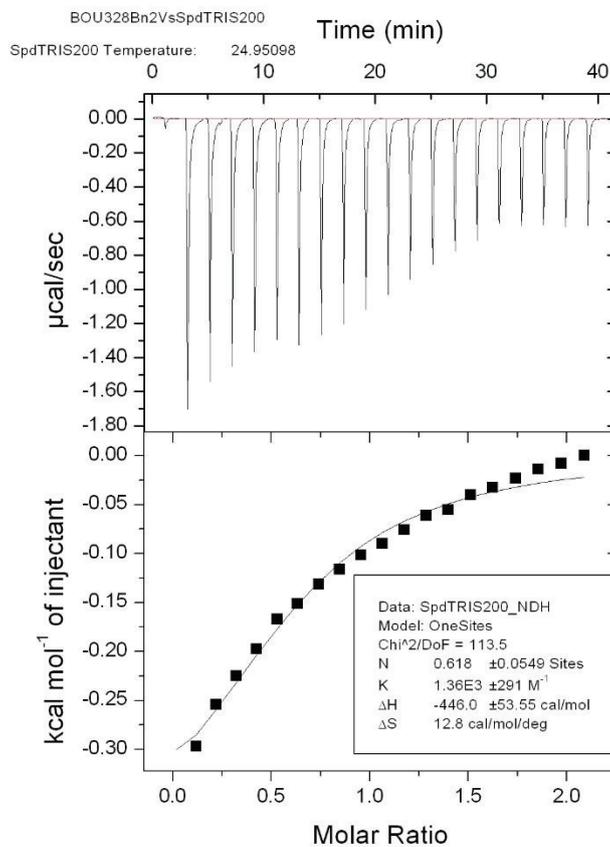
Scheme 113: Thermogram of I-15₂ (1 mM) and I-7 (10 mM) in Tris 200 mM pH 7.4



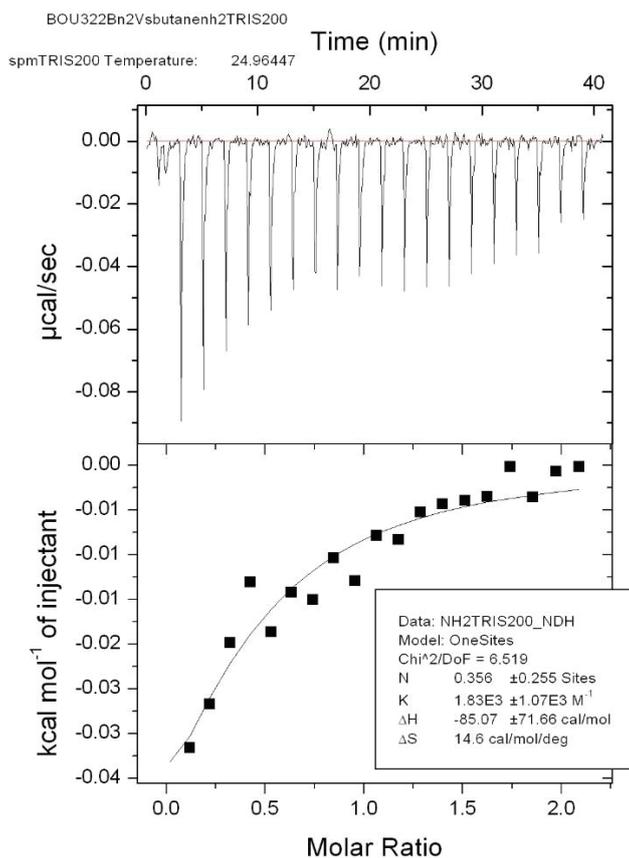
Scheme 114: Thermogram of I-15₂ (2.5 mM) and I-7 (25 mM) in Tris 200 mM pH 7.4



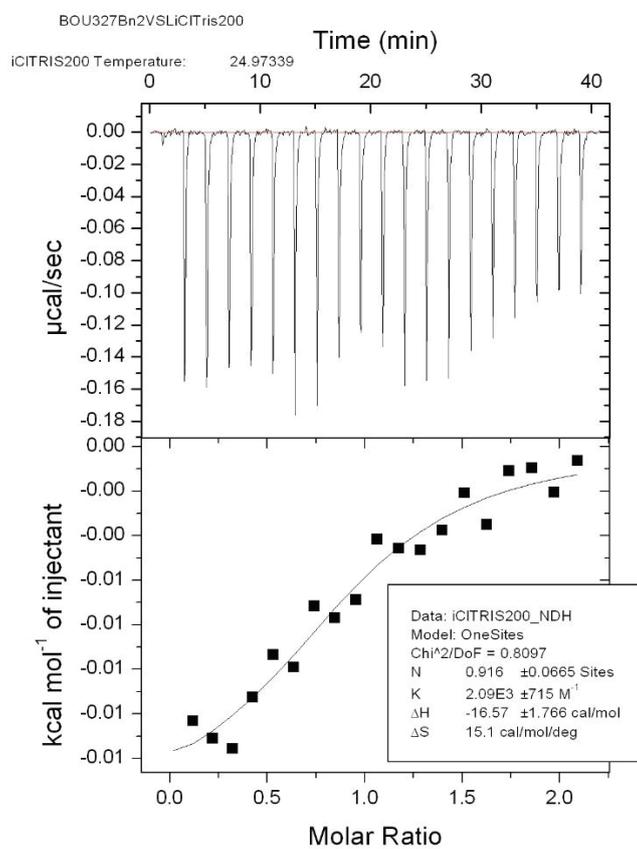
Scheme 115: Thermogram of *I-15₂* (1 mM) and *I-8* (10 mM) in Tris 200 mM pH 7.4



Scheme 116: Thermogram of *I-15₂* (2.5 mM) and *I-8* (25 mM) in Tris 200 mM pH 7.4



Scheme 117: Thermogram of *I-15₂* (1 mM) and *I-10₄* (10 mM) in Tris 200 mM pH 7.4



Scheme 118: Thermogram of *I-15₂* (2.5 mM) and LiCl (25 mM) in Tris 200 mM pH 7.4

i. Appendix I-IX: NMR experiments of I-105 / I-105₄

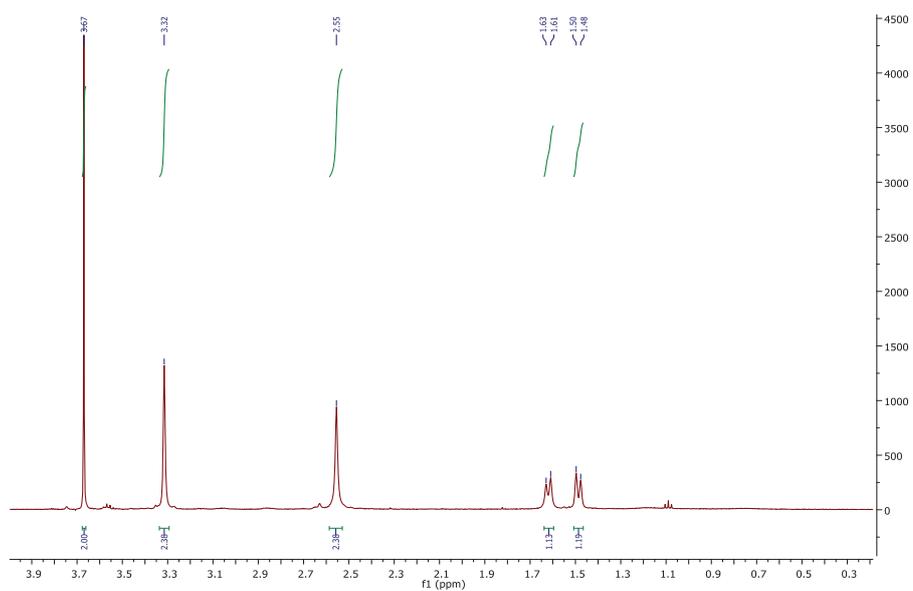


Figure 19 : ¹H NMR of the DCL of I-105 at 5 mM in ND₃CD₃COOD at 50 mM without pH adjustment after 0 mn of stirring. Internal standard dioxane at 3.67 ppm.

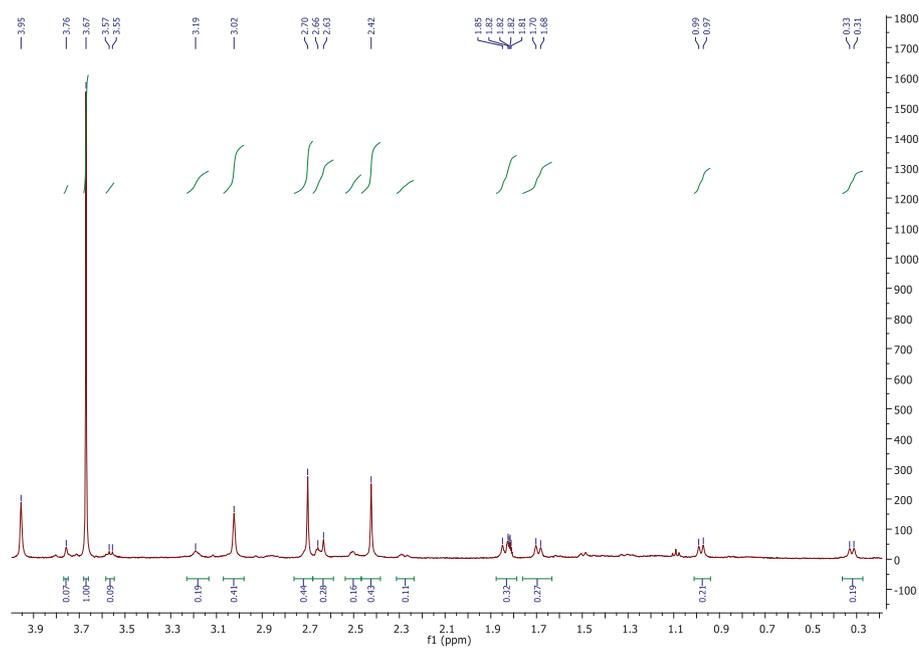


Figure 20 : ¹H NMR of the DCL of I-105 at 5 mM in ND₃CD₃COOD at 50 mM without pH adjustment after 24 hours of stirring. Internal standard dioxane at 3.67 ppm.

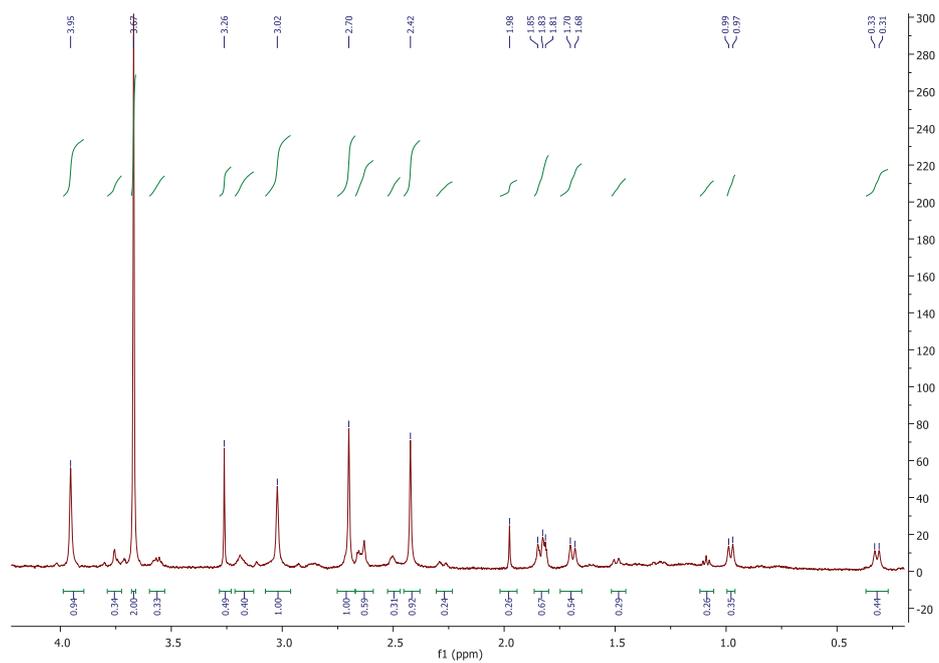
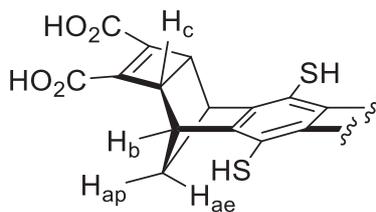
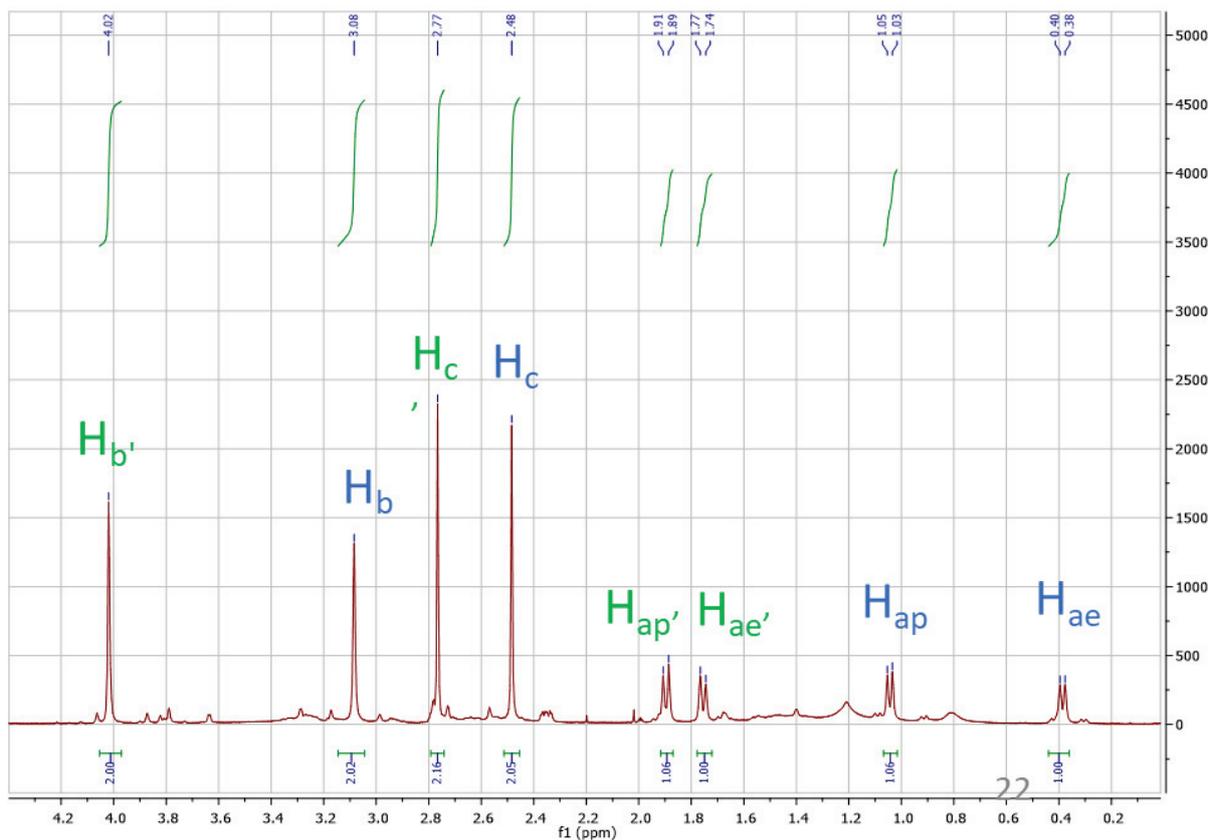


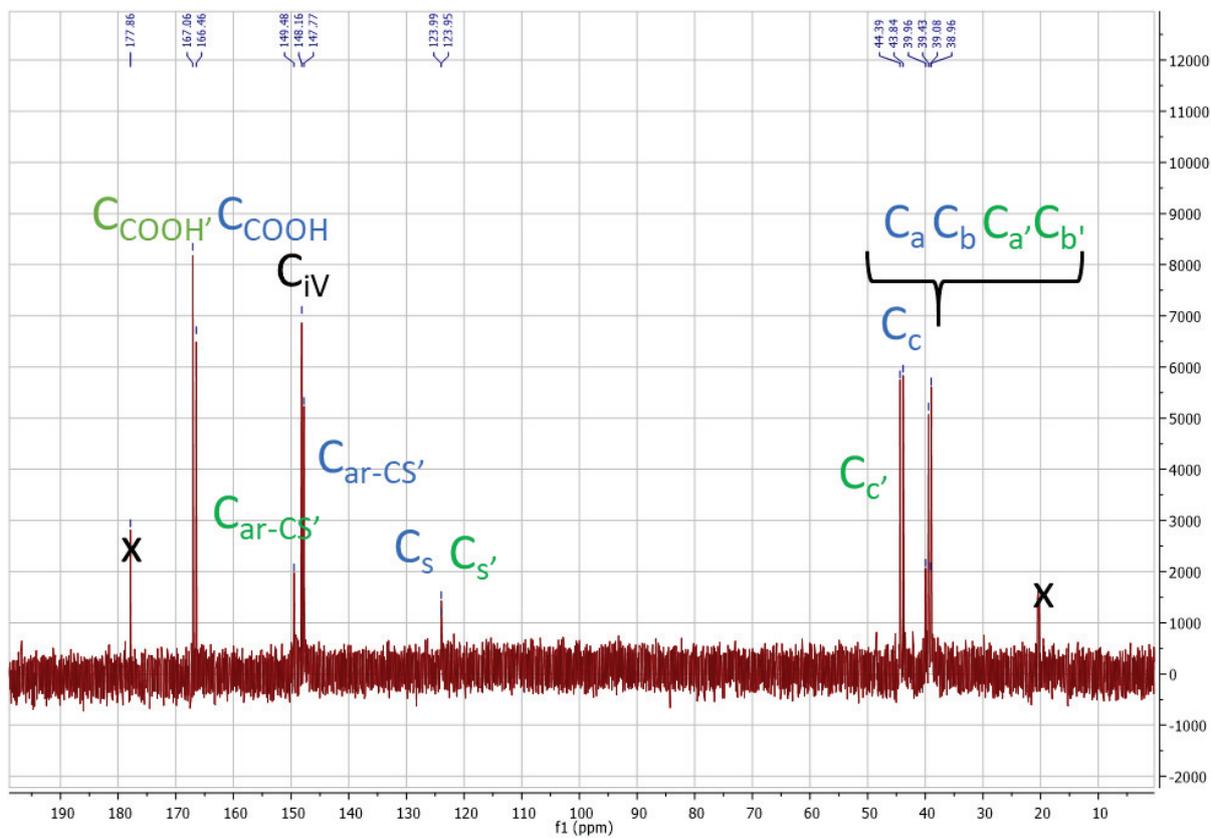
Figure 21: ^1H NMR of the DCL of **I-105** at 5 mM in $\text{ND}_3\text{CD}_3\text{COOD}$ at 50 mM without pH adjustment after 48 hours of stirring. Internal standard dioxane at 3.67 ppm.



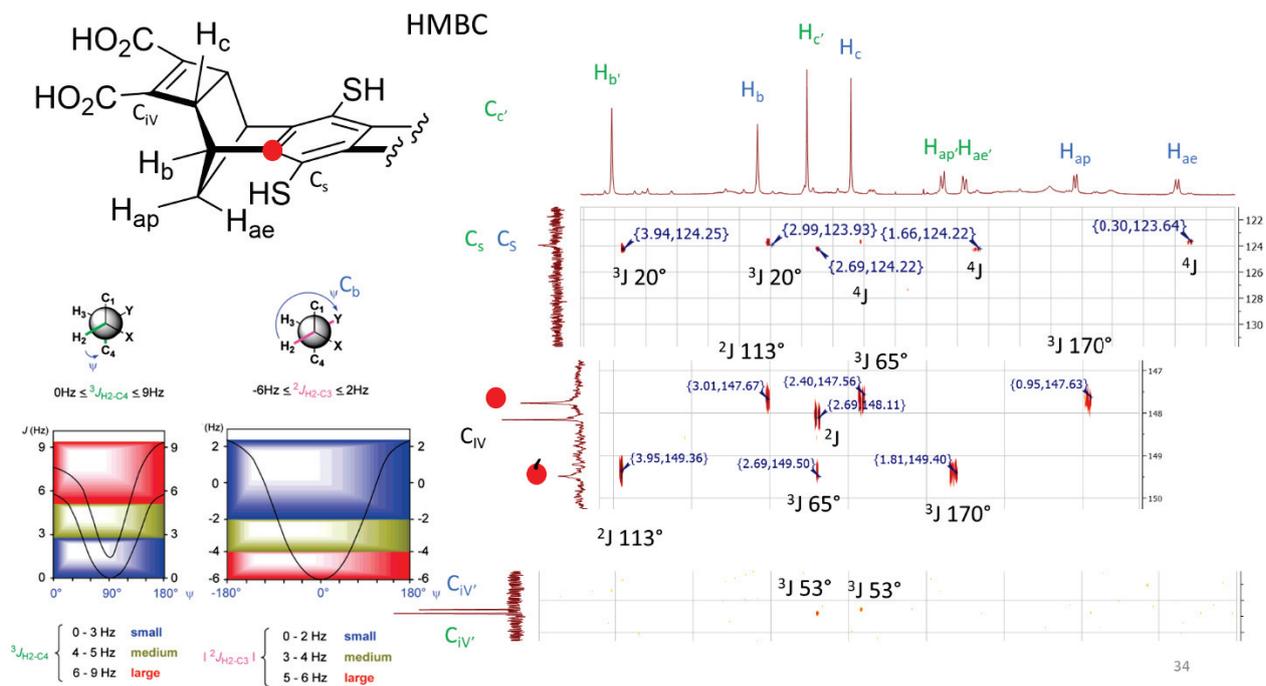
Scheme 119 : Repetitive unit of **I-105**. 4 protons a, b, c, ae, ap are highlighted. p axial, e equatorial



Scheme 120: ^1H NMR spectrum of I-104a



Scheme 121: ^{13}C NMR spectrum of I-104a



Scheme 124: HMBC NMR spectra of **I-104**

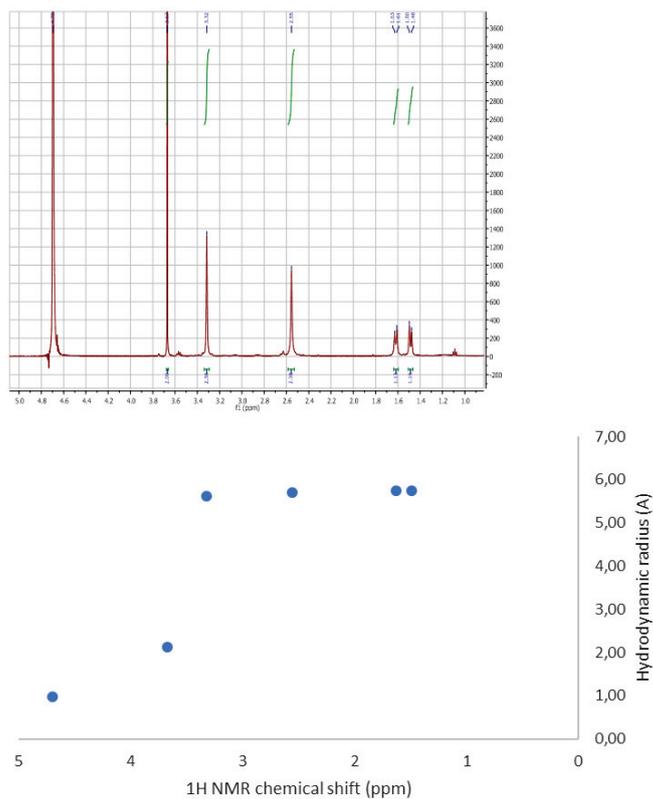


Figure 22: DOSY experiment of **I-105** dissolved in 50 mM of ND_3CD_3COOD without pH adjustment with dioxane

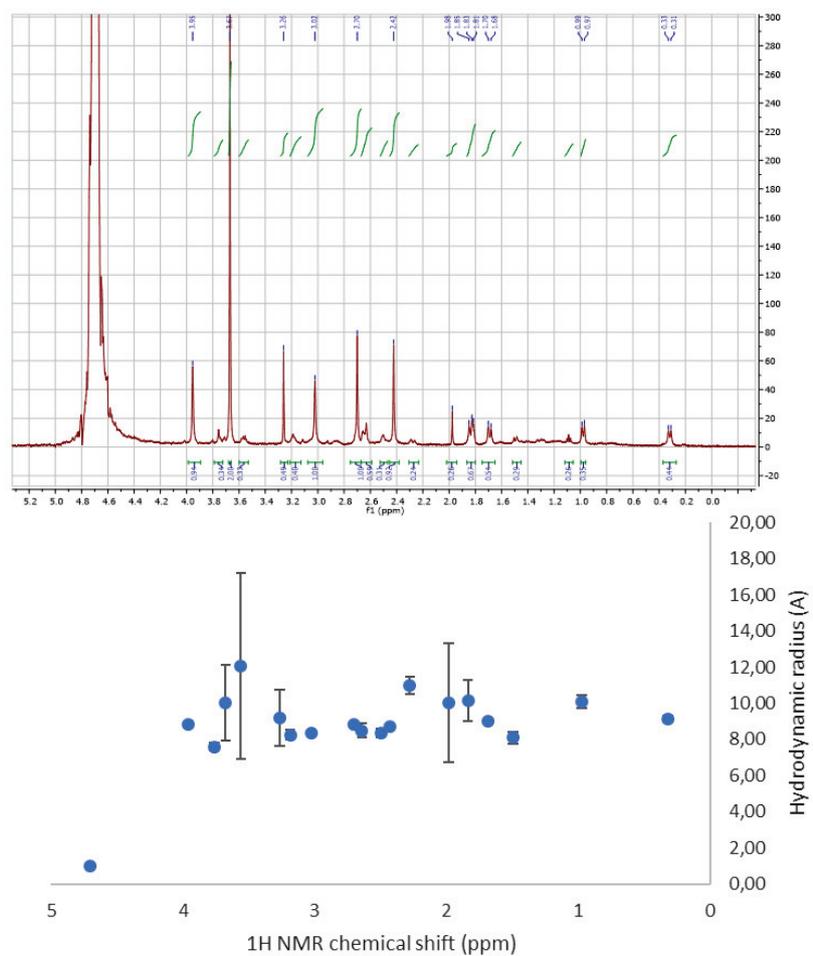
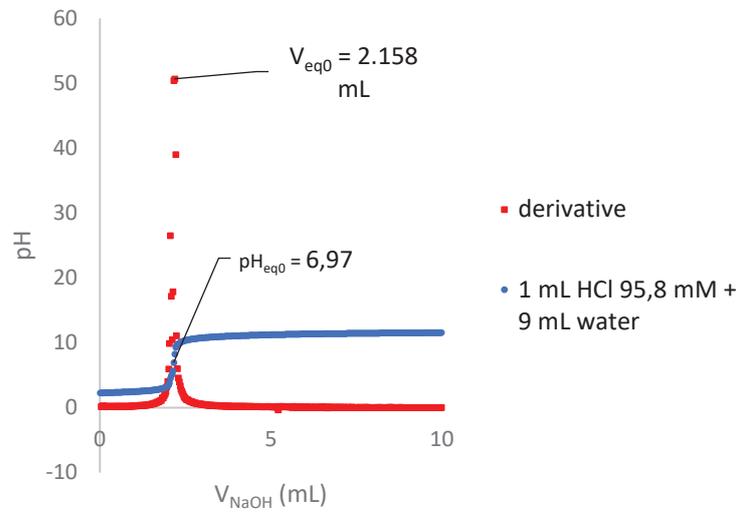
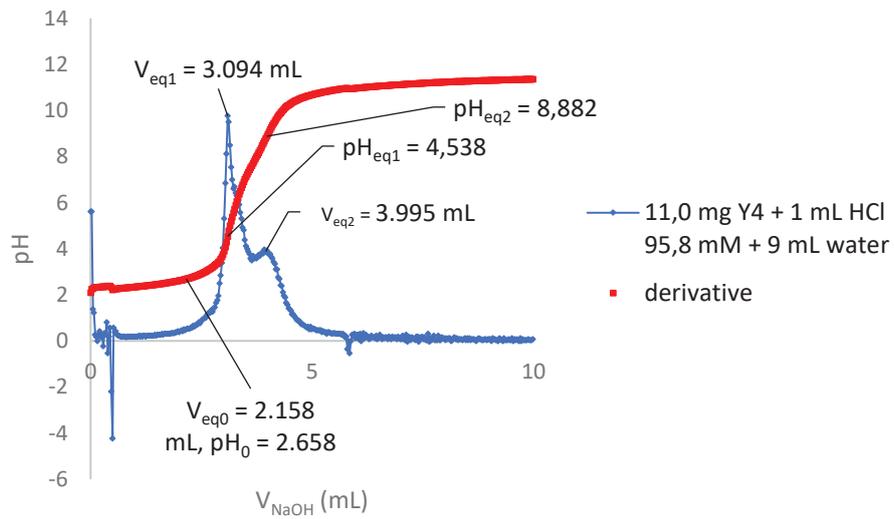


Figure 23 : DOSY experiment of *I-105* dissolved in 50 mM of ND₃CD₃COOD without pH adjustment after 48 hours (mainly *I-105*₄)

j. Appendix I-XI: pH titration of **I-105₄**

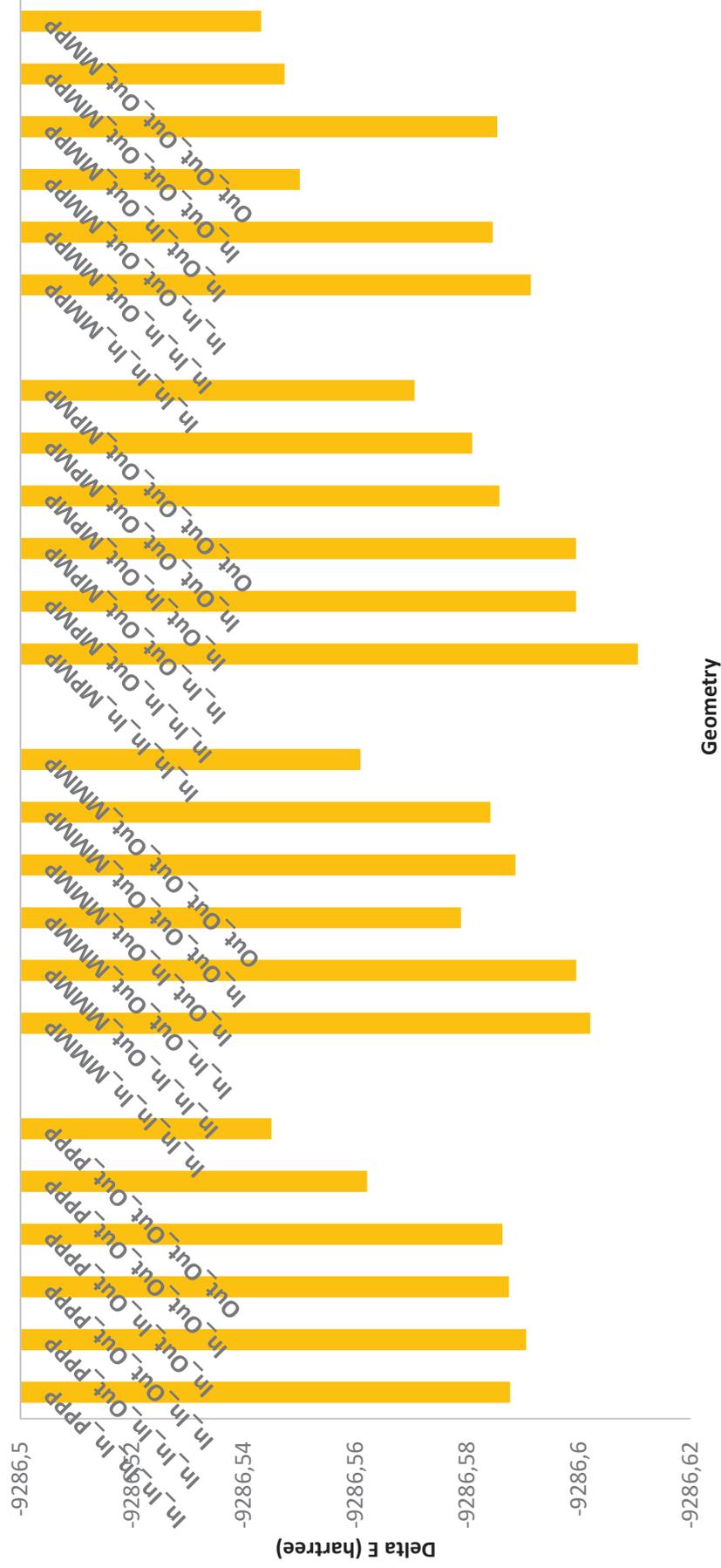


Scheme 125: Titration of **I-105₄** (2)



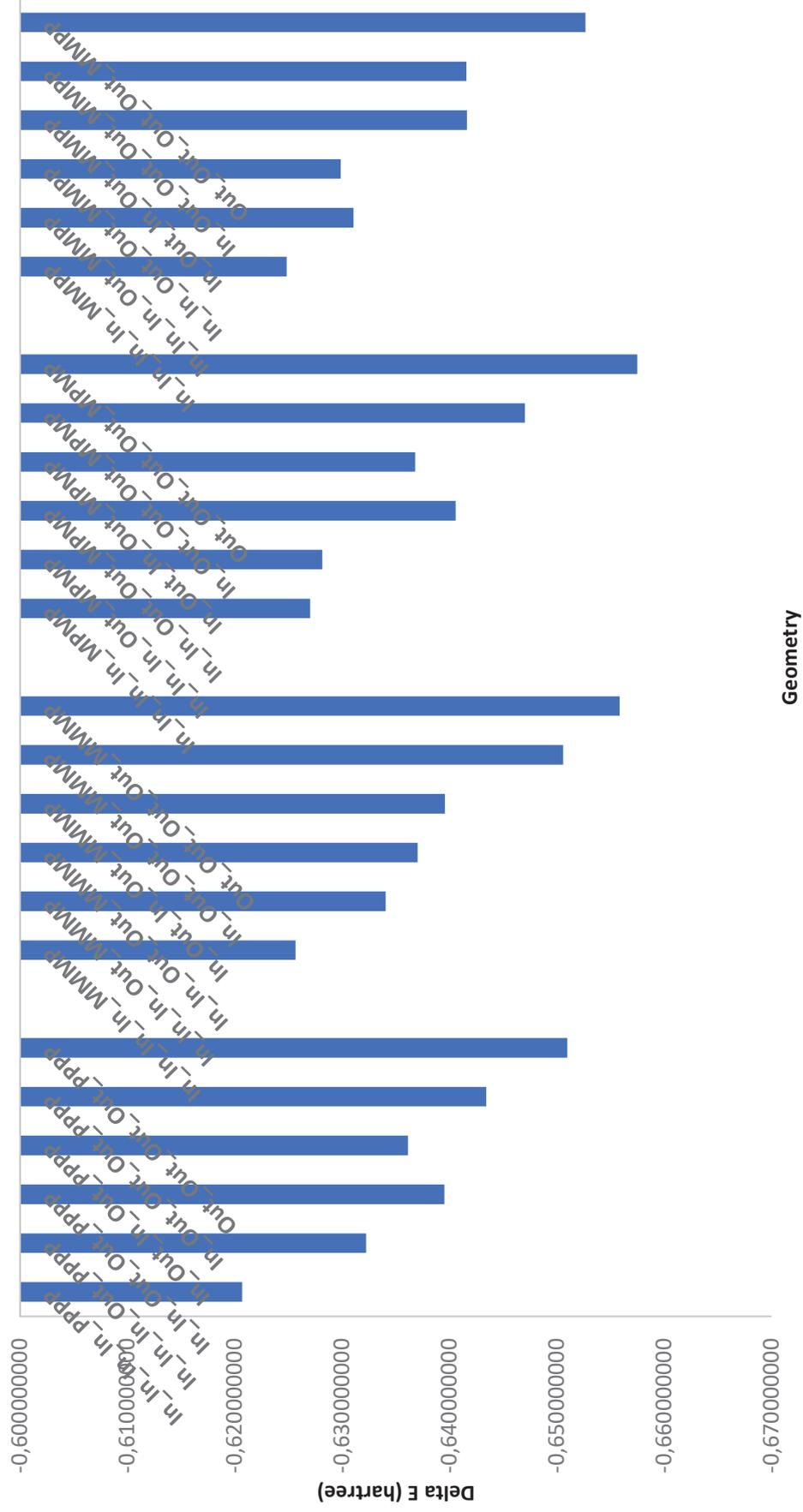
Scheme 126: Titration of **I-105₄** (3)

Energy, B3LYP-D3BJ



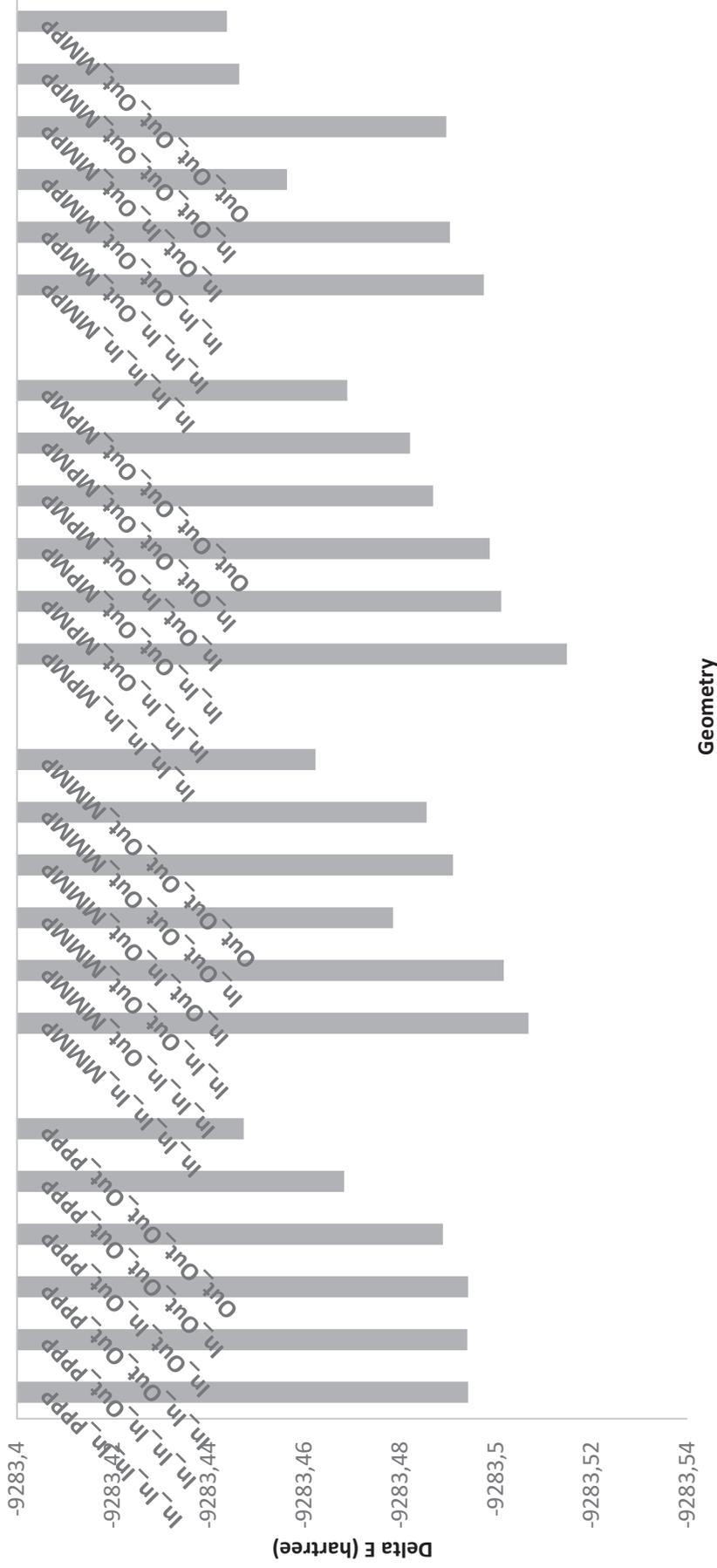
Scheme 128 : Energy of the 24 possible diastereoisomers of 1-105a, SP with B3LYP-D3BJ

Dispersion contribution (B3LYP-D3BJ)



Scheme 129: Dispersion energy contribution of the 24 possible diastereoisomers of I-105₄ SP with B3LYP-D3BJ

Energy, M062X



Scheme 130: Energy of the 24 possible diastereoisomers of I-105a, SP with M062X

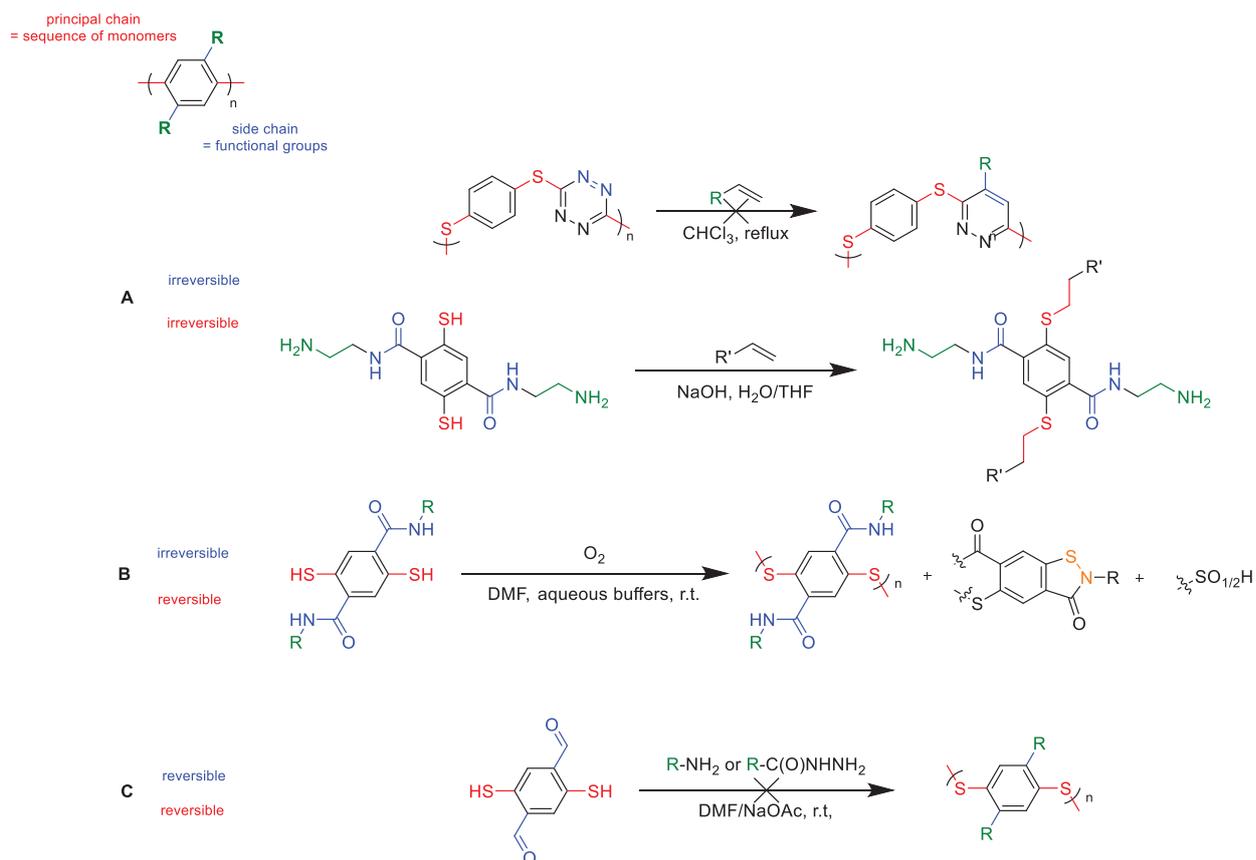
Geometry	E hartree B3LYP	E hartree B3LYP-D3BJ	E hartree m062X	dispersion hartree B3LYPD3BJ	Delta E kcal/mol B3LYP	Delta E kcal/mol B3LYP-D3BJ	Delta E kcal/mol M062X	Delta disp kcal/mol B3LYP-D3BJ
In_In_In_In_PPPP	-9285,966998	-9286,587673	-9283,49424	-0,620675166	14,36740731	14,36740731	12,9367986	3,979039434
In_In_In_Out_PPPP	-9285,958322	-9286,590536	-9283,494039	-0,632214707	15,8328177	12,57067819	13,06265191	-3,262132169
In_In_Out_Out_PPPP	-9285,947947	-9286,587474	-9283,494223	-0,639526735	22,34305933	14,49259546	12,94762942	-7,850498953
In_Out_In_Out_PPPP	-9285,950216	-9286,586343	-9283,488912	-0,636127292	20,91924028	15,20185062	16,27982409	-5,717316364
In_Out_Out_Out_PPPP	-9285,918636	-9286,562066	-9283,468346	-0,643430116	40,73622246	30,43631291	29,1857304	-10,29990799
Out_Out_Out_Out_PPPP	-9285,893989	-9286,544944	-9283,447355	-0,650955773	56,20247421	41,18014151	42,35720501	-15,02232906
In_In_In_In_MMMP	-9285,976408	-9286,602063	-9283,506804	-0,625654418	4,483323199	5,337620907	5,052932398	0,854511689
In_In_In_Out_MMMP	-9285,965457	-9286,599519	-9283,501659	-0,634061676	11,35524876	6,934124302	8,281311898	-4,421123077
In_In_Out_Out_MMMP	-9285,941837	-9286,578875	-9283,478489	-0,637037952	26,17730553	19,88853082	22,82033678	-6,288764353
In_Out_In_Out_MMMP	-9285,949087	-9286,588648	-9283,491021	-0,639561384	21,62791813	13,7556734	14,95658853	-7,872241529
In_Out_Out_Out_MMMP	-9285,933553	-9286,584147	-9283,485554	-0,650594028	31,3755252	16,58019406	18,38732102	-14,79533057
Out_Out_Out_Out_MMMP	-9285,905018	-9286,560853	-9283,462324	-0,65583438	49,28141466	31,19745055	32,96402785	-18,08370143
In_In_In_In_MPMP	-9285,983553	-9286,610569	-9283,514856	-0,627016169	0	0	0	0
In_In_In_Out_MPMP	-9285,971332	-9286,59949	-9283,50114	-0,628157224	7,668749674	6,952485231	8,606964227	-0,716022853
In_In_Out_Out_MPMP	-9285,958917	-9286,59949	-9283,498739	-0,64057302	15,45921736	6,952158926	10,11332588	-8,507052981
In_Out_In_Out_MPMP	-9285,948968	-9286,585768	-9283,486862	-0,636800109	21,70237213	15,56261838	17,56662644	-6,139515548
In_Out_Out_Out_MPMP	-9285,933809	-9286,58088	-9283,482068	-0,647022999	31,21502082	18,62987854	20,57472501	-12,55447639
Out_Out_Out_Out_MPMP	-9285,913091	-9286,570579	-9283,468946	-0,657487829	44,21559321	25,09433081	28,8092247	-19,12125663
In_In_In_In_MMPP	-9285,966575	-9286,591403	-9283,49751	-0,624827673	10,65374334	12,02703533	10,88464801	1,373301591
In_In_In_Out_MMPP	-9285,953531	-9286,584565	-9283,490425	-0,631034226	18,83908393	16,31770684	15,33043987	-2,521369253
In_In_Out_Out_MMPP	-9285,920276	-9286,550007	-9283,456387	-0,629860562	39,70708178	38,00308601	36,68975255	-1,784883566
In_Out_In_Out_MMPP	-9285,943755	-9286,585375	-9283,489661	-0,641619512	24,9735415	15,80980065	15,80985085	-9,163736904
In_Out_Out_Out_MMPP	-9285,905745	-9286,547307	-9283,446403	-0,641562413	48,82521525	39,69746206	42,95483878	-9,127906801
Out_Out_Out_Out_MMPP	-9285,890411	-9286,543084	-9283,443814	-0,652673738	58,44720119	42,34729036	44,57969305	-16,10036842

Table 14 Set of energetical data for the 24 possible diastereoisomers of I-105₄

Chapter 2: Towards macrocycles dedicated to the selective recognition of therapeutic proteins

Abstract

PTMs are ubiquitous chemical transformations occurring on proteins after their translation by ribosomes. They are involved in many biological processes and contribute to the proteomic diversity. Nevertheless, they can be accidentally happened during the production of therapeutic protein, affecting their biological activity. Modified proteins are challenging to detect and purify. Hence, the development of selective receptors of the PTMs substrates and products is needed. In this perspective we chose to explore the ability of cyclophanes in selectively forming inclusion complex with modified proteins. Their hydrophobic cavity can bind to the alkyl chains of amino acids and are surrounded by two rims equipped with polar groups, which may bind to specific amino acid of the protein side chains. While there are a few examples of efficient and selective recognition of alkylated proteins by both static and dynamic cyclophanes, we herein focused on a new challenge: the selective recognition of products or substrates of methionine oxidation and primary amide hydrolysis (deamidation). An analysis of the binding pockets of enzymes which achieve this task in vivo provided candidates of functional groups of interest to equip the rims of the cavitands. Three alternative pathways were explored for the synthesis of the targeted cyclophanes bearing such side chains A) the cyclo-oligomerization of pre-functionalized monomers using irreversible bonds, B) the cyclo-oligomerization of either pre-functionalized monomers using a reversible linkage or the cyclo-oligomerization of unfunctionalized monomers using a reversible linkage followed by the post-functionalisation of the resulting cavitand, and C) the cyclo-oligomerization and monomers functionalization using two orthogonal reversible linkages. From A to C, the virtual combinatorial, i.e. the theoretical amount of species which may be obtained during a single synthetic procedure increase while the analysis of the mixture become more tedious. Regarding path A, the synthesis of corona[*n*]arenes and thioether-bridge cyclophanes by nucleophilic substitution could be validated on models but should be further optimised. The path B provided cyclophanes bearing reversible disulfide bonds between the aromatic units. However, in the conditions investigated for the cyclo-oligomerization, side reactions competed with disulfide formation in the case of 1,4-bisthiophenols bearing basic side chain in ortho 2,5 positions (amine for instance). As a matter of fact, sulfenyl amides, sulfinic and sulfenic acids were observed instead and the oligomers were limited to linear dimers. Finally, the path C led to complex mixtures and analysis. The 4-bisthiophenol bearing aldehyde in 2,5 position of the aromatic core was highly reactive and optimal conditions were not found to smoothly obtained an a cyclo-oligomerisation by disulfide bond and a functionalisation by acylhydrazone bond formation.

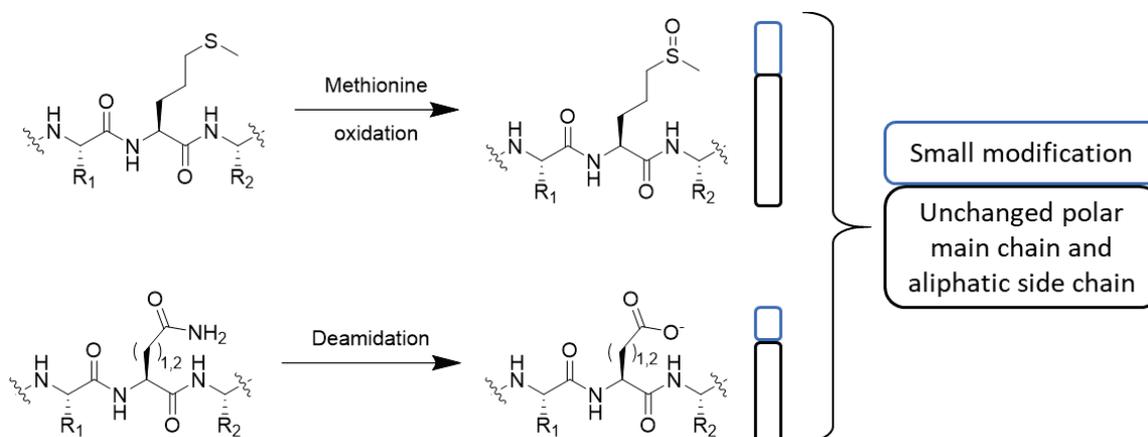


Scheme 132 : Synthesis of cyclophanes dedicated to the recognition of proteins that were affected by PTM from their native counterparts, thanks to side chains bearing specific groups R (green). They were 3 strategies, following the connection of side chains to aromatic monomers (blue) and the connection between these monomers (red). A) the cyclo-oligomerization of pre-functionalized monomers using irreversible bonds or synthesis of the cyclophanes by nucleophilic substitution or addition, B) the cyclo-oligomerization of either pre-functionalized monomers using disulfide bond formation or the cyclo-oligomerization of unfunctionalized monomers using a reversible linkage followed by the post-functionalisation of the resulting cavitand, and C) the cyclo-oligomerization and monomers functionalization using orthogonal reversible disulfide and acylhydrazone or imide bonds

1. Design of the macrocycles

a. Introduction and context

The goal of this chapter was to design and synthesize artificial supramolecular receptors which would preferentially bind either a native therapeutic proteins synthesised by recombinant technologies or the product resulting from one of the following post-translational modifications (PTMs): methoxidation or deamidation (*Scheme 133*). A PTM refers to the addition or transformation of a functional group covalently to a protein after its translation by the ribosome but it may also refer to the proteolysis and the folding of a protein²²⁵. Modifications occurring after ribosome translation can be decomposed into several categories: chemical alteration of amino acids, addition of functional groups, ligation to other proteins such as Ubiquitin (a marker for degradation) and chemical modifications, such as disulfide formations, proteolysis or racemisation²²⁵. Most PTMs are covalent modifications and result from enzymatic actions. The most frequent PTMs are, by order of frequency, phosphorylation, ubiquitination, succinylation, methylation, malonylation, and glycosylation²²⁶. These modifications are ubiquitous transformations occurring on proteins after their translation by ribosomes. Many mature proteins undergo PTMs before undertaking any function. Biologically programmed PTMs are an incredible source of proteomic diversity²²⁷ and are involved in many biological processes such as gene expression or metabolic regulation.²²⁸ This explains why defects in PTMs can lead to serious diseases²²⁹. More than 200 PTMs are known²³⁰. Among them, deamidation and methoxidation (or sulfoxidation, or methionine oxidation) are not listed in the first 30 ones. Deamidation refers to the hydrolysis of the primary amide ending group of the side chain of asparagine or glutamine residues leading to aspartic or glutamic acid²³¹, while methoxidation refers to the oxidation of the sulfur atom of the side chain of methionine residues.

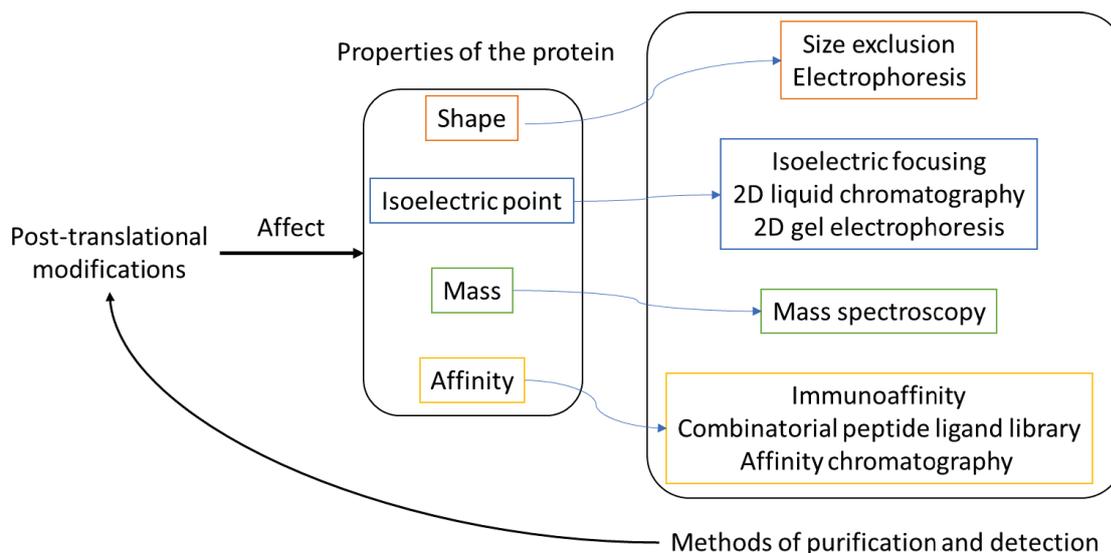


Scheme 133: Presentation of the targeted PTMs (methionine oxidation and deamidation) affecting methionine, glutamine and asparagine amino acids. R_1 , R_2 : any amino acids. The chemical change occurs on the side chain of the peptide, while the aliphatic part of the side chain and main chain remained unchanged. PTMs impact on folding was omitted for clarity.

Recombinant technology is currently participating to the rise of biomedicine drugs. These drugs are produced in living systems such as yeast or mammalian cells. This type of production provides access to complex biomolecules in their active form. One of the advantages is the reduced steps of purification compared to classical synthetic chemistry²³². Among biomedicine drugs, therapeutic proteins are defined as modified human peptides²³³. They show potential to treat particular diseases such as cancer and have a high biocompatibility in the human body²³⁴. However, recombinant technologies suffer

from undesired PTMs occurring mainly after the biosynthesis during the cellular lysis and purification steps for instance, alter the drug efficiency²³⁵. Currently, there is a lack of technologies that enable the detection and the purification proteins affected by PTMs^{235b}.

The methods of analysis and of purification of proteins that are affected by PTMs rely on classical biochemical techniques (Scheme 134). The separation of modified proteins from their native counterparts is based on the difference of physical properties such as ionization potential (pI) or molecular weight²³⁶. PTMs generally affect the shape of proteins that is why size exclusion chromatography and gel electrophoresis can detect modified proteins, even though they are not adapted for low molecular weight analytes²³⁷.



Scheme 134: Purification and detection of modified proteins affected by PTMs from their native counterparts. PTMs affect the properties of proteins. Each kind of modified properties can be the basis for the biological techniques of detection or purification.

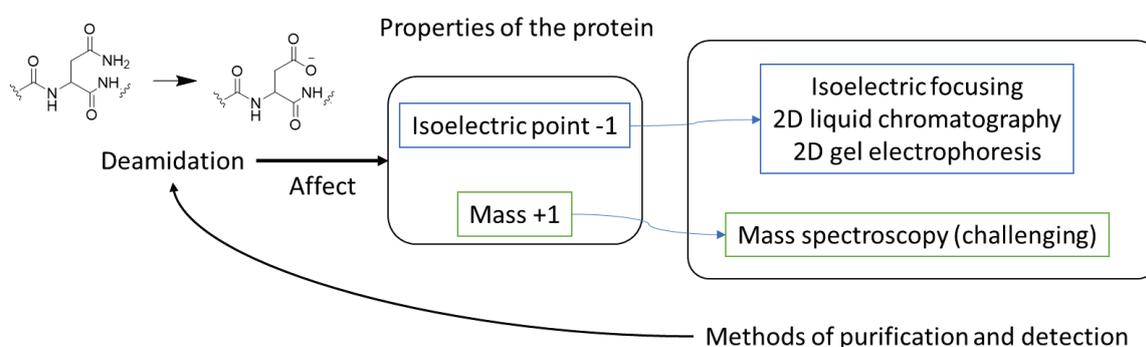
PTMs also usually modify the net charge and the iso electric point of proteins. This feature is exploited by 2D liquid chromatography, which is the improved version of 2D gel electrophoresis²³⁸, and by iso electric focusing, usually with a low resolution in the case of PTMs²³⁹. However, PTMs do not usually affect the shape of protein to a sufficient extent to afford a effective separation when the afore mentioned techniques are used.

Another approach consists in using (if any) new binding properties of proteins to perform separation based on affinity. The sample is analysed through a solid support (sugar, silica for instance) loaded with compounds displaying affinity for modified proteins²⁴⁰. The native protein is then eluted while the others are retained. Similarly, immunoaffinity chromatography is a separation method based on specific and reversible interactions between the proteins and a specific antibody column²⁴¹.

Being able to discriminate two proteins which differ from a single aminoacid (modified vs unmodified/native) is a real challenge in molecular recognition, which is defined by the specific interaction between two or more molecules through complementary noncovalent bonding²⁴². Yet, it would also meet a demand from the pharmaceutical industry. In fact, UCB Pharma developed an antibody drug CDP791 for applications in oncology. The candidate met its first objectives in a phase I study with patients suffering from refractory solid tumours. Then, in a phase II clinical study, the CDP791 was active on patients with lung cancer, in association with Paclitaxel. This drug is composed

of the Pegylated di-Fab antibody, blocking the angiogenesis regulator VEGFR-2²⁴³, a vascular endothelia growth factor receptor-2²⁴⁴ responsible for tumour vasculature. CDP791 is produced by recombinant technology²⁴⁵ and is mainly affected by two undesired and artificial PTMs, deamidation and sulfoxidation²⁴⁶, responsible for diseases²⁴⁷. UCB Pharma has developed some analytical techniques enabling to detect the presence of simple or multiples PTM, but it still lacks straightforward means of purifying the native proteins from their modified counterparts.

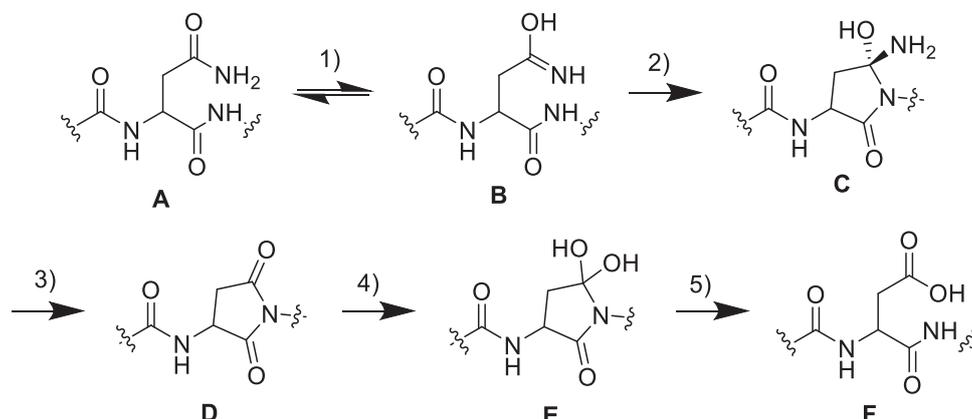
Each deamidation occurring on a residue gives an additional negative charge at the peptide (the carboxylic acid is deprotonated at physiological pH). Hence, methods of purification involving physical properties such as iso electric point are the most used to detect and purify modified proteins from their native analogues (Scheme 135). Other methods rely mainly on tandem mass spectrometry²⁴⁸. The deamidation changes the mass of the residue by +1 Da, which is challenging for mass analysis, detecting a m/z ²⁴⁹. New combinations of separation techniques and mass analysis are continuously proposed to improve the separation and the purification of protein that are affected by deamidation²⁵⁰. Even machine learning is now used to predict this PTM²⁵¹. Deamidation can occur up to 5 % in the Fab antibody responsible for the activity of CDP791²³.



Scheme 135 : Purification and detection of proteins affected by deamidation. Deamidation affects the iso electric point of proteins. Thus, techniques of purification are mainly based on iso electric techniques. The PTM also changes the mass from +1 Da, which is challenging for MS.

Deamidation is crucial for protein activity²⁵². Indeed, the backbone of the protein is changed by the addition of a negative charge. Deamidation is sometimes referred to a molecular clock²⁵³. This reaction proceeds through a succinimide or glutarimide intermediate. The mechanism is illustrated below with an asparagine²⁵⁴ (Scheme 136). First, the tautomeric form **B** undergoes a cyclisation with the neighbouring amide bond, forming a succinimide intermediate **D** after the release of an ammoniac moiety. A geminal diol **E** is the product from the nucleophilic attack of a water molecules. Finally, the cyclic intermediate **E** splits up, giving the carboxylic acid **F**. The rate determining step is the cyclisation²⁵⁵. A direct mechanism of hydrolysis was also proposed and modelised by DFT²⁵⁶. The process depends on the solvent, the pH, the temperature, the media and the protein shape²⁵⁷. The rate increases with the pH and the temperature. It also increases when the neighbouring residues are not sterically hindered, facilitating the cyclic intermediate formation. Water assistance is important in the kinetic²⁵⁸. In physiological conditions, the half-life time of asparagine and glutamine is on average around 500 days but can decrease to 1 day in specific conditions²⁵⁹. The deamidation occurs mainly at

the surface of the proteins because of kinetic considerations. Indeed, intern sites of deamidation are buried, limiting the diffusion of water.

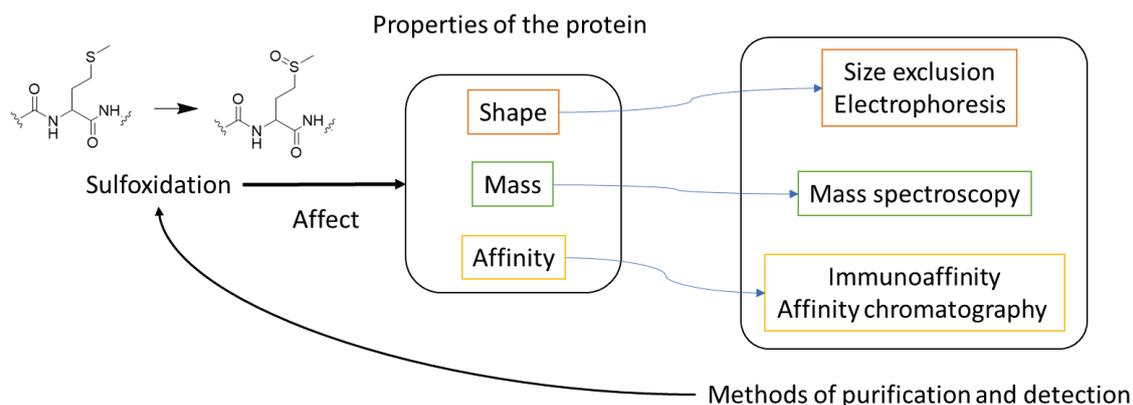


Scheme 136 : Mechanism of deamidation for an asparagine. 1) tautomerization, 2) formation a cyclic intermediate, 3) Release of ammoniac giving a succinimide, 4) formation a geminal diol by a nucleophilic attack of water on the succinimide, 5) Breaking of the 5-membered ring. The protonation/deprotonation steps are omitted for clarity.

Similarly, oxidation of methionine occurs mainly at the surface of the proteins. Methionines are transformed into methionine-sulfoxide and could even be oxidised to methionine sulfone²²⁵. The conversion of methionine into a sulfoxide creates a stereocenter²²⁶. Reactive oxygen species such as triplet dioxygen or hydroperoxides²⁶⁰ are responsible for this oxidation, assimilated to oxidative stress. Oxidative stress is a perturbation in the balance between the production of reactive oxygen species and antioxidant defenses²⁶¹. The number of methionine sulfoxide in a protein is a metric of its oxidation level and is used as a point of control in manufacturing processes^{235b}. A high level of oxidation level in human proteins is linked to diseases such as Alzheimer^{247c}. The mechanism is described as not universal and complex²⁶⁰. Under physiological conditions, the transport of oxidant molecules to the methionine residue governs the kinetic²⁶². Sulfoxidation affects the protein activity and is largely found in ageing cells²⁶³. Its impact is complex and various. For instance, methoxidation can improve the state of some autoimmune diseases. Indeed, inflammation states release highly oxidant species, inducing oxidation. The oxidised proteins behave as antibodies against the antigen in the corresponding autoimmune affection²⁶⁴. Thus, methionine residues act as antioxidants and protect the other residues from oxidation²⁶⁵. However, sulfoxidation has negative effects most of the time. Immunoglobulin antibodies are impacted by methionine oxidation because of a change of conformation. These deviations can highly increase their intrinsic cytotoxicity and modify their activity²⁶⁶. Methionine oxidation also induces phosphorylation at the neighbouring positions²⁶⁷ and increases the proteolytic rate of the whole protein²⁶⁸.

The sulfoxidation affects the shape of the proteins but does not modify their overall charge. Hence, electrophoresis methods are not adapted to purify modified proteins from their native counterparts. Only affinity chromatography is efficient and based on the change of hydrophilic/hydrophobic properties²⁶⁹. However, this method is adapted when the oxidation level is high and when the shape of the protein is significantly changed. The detection of proteins modified by sulfoxidation is usually performed by liquid chromatography coupled with MS analysis. This method suffers from artefacts and a long time of analysis. Indeed, the whole protein should be cleaved, increasing the time of analysis

and creating artefacts from the proteolysis. To tackle this issue, a method involving partial digestion was developed²⁷⁰. Raman spectroscopy can be used to detect the newly formed S=O bonds and the modified vibration of the C-S bond²⁷¹. Affinity methods involving specific antibodies of methionine sulfoxide were also developed²⁷². All the methods rely on the performance of the mass spectrometry²⁷³. Another original approach inserts a fluorogenic moiety into a methionine sulfoxidase sequence. That intracellular sensor is quickly able to detect a change in the oxidation level²⁷⁴. Finally, machine learning approaches are now employed to predict methionine oxidation²⁷⁵.



Scheme 137 : Purification and detection of proteins affected by methionine oxidation. Methoxidation affects the shape, the mass and the affinity of proteins. Thus, techniques of purification are mainly based on affinity chromatography and liquid chromatography. The PTM does not change the iso electric point, but the mass of the protein. MS is particularly efficient to detect that PTM.

As previously mentioned, PTMs contributes to the proteome diversity and to a variety of functions in the human body. However, deregulation in the number of PTMs affected a given protein is linked to many diseases such as cancer. That is why a regulatory process involving enzymes and reader proteins exists and interacts with modified proteins in order to limit the level of PTMs. A reader is a molecule or a protein domain that recognises and bind with the amino acid affected by PTMs, often by the formation of complexes. Similarly, specific enzymes react with the modified moiety of the side chain of the peptide, restoring its native form. This mechanism is found in histone tails through the dynamic addition, removal and addition of PTMs, regulating gene transcription of the histone code. The PTMs involved are mainly methylation of lysine and arginine.

PTMs are inevitable in recombinant technologies. Biomedicine needs improvements for the detection and separation of modified therapeutic proteins. Currently, the analytical tools mainly rely on MS and no straightforward separation exists to date.

b. Molecular recognition of lysine methylation

Products and substrates of deamidation and methoxidation have never been targeted by supramolecular receptors, at the opposite of other specific moieties of complex biomolecules^{11,35} and modified proteins by PTMs like lysine methylation (Scheme 138). The methylation of lysine or arginine is a widespread natural PTM that is usually controlled by methyltransferases and lysine demethylases. Lysine methylation mainly occurs on histones and transcription factors on chromatin. Chromatin is a complex between DNA and histones that forms chromosomes within eukaryotic cells. Hence, lysine methylation influences biological processes such as cellular responses²⁷⁶. Many supramolecular receptors forming a complex of inclusion with the side chains of lysine peptides were proposed in order

to bind to methylated lysine, mainly by cation – π and electrostatic interactions²⁷⁷. The binding of methylated arginine is more complex, as the corresponding guanidinium charge is more diffuse²⁷⁸. Water-soluble sulfato-calixarene **II-2** bound with short peptides containing a trimethylated lysine (K_{me^3}) with K_a in the range of $10^6 M^{-1}$ (ITC titration) (Scheme 138)²⁷⁹, more efficiently than a natural protein CBX7 dedicated to the recognition of trimethylated lysine selectively with lysine. The binding was based on electrostatic interaction between the anionic sulfate and the cationic ammonium. The interaction was reinforced in the case of the second calixarenes **II-3** by cation- π interactions with the aromatic ring on the upper rim. These calixarenes were successfully grafted on agarose and used in affinity chromatography, the eluting order of peptides following their lysine methylation state²⁸⁰. This example highlighted that cavitands were adapted to recognise a specific moiety of the side chains of a peptidic sequence.

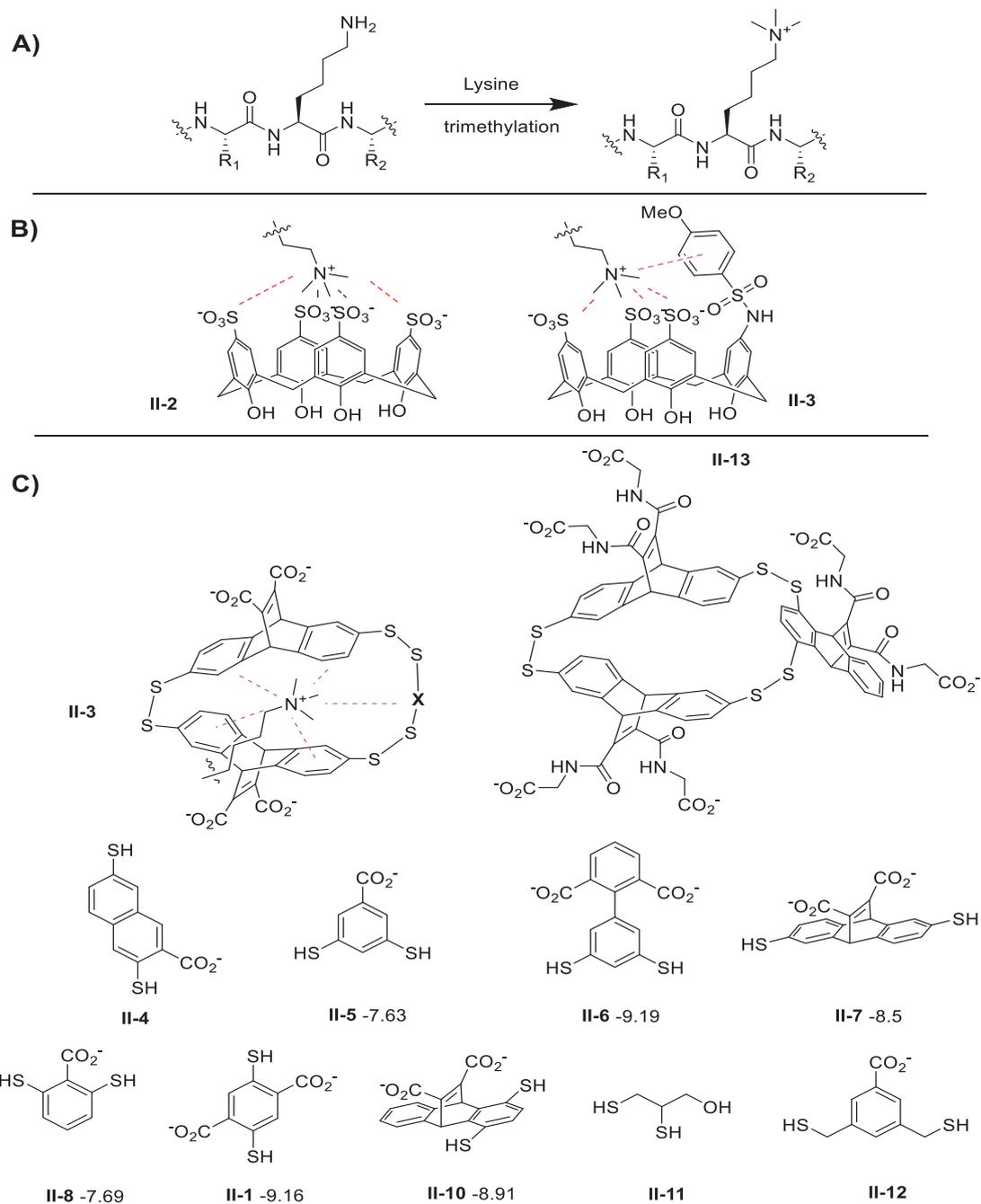
The analysis of methyl lysine and dimethyl arginine reader proteins such as malign brain tumour domains guided the design of other receptors²⁸¹. This approach illustrates the scientific approach for the design of synthetic receptors by the analysis of the binding pockets of peptides which bind to the modified moiety. Methylation of lysine does not affect the charge of the residue but removes a proton from the ϵ -amine, increasing its hydrophobicity and decreasing its ability to be a donor of H-bonds.

Waters *et al.* studied the X-ray structure of readers proteins of lysine methylation. They found that trimethylated-lysine K_{me^3} bound with hydrophobic pockets composed of aromatic residues (typical composition: 2-4 tyrosine, tryptophan, phenylalanine) thanks to cation- π interactions with the NCH_3 groups. The selectivity for trimethyl-lysine was provided by the reduction of desolvation of the ammonium K_{me^3} inside the hydrophobic pocket, compared to the highly solvated NH_3^+ . The selectivity for lower methylated K_{me^2} and K_{me} was achieved by 1) providing steric hindrances, preventing K_{me^3} to be efficiently recognised, 2) a change of an aromatic residue for an acidic one, developing salt bridges with available hydrogen of lower methylation states.

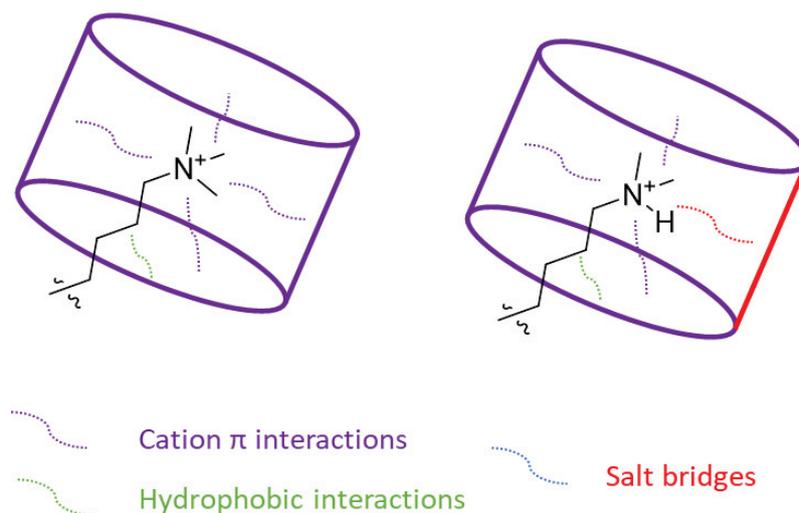
Dougherty *et al.* demonstrated that the cation- π interactions drove the recognition of cationic guests by cyclophanes with a limited influence of hydrophobicity^{22,278}. Otto and Sanders developed similar disulfide bridge cyclophanes through dynamic combinatorial chemistry for the recognition of various cationic guests in water³⁹ⁱ. Waters *et al.* based their design of synthetic receptors of proteins that are affected by lysine methylation on these analysis and results^{79d,80,282}. They used a combination of water soluble aromatic polythiols with carboxylic acids to generate cyclophanes with negatively charged exteriors for solubility and reinforced binding with ammonium, aromatic interiors for cation- π interactions and hydrophobic interactions with methylated ammonium and disulfide bonds between the monomers. The reversible disulfide exchange enabled a screening of the receptors for different methylation states on short peptides.

All the receptors generated from DCLs were cyclic heterotrimers composed of 2 different units (Scheme 138). The **II-3** receptors bound to all methylation states, with association constants between 10^5 and 10^7 . A high selectivity for short model peptides with trimethylated lysine compared to the native models was reported. **II-9** showed a selectivity of 70 (ratio between the association constants with $Ac-K_{me^3}G-NH_2$ and $Ac-K_{me^0}G-NH_2$) for instance^{79d,80,282}. The study highlighted that the $N-CH_3$ groups were buried inside the cavity and bound with the receptors by cation- π interactions (Scheme 139). The hydrophobicity participated in the selectivity over unmethylated states thanks to their higher cost of desolvation. In addition, aromatic moieties interacted with aliphatic side chain of lysine residues. In addition, similarly to the readers proteins, selectivity for K_{me^2} over K_{me^3} was obtained by bringing

carboxylate moieties near the hydrophobic cavity, which created salt bridge with K_{me}^2 while K_{me}^3 could not developed H-bonds.



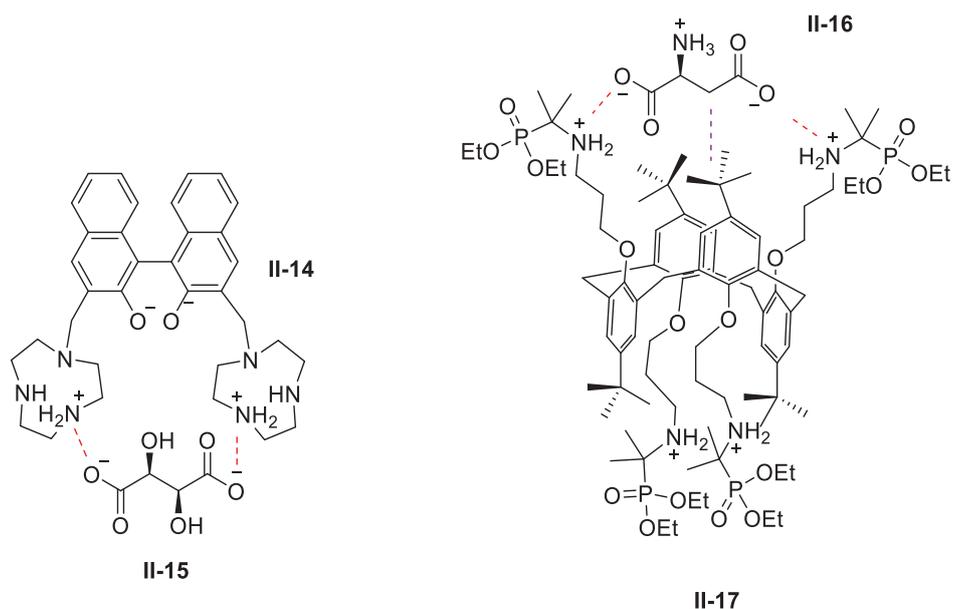
*Scheme 138: Examples of architectures that recognised selectively the products or substrates of lysine trimethylation. A) Representation of the lysine trimethylation R_1 , R_2 : any amino acids. B) Examples of cavitands **II-2** and **II-3** that recognised the products of lysine trimethylation $H3K27me^3$ more efficiently than a natural reader protein of trimethyl-lysine CBX7, thanks to electrostatic interactions (red) and cation- π interactions (purple). C) Receptors synthesised from self-assembly for selective recognition of Kme^3 by cation- π interactions (purple), identified from DCLs (5 mM, 50 mM borate buffer, pH 8.5, template Ac- Kme^3 GGL-NH₂). ΔG of association with a short peptide (Ac- Kme^3 GGL-NH₂ or Ac- Kme^3 GL-NH₂) including a Kme^3 mutation in kcal.mol⁻¹ determined by ITC when available. Top left: generic receptor backbone, top right: example **II-13** of post modification with a glycine spacer. bottom, chemical structure of the third building block **X**.*



Scheme 139 : Recognition of methylation of lysine (left, K_{me^3} , right, K_{me^2}) by mercaptophanes of Water et al, through inclusion complex driven by cation- π interactions and hydrophobic interactions. Lower desolvation cost of the ammonium was at the origin of the selectivity for trimethylated lysine. The orientation of carboxylic acids inside the cavity developed selective salt bridges with lower methylations states.

From that example, cavitands and DCC are adapted to the recognition of modified proteins and the analysis of PTM readers provides useful information for the design. In addition, carboxylic acids solubilised the receptors without interfering with the process of recognition. Finally, proteins are more and more directly used as template in DCLs²⁸³. The use of small peptides bearing mutations will be useful to select a receptor from DCLs.

Deamidated moieties, i.e. aspartate, isoaspartate and glutamate residue bear oxyanionic moieties which are known to be challenging to recognise in water²⁸⁴. Indeed, the negative charge of anions is general more diffuse in water than cationic charges of cations because of their larger size²⁸⁵. Over the last decades, important milestones were achieved, involving the understanding of the influence of anion- π interactions²⁸⁶, H-bonds and halogen bonds, on recognition of anions in water²⁸⁷. Few examples of receptors reported a recognition of carboxylate ions such as short peptides or tartrate-like ions. In those examples, ammonium groups were responsible for salt bridges with the carboxylate moieties. For instance, a Binol-based receptor **II-14** bound with **II-15** in aqueous buffers with a K_a of $10^{5.8} \text{ M}^{-1}$ mainly through salt bridges²⁸⁸ (*Scheme 140*). That highlighted the importance salt bridges and H-bond donors like amines and amides for the recognition of deamidated proteins. Calixarenes were able to recognise carboxylic acids ($\log K_a$ in the range of 2.1-6.1, determined by UV titration)²⁸⁹. For instance, **II-17** formed an inclusion complex with aspartic acid **II-16** ($\log K_a = 5.7$) (*Scheme 140*)²⁹⁰. In addition, a conical calix[4]arene with four cationic guanidiniomethyl groups at the upper rim and hydrophobic loops at the lower rim recognised the four monomers of the modified p53-R337H proteins; by both salt bridges between the cationic amines in solutions with the carboxylate moieties, and by weak hydrophobic interactions between the side chain of the amino acid and the hydrophobic interior, with an association constant of 10^4 M^{-1} ²⁹¹. Cucurbituril could also bind to aspartic acid but in a lesser extent than other amino acids²⁹². The hydrophobic cavity of the cavitants play a role in the recognition by modifying the local dielectric environment²⁹³.



Scheme 140: Various hosts of carboxylate moieties. Left: Binol-derivative **II-14** binding tartrate ion **II-15** by salt bridges (red), right: alternate calixarenes **II-17** binding aspartic acid **II-16** thanks to salt bridges (red) and weak hydrophobic interactions (purple).

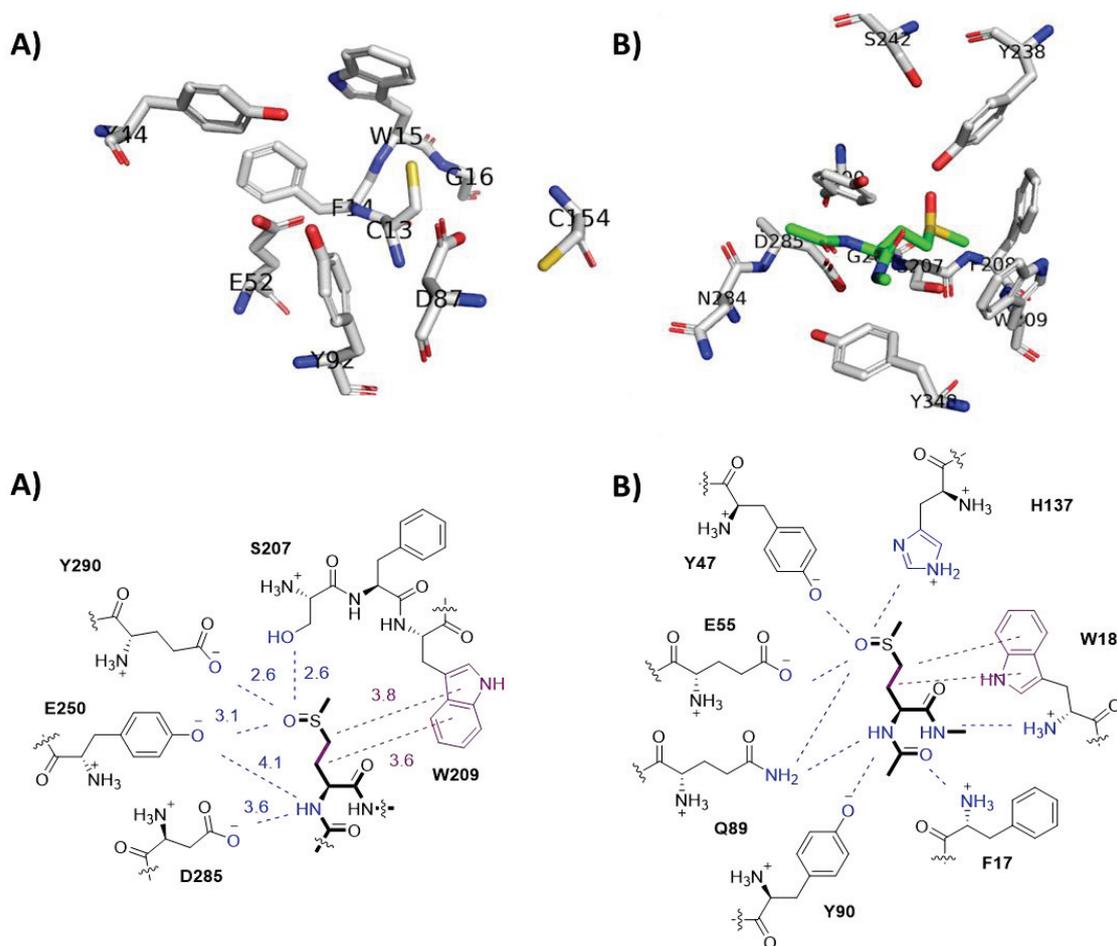
In a nutshell, the design of synthetic receptors can be inspired by the analysis of readers proteins and enzymes, as illustrated by the example of the recognition of lysine methylation by Water *et al.* The use of dynamic combinatorial chemistry enables an efficient screening of building blocks incorporating the chemical functions identified from the analysis of enzymatic pockets. Therefore, the use of disulfide-bridge cavitants seemed particularly adapted to the recognition of modified proteins. The binding by inclusion complex led to important affinity and participated in the selectivity, by reinforced hydrophobic effects and dispersive interactions for instance, compared to linear receptors. Given that numerous enzymes linked to deamidation of methoxidation were crystallised, the same approach could be translated for our goal. Nevertheless, the use of cavitands bounding by inclusion complex would be limited to the surface of proteins, but most of the modified proteins that are affected by the two targeted PTMs are located at the surface of the proteins.

c. Analysis of enzymes

Therefore, following the analysis of Waters *et al* regarding lysine methylation, enzymes linked to deamidation and methionine oxidation were analysed in order to understand the driving forces of the recognition.

Regarding the deamidation, a protein and a deamidase of prokaryotic ubiquitin-like protein and their substrates, a C-terminal glutamic acid, highlighted in bold were analysed (Scheme 141). Deamidase are enzymes that cleave the amide bond of amino acid. The active pockets of two enzymes, an aspartate binding protein and a gave substantial elements. The binding pockets were mainly composed of H-bonds donors such as tyrosine, serine, highlighted in blue, and functional groups salt bridges, such as guanidine, histidine, coloured in red. In the first example (A), a periplasmic L-aspartate/L-glutamate binding protein from *Shigella flexneri* complexed with an L-glutamate²⁹⁴, 6 H-bonds (2 serines, 2 tyrosines, 1 aspartic acid) and 4 salt bridges (3 guanidines, 1 histidine) established interactions with the substrate at a distance of 3 Å in average. The binding pocket of the second example (B)²⁹⁵ was composed of 5 functional groups for salt bridges (4 histidine, 1 guanidine) and 3 H-bonds donors (2

the residue (1 tyrosine, 1 aspartic acid at 3.8 Å). In addition, a triptophane residue interacted with the alkyl moieties of the side chain at 3.6 Å. The second enzymatic site (B)²⁹⁸ of mycobacterium tuberculosis methionine sulfoxide reductase A was buried inside the protein mycobacterium tuberculosis methionine sulfoxide reductase A and similarly organised. 6 H-bond donors (1 histidine, 2 tyrosines, 1 glutamatic acid, 1 glutamine, 1 phenylalanine) and an aromatic tryptophane paved the cavity. Methionine was known to pack with aromatic residues such as tyrosine and tryptophan by S – π , hydrophobic and dispersive interactions, which were increased when methionine was oxidized, by introducing a dipole-quadrupole interaction²⁹⁹. Other analysis of crystals of methionine reductases²⁵ were consistent with this analysis^{297d,300}.



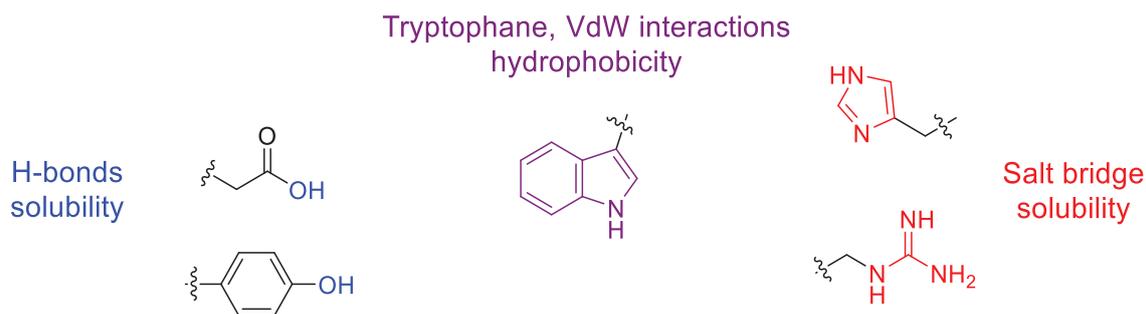
Scheme 142: Binding pockets of three methionine sulfoxidase enzymes²⁶. Top: PDB 3D view, bottom: 2D representation and analysis. One letter code as amino acid identifier, residue number from corresponding PDB, generated with Pymol¹⁹⁸ software from PDB native structures²⁷. Substrate: methionine which is the product of the methoxidation of lysines. The binding pockets are mainly composed of H-bonds donor, highlighted in blue, and aromatic developing a hydrophobic area, coloured in purple. The distances are given, if measurable thanks to the PDB data. Protonation states at physiological pH.

25 A) <https://www.rcsb.org/structure/3CXK>, B) <https://www.rcsb.org/structure/1L1D>

26 A) <https://www.rcsb.org/structure/1FF3>, B) <https://www.rcsb.org/structure/1FVA>, C) <https://www.rcsb.org/structure/1NWA>

27 A) <https://www.rcsb.org/structure/2vha>; B) <https://www.rcsb.org/structure/4B0R>

The analysis of the previous proteins readers and enzymes revealed the nature of the functional groups that may be relevant in the design of cyclophanes binding to modified protein suffering from deamidation and sulfoxidation. Carboxylic acids contributed to the solubility of the proteins and were engaged in H-bonds. An hydrophobic environment, induced by the presence of tyrosine and en tryptophane residues, induced a quadrupole- π interactions with the S=O bond and interacted with the alkyl part of the side chains. Phenols were widely present with tyrosine residues in the analysed binding pockets because they are strong donors of H-bonds and contributed to the hydrophobicity of the pockets. Regarding the deamidation, numerous salt bridges were developed by amines that are protonated in solution, such as guanidine and histidine.



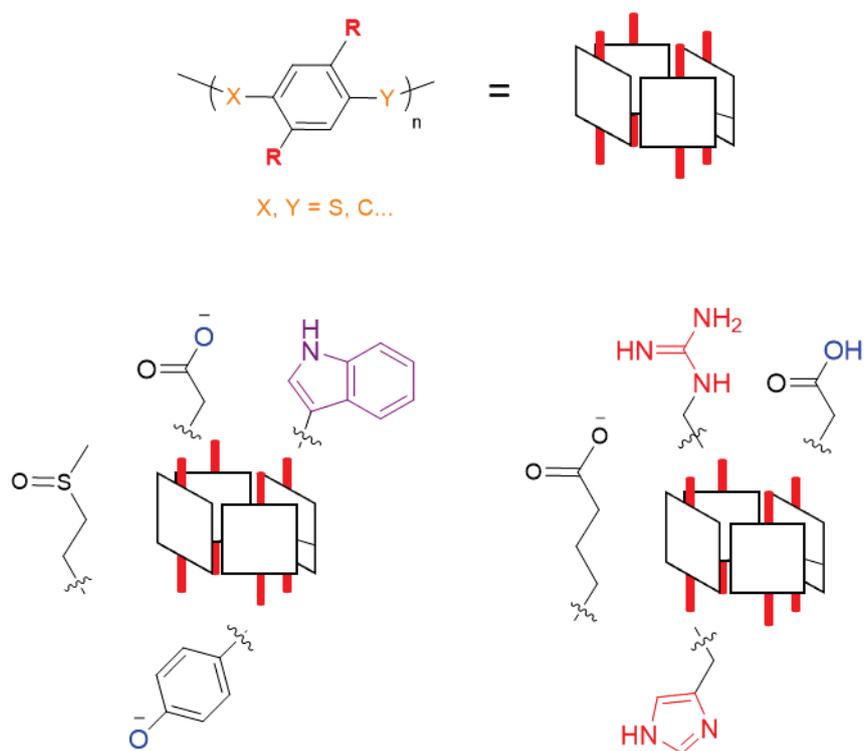
Scheme 143: Functional groups identified from the readers of proteins affected by deamidation and methoxydation

At this point, the key functional groups and mechanism of action involved in the binding with modified proteins were identified. Their spatial arrangement should now be proposed in the design of synthetic receptors.

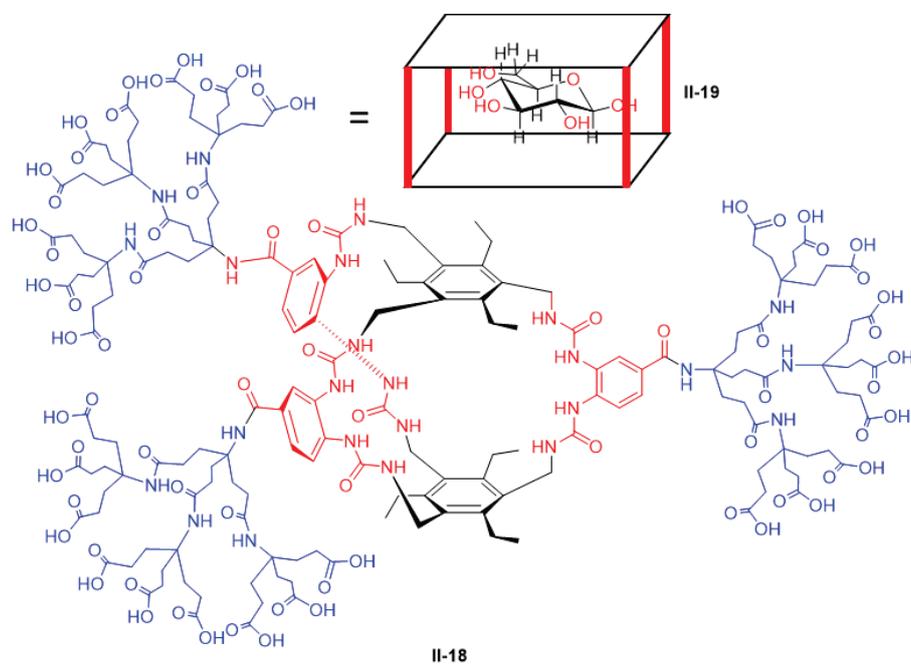
d. Proposition of design

Therefore, we propose the following design: an inclusion complex with the residue affected by the PTM will be obtained with a cavitant, especially a cyclophane, a which involves the repetition of aromatic units (Scheme 132). Those units should provide an hydrophobic environnement in the equatorial region of the cavitand, which may display an affinity for the apolar aliphatic chain of the aminoacid to be recognised on the protein side chains. In addition they will enhance electrostatic interactions between the modified moieties and their polar rims bearing the functional groups previously identified.

The opposite configuration can also be explored with polar equatorial sites and hydrophoc heads and was efficient for the recognition of acetylcholine³⁰¹ or chlordecone by hemicryptophanes³⁰². In the following example, host **II-18** (Scheme 145) bound glucose **II-19** with a K_a of $1.8 \cdot 10^4$, hundred times stronger than other simple sugars. This receptor contained aromatic electron rich heads (black) and tailored interactions (red) with equatorial groups for glucose³⁰³. The size was anticipated, secondary interactions³⁰⁴ and solubility were provided by dendritic side chains (green).



Scheme 144 : Representation of the targeted cyclophanes. The principal chain is composed of an aromatic core obtained by a sequence of monomers. Each monomer carries polar heads bearing functional groups dedicated to the recognition of modified proteins (left, bottom: case of methoxidation, right, bottom, case of deamidation). The monomers are bridged by aliphatic chain including heteroelements such as sulfur.



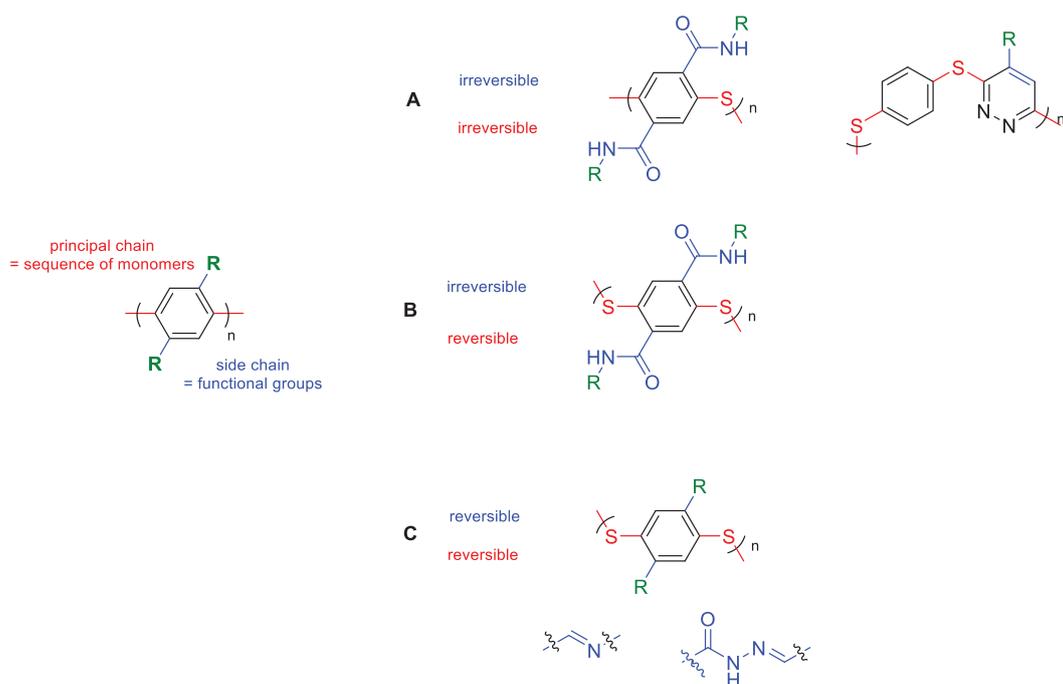
Scheme 145 : Biomimetic receptor **II-18** for carbohydrates, especially glucose **II-19**. Solubilising and secondary interaction by blue dendrimeric chain, tailored interactions in red with polar equatorial groups, tailored apolar interactions in black with aromatic core and axial groups.

The first configuration (polar heads with tailored functional groups and hydrophobic equatorial site) was chosen because of our expertise on that sequence with bisthiophenols^{40a} and given that it was efficiently realised by Waters *et al.* in their recognition of lysine methylation. Similarly to the peptidic terminology, the main or principal chain of the cavitands are defined as the sequence of monomers leading to the aromatic core, and the side chains are defined as the hydrophilic functional groups.

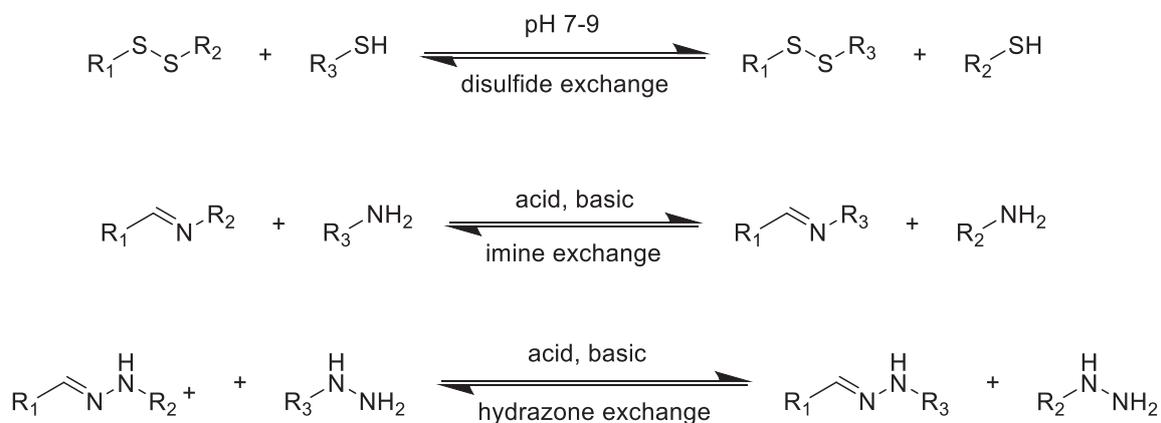
Three strategies were explored regarding the synthesis of the cavitands, ie the assembly of the main and side chains either by reversible (DCC) or irreversible connecting chemistry (classical synthetic chemistry) (Scheme 146):

- **Strategy A:** The aromatic monomers are connected by irreversible bonds thanks the nucleophilicity of thiols. The functional side chains are connected by irreversible bonds to the macrocyclic scaffold. Corresponding static cyclophanes will be obtained by the nucleophilic substitution of 1,4-bisthiophenols with bromides derivatives in the case of thioether bridge cyclophanes. In addition, the aromatic nucleophilic substitution between 1,4-bisthiophenols and a 3,6-dichloro-1,2,4,5-tetrazine will give access to the family of corona[n]renes. The disadvantage is the design a priori: the cavitands are synthesised and then their binding with modified proteins will be evaluated. They are not selected, like the calixarenes described previously (Scheme 140).
- **Strategy B:** The monomers are connected by reversible bonds while the side chains are connected by irreversible bonds to the main macrocyclic scaffold using amide bonds. This design enables a selection process to occur. Strategy B can be divided into B1 and B2. B1 refers to the introduction of the functional groups irreversibly before making the principal chain by reversible bonds. That strategy was widely employed for static cavitands, for instance with hemicryptophanes³⁰⁵ and also for carboxylate aromatic bistiols^{79e}. Alternatively, B2 refers to the formation of the principal chain by reversible bonds before the introduction of the functional groups on the heads. That strategy was employed in different areas of supramoleclar chemistry, from self-assembled nanostructures³⁰⁶ to pillar[n]arenes³⁰⁷. II-13 is a relevant example regarding the recognition of methylated lysine (Scheme 138). As a matter of fact, the principal chain was self-assembly by disulfide bond formation from thiols, and after, a post modification by the formation of amide bonds introduced the glycine spacer, contributing at both recognition and solubilisation. A glycine moiety could then receive a pegylated biotin linker efficiently used for grafting on silica support^{79e,308}. The amide bond as connector and the disulfide bond as reversible link are hence validated, which is reinforced by the expertise of the group in the self-assembly of thiols^{40a}. In our design, the monomers will be connected by disulfide bond formation during DCLs, while the connection with side chains will be obtained mainly by amide bond formation.
- **Strategy C:** The monomers and the side chains are assembled by reversible bonds. This design explores the largest combinatorial space but the complexity of the mixtures potentially generated may prove to be tedious to analyse. DCLs based on simultaneous exchanges reaction were well documented. They relied on a variety of reversible bonds such as imines^{37,57m,309}, disulfides, hydrazones^{47b,310}, acylhydrazones^{283d,311}, dithioacetals, acetals³¹², adducts of Diels-Alder^{38b}, products of metathesis^{60a} and many others^{39a,309c,313}. Regarding the recognition of modified protein or simplest target, the simultaneous use of two reversible bonds in DCLs to get a selection of one macrocyclic receptor is not documented. However, the imine bond, as well as the hydrazone and acylhydrazone could be combined efficiently with

disulfide bonds^{39c,f,g,314113d315}. The two reversible bonds can contribute simultaneously to the self-assembly, or alternatively following the conditions. For instance the disulfide exchange is frozen at acidic pH in which operates the hydrazone exchange. Examples of 3³¹⁶ and 4³¹⁷ simultaneous exchanges were documented with disulfides but were considered too complex here. Consequently, the reactions for the reversible connections between the monomers will be the thiol-disulfide, imine and (acyl)hydrazone exchanges (Scheme 146). In addition, the three last involve amines as building blocks, which are relevant functional groups for the creation of salt bridges with glutamate/aspartate residues.



Scheme 146: Retrosynthetic analysis for the design of cavitants dedicated to form inclusion complex with modified proteins. Three strategies were proposed in order to build the connection between the monomeric units forming the principal chain (red) and the connection between the principal chain and the side chain bearing the functional groups (blue). The strategies differ from their level of dynamic connections. R: functional groups dedicated to the recognition of modified proteins.



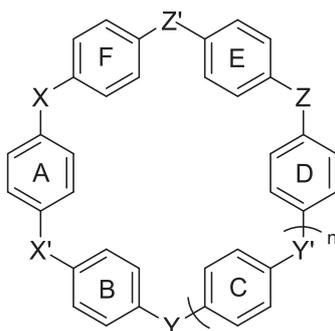
Scheme 147: Reversible reactions explore for the assembling of the monomers and side chains into functionalized cavitand

The monomers will be mainly 1,4-bisthiophenols given our expertise in their self-assembly through DCC. The 3 strategy of design will be experimentally explored and are detailed in this chapter. Computational chemistry will be employed in case of synthesis issues in order to understand the involved reactivity.

2. Synthesis of corona[n]arenes out of framework of DCC

This subchapter reports the synthesis of corona[n]arene, a class of cavitands corresponding to the strategy **A** described in the introduction (Scheme 146).

Macrocycles are versatile receptors for supramolecular chemists, as attested by the wide range of action of calixarenes^{289d} and pillararenes^{14a,318}. Over the last decade, a diversification of their structure and their use could be observed³¹⁹. This activity in the field of synthetic macrocyclic chemistry increased the potential of static cavitands for applications such as the recognition of modified proteins. Synthetic macromolecules, obtained out of the framework of DCC remain potential receptors for proteins that are affected by PTMs. In the early 2010s, a new family of macrocycles was reported : corona[n]arenes (Scheme 148). They are composed of *para*-(het)arylenes and heteroatoms alternatively³¹⁹.

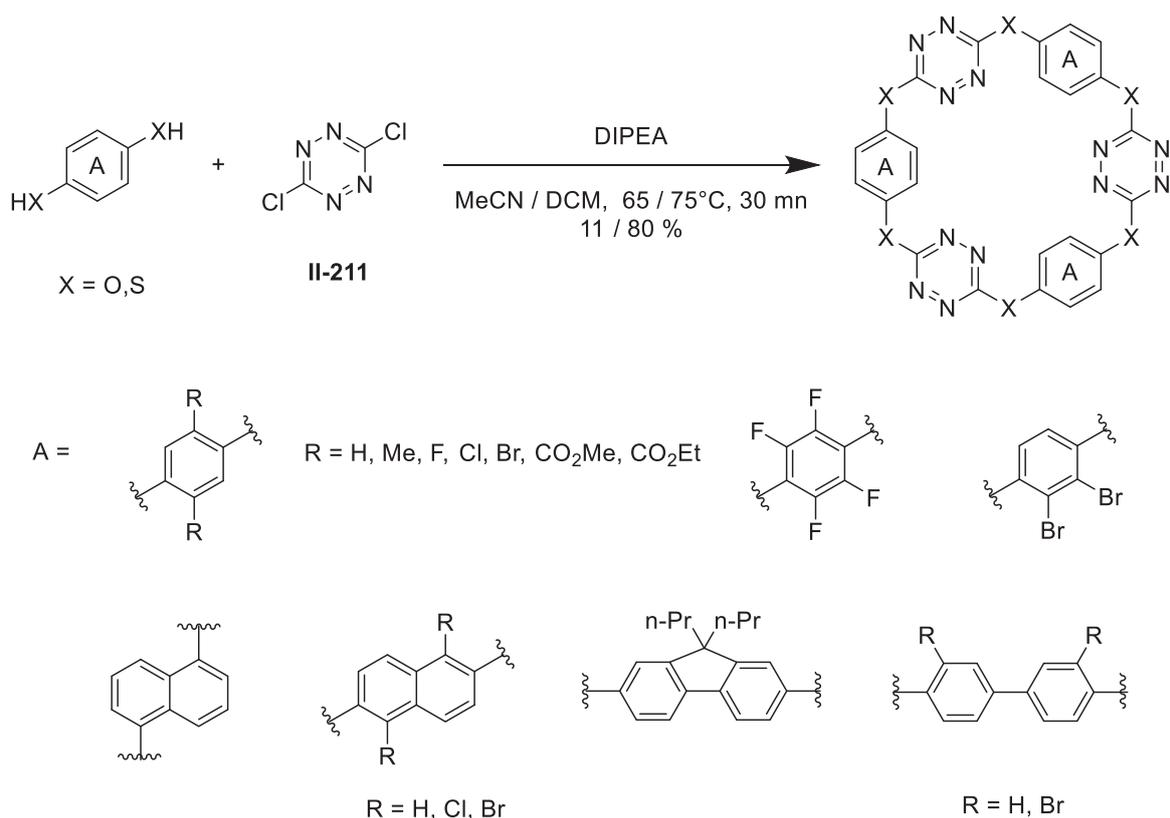


Scheme 148: Structure of corona[n]arenes. X,X',Y,Y',Z,Z' = heteroatoms, A,B,C,D,E,F = (hetero)aromatics

Based on the literature, developed below, this family of cavitands was chosen to be used in our desired applications of recognition of modified proteins, particularly corona[n]arene[n]tetrazines and pyradizines.

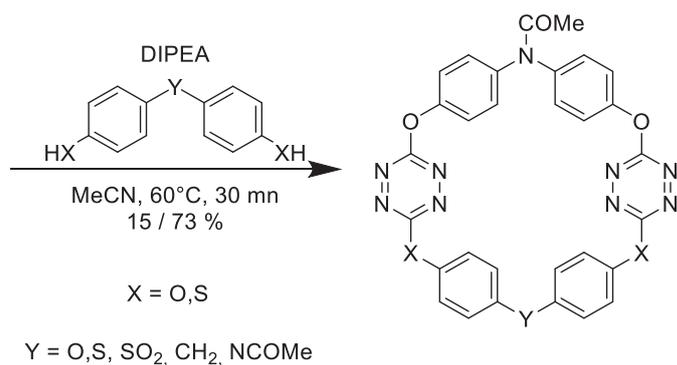
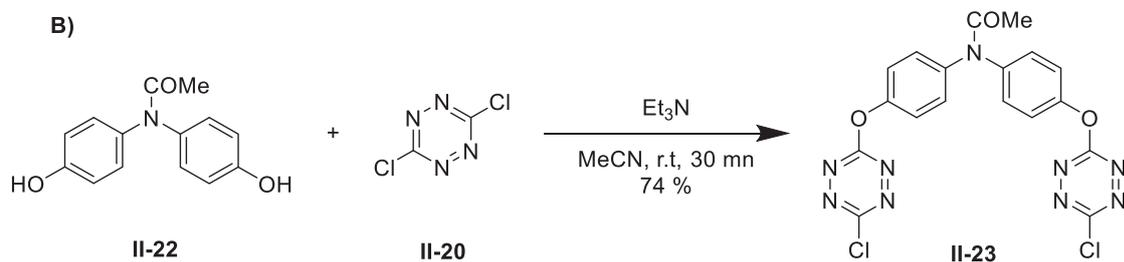
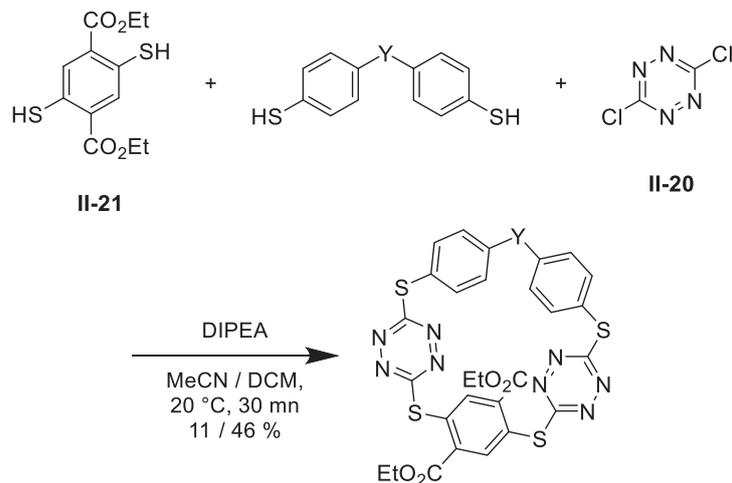
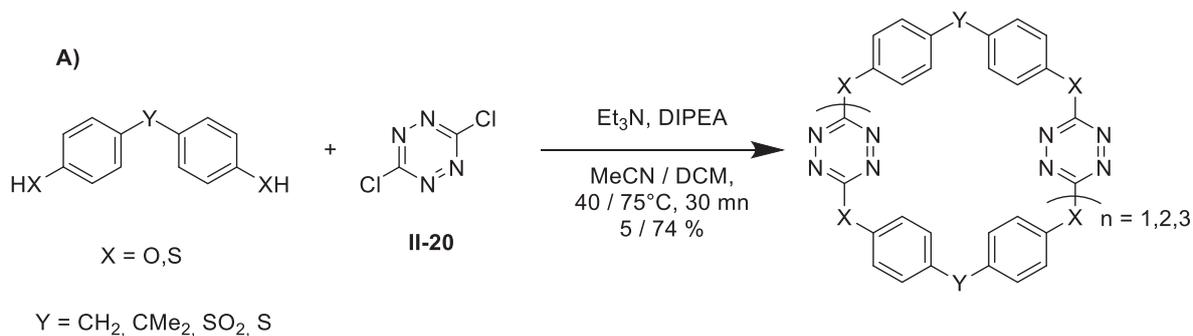
a. Access and applications of corona[n]arenes

This bibliographic subpart gives elements on that family of cavitands and summarises the current available literature on its synthesis. Following the work of Wang *et al.*, Corona[n]arene[n]tetrazines are formed by S_{NAr} between aromatic dithiols or diphenols and **II-20** 3,6-dichloro-1,2,4,5-tetrazine³²⁰. The reaction is quick (30 mn) and presented an interesting scope of aromatics (Scheme 149). Indeed, different backbones were accessible with various electronic properties. For instance, electronically deficient corona[n]arene were obtained from aromatic halides. That is why that family belongs to the strategy **A** of design, the main chain being obtained by the S_{NAr} (Scheme 146). The major formation of one object over all the combinations between the two initial reactants was remarkable. Most of the time the yields were satisfying. The major products corona[3]arene[3]tetrazines were formed thanks to low concentration (5 mM) and dropwise addition during the process. No links between the structures of the aromatic starting material and the yields were pointed out.

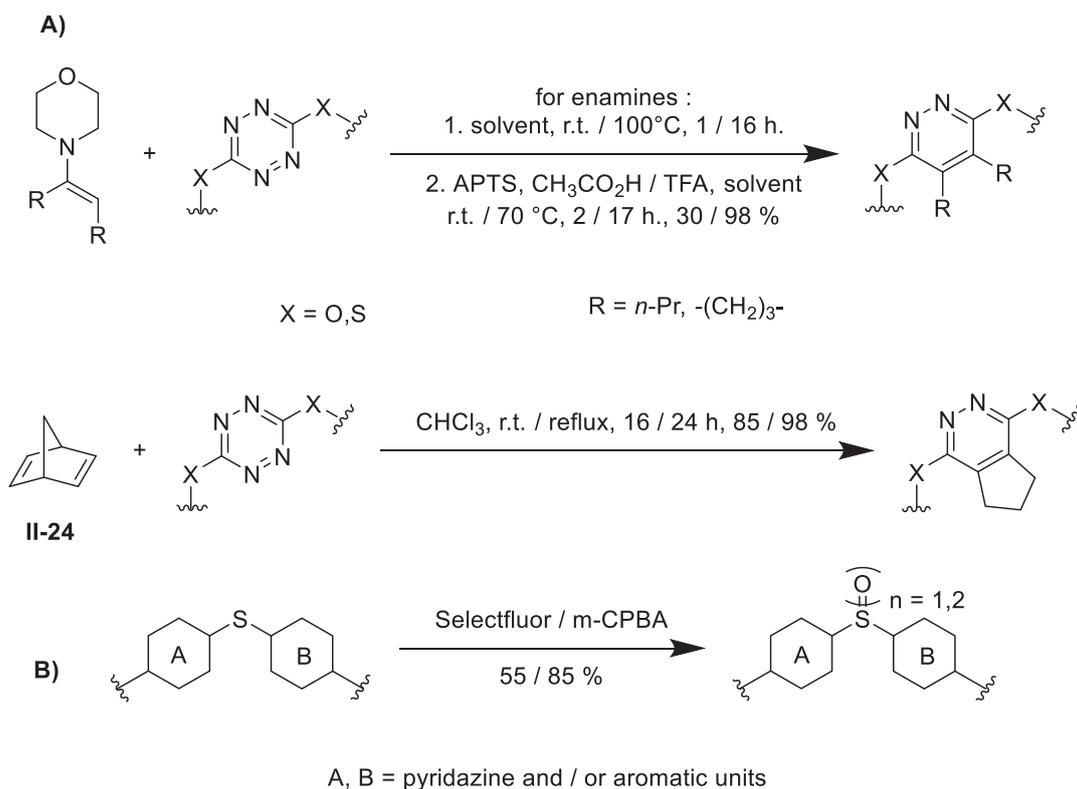


Scheme 149: Reported synthesis of corona[3]arene[3]tetrazines

In addition, symmetric and disymmetric corona[3]arene[2]tetrazines and corona[4]arene[3]tetrazines could be synthesised from a one (A) or two-step (B) procedure³²¹. The real potential of such backbones lies in the electronically deficient tetrazine moiety, which could undergo an inverse electronic Diels-Alder reaction (iEDDA) / rétro Diels-Alder sequence^{206,322} as a diene. This paved the way for a further post-functionalisation (*Scheme 151*), well documented in chemical biology³²³. Regarding the previously introduced macrocycles, few examples of [4+2] iEDDA with alkenes were reported^{320b-d,321c,324}. Electron rich enamines and strained **II-24** norbornadienes, employed as dienophiles, reacted with most of the compounds, leading to corona[*n*]arene[*n*]pyridazines. The oxidation of the sulfur bridges thanks to Selectfluor or *m*-CPBA gave access to sulfone and sulfoxide with good yields. This oxidation depleted the electronic density in the macrocycles^{320c} and increased the solubility in water.



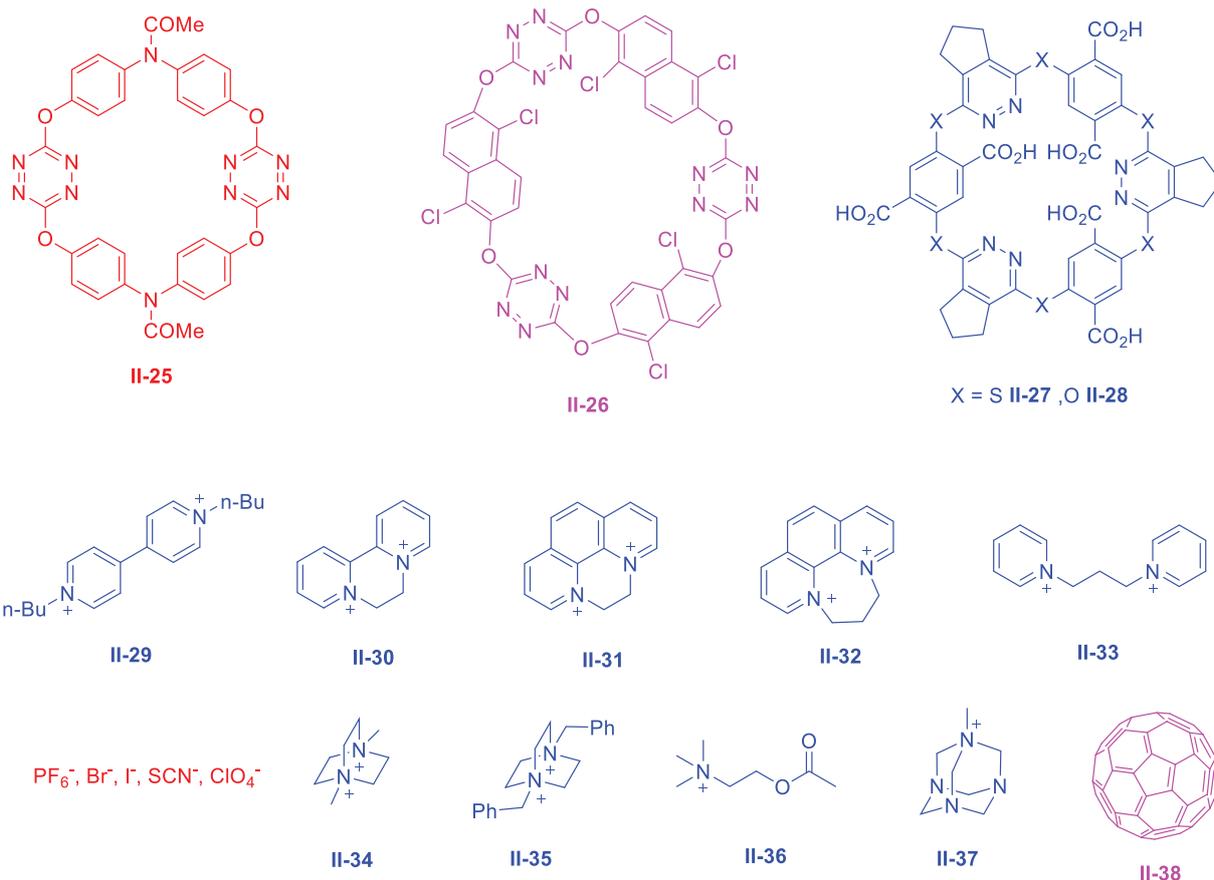
Scheme 150: Reported synthesis of dissymmetric corona[3]arene[2]tetrazines and corona[4]arene[3]tetrazines. A) one pot procedure, B) 2-steps procedure



Scheme 151: Reported post-conversion A) of corona[n]arene[n]pyridazines from corona[n]arene[n]tetrazines and further oxidation B)

Due to their variety of electronic properties, shapes and functionalisations, they were receptors for anions, cations, and neutral molecules. Indeed, corona[n]arene[n]tetrazines such as **II-25** were electron-poor cavities and formed complexes with anions (*Scheme 152*). These complexes had various geometries and shapes in organic media. They were mainly formed because of anion- π interactions and weak H-bonds^{320a,321a}. A specific recognition of fullerenes was reported for a O₆-corona[3]arene[3]pyridazines **II-26**^{320d}. **II-38** C₇₀ bound with **II-26** with an association constant K_a of $3.98 \cdot 10^4 \text{ M}^{-1}$ in toluene (fluorescence titration), while C₆₀ did not interact. The authors suggested an electronic rather than a steric effect to explain the selectivity. In contrast, corona[n]arene[n]pyridazines were electron-rich cavities and hosts for cationic guests such as ammoniums. They were receptors of cations without further modifications in a mixture of acetonitrile and dichloromethane with various ammonium ions (K_a in the range of 10^{2-5} , RMN, ITC, UV titrations)^{320b,c,324-325}. Terephthalate esters could be hydrolysed by lithium hydroxide^{320b,c,324-325}. Hence, the corona[n]arene[n]pyridazines **II-27/28** became water-soluble and bound with cations **II-29-37** with K_a in the range of 10^{3-5} M^{-1} (RMN, UV titrations) (*Scheme 152*).

All things considered, this family was efficient receptors of anions, and of cations in water. Even though most of the ammoniums used were aromatic ones, these justified the investigation of these new objects. The iEDDA / rétro Diels-Alder was an interesting entry point for functionalisation.



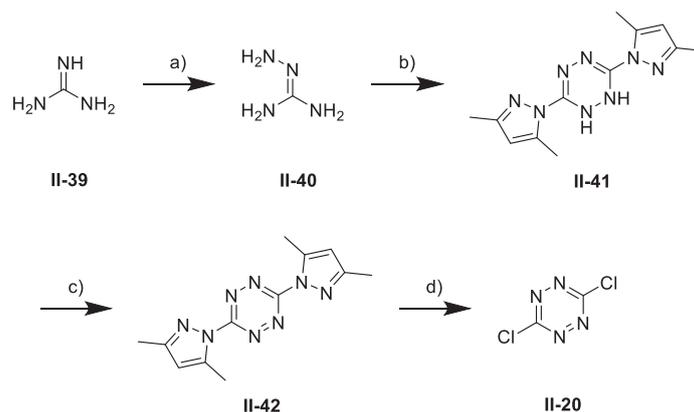
Scheme 152: Overview of the binding ability of corona[n]arenes. Red: corona[3]arene[3]tetrazines for inorganic anions; Purple: corona[3]arene[3]pyridazines II-26 binding with C₇₀ II-38; Blue: Water-soluble corona[3]arene[3]pyridazines II-27/28 for cations II-29-37 with examples of ammonium guests

Corona[n]arenes constitute a recent family of macrocycles. They perform host-guest chemistry in water and are suitable for functionalisation.

b. Synthesis of corona[n]arenes

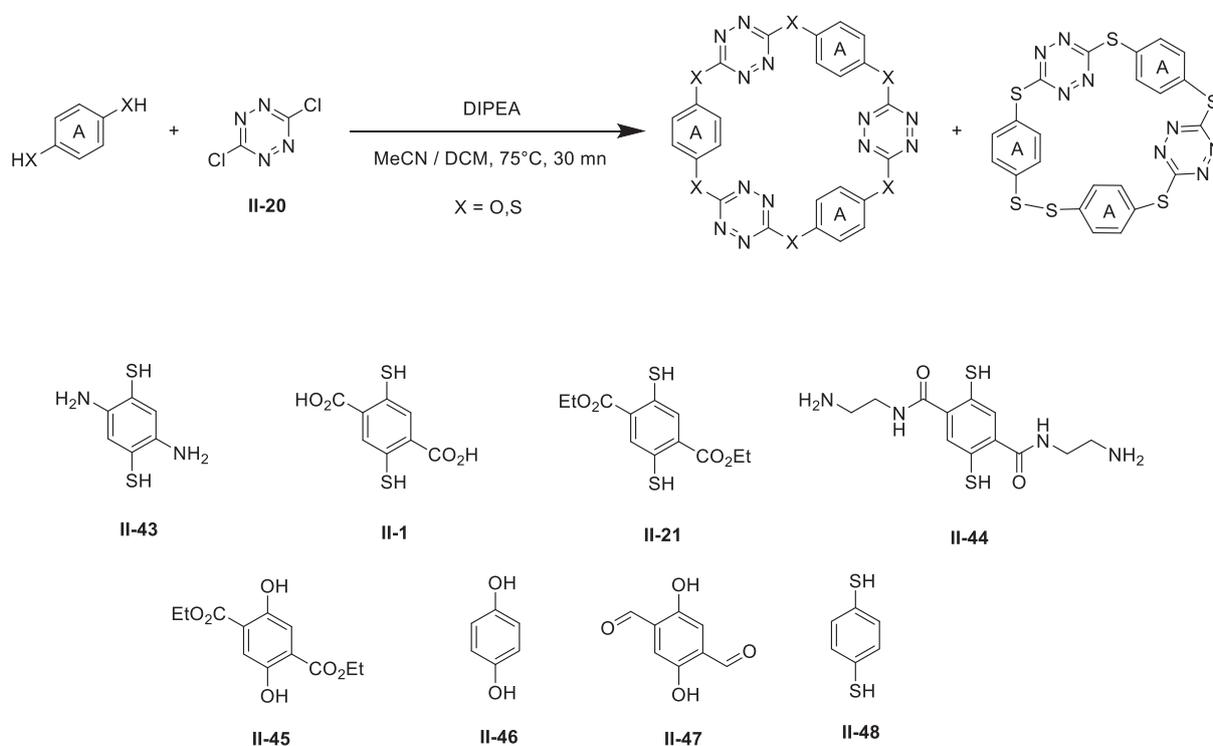
i. Synthesis of corona[n]arene[n]tetrazines

The synthesis of corona[n]arenes for the recognition of modified proteins was investigated. A scope of the aromatic units used in the S_{NAr} process and a variation on the partner of iEDDA was identified as the most relevant ways to generate new macrocycles. II-20 is commercially available, but a 3-step synthesis from II-39 guanidine hydrochloride²⁰⁷ was considered in order to ensure the supply (Scheme 153). II-39 was mixed with hydrazine giving II-40 with a quantitative yield. Then, II-41 was obtained with a yield of 80 % thanks to the reaction of 2,4-pentadione with II-40 in water. Oxidation gave the aromatic II-42 with a yield of 74 %. Finally, a one-pot procedure involving hydrazine and then TCCA gave the desired product II-20 with a yield of 29 %. The overall yield was 17 % in 3 steps, which was far less than the 78 % described. This poor yield was mainly due to additional purification. The authors mentioned that II-20 could be altered by acidic conditions. Purification on silica gel could be responsible. The molecules involved in that synthetic plan contained an important N/C atomic ratio, which is a source of explosion. In addition, the results obtained from the synthesised II-20 batch were different from the commercial one. Thus, the commercial batch was preferred for safety reasons and repeatability.



Scheme 153: Synthesis of II-20 a) Hydrazine, dioxane, reflux, 2h, 99 %; b) 2,4-pentadione, water, 70°C, 16 h, 80 %; NaNO₂, AcOH, DCM, r.t., 4 h, 74 %; d) 1. Hydrazine, MeCN, r.t., 20 mn, 2. TCCA, MeCN, r.t., 30 mn, 29 %

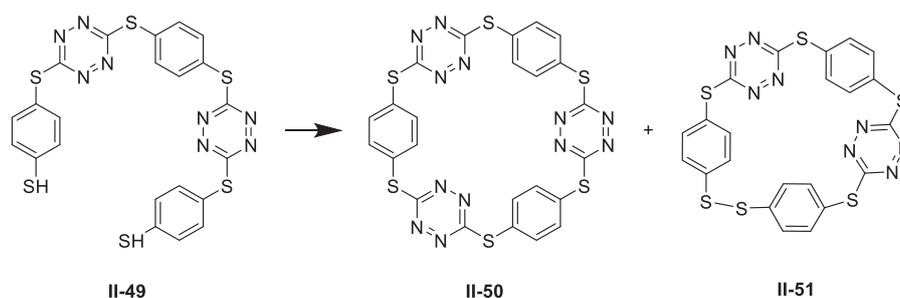
Various 1,4-bisthiophenols were engaged with **II-20**. The reported units were used^{320a-c,324}, as well as other backbones. These different structures involved chemical functions such as amines, aldehydes and carboxylic acids. The aim was to explore the chemical tolerance of the reaction and the direct access to corona[n]arenes[n]tetrazines with various electronic properties. The heteroatoms bridging the arenes were mainly sulfur but also oxygen. The reaction conditions of the scope were the following one: the two chemicals were separately added via a funnel at the same rate over 30 mn in a mixture of degassed solvents, usually MeCN / DCM. DIPEA was present, acting as an acidic scavenger and nucleophilic activator. The mixture was maintained at 75 °C during the process. This protocol enabled low working concentration, which should prevent the polymerisation and favour small and cyclic oligomers. This was important because the kinetic of substitution with tetrazine was known as fast³²⁶. The first targeted corona[n]arenes were the reported corona[3]arene[3]tetrazines^{320a-c,324}.



Scheme 154: Synthesis of corona[3]arene[3]tetrazines

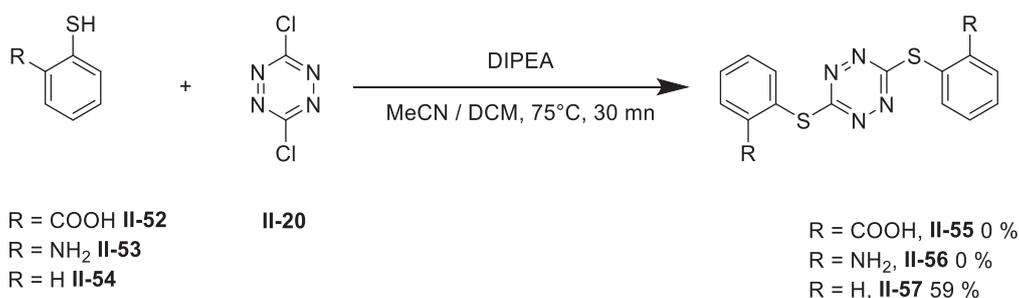
The replication of the results was difficult (*Scheme 154*). Indeed, the S_{Nar} with **II-46** hydroquinone gave a yield of 1 %, and the reaction between **II-20** and **II-21** as well as **II-45** did not give the desired macrocycles, despite reported yields of 24 %^{320a}, 58 %^{320b} and 65 %^{320a} respectively. The authors highlighted that the yield for the reaction with hydroquinone could be improved up to 34 % under argon. However, working in inert conditions had no impact on the yields. On the other hand, the reaction with **II-48** 1,4-dithiophenol gave the corresponding **II-50** with a yield of 29 % (*Scheme 155*) (reported 45 %)^{320c}. Several modifications of the protocol (temperature, addition time, order of addition) were not successful for increasing the yield. The solvents used in the literature were at UPLC grade and further distilled, while anhydrous solvents from Sigma-Aldrich were used for this study. At first glance, the presence of water was not an issue as **II-20** was soluble in water and remained stable over 1 hour at room temperature (UV, ¹H NMR). Most of the loss were probably due to the chromatographic separation. Indeed, the authors insisted on the fact that the tetrazine unit was not stable in acidic conditions such as silica gel. Furthermore, a lot of different species could be formed over the addition. Therefore, the convergence in one macromolecule was not trivial. Polymers could be easily formed, which must be the dominant path and could also explain the poor or the absence of yield. In order to enlarge the scope, **II-47**, **II-1**, and **II-44** were tested as partners. Unfortunately, the corresponding corona[n]arene[n]tetrazines could not be detected (ESI-MS analysis, ¹H NMR of the crude and isolated species) in all the cases. No identification could be attributed, suggesting a formation of polymers. **II-43** was known to easily aggregate in solution^{40a} and precipitation occurred immediately upon addition. The precipitate remained insoluble and was not further analysed. Regarding **II-1**, the poly-functionality and the protonation state could be a problem for the nucleophilic substitution. The literature did not contain examples of similar S_{Nar} with unprotected carboxylic acids.

The authors proposed a pentameric intermediate **II-49** (*Scheme 155*) during the S_{Nar} . The two terminal aromatic units could be then capped by a **II-20** molecule. Alternatively, an intramolecular oxidative coupling of the two thiols into a disulfide bond could give **II-51**^{320c}. It was reported as a marginal process. Thus, undegazed solvents were used.



Scheme 155: Pentameric intermediate involved in the formation of II-50 and II-51

The literature reported a maximal yield of 3 % for such S_6 -corona[3]arene[2]tetrazines, but we obtained 49 % with dithiophenol. The S_6 -corona[3]arene[3]tetrazines were difficult to synthesise with a poor repeatability. Hence, a simpler version of the reaction involving similar mono-thiols was examined (*Scheme 156*). With the same conditions, the product of the reaction **II-57** between **II-20** and **II-54** was synthesised with a satisfying yield of 59 %. However, in the cases of **II-52** salicylic acid and **II-53** amino-thiophenol, the substituted tetrazine **II-55** and **II-56** were not obtained, similarly to the scope with aromatic dithiols.



Scheme 156: Synthesis of substituted tetrazine

Finally, **II-20** was mixed with **II-4₄** and DTT. The idea was to cleave a disulfide bond of **II-4₄** with DTT, enabling the S_{Nar} and the formation of a derivative of **II-4₄**. Yet, despite different conditions, the S_{Nar} did not take place (HPLC and ESI-MS analysis control). Only the expected opening of the macrocycle was detected.

In a nutshell, the synthesis of corona[*n*]arene[*n*]tetrazines was less successful than expected compared to the literature. Indeed, only 2 corona[*n*]arene[*n*]tetrazines were synthesised with an acceptable yield, without chemical functionalities. The poor reproduction of the reported examples remained unsolved (one of the compounds was recently commercialised²⁸). The scope of aromatic units was not enlarged by this study. The sensibility to acidic media for the tetrazine moiety and the hypothetical quick formation of polymeric compounds were incriminated. Access to a disulfide S₆-corona[3]arene[2]tetrazines was also optimised.

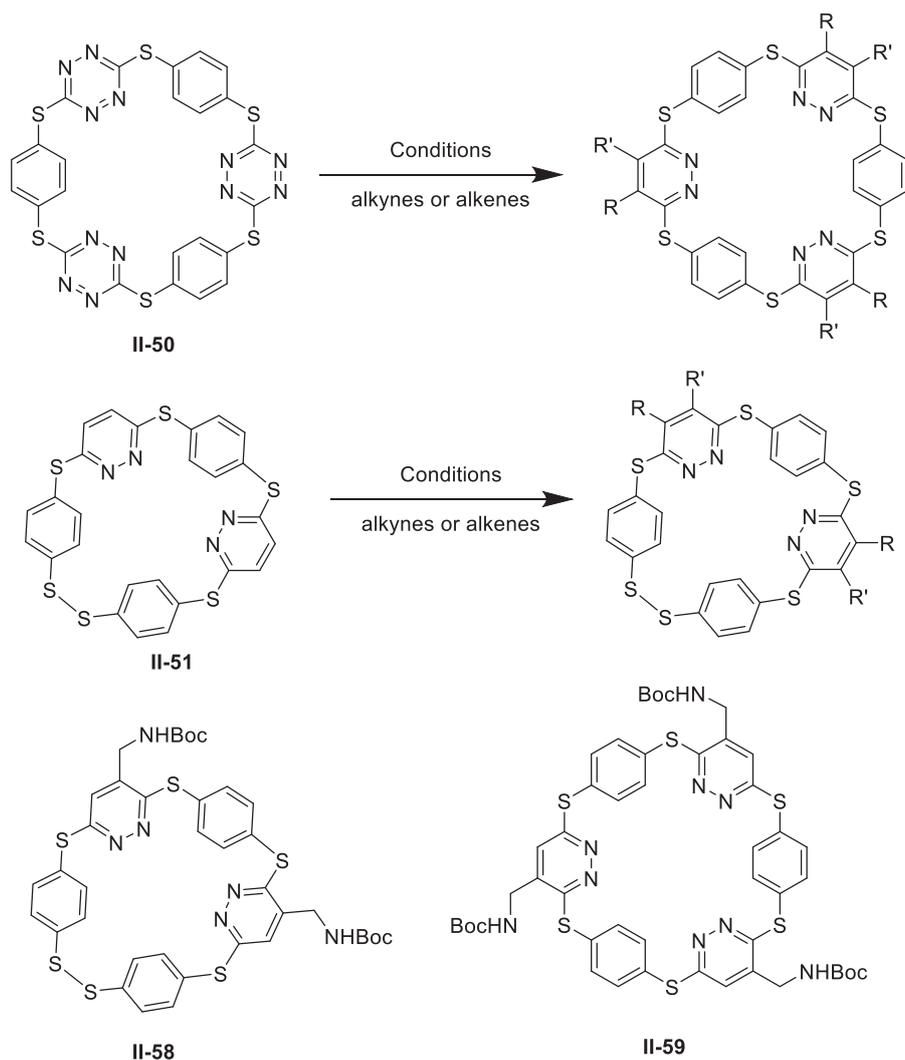
*The corona[*n*]arene[*n*]tetrazines synthesis was poorly reproducible. Only a simple corona[*n*]arene was obtained. This reduced the possibility of functionalisation.*

ii. Synthesis of corona[*n*]arene[*n*]pyridazines

Then, the formation of corona[*n*]arene[*n*]pyridazines from corona[*n*]arene[*n*]tetrazines was considered. The transformation occurred by an iEDDA with a dienophile, followed by a retro-Diels-Alder with a release of N₂. Finally, a rearomatisation took place spontaneously with heat or naturally present oxygen. The intention here was to develop a robust methodology with amino protected dienophiles in order to access precursors of cationic macrocycles. Enamines were not investigated because their synthesis seemed to require more time than amino protected alkenes and alkynes. Different dienophiles and conditions were screened, including the reported one^{320c} (Scheme 157, Scheme 158).

Firstly, attention was turned on a simple tert-butylallylcarbamate **II-60**. The sequence was conducted in refluxed chloroform because the molecules were electronically similar to norbornadiene. The two desired **II-58** and **II-59** were observed by HRMS analysis but could not be isolated. A change of colour was progressively observed over 24 h (red, tetrazine moiety, to orange, pyridazine moiety) and indicated the end of the sequence. The reaction was followed by ESI-MS analysis and also by ¹H NMR.

28 <https://www.jk-scientific.com/Product/ProductDetails/12802?categoryid=1109> 02/07/19, 100 mg 479 €



Scheme 157: Sequence iEDDA / rétro Diels-Alder for 1,4-dithiophenol corona[n]arene[n]tetrazines and a resulting S_5 -corona[3]arene[2]pyridazines (R, R' positions could be inverted on the products)

Entry	Diene	Dienophile	Conditions
1	II-60	II-50	CHCl ₃ , reflux, 24
2		II-51	Solvent, T, t
3	II-61	II-50	CHCl ₃ , reflux, 24 h
4		II-51	
5	II-62	II-50	DMF, microwave, 160°C, 30 mn
6		II-51	
7			
8	II-63	II-20	CHCl ₃ , reflux, 24 h

Scheme 158 : Scope for $S_{5/6}$ -corona[3]arene[2/3]pyridazines from $S_{5/6}$ -corona[3]arene[2/3]tetrazines

The crude mixture corresponding to the entry 1 contained various species, including partial addition / retro-addition as well as rearomatised and non-aromatic macrocycles, in several conditions of reaction time and temperature. An optimisation on the solvent (DMF, toluene, chloroform) and the temperature (90, 100, 140°C) was conducted for the entry 2. Chloroform at 90°C for 3 days was the most promising condition, leading to a major II-59 in ¹H NMR and ESI-MS analysis., without efficient

purification. Some rules reported the kinetic tendency for the reactivity of dienophiles and tetrazines^{323a}. The dienophile seemed to impose the poor reactivity. A metathesis with Grubbs-I^{322d,e,i} gave the dimer **II-61** of **II-60** with a correct yield of 44 %. Although bulkier, **II-61** was less complex as no regioisomer could be formed. Unfortunately, none of the product or intermediate of the sequence was observed (entries 3 and 4). Quickly, a black precipitate was formed. Carbamate protective groups were known to degrade at high temperature³²⁷. Diels-Alder reactions with **II-62** propargyl amine were known^{322h,n} and were investigated with corona[*n*]arene[*n*]tetrazines. On both **II-50** and **II-51** (entry 5 and 6), microwave activation did not give the corresponding product. ESI-MS analysis and ¹H NMR did not solve the side product structure. Once again, the poor reactivity of the alkynes seemed to be responsible for the thermal degradation. An addition of a Lewis acid^{323b,328}, BF₃·(OEt)₂, did not change the composition of the crude. Regarding the smaller substituted tetrazine **II-20** (entry 7), the conditions did not led to the identification of any product of the sequence. **II-62** was protected in **II-63** with a Boc protecting group^{322a}, with a yield of 97 %. Unfortunately, with **II-63** (entry 8), no desired products were detected by ESI-MS analysis.

To sum up, the formation of corona[*n*]arene[*n*]pyridazines from corona[*n*]arene[*n*]tetrazines was unsuccessful. Without repeating the reported examples^{320c}, the use of amino-protected dienophiles failed to access new corona[*n*]arene[*n*]pyridazines. The synthesis of specific enamines could be a solution, as the sequence iEDDA / retro Diels-Alder seemed easier for enamines^{320b-d,321c,324}. As the two synthetic steps to corona[*n*]arene[*n*]pyridazines were difficult to control and to expand to other molecules, a DFT study was developed. The goal of this study was to gather information on the two steps.

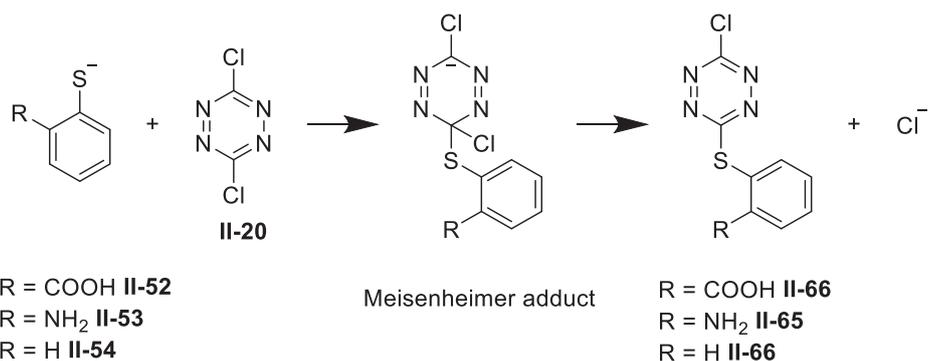
*The formation of corona[*n*]arene[*n*]pyridazines bearing chemical functions that are precursors of cation in solution encountered synthetic issues.*

c. Theoretical investigation of the corona[*n*]arenes synthesis

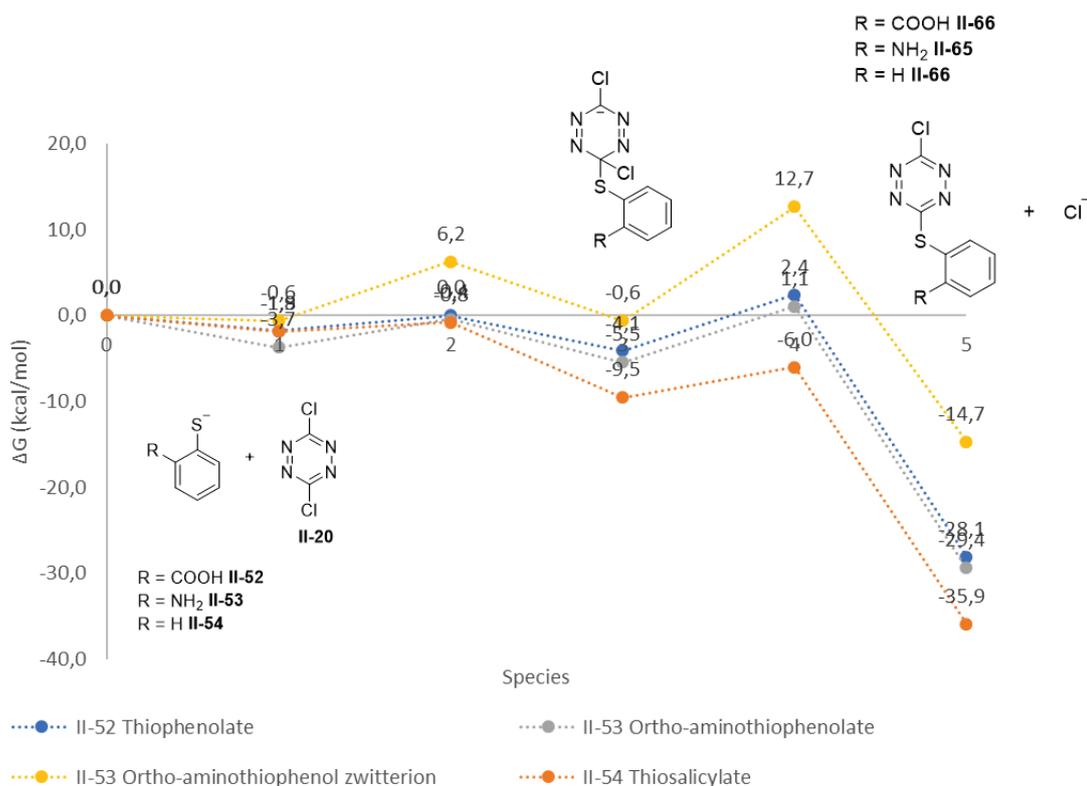
The computational details are available in the experimental section. Computational resources were provided by Lionel Perrin (ITEMM, ICBMS, Lyon 1 University). Firstly, the S_NAr was modelled.

i. Modelling of the S_NAr reaction

A DFT modelling was employed to understand the influence of electronic substituents and the protonation state on the reaction (Scheme 159, Scheme 160). Indeed, the two steps of the S_NAr mechanism could involve a tetrameric intermediate named Meisenheimer complex. The rate determining step was commonly accepted to be the formation of the complex. However, recent studies showed that the Meisenheimer complex could be stabilised and easily formed³²⁹, changing the paradigm on this mechanism. Thus, following the electronic properties of the nucleophile, the determining rate step could change. In our case, it was the elimination of the chloride. Regarding mono aromatic thiols, **II-52** had the lowest transition state in energy. Regarding **II-53** the zwitterionic form had a lowest transition state in energy than anionic **II-52**, suggesting that the protonation state influenced the energy.

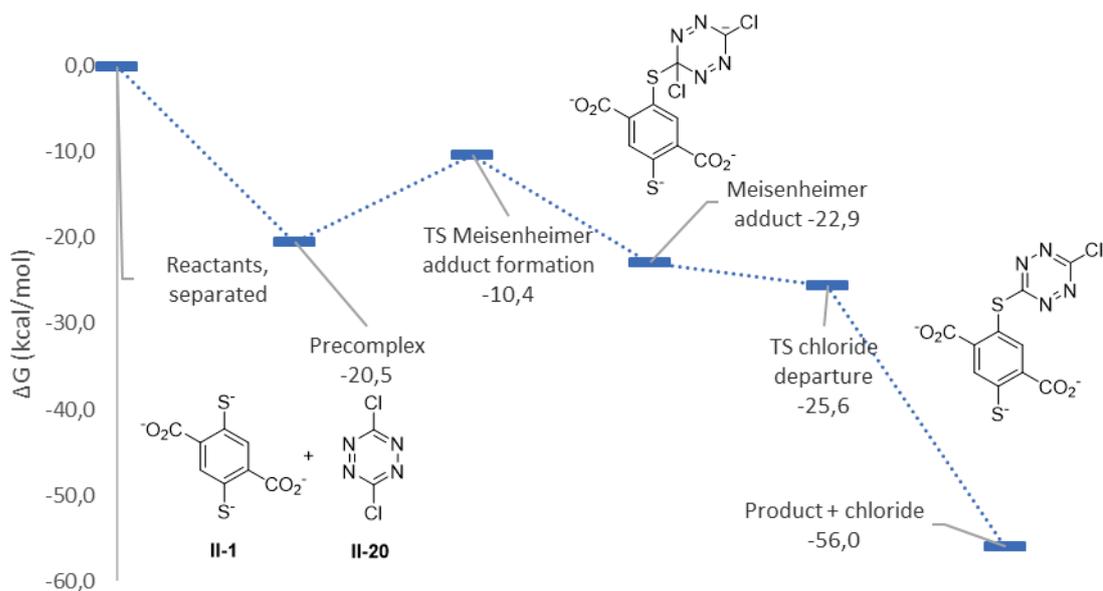


Scheme 159: Reaction modeled by DFT: S_{NAr} between **II-20** and aromatic thiolates **II-52-54** through a Meisenheimer adduct



Scheme 160: Computed theoretical data for the S_{NAr} of **II-2** with various aromatic thiolates. 0 **II-20** and thiolate. Computed free energies in kcal.mol⁻¹ vs reaction coordinates: 1 pre-complex, 2 TS Meisenheimer adduct formation, 3 Meisenheimer adduct, 4 TS chloride departure, 5 product and chloride

Nevertheless, modelling for **II-1** revealed an opposite trend (Scheme 161) since the mechanism seemed to process through a concerted mechanism. Indeed, the second transition state was found lower in energy than the Meisenheimer adduct. It was reported for some prototypical S_{NAr} ^{329b,c,330}. The energetic cost of the whole S_{NAr} was low and could explain a polymerisation, avoiding the formation of the desired corona[3]arene[3]pyridazines.



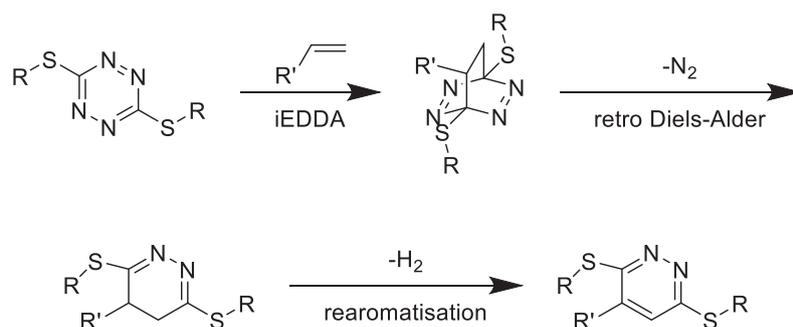
Scheme 161: Computed energetical data for the S_{NAr} between **II-20** and **II-4**. Computed free energies in kcal.mol⁻¹ vs reaction coordinates

To conclude, electronic effects on the substituent seemed to change significantly the energetical landscape of the S_{NAr} . The activation energy could be very low, explaining the fast kinetic of reaction. Hence, the polymerisation could be difficult to avoid. This was coherent with the experimental observation. In addition, this study revealed that the S_{NAr} could take place by two steps or by a concerted process. There is a current debate on whether the Meisenheimer complex is an intermediate or a transition state³³¹. The observation of the intermediate, thanks to ¹³C NMR, could provide an experimental support to our computational remarks. The impact of the substituent could be determined by isotopic kinetic study.

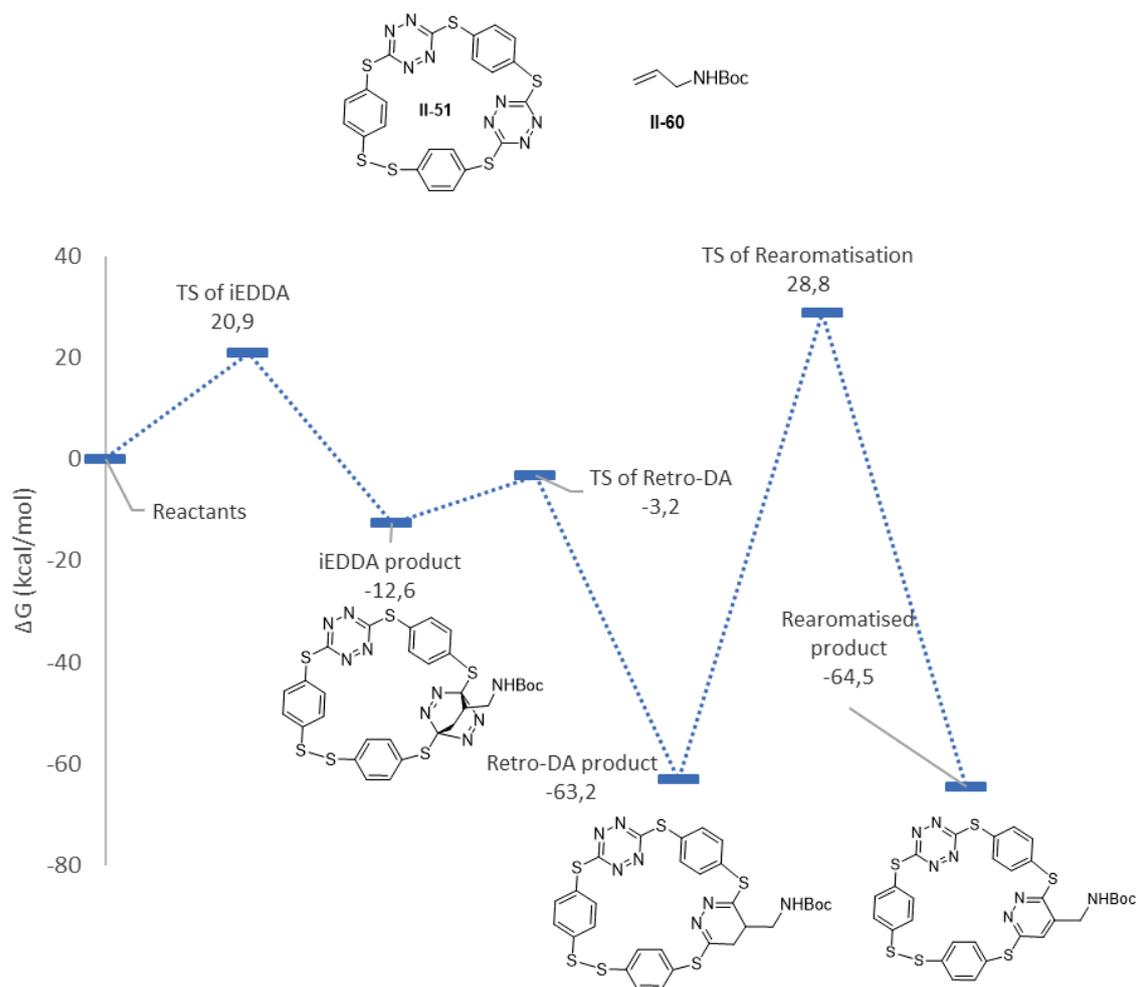
*DFT showed that the substituent and the protonation state of thiolate impacted the thermodynamic of the S_{NAr} with **II-20**.*

ii. Modelling of the iEDDA sequence

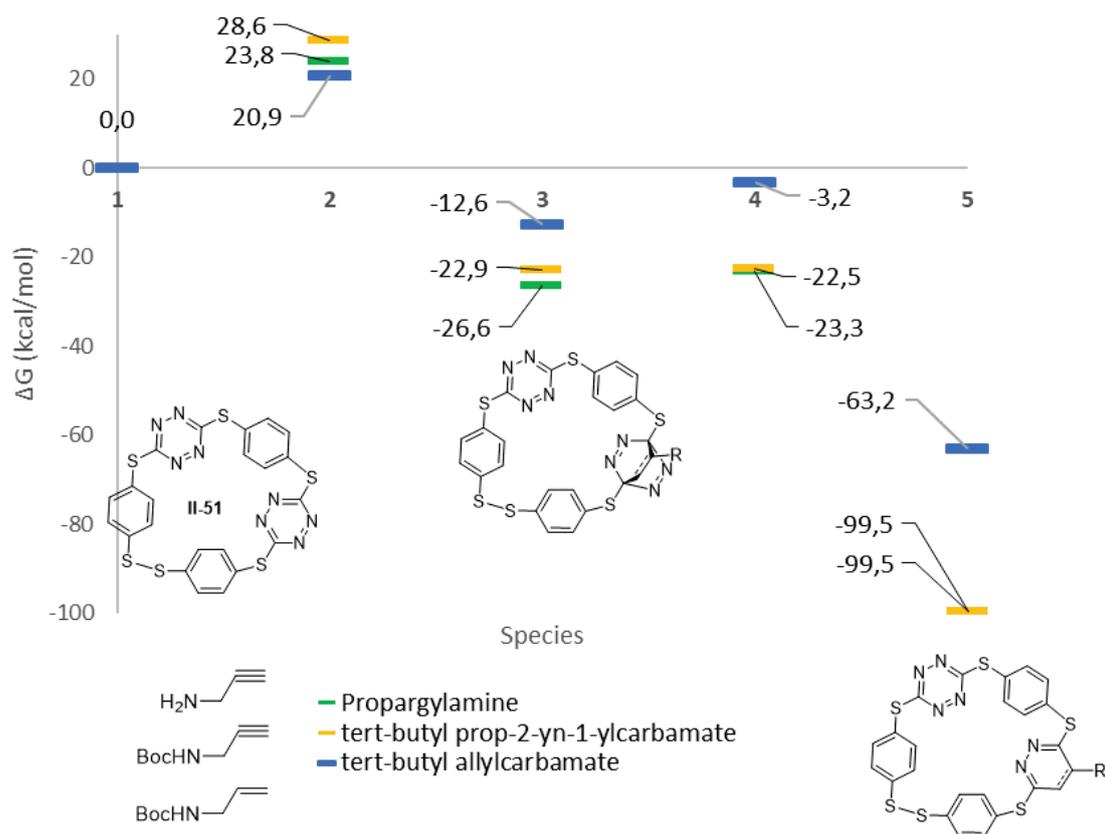
DFT predicts efficiently the reactivity of tetrazines in iEDDA^{322j,332}. Thus, it was used here to obtain further information on the process for our system. The sequence iEDDA / retro Diels-Alder between **II-51** and *tert*-butylallylcarbamate **II-60** was chosen as a model (Scheme 162, Scheme 163, Scheme 164). Indeed, this system was representative and included the macrocyclic conformation. Two tetrazine moieties were available for the sequence. Only one sequence of addition elimination was studied. The thermal rearomatisation was found to be the determinant step and showed a high energy of activation. The first step also required a non-negligible activation energy, which was coherent with the slow conversion of the starting material. It suggested to stop the reaction at the retro-Diels Alder product, and then to use an oxidative agent for instance. The alkynes had a slightly higher activation energy than the alkene for the iEDDA, while their respective products were lower in energy.



Scheme 162: From corona[n]arene[n]tetrazines to corona[n]arene[n]pyridazines, steps of the sequence



Scheme 163 : Energetical landscape of one sequence iEDDA/retro-DA between **II-51** and **II-60**. Computed free energies in kcal.mol⁻¹ vs reaction coordinates



Scheme 164: Energetical landscape of the sequence iEDDA / retro-DA for various partner with **II-51**. Computed free energies in kcal.mol⁻¹ vs reaction coordinates: 1. Reactants, 2. TS iEDDA, 3. iEDDA product, 4. Retro-DA TS, 5. Retro-DA product

The two DFT modellings provided information on the problematic reactions and were coherent with the experimental observations. The S_{NAr} was found to be dependent on the chemical group beared by the aromatic thiols. This reaction had a moderate or low energetic cost. That was coherent with the difficulty to obtain small oligomers. Besides, the sequence of iEDDA / rétro Diels-Alder was limited by the rearomatisation step, inducing harsh conditions that are not supported by the carbamate moiety. To complete the data obtained so far, a cristallographic structure of **II-51** was discussed.

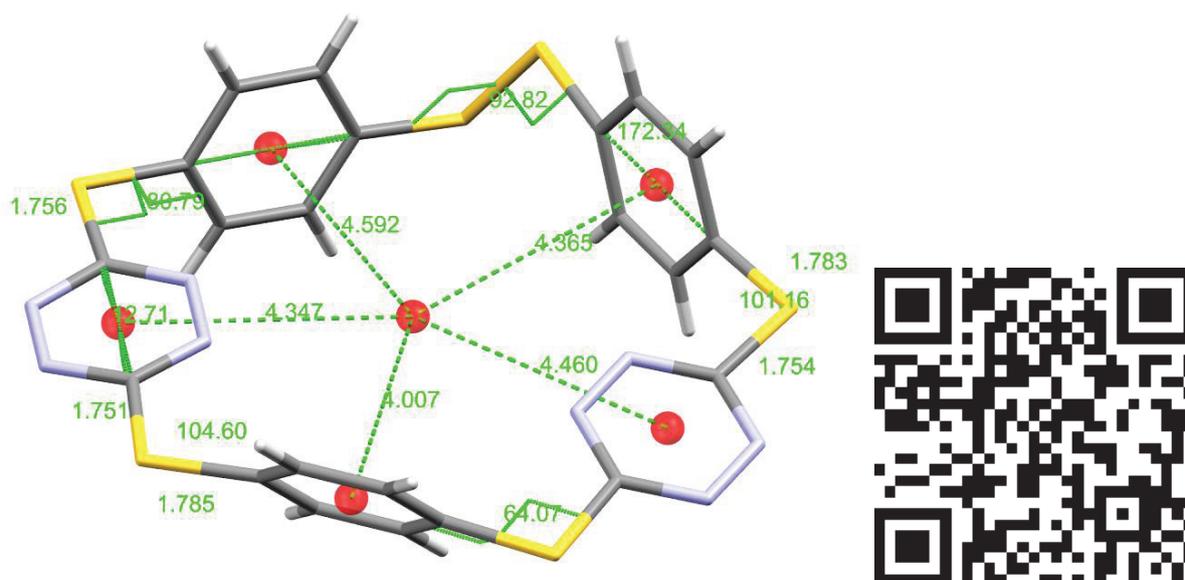
DFT indicated that the rearomatisation should be complete by a separate step and that the iEDDA required an important activation energy.

d. Crystallographic structure of **II-51**

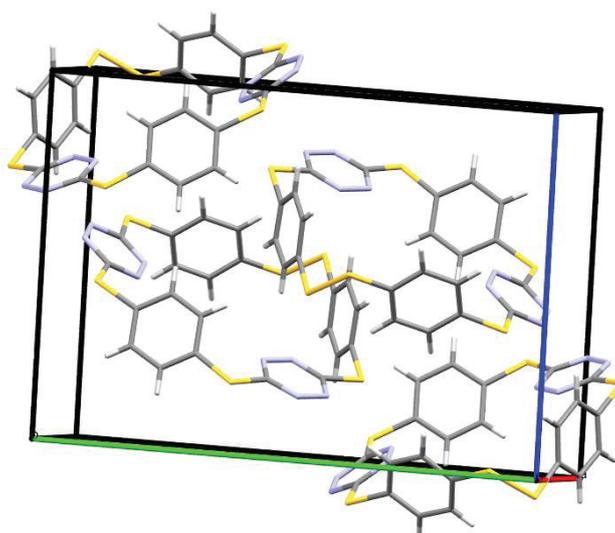
Macrocyclic conformation of **II-51** in the solid state was determined by X-ray²⁹. **II-51** co-crystallised with a dichloromethane molecule (Scheme 165). Firstly, the two tetrazine aromatic units almost formed a plan, and another plan was almost formed by the three benzene units. The dihedral angles between these plans were 78.8° and 64.7°. It was remarkable to observe a macrocycle of this size bearing two different aromatic parts, forming two plans. The angles between the aromatics were 80.8°, 104.6°, 101.2° and 101.2°. The last three angles were similar with other thioethers³³³ and with similar macrocycle^{320c,321}. The first one was slightly lower and was the sign of the constraint induced by a 5-membered ring, bearing two electronically different aromatics. This feature was not reported for

29 Structure resolved by Dr. Erwann Jeanneau, ISA, Lyon. Available details in appendices.

corona[*n*]arenes. Indeed, crystallographic structures of undesired disulfide products were never described. Furthermore, the bond lengths of S-C_{tetrazine} are slightly shorter than S-C_{benzene}, typically with 1.751 and 1.785 Å. It suggested that the sulfur bridging atoms formed a better conjugation with the tetrazine ring, which was electronically deficient. The cavity size was defined by the average distance between the centroid of the benzene units. Thus, it was 7.275 Å. The disulfide dihedral angle was 92.8°, which was nearly ideal (90°), and comparable with similar cyclophanes^{114b}. **II-51** crystallised in P21/c space group. The main interactions in the packing unit were not clear. Dispersive Van der Waals interactions might be responsible for the arrangement (*Scheme 166*). Finally, the conformation of the disulfide bond was *M* or alternatively *P* in the packing unit. There was no *pR pS* configuration because there was no substituent on the aromatic units.



Scheme 165: X-ray structure for II-51 with distances (Å), angles (°), and torsions(°) of interest (centroid of the aromatics in red, in the middle centroid of the aromatic centroids), snapshot from Mercury software¹³²



Scheme 166: Packing unit for II-51, snapshot from Mercury software¹³²

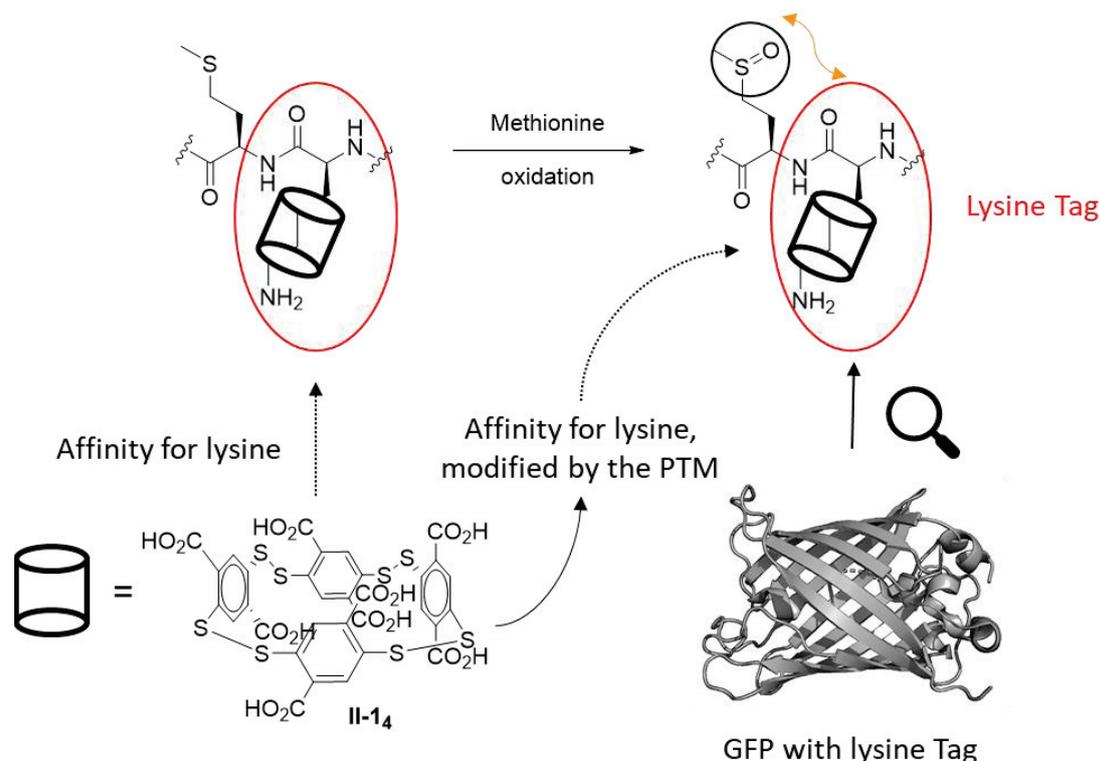
Crystallographic structure of **II-51** was the first of this class of compounds and revealed a shape with two plans.

e. Conclusion

Corona[n]arenes appeared to be appropriate backbones to develop static receptors thanks to post functionalisation by the sequence iEDDA / retro Diels-Alder following the strategy **A** of cavitand design. Firstly, the synthesis of reported corona[n]arenes was difficult. The reported examples as well as others in order to expand the initial scope were not obtained. Only a simple S_6 -corona[3]arene[3]pyridazines gave a satisfying yield. An improved protocol for a new S_6 -corona[3]arene[2]pyridazines was proposed and a crystallographic structure was resolved. Amino protected alkenes were unsuccessfully screened in the sequence iEDDA / retro Diels-Alder. Even though the desired products were clearly identified by ESI-MS analysis for one alkene, their isolation failed. A DFT study gave some insights into the mechanism of the S_{Nar} and the sequence and was coherent with the experimental data. An optimisation of the synthesis could be done in a future work. Furthermore, the repetition of the reported examples should be managed before further tries to develop the scope. The family of the corona[n]arenes remained interesting for the host guest chemistry in water, upon the conditions of a variation on their chemical functions to introduce the functional groups dedicated to the recognition of modified proteins.

3. Introduction of a Lysine-tag on proteins and binding with **II-14**

The aim of this short subchapter was to assess the ability of **II-14** to bind by an inclusion complex with Lysine-tagged Green fluorescent proteins (GFPs). (Scheme 167).



Scheme 167 : The goal of this part was to evaluate the binding ability of **II-14** with Lysine tagged **II-14** has affinity for lysine and binds with lysine derivatives by inclusion complex, which will be influenced by the presence of a neighbouring oxidised methionine (orange). PDB of the GFP represented in cartoons (Pymol software¹⁹⁸): <http://www.rcsb.org/structure/4KW4>

The 1,4-dithiophenol-2,5-terephthalic acid **II-1** self-assembled into a homochiral cyclic tetramer **II-1₄** through DCC at 4 mM in Tris buffer (200 mM pH 7.4) under templating by spermine. This disulfide bridge cyclophane corresponded to the strategy of design **B**, because the monomer units were connected by reversible bond through DCC, even though the monomers did not bear side chains on their 2,5 positions of the aromatic core. The aptitude of **II-1₄** to selectively recognize Lysine residue in N-terminal position on peptides of increasing chain length was explored by a previous PhD student^{82a} : **II-1₄** showed sub millimolar constants of association for N-Lysine peptides. The affinity was modulated regarding the nature of the C-Terminal amino acids (range from 1.34×10^5 to 4.26×10^3 M⁻¹ determined by ITC). Interestingly, a 1.6 fold increase of selectivity was obtained for Lys-MetO ($K_a = 9.95 \times 10^4$ M⁻¹) compared to Lys-Met ($K_a = 6.38 \times 10^4$ M⁻¹). This selectivity for the oxidised product was explained by the formation of favourable sulfoxide-aromatic interactions thanks to quadrupole/ π interactions. In addition, lysine derivatives led to distinct induced circular dichroism (ECD) outputs in buffered water upon binding with the receptor, because of the asymmetric deformation of its cavity⁴¹. Therefore, it would be possible to detect residues that are affected by deamidation or methoxidation in the following amino acid is a lysine. Bradykinin, an octa-peptide, was modified with a Lysine tag in N-terminal positions. The binding of **II-1₄** by an inclusion complex with Bradykinin increased by 10 fold ($K_A = 1.18 \times 10^5$ M⁻¹) compared to the native octa-peptide, which was barely attenuated by the presence of an extra N-terminal methionine.

In the continuation of this work, fluorescent GFPs were used as models of Lysine tagged therapeutic proteins, for their stability and ease of modification³³⁴. The aim was to use **II-1₄** for the extraction of Lysine tagged therapeutic proteins. Given that the neighbouring amino acids of Lysine led to different induced circular dichroism, the presence of oxidised methionine would be detected by EDC and would be quantified. In this perspective, three GFP mutants were designed (*Table 15*) : 1) MKQ: mutant incorporating in N-terminal position lysine preceded by methionine, masking a poorly stable terminal lysine by a residue with poor effect on the binding with **II-1**. 2) MAA: a negative control tagged with alanine which did not interact with the receptor. 3) ΔK : similar negative control without the C-terminal modification.

Measures of the affinity between **II-1₄** and GFPs were realised³⁰. The proteins were prepared by recombinant technology, production in a fermenter and purification by IMAC column. The μ -FFE³³⁵, ITC and Octet® (based on the optical interference pattern of white light reflected from two surfaces - a layer of immobilized protein and an internal reference layer) were tested to investigate the interactions between the proteins and the macrocycle. A discrimination between the GFPs in favour of MKQ was expected.

GFP mutant	"MKQ"	"MAA"	" ΔK "
C-terminal	6 H, C-terminal		\emptyset
N-terminal	MKQ	MAA	
Interest	Lysine specificity positive control	Lysine specificity negative control	Observation of a potential second binding site

Table 15: Details on the GFP mutants

³⁰ Experiments done during a period at the CIP, University of Liège, Belgium

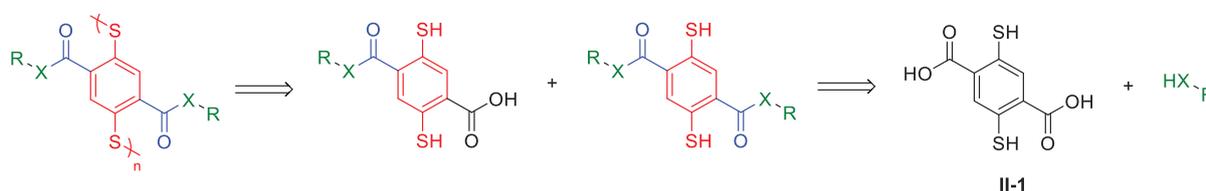
GFP mutant	“MKQ”	“MAA”	“ΔK”
N (1:1 model)	0.736 +/- 0.02	0.473 +/- 0.04	0.577 +/- 0.072
K (L.mol ⁻¹)	6.19 10 ³ +/- 578	1.04 10 ⁴ +/- 2250	1.80 10 ³ +/- 196
ΔH (cal.mol ⁻¹)	-1919 +/- 89	-1644 +/- 182	-3357 +/- 523
ΔS (cal mol ⁻¹ .K ⁻¹)	10.9	12.9	3.64

Table 16: ITC experiments between II-1₄ (5.6 mM) and GFPs (0.56 mM) in Tris 100 mM pH 8.0

The samples containing a GFP and **1-1₄** showed a spray behaviour, typical of a high conductivity in the media. That did not discriminate the 3 GFPs. The Octet® analysis gave the same results for MKQ and MAA and showed a change in the protein conformation resulting from the interaction with the receptor. Finally, the ITC experiments showed a possible submolar affinity for each GFP (Table 16). There was no discrimination between the GFPs. The thermodynamic values obtained from the titrations were not robust and were not replicated. In a nutshell, there was not discrimination regarding the binding of **II-1₄** with the three GFPs. So far, the use of a lysine tag increasing the affinity between the receptor and proteins was only effective on dipeptides and Bradykinin. The GFPs were complex proteins which could interact by many ways to **II-1₄**.

4. Introduction of side chains on a carboxylate aromatic 1,4-bisthiophenol

The aim of this subchapter was to develop a synthetic method to introduce, two side chains bearing H-bond donating groups on a 1,4-dithiophenol backbone. In this perspective and given the experience on the group on 2,5-dicarboxy-1,4-dithiophenol **II-1**, this was targeted through amide coupling on position 2 and 5 and referred to the strategy **B1** (Scheme 168). In addition, the synthesis of **II-1** can be performed in high yields at multi gram scale, which validates that choice of starting material.



Scheme 168 : Retrosynthetic analysis of tailored monomers for the strategy **B1** based on **II-1**. On the 2,5 positions, an amide formation will connect the side chain (blue) to the principal chain (red). X: NH, O. R: functional groups (green) for the recognition of modified proteins, such as alcohols, amines, especially primary amines, and guanidines for instance. Cyclo-oligomerization through disulfide bridge formation is intended to be the last synthetic step.

The targeted monomers will present two functional groups on their side chains bearing H-bond donating end group but can also contain only one functional groups and a free carboxylate moiety. Before the amide coupling, thiolphenol moieties should be first protected. Trityl and cyanopropyl protecting groups were investigated. The two next subparts discuss the synthesis of symmetric monomers such as **II-69**.

a. Protection of thiols by the trityl group

This subpart presents the connection of side chains to the aromatic core of **II-1** following a strategy of trityl protection (Scheme 169). This protection of thiols is well documented³³⁶. Among the available protecting groups, Trityl, or triphenylmethyl is an example of a protection by a thioether moiety³³⁷.

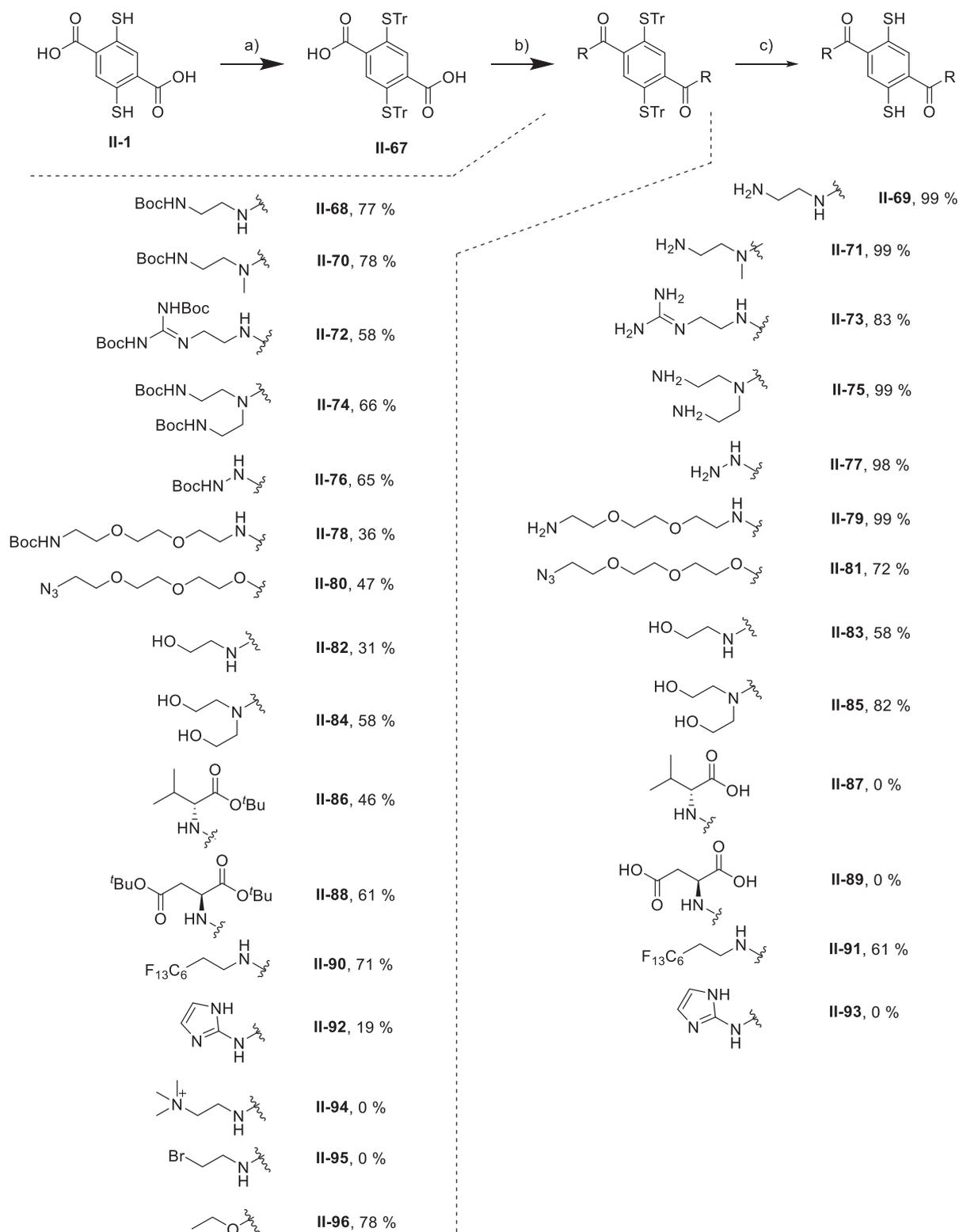
The protecting group is installed by a nucleophilic attack of the thiol on trityl chloride or on trityl cation in the case of trityl alcohol and an acidic assistance^{79e}. The deprotection occurs under mild acidic conditions (r.t., 1 h.). In such conditions a carbenium ion is generated and trapped by a silane^{79b,338}. The strategy was inspired by a similar work of protection on aromatic thiols^{79e}.

Firstly, the two thiol moieties of **II-1** were protected by a nucleophilic substitution with trityl chloride under basic conditions with an excellent yield of 99 %, comparable with a similar aromatic bithiophenol³³⁹. No chromatographic purification was required. Then, the formation of the amide bond was developed starting from **II-67**^{79e,340}. After completion of the reaction and purification, a deprotection of the thiols involving triethylsilane, a cationic scavenger and TFA was followed^{39b,340}. Various primary and secondary polyamines were used for the amide bond formation. Before, the amines were protected by a Boc group, which can be cleaved in acidic deprotection³⁴¹, similarly to the trityl protecting group. When not commercially available or too expensive, protected amines were synthesised with excellent yields compared to the literature. The details are given in appendices (Appendix II-II: Reaction conditions).

As previously discussed, the side chain of the cavitands will contain functional groups such as amines, especially guanidine, imidazole and primary amines. Consequently, various protected amines were coupled on scaffold to yield precursors of cavitands devoted to the recognition of electron-rich peptide sequences. Regarding an ethylene diamine derivative **II-63**, the amide bond formation on the positions 2,5 of the aromatic scaffold gave a satisfying yield of 77 % (**II-68**). It was followed by an effective deprotection of both trityl and Boc protection groups with a yield of 99 % (**II-69**). The scope of the side chains containing amines was then enlarged. The goal was to dispose of different side chains regarding their length and charge. A guanidine derivative **II-73** was considered to provide a delocalized positive charge and a specific H-bond pattern. The DETA derivative **II-65** was synthesised to increase the number of cationic charges borne by the side chain. A pegylated amine **II-79** was envisaged to longer chain. Another methylated derivative of ethylenediamine **II-71**, was introduced to overcome some potential cross-reactivity between the nitrogen of the amide bond and the thiophenol moiety, developed in a next part.

Phenols were previously identified as functional groups of interest that is why simple primary alcohols were introduced in the scope of side chains. The corresponding 1,4-bithiophenols will also be a reference in the cross-reactivity previously evoked, regarding the influence of the charge (neutral alcohols vs cationic amines) on the process. 2-aminoethan-1-ol and 2-aminopropane-1,3-diol, N-protected with a Boc group, were efficiently introduced and deprotected with the trityl group into the molecules **II-83** and **II-85**. A 1,4-bithiophenol **II-81** contained a reactive end group azide which may either be coupled to a probe or a solid support by copper-catalysed click chemistry. The starting azide **II-97** was obtained by a nucleophilic substitution of 2-(2-(2-chloroethoxy)ethoxy)ethan-1-ol with sodium azide³⁴² with a yield of 98 %. A 1,4-bithiophenol **II-77** bearing hydrazide function was synthesised in order to form acylhydrazones with carbonyl compounds following the strategy **B1**, **B2** and **C** of cavitand design. Nevertheless, an unexpected attack of the free hydrazide by its terminal NH₂ groups on the *tert*-butyl cation was observed in ESI-MS analysis and ¹H NMR, despite the use of trimethylsilane as scavenger of cations. The duration of the deprotection of the Boc protection group was hence limited to 15 mn in order to reduce the extent of this unwanted phenomenon. The molecule **II-96** would be useful for a further discussion. Finally, a 1,4-bithiophenol **II-91** bearing a lipophile fluorinated aliphatic chain was synthesised. The corresponding cyclo-oligomers could therefore be

amphiphilic and potentially used as detergents for the extraction and stabilisation of membrane proteins such as GPCR³⁴³.



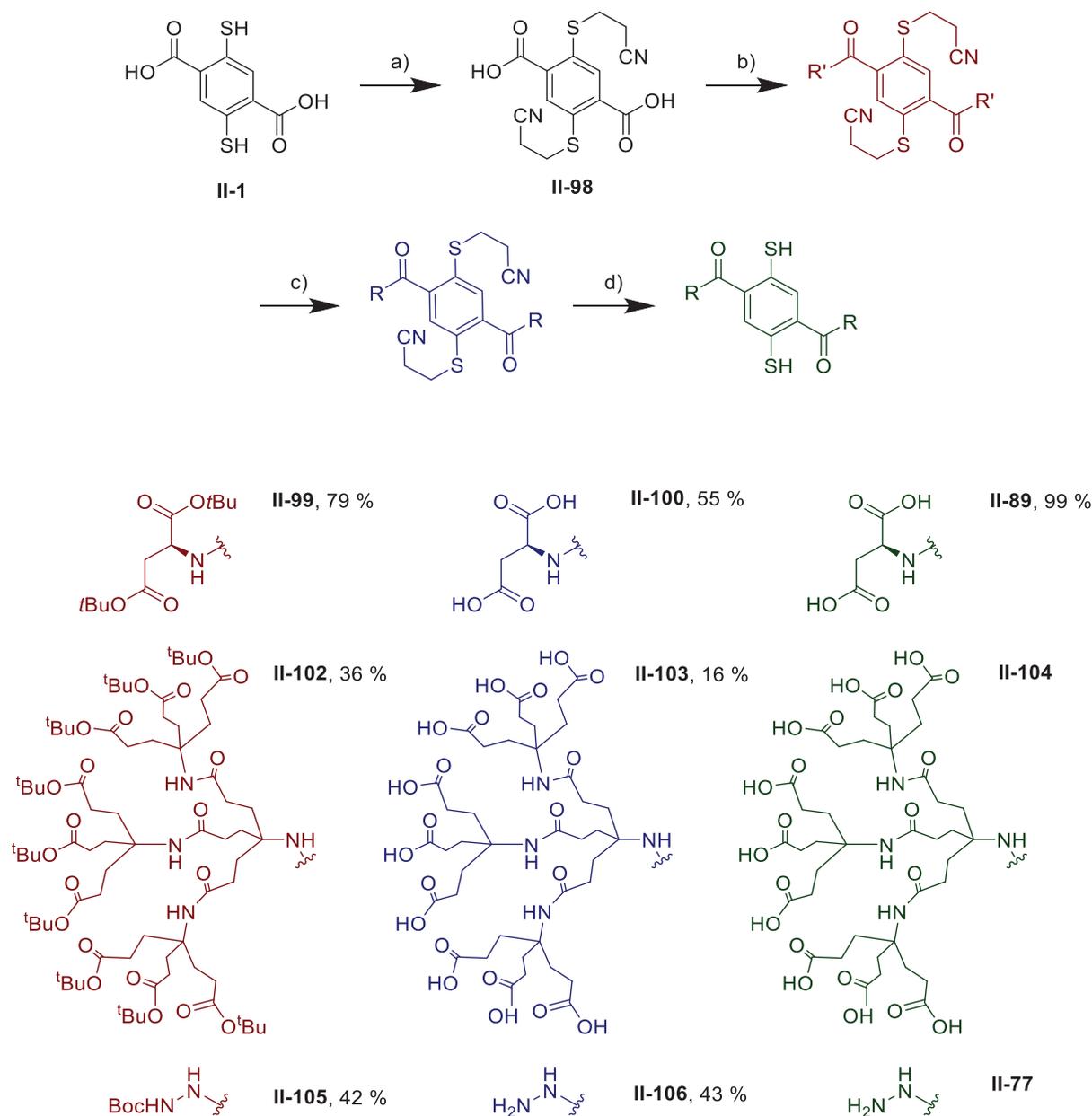
Scheme 169 : Synthesis of amide derivatives of building block II-1 involving protection of the thiol moiety with trityl group. a) Tritylchloride, DMF, r.t., 48 h., 97 %; b) HOBt, EDC, DCM, r.t., 48 h; c) TFA, Et₃SiH, DCM, r.t., 3 h. Scope for II-1 pre-functionalisation with trityl protection.

The yields of the amide bond formation were in the range of [31 – 78] %, which was coherent with the literature^{79e,340}. The loss of matter was due to a partial conversion on a product with only one amide bond and by-products of the activation cascade (attested by ESI-MS analysis). The yields of the second step lied in the range of [58 – 99] %, which was coherent with the literature^{97a,98}. In this case, the conversions were total, but a part of the product could be lost in the purification process. The purification was most of the time unnecessary. No trend could be drawn between the nature of the structure and the overall yield. 1,4-bisthiophenols, protected with a trityl group, bearing quaternary ammonium (**II-94**) and bromide (**II-95**) could not be synthesised. The nucleophilic attack of the thiols on the C_{Br} could compete the reaction. In addition, a 1,4-bisthiophenol, protected with a trityl group, bearing an imidazole (**II-92**) but the free thiols could not be obtained. **II-1** was detected in the crude mixture of the thiol deprotection by ESI-MS analysis, suggesting a degradation of the imidazole moiety. Eventually, 1,4-bisthiophenols, protected with a trityl group, bearing derivatives of valine and aspartic acid were synthesised (**II-86**, **II-88**). However, the cleavage of both esters and trityl groups failed. In fact, for both compounds, the hydrolysis of esters required a long time of reaction, superior to 24 hours. The reactions were monitored by HPLC and ESI-MS analysis and clearly showed the fast trityl cleavage and the slow conversion of esters into carboxylic acids. It seemed that the more *tert*-butyl groups were hydrolysed, the less the products were soluble in the TFA / DCM media, which slowed down the kinetic of ester hydrolysis. Therefore, side reactions such as a thiol oxidation and nucleophilic attack of thiols on free carboxylic acid competed and were detected by ESI-MS analysis. The optimisations, including solvent composition, temperature, and equivalent of scavengers of cations did not solve the issue. In a nutshell, the trityl strategy succeeded in the synthesis of a large scope of 11 1,4-bisthiophenols bearing a variety of functionalities. Nevertheless, the synthesis of the corresponding valine and aspartic acid derivatives failed, calling for an alternative protecting group. This one should be orthogonal with the hydrolysis of the esters. The cyanopropyl looked appropriate.

Trityl protection was an effective strategy to turn from II-1 to 1,4-bisthiophenols bearing amines, alcohols, and other neutral chemical moieties. However, the strategy failed for analogues with anionic carboxylates groups.

b. Protection of thiols by the cyanopropyl group

This subpart presents the connection of side chains to the aromatic core of **II-1** following a strategy of cyanopropyl protection. Cyanopropyl is used as a protecting group of thiols^{336,344}. The group is introduced by nucleophilic attack of the thiol on cyanopropylbromide. The deprotection occurs in mildly harsh basic conditions. The synthesis of 1,4-bisthiophenols bearing derivative of amino acids such as aspartic acid will be obtained with 4 steps. Firstly, **II-1** was protected with the cyanopropyl group by substitution with cyanopropyl bromide³⁴⁵ (**II-98**) before the amide bond formation^{79e} (Scheme 170). A deprotection with TFA and Et₃SiH³⁴⁶ will first cleave the *tert*-butyl esters. Et₃SiH was known to be a scavenger of carbocations³⁴⁷. The last step will be the basic hydrolysis³⁴⁸ of cyanopropyl group to release the free thiols.



Scheme 170: Synthesis of II-1 derivatives with cyanopropyl protection. R: group with free carboxylic acids, R' group with protection with ester of carboxylic acids. a) 3-bromopropanenitrile, NaH, THF, reflux, 48 h, 94 %; b) HOBT, EDC, DIPEA, L-Aspartic acid di-tert-butyl ester hydrochloride, DCM, r.t., 48 h, 78 %; c), TFA, Et₃SiH, DCM, r.t., 16 h; d) CsOH, THF/MeOH, r.t., 3 h.

Scheme 171 : Scope for II-1 pre-functionalisation with cyanopropyl protection

A 1,4-bis(2-mercaptoethyl)benzene-2,5-dicarboxylic acid, protected with a cyanopropyl group, bearing aspartic acid protected with *tert*-butyl esters was synthesised with a satisfying yield of 78 % (**II-99**), slightly higher than the reported example^{79e}. Then, acidic conditions released free carboxylic acids **II-100** with an unsatisfactory yield of 55 % compared to the literature³⁴⁶. A part of the product was lost in the purification. Finally, the basic conditions involving CsOH gave the S-deprotected 1,4-bis(2-mercaptoethyl)benzene-2,5-dicarboxylic acid **II-89** with a yield of 99 %. Acidic precipitation and washing removed caesium salts and no further purification was required. Then, a

dendrimeric chain^{303g,349} was employed³¹. Each carboxylic acid could serve for recognition, secondary interactions or solubility^{303g}. A 1,4-bisthiophenol **II-102**, protected with a cyanopropyl group, bearing the dendrimeric chain protected with *tert*-butyl esters was synthesised with a low yield of 36 %^{79e}. Then, the acidic deprotection^{303g} released the carboxylic acids of the dendrimeric chains (**II-103**) with a poor yield of 16 %. The conversion was partial, attested by HPLC and ESI-MS analysis. The long reaction time (24 h) cleaved a part of the amide bonds that was previously formed. At the opposite, a similar deprotection of *tert*-butyl esters on the same dendrimer required only 30 mn^{303g}. Finally, the basic conditions³⁴⁸ could not be tested because of few remaining materials. Even though the synthesis of **II-77** could be obtained by the trityl strategy, the synthesis of 1,4-bisthiophenols, protected with cyanopropyl groups, bearing hydrazide protected with carbamates was synthesised with a yield of 42 % (**II-105**), lower than the same reaction with thiols protected with trityl group. Then, the acidic hydrolysis of carbamate moieties released the hydrazide moieties (**II-106**) with a yield of 43 %, which was lower than expected³⁴⁶. The deprotection of thiols was not assessed.

In a nutshell, the thiol protection with cyanopropyl group was complementary to protection with trityl group. Thanks to this strategy, a 1,4-bisthiophenol bearing aspartic acid **II-89** was obtained.

Cyanopropyl protecting group complemented trityl group. The access of anionic chemical functions was possible in contrast with the trityl group.

c. Deadlock route to asymmetric monomers

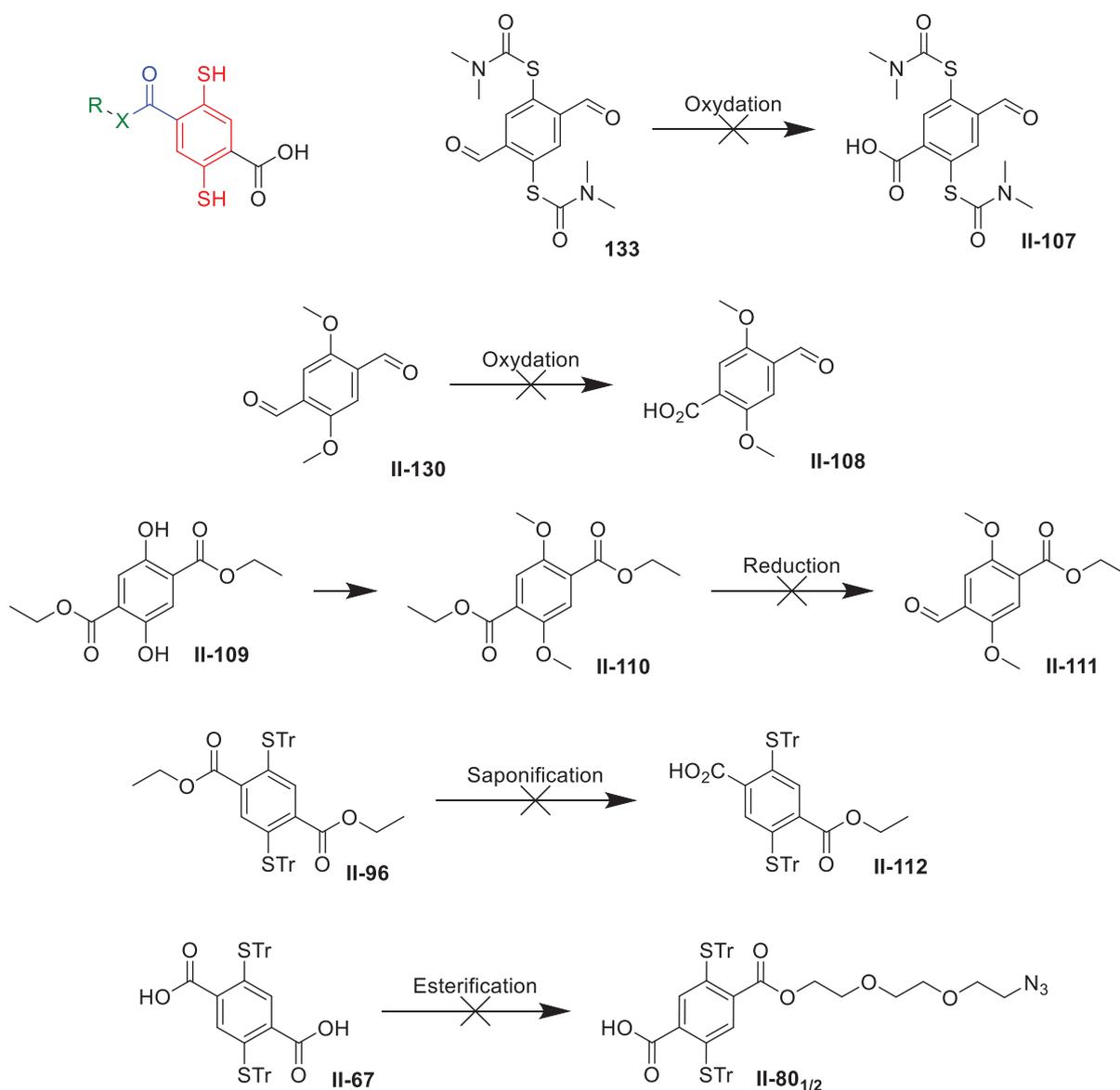
The subpart discusses the synthesis of asymmetric 1,4-bisthiophenols. Several strategies were followed: oxidation, reduction and saponification of symmetric 1,4-bisthiophenols. The idea was to find a condition where the reaction proceeds only on one of 2,5 position of the aromatic core. In each case, the media was diluted, and the reactant was slowly added at 1 equivalent (*Scheme 172*).

The first attempt was the oxidation of the aldehydes of **II-133**, which was an intermediate of **II-138**. Many conditions were tested³⁵⁰ (Appendix II-II: Reaction conditions). The main weakness of this strategy was the very low solubility of the starting material in most of the solvents. All attempts were unsuccessful to access the target **II-107**. The aldehyde function might be deactivated and sterically hindered. Moreover, ESI-MS analysis revealed that the thiol function was oxidised before the carbonyl function. Another intermediate **II-130** of **II-138** with less steric hindrance was used. The oxidation was difficult to control. The tested conditions^{350b,e} (Appendix II-II: Reaction conditions) systematically lead to a complete oxidation into two carboxylic acids (ESI-MS analysis and ¹H control).

Then, **II-110** was obtained through one step from **II-109** with a yield of 99 %³⁵¹. The reduction of both esters was detected by ESI-MS analysis in several mild conditions³⁵². A saponification on the 1,4-bisthiophenol bearing ethyl ester on the 2,5 position of the aromatic core (**II-96**) was investigated³⁵³. First tries did not give the desired target **II-112**. **II-96** showed poor solubility in common solvents. Only DCM and chloroform could be used, which tends to complicate the saponification. One equivalent of hydroxide salts was added in a 10 mg / mL solution and refluxed overnight, without conversion, as attested by ESI-MS analysis and ¹H RMN. Finally, a peptide-like coupling^{79e} with **II-67** in diluted media and by a slow addition gave **II-80** and unreacted **II-67**, without further optimisation. The structures were attributed by ESI-MS analysis. To sum up, the obtention of an asymmetric 1,4-bisthiophenols was

31 Provided by Pr. Antony Davis, Bristol University

unsuccessful. Further investigations could be done, involving flow chemistry or very slow addition of reactants.



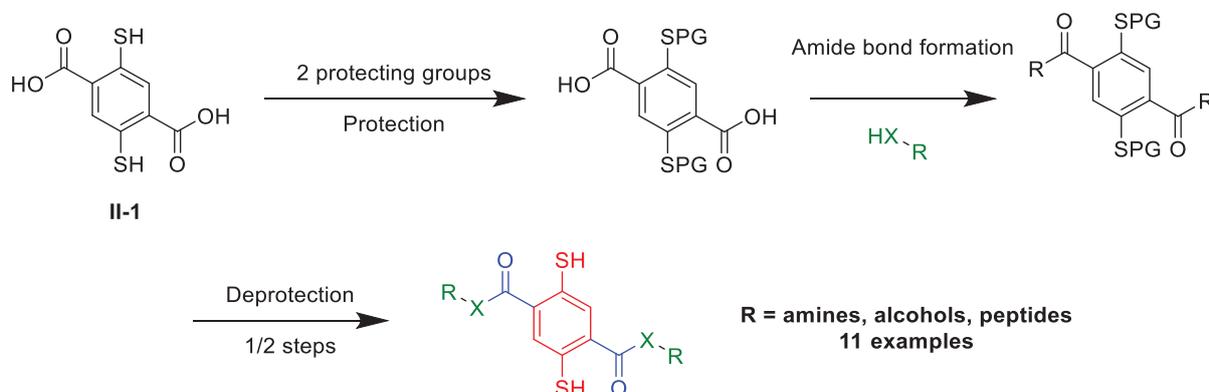
Scheme 172: Different attempts of desymetrisation in order to obtain asymmetric functionalised carboxylate 1,4-bisthiophenols (general structure top left: R: functional groups for the recognition of modified proteins. X: NH, O). Cyclo-oligomerization through disulfide bridge formation is intended to be the last synthetic step. green functional groups, connection between the side chain and the principle chain (blue), red aromatic backbone and connection between the monomer for the principle chain). Detailed conditions available in appendices.

The obtention of asymmetric building blocks was not achieved. The employed reactions could not be controlled. Optimisation and very slow addition could lead to a target.

d. Conclusion

The synthesis of 11 1,4-bisthiophenols bearing side chains on the 2,5 positions of the aromatic core with functional groups was successful thanks to a strategy of protection, peptide-like coupling and deprotection in one or two steps (Scheme 168) from II-1. The cyclo-oligomerisation will come later and is supposed to form tailored cavities for the recognition of modified proteins. Two complementary

protection groups were used following the nature of the desired chemical functions of the side chains. Trityl groups were efficient for the introduction of amine and alcohol moieties while cyanopropyl was more adapted to carboxylate moieties. That synthetic path suffered from few failures and showed an appreciable tolerance for various chemical functions. The synthetic road to asymmetric monomers was more complex because of the difficulty of performing a reaction on one in two positions. These new monomers will serve the design of cavitands by the strategy **B1**, which is exposed in the next part.

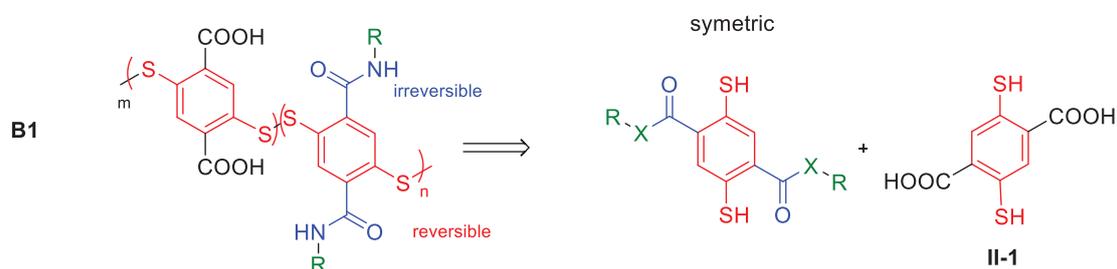


*Scheme 173 : Recap of the synthetic approach for functionalised 1,4-bisthiophenols involved in the strategy **B1** of cavitands. R: functional groups for the recognition of modified proteins. X: NH, O. Two protecting groups (cyanopropyl and trityl). Connection between the monomeric units forming the principal chain (red) and the connection between the principal chain and the side chain bearing the functional groups (blue). R: functional groups dedicated to the recognition of modified proteins.*

*Pre functionalisation strategy succeeded in the access of a new series of 1,4-bisthiophenols that derived from **II-1**. Trityl and cyanopropyl protecting groups were complementary. The path to asymmetric building block was still locked.*

5. Setbacks to the connection of the functionalised monomers into a principal chain

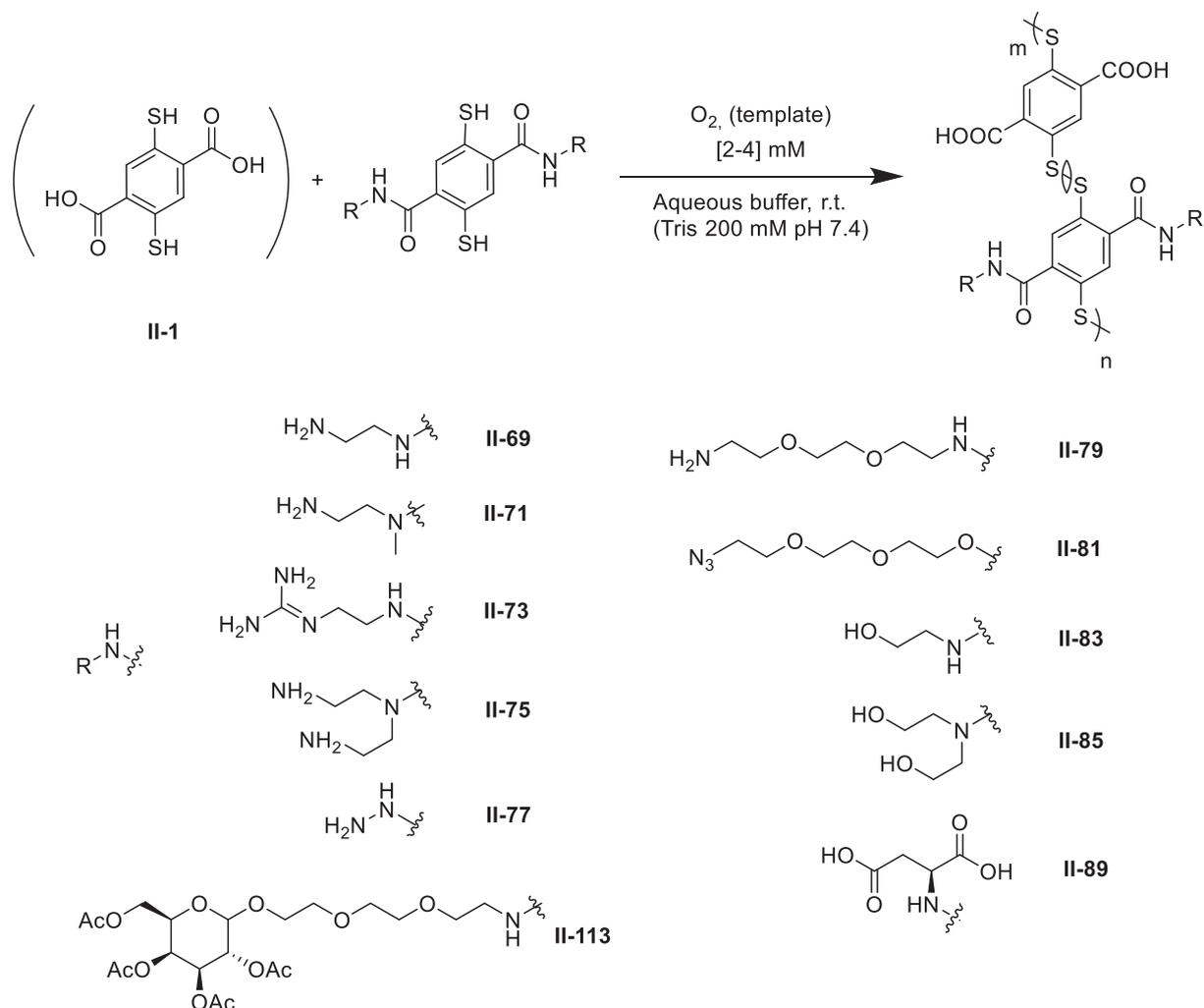
The aim of this subchapter was to connect the previously synthesised functionalised monomers by disulfide bond in DCLs in order to form the principal chain of the cavitands following the strategy **B1**. (Scheme 174)



Scheme 174 : Retro-synthetic analysis for the formation of hetero and homo oligomers starting from a simple monomer and functionalised monomers bearing functional groups (green) dedicated to the recognition of modified proteins. Connection between the aromatic monomers by disulfide bond formation thanks to DCLs (red), side chain connected to the principal chain by irreversible amide bond (blue)

- a. Formation of reversible disulfide linkage between 2,5-functionalised-1,4-bisthiophenols through DCC

Most of the previously synthesised 1,4-bisthiophenols with amide bond and functionalisation on the 2,5 positions of the aromatic core were studied individually and with **II-1** in DCLs. The building blocks were typically dissolved between [2-4] mM in ammonium acetate or Tris buffer (pH 7.4, 200 mM). The vials were left open, stirred and the mixtures were analysed by HPLC, ESI-MS analysis or LC-MS. For most of the building blocks (except **II-89**, **II-81**), the chromatographic separation of the oligomers produced by HPLC was only partial despite a thorough screening of the elution condition. As a consequence, monitoring of the libraries was most achieved by ESI-MS analysis, and when possible by HPLC. The ESI-MS analysis spectra were recorded after a few minutes and 6 hours. Analysis of the DCLs and HRMS data are detailed in appendices (Appendix II-IV: Additional HRMS or LC/MS data).



*Scheme 175: DCLs combining **II-1** and selected functionalised 1,4-bisthiophenols in the 2,5 positions by amide bond with side chains.*

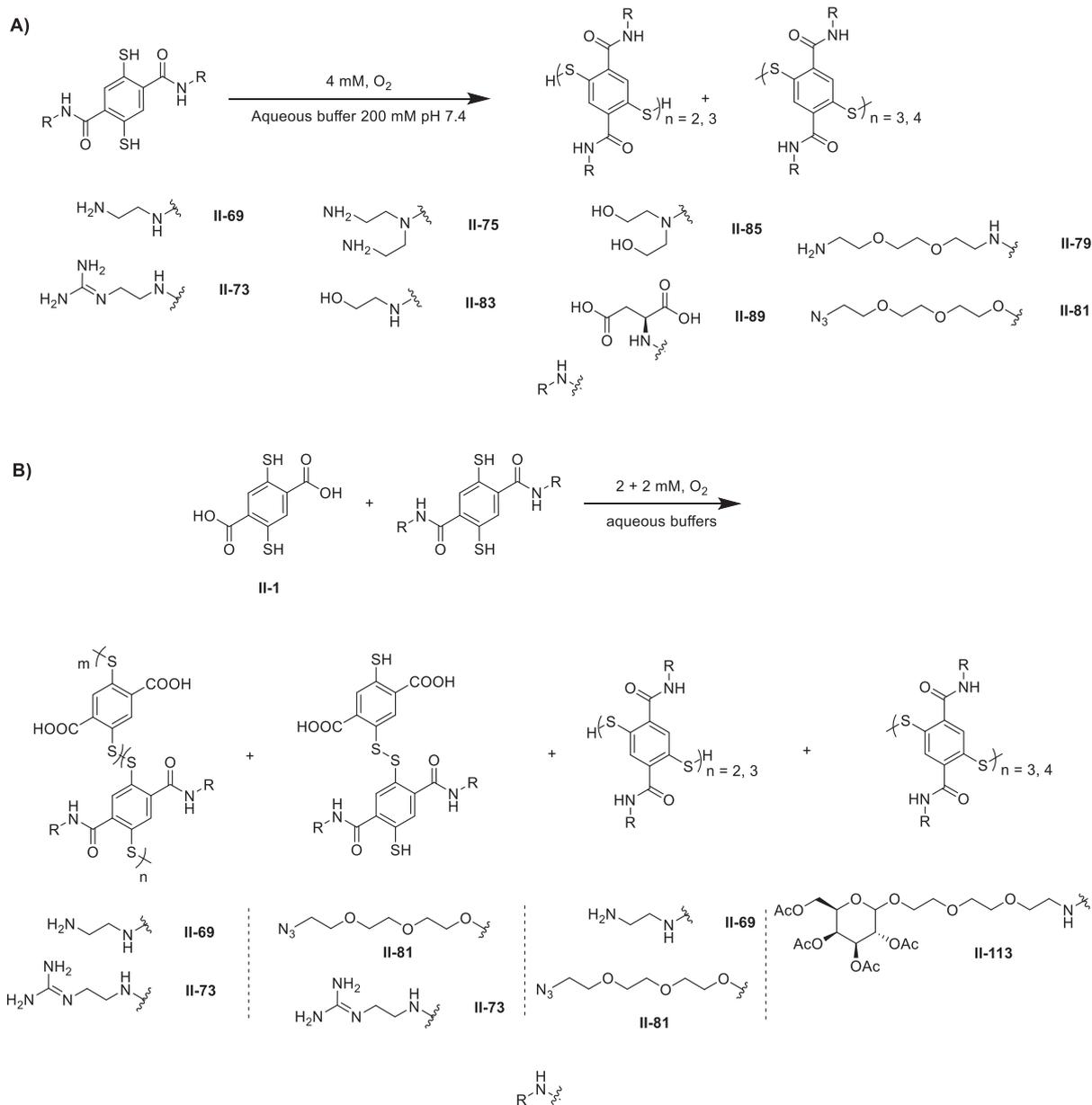
The main trends in terms of reactivity are listed below: disulfide bond formation or oligomerisation, NS coupling or overoxidation of thiols into sulfenyl amide, overoxidation of thiols into sulfenyl acids, hydrolysis of amide bond connecting the side chains to the principle chains, transamidation, reduction of azide and influence of templating in the cyclo-oligomerisation.

- Oligomerisation/Disulfide bond formation

Two different types of DCLs were studied: the oligomerisation of 1,4-bisthiophenols with functionalisation in the 2,5 positions individually, and with **II-1**. Regarding the DCLs of functionalised 1,4-bisthiophenols, the formation of disulfide bond was observed in all aqueous buffers (200 mM, pH 7.4 : Tris, ammonium acetate, citrate, phosphate) and for each R group (**II-69**, **II-71**, **II-73**, **II-75**, **II-79**, **II-81**, **II-83**, **II-85**, **II-89**) at 6 hours. However, the disulfide-containing species were limited to linear dimer, except for **II-77** (*Scheme 176*). Besides, cyclic and linear trimers were observed in the DCLs of **II-85**, **II-73**, **II-79** and **II-89**. For **II-79**, at 48 hours, HPLC and ESI-MS analysis indicated that a cyclic tetramer (**II-79**)₄ was the main species which was coherent with what was observed for the DCL of **II-113**³², a 1,4-bisthiophenol with a similar length of side chain, leading spontaneously to a cyclic tetramer^{82c}. Cyclic (**II-89**)₂ and linear H(**II-89**)₂H tetramers and pentamers (**II-89**)₅ were also detected for **II-89**. No clues of polymer formation could be collected. Linear dimers dominated the generated dynamic covalent systems in terms of disulfide species. Electrostatic repulsion between cationic side chains of 1,4-bisthiophenols bearing amines might explain the difficulty to generate longer oligomers, although **II-1** self-assembled into a cyclic tetramer in Tris buffer (at 4 mM, buffer at 200 mM, pH 7.4). The nature of the side chains of the 1,4-bisthiophenols clearly impacted the thiol reactivity. While for **II-1** the reactivity was dominated by disulfide bond formation, the reactivity of functionalised 1,4-bisthiophenols in the 2,5 positions was not principally disulfide bond formation.

Regarding the DCLs of functionalised 1,4-bisthiophenols dissolved in aqueous buffers with **II-1**, cyclohetero-oligomers such as cyclic tetramers (**II-1**)₂(**II-73**)₂ and (**II-1**)₃(**II-73**) were detected by ESI-MS analysis. Ineffective HPLC separation prevented quantitative assessment. Similar oligomers between **II-1** and **II-69** were observed in the corresponding DCL. **II-1** and **II-113** evolved distinctly by a self-sorting process. Finally, **II-1** and **II-81** only self-assembled into linear dimer (**II-1**)(**II-81**). Consequently, linear heterodimers dominated the generated dynamic covalent systems in terms of disulfide species. Similarly with the previous case, the building blocks formed disulfide bonds, but it was not the only dominant reactivity in the conditions of these DCLs.

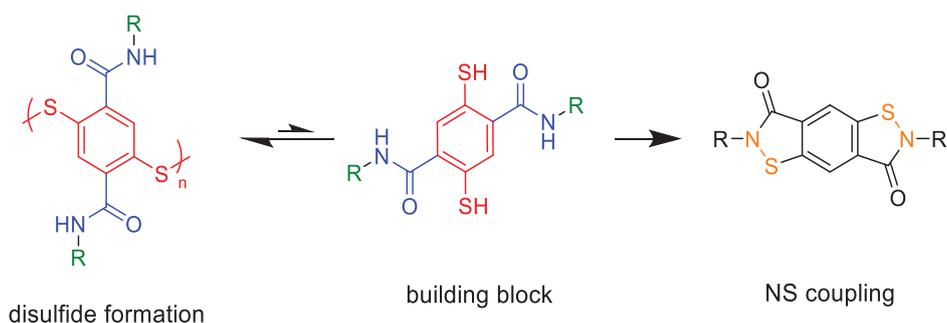
32 **II-113** is a building block applied on lectin recognition(c) Pascal, Y., Université Claude Bernard Lyon 1, 2018.. Sample provided by Dr. Pascal. Should be considered only for combination with **II-1**, not for the whole study.



Scheme 176: Disulfide bond formations for DCLs II-1 and 1,4-bisthiophenols bearing side chain by amide bond connection in the 2,5 positions of the aromatic core. A) DCL involving one 1,4-bisthiophenol, B) DCL involving two building blocks: II-1 and a functionalised 1,4-bisthiophenol

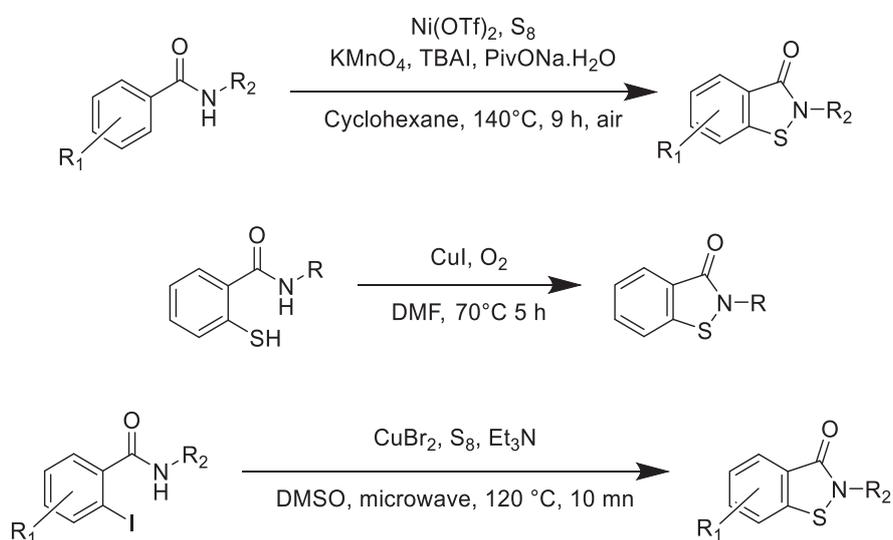
- NS coupling/formation of sulfenyl amides

During the DCLs, an unexpected phenomenon occurred at first stages, consisting in the oxidation of thiols into sulfenyl amides, (or benzo[d]isothiazol-3(2H)-ones or benzoisothiazolones) preventing from a straightforward disulfide bond formation. The process was named NS coupling as opposed to S-S coupling (disulfide bond formation) (Scheme 177). The following short paragraph gathered the bibliographic data on this chemical function.



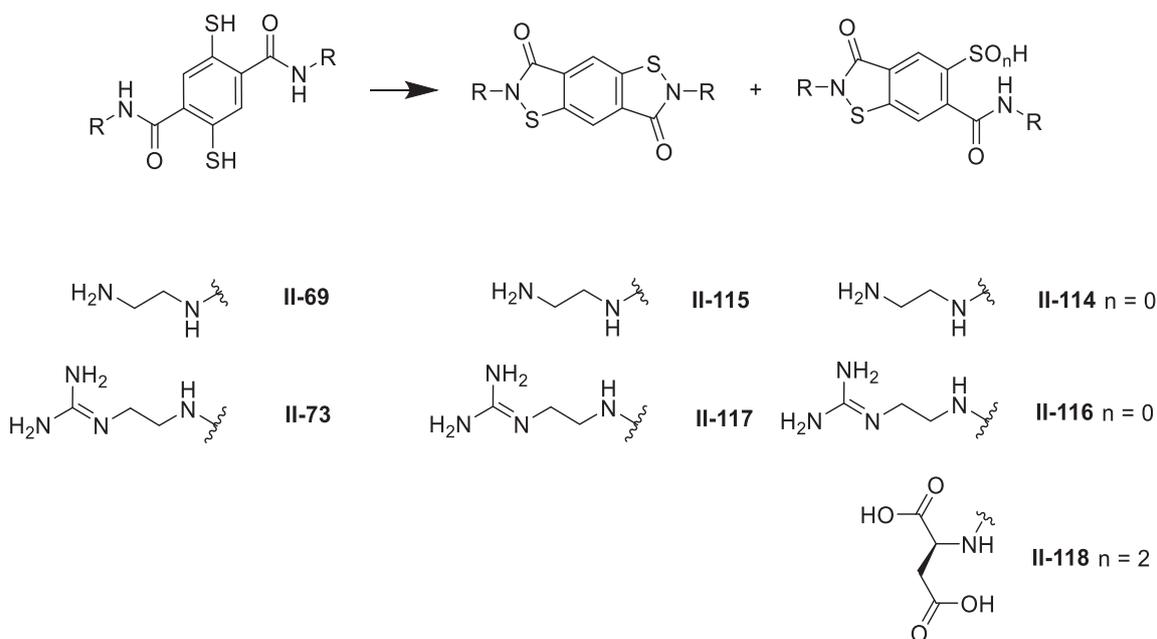
Scheme 177: Presentation of the NS coupling: instead of going through the disulfide bond formation, a benzothiazolone could be formed

The sulfenyl amide moiety has medicinal activity. Indeed, benzothiazolones are potent antimalarial drugs³⁵⁴ and inhibitors of PHOSPHO1³⁵⁵. Their synthesis are mainly based on organometallic chemistry: N/S oxidative dehydration³⁵⁵⁻³⁵⁶, Ni-catalysed carbon sulfuration and annulation³⁵⁷ and a domino Cu-catalysed Aryl-I thiolation³⁵⁸ (*Scheme 178*).



Scheme 178: Examples of organometallic routes to benzothiazolones. Top: Ni-catalysed carbon sulfuration and annulation, middle: N/S oxidative dehydration and bottom: domino Cu-catalysed Aryl-I thiolation. R, R₁, R₂ various chemical groups

In our project, the formation of sulfenyl amide prevented the formation of disulfide bond. As a matter of fact, benzothiazolones were formed for a large majority of the previously synthesised 1,4-bisthiophenols. Even **II-81**, having an ester instead of an amide bond, was concerned. Examples of NS species are represented in *Scheme 179*. In this case it was not properly a benzothiazolone. There is no generic name, a simple similar structure is named 3*H*-benzo[*c*][1,2]oxathiol-3-one. The proportions were not accessible.



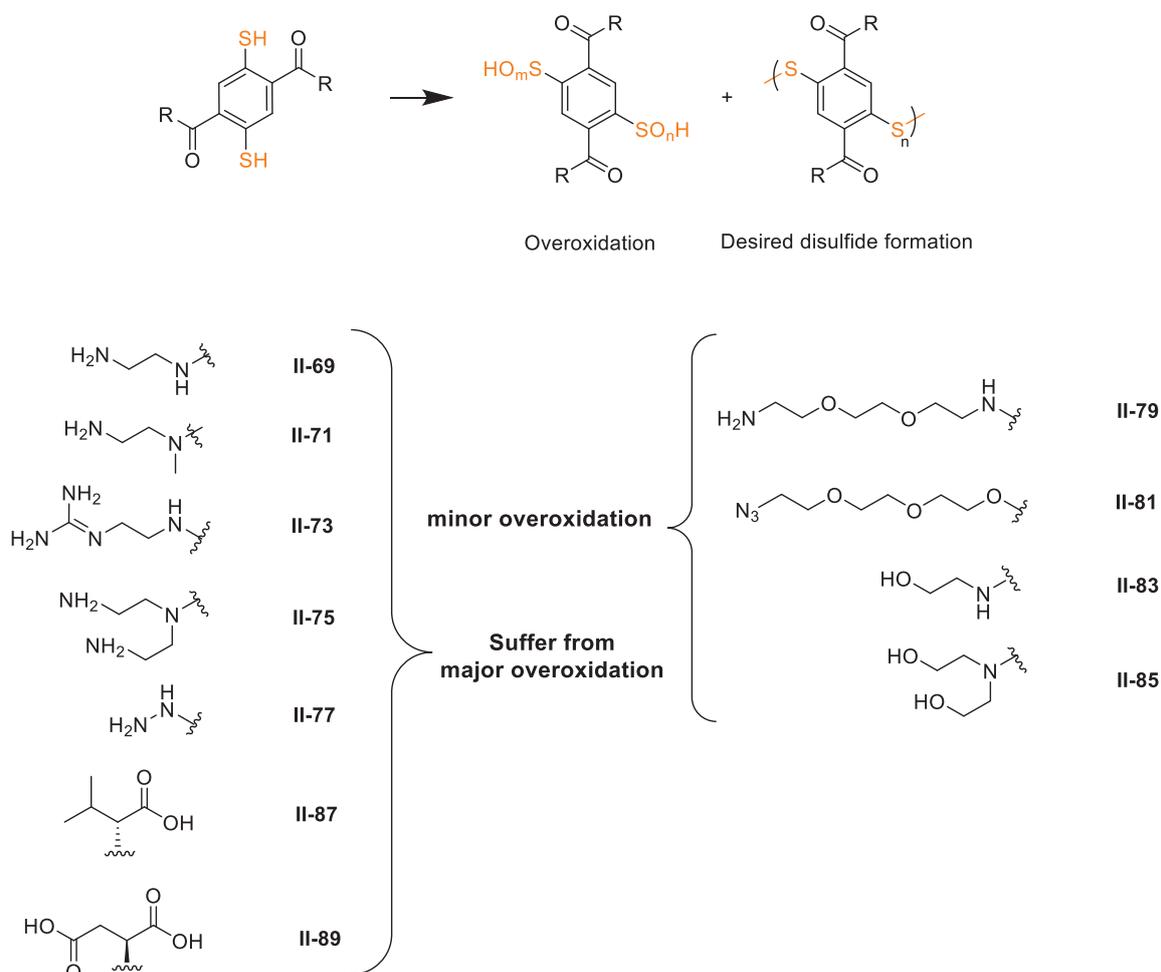
Scheme 179: Examples of observed benzoisothiazolones in DCLs by ESI-MS analysis (II-114, II-116, II-117), HRMS (II-115), or LC/MS (II-118). R: functional groups dedicated to the recognition of modified proteins. No yields were available.

The NS coupling occurred mainly on 1,4-bisthiophenols, but also affected the ending thiols of linear oligomers of 1,4-bisthiophenols. For instance, all the linear di- tri- tetra and pentamers of **II-89** were detected in the LC-MS analysis of its DCL as well as their sulfenyl amide counterparts in one or two ending thiols. Usually, all the available thiols underwent the NS coupling. At short times (1 hour, DCL of **II-89**), some cases of one to two NS coupling on 1,4-bisthiophenols were identified. The formation of the N-S bond was irreversible, hence trapping the species in a kinetic well, impacting on the DCC based on the reversible thiol-disulfide exchange. Regarding 1,4-bisthiophenols bearing basic amine in their side chains at the 2,5 positions (**II-69** for instance), the NS coupling was immediate, and the resulting products were insoluble in water and precipitated. The process took place spontaneously in water, which was in contrast with the previously described harsh conditions find in the literature³⁵⁴⁻³⁵⁸. This chemical path was imputed to the short distance between the nitrogen and the sulfur atoms found in our backbone of 1,4-bisthiophenols with amide bond in the 2,5-positions of the aromatic core. It was probably favoured by the electrostatic interactions between thiolates and ammoniums, preventing from the disulfide bond formation. This unexpected reactivity would eventually lead to an expansion of the available benzoisothiazolones for potent medicinal applications. Strategies to avoid the NS coupling are presented in the next part (Overcoming the NS coupling).

- Overoxidation of thiols into sulfe/inic acids

In the typical DCL conditions used for **II-1**, i.e. aqueous buffer, pH 7.4 at 4 mM, oxidation of thiols went through the reversible disulfide bond formation^{82a}. However, thiols could be oxidised into sulfenic acid (RSOH), rather unstable, and sulfinic acid (RSO₂H)^{62,359} (Scheme 180). If the thiol became a sulfe/inic acid, a disulfide bond could not be created. The DCLs of the 1,4-bisthiophenols suffered in a large extent of overoxidation, apart from **II-85**, **II-81**, **II-79**. The thiols of 1,4-bisthiophenols and their corresponding linear oligomers were partially converted into sulfinic acid. Sulfenic acids were presumed to be unstable. Consequently, when two overoxidations were detected (+16/32) in ESI-MS

analysis, one sulfinic acid was assigned rather than two sulfenic acids. Surprisingly, the overoxidation occurred quickly, at short times (few minutes). Building blocks with long side chains such as **II-79** were not affected by overoxidation. Therefore, the oxidation might be favoured by intramolecular H-bond between the resulting sulfenic acid and the carboxylic acids in the case of **II-89**. In other cases, such as **II-69**, the oxidation might be favoured by a salt bridge between ammonium groups and sulfenic acids. This hypothesis was supported by examples of supramolecular effect on redox potential³⁶⁰. Indeed, the local electrostatic environment (protein charge state) can change the standard potential of chemical functions or groups (a ferrocene for instance)³⁶¹. The determination of standard potentials by electrochemistry could support the hypothesis. Combined with the NS coupling, this undesired reactivity of the thiols restrained the oligomerisation process.



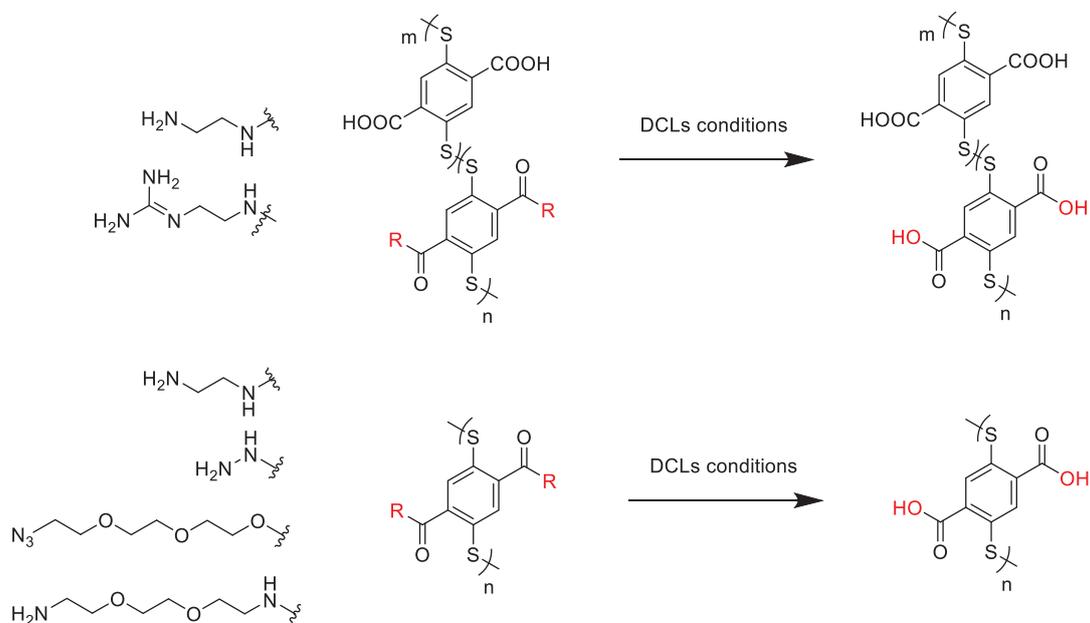
Scheme 180: Oxidation of thiol moieties in the DCLs of functionalised 1,4-bisthiophenols: the oxidation could lead to the desired disulfide bond formation or to undesired sulfenic ($m/n = 2$) and sulfinic ($m/n = 1$) acids.

Monomers can also be connected by disulfide bond and then a terminal thiol can be overoxidised (not represented here). R: functional groups for the recognition of modified proteins.

- Hydrolysis of amide bond

In addition, a hydrolysis of the amide bond connecting the side chain to the principle chain occurred (*Scheme 181*). This process was most of the time partial and minor, occurring only on one position of the building block in an auto-assembly (linear or cyclic). However, it was more present for macrocycle

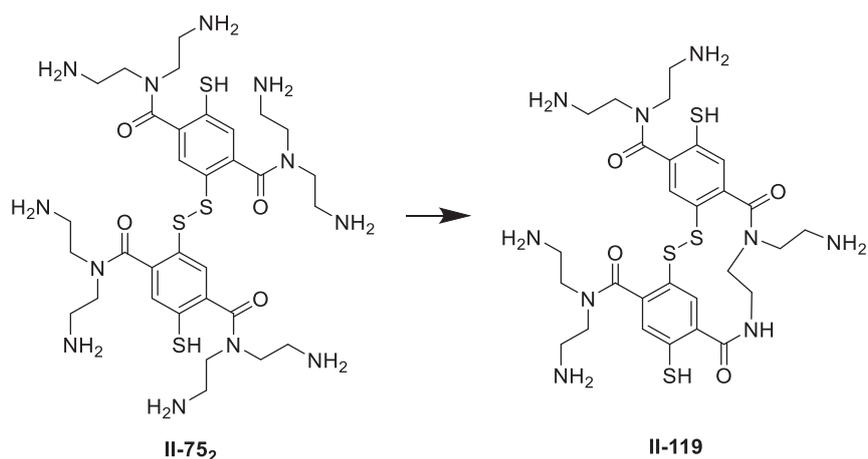
of **II-77**. Indeed, most of the detected auto-assembly suffered from numerous hydrolysis. For instance, a $(\text{II-77})_2(\text{II-1})_2$ could be observed even though no **II-1** monomer was present in the DCL. Amides are one of the most abundant and stable chemical bonds that are found in nature³⁶². Furthermore, minor hydrolysis of cyclic hetero tetramers between **II-69** / **II-73** and **II-1** monomers were detected by LC/MS analysis. The hydrolysis rate of amide was found to be similar between pH 5 – 9³⁶³, including the conditions used in this study. In these, the half-life of the bond should be 267 years³⁶⁴. Besides, the amide bond could be hydrolysed through biological activation and electrophilic assistance³⁶⁵, such as nicotinamidase, serine and cysteine proteases, metallo-endo/exopeptidase, or flavoenzyme. Therefore, the observed hydrolysis should be assisted by the neighbouring cationic side chains. Indeed, no hydrolysis on the free building blocks was detected. Hence, the proximity of terminal amines may be responsible. The resulting hydrolysed oligomers might be stabilised by interactions between cationic charges and the resulting carboxylate groups. In addition, the PEG derivatives were all affected. The combination of amide hydrolysis and PEG degradation in aqueous solution could explain the situation. In fact, although stable in aqueous media, PEG chains could be cleaved in not harsh conditions in water³⁶⁶. No mention of similar reactivity was found in the literature. This process being marginal, it was not a major concern.



Scheme 181: Amide hydrolysis in macrocycles, observed in DCLs by ESI-MS or LC/MS analysis. The hydrolysis occurred on all the size of macrocycle, from linear dimers to cyclic tetramers.

- Transamidation

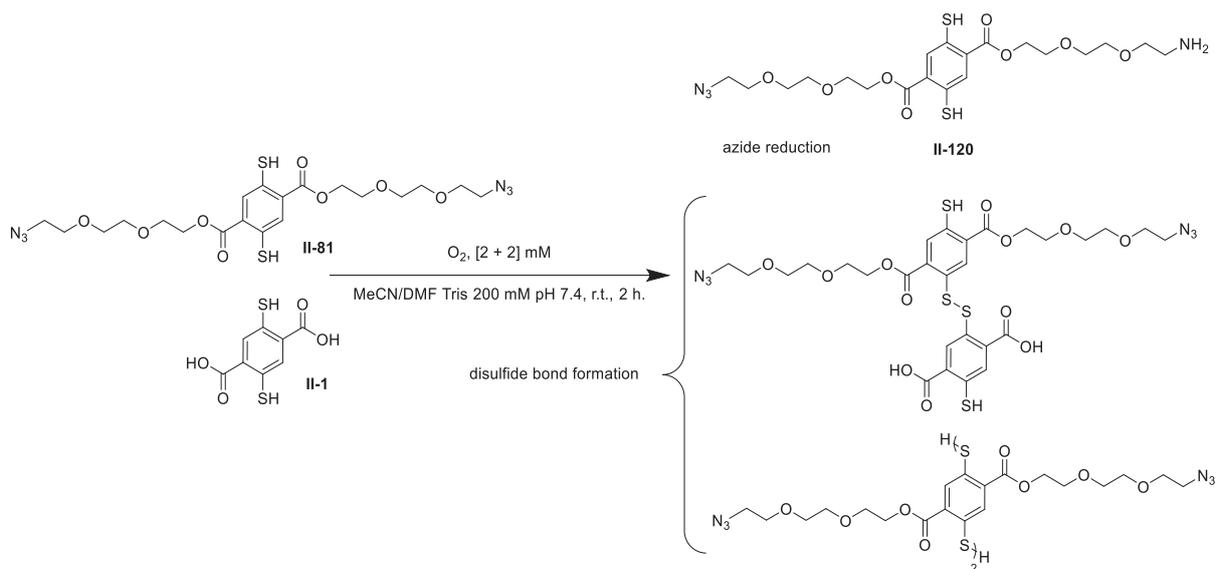
The DCLs of **II-65** showed another unexpected reactivity. The linear dimer underwent a transamidation, expelling a DETA moiety (Scheme 182). This reactivity was clearly minor and limited to **II-65**. Transamidation could play an important role in biology and was mainly obtained through harsh conditions in synthetic chemistry³⁶⁷. The observation of a spontaneous transamidation was therefore unexpected. It could be explained by the decrease of steric hindrance obtained by the expel of a DETA side chain. Furthermore, no similar transamidation was observed for other building blocks, DETA being the bulkiest one.



Scheme 182: Transamidation of a linear dimer of II-65, observed by ESI-MS analysis

- Azide reduction

Click chemistry between terminal alkynes and azides is a useful tool for the construction of silica based functionalised material³⁶⁸. In this context, the incorporation of azide moiety in macrocyclic receptors could lead to material which may be used as stationary phase for affinity chromatography by CARMAPHARM partners of UCB Pharma. This justified the choice of equipping the 1,4-bisthiophenol precursors of cavitands with a terminal azido function in **II-81** and the attempts of heterocyclo-oligomerisation with **II-1**. In 2 hours, the solution became colourless and a precipitate appeared. The use of an organic cosolvent (MeCN or DMF in 1:1 volume with Tris buffer 200 mM pH 7.4) did not help to keep a homogenous mixture. LC/MS analysis of the supernatant confirmed the structure of several species formed in solution. Among them, as previously discussed, the dimers **II-1/II-81** and **II-81/II-81** were identified (Scheme 183). The solid and two species in solution remained unidentified by ESI-MS analysis. Moreover, in solution, a linear dimer composed of a **II-1** unit and a **II-81** unit with one azide reduction in an amine was observed (Scheme 183). The corresponding free building block **II-62** was also detected in minor proportion.



Scheme 183: Species detected by LC/MS analysis in the supernatant of the DCL between II-81 and II-1 :Reduction of an azide function, formation of homo- and linear heterodimers by disulfide bond formation

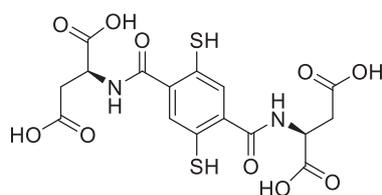
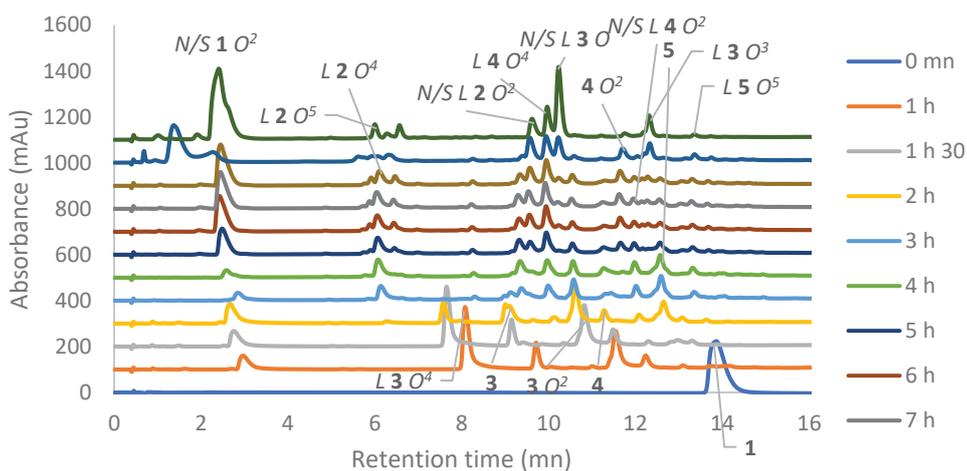
Hence, there was no denying that the azido function was not compatible with thiols in this example. Indeed, an oxidoreduction process took place between the two functions. This reactivity existed in the literature³⁶⁹. This incompatibility could be possibly tackled by monitoring the potential of the library. Besides, the poor solubility of **II-81** was a problem for the DCC. The strategy of click chemistry with CuAAC could be replaced by thiol-ene chemistry for instance for a further grafting on solid support³⁷⁰.

- Template

In order to avoid the other reactivities than the disulfide bond formation regarding the oligomers of the functionalised 1,4-bisthiophenols and in order to search for a possible macrocyclic amplification templates were used at 0.25 equivalent per building blocks in the same conditions than the previous DCLs (Tris or ammonium acetate buffer, 200 mM pH 7.4, 2-4 mM, r.t.).

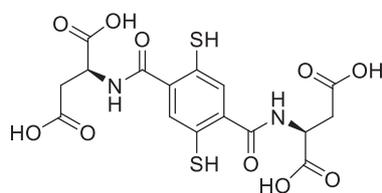
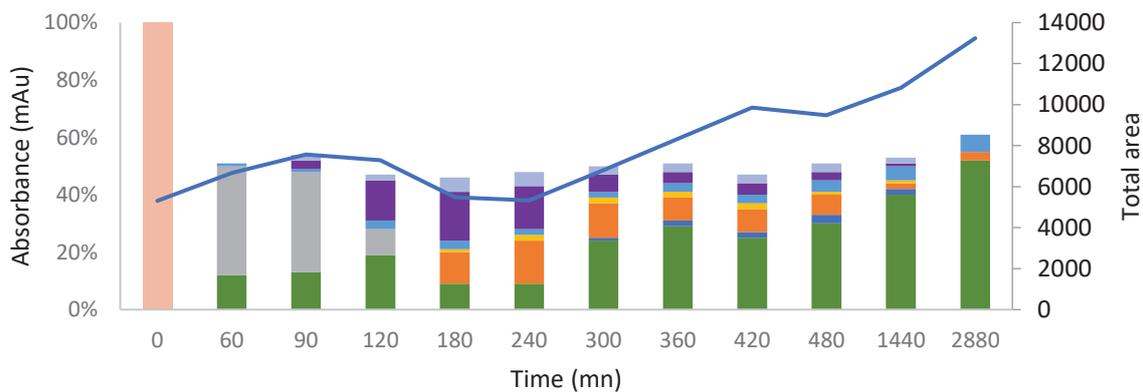
In the case of the 1,4-bisthiophenols bearing basic amino side chains in the 2,5 positions (**II-69**, **II-73** and **II-65**), dicarboxylic acids (such as succinic acid, oxalic acid, and H-Glu-Glu-OH) and Fondaparinux, an anionic pentasaccharide did not prevent the overoxidation of thiols into sulfenic acids and the formation of sulfenyl amides. In fact, ESI-MS analysis showed NS coupling and overoxidation products.

In the case of **II-89**, the same template used for the total amplification of **II-14** was employed. In addition, a longer spermine analogue, *N*¹,*N*¹-(hexane-1,6-diyl)bis(propane-1,3-diamine), was chosen as the carboxylic acids of **II-89** were more distant from the aromatic core compared to **II-1**. Finally, ammonium acetate was used as both template and buffer (200 mM, pH 7.8). Interestingly, the templates similarly favoured the cyclo-oligomerisation of **II-89** compared to the DCL without template, without preventing the undesired reactions. An HPLC optimisation afforded a suitable separation with LC/MS analysis of the different linear- and cyclic di-, tri-, tetra- and pentamers of **II-89** and their overoxidised counterparts, including sulfinic acids and sulfenyl amides (Case of spermine, *Scheme 184*, *Scheme 185*, *Scheme 186*). The total area usually was a metric of the quantity of matter of 1,4-bisthiophenols and their corresponding oligomers in our DCLs, because of the additive Beer-Lambert law, linking the intensity of the absorbance with the quantity of matter of each compounds following their molar extinction coefficient λ ¹²⁶. In a DCL composed of one 1,4-bisthiophenol such as this one, we usually applied the hypothesis that each compound had the same λ given that their chemical moieties involved in the absorbance were not changed. Regarding the case of the formation of benzothiazolones, an increase of the total area could be seen. That could be explained by the cyclic sulfenyl amide, which could slightly increase the molar extinction coefficient of the corresponding compounds. Therefore, the quantitative data should be carefully used. Only the isolation of a specie containing a sulfenyl amide and the corresponding specie with the free thiols and the determination of their λ would verify the hypothesis.



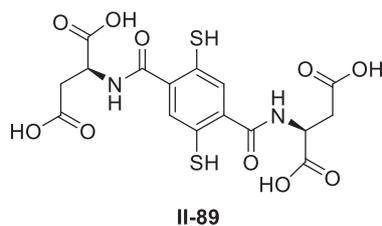
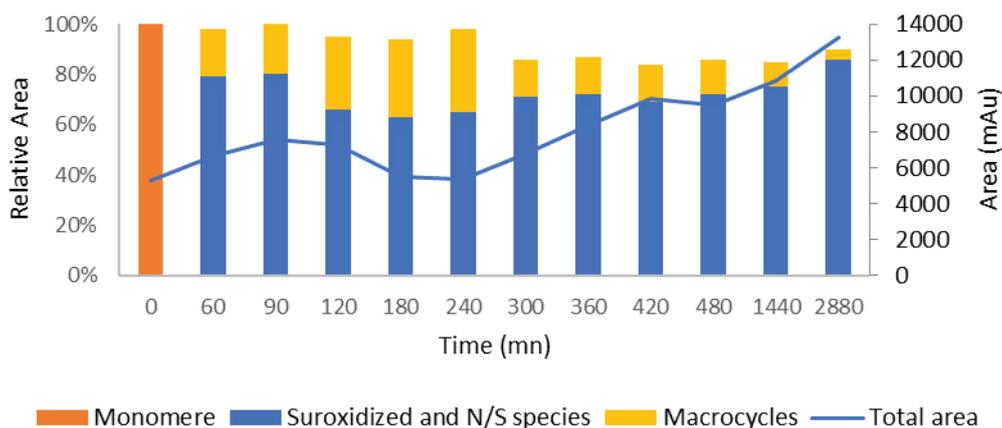
II-89

Scheme 184: Evolution of **II-89** and spermine (4 + 1 mM) in Tris buffer 200 mM pH 7.4, method F. N/S: sulfenyl amide formation, Oⁿ: n overoxidation of the sulfur atoms, 1,2,...: cyclic oligomer ring size



II-89

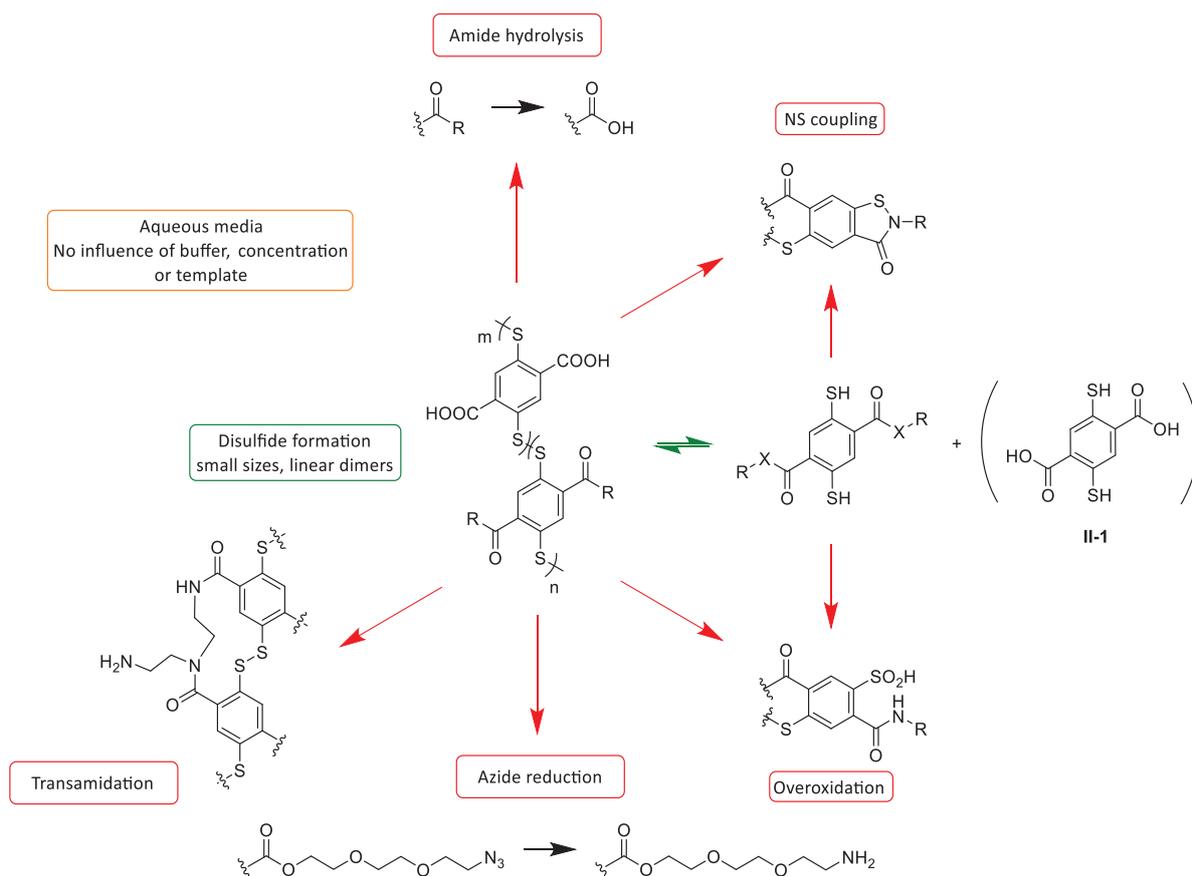
Scheme 185: Population evolution for the DCL of **II-89** and spermine (4 + 1 mM) in Tris buffer 200 mM pH 7.4. L linear, N/S: sulfenyl amide formation, O_n: n overoxidation, 1,2,...: cyclic oligomer size (1)



Scheme 186: Population evolution for the DCL of II-89 and spermine (4 + 1 mM) in Tris buffer 200 mM pH 7.4 (2)

The overoxidation and the NS coupling occurred immediately, on a linear dimer. The proportion of overoxidised and species with sulfenyl amides was in average 80 %. The kinetics of disulfide bond formation was fast (total consumption of **II-89** in one hour), compared to **II-1** (total consumption of **II-1** in one day)¹⁶. Usually, the kinetic of disulfide bond formation increased with low pK_a of thiol^{128a,371}. Generally, anionic template seemed inefficient to amplify the proportion of oligomers of 1,4-bis(2-thiophenols) bearing basic amino side chains in the 2,5-positions, while cationic amino templates slightly stabilised the oligomers of **II-89**. This is not surprising since the effect of a template was supposed to be thermodynamic, while the undesired phenomena were kinetically fast. Indeed, the template aims to minimise the energy of the whole system by the creation of a thermodynamic well³⁷.

To conclude, the series of functionalised 1,4-bis(2-thiophenols) showed unexpected reactivity in the DCLs conditions (2-4 mM, Tris or ammonium acetate buffer, 200 mM pH 7.4) (Scheme 187). Even though disulfide bond formation occurred, mainly leading to linear dimers, the library suffered from kinetics traps. Oxidation of thiols into sulfenyl acids and sulfenyl amides occurred quickly and competitively with disulfide bond formation. In addition, other side reactions were observed, such as amide hydrolysis, transamidation as well as azide reduction. The improvement of the HPLC separation is a *sine qua non*-condition to obtain additional information on these processes. The first DCLs were disappointing, and no quantitative cyclo-oligomerisation could be obtained. This clearly prevented us from going to the next step which is the use of a PTM mimic as a template. In the light of these experimental data, an evaluation of the mechanism for N/S coupling was necessary.



Scheme 187: Desired (green), undesired and unexpected (red) reactivity of the functionalised 1,4-bisthiophenols in aqueous DCs ($X = N, O$)

Functionalised 1,4-bisthiophenols mainly transformed into overoxidised sulfenic acids and sulfenyl amides upon dissolution at millimolar concentration. The proximity of the thiols and the amide moieties in ortho on the aromatic core might be responsible. The disulfide bond formation occurred but was limited to the formation of linear dimers.

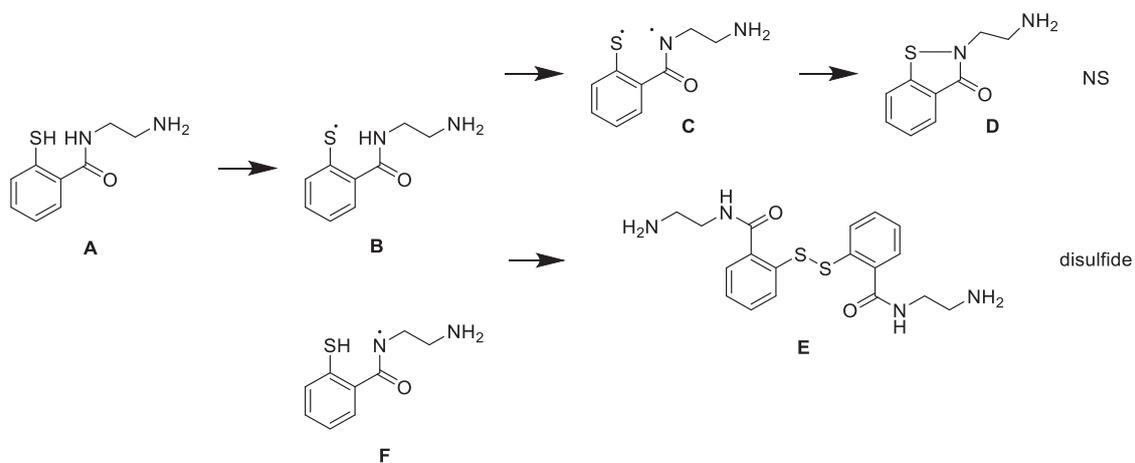
b. Rationalisation of the oxidative NS coupling

Different reaction pathway may lead to the NS oxidative coupling. A radical mechanism was postulated and investigated by DFT while a recent electrophilic mechanism was found in the literature and was developed in a dedicated subpart.

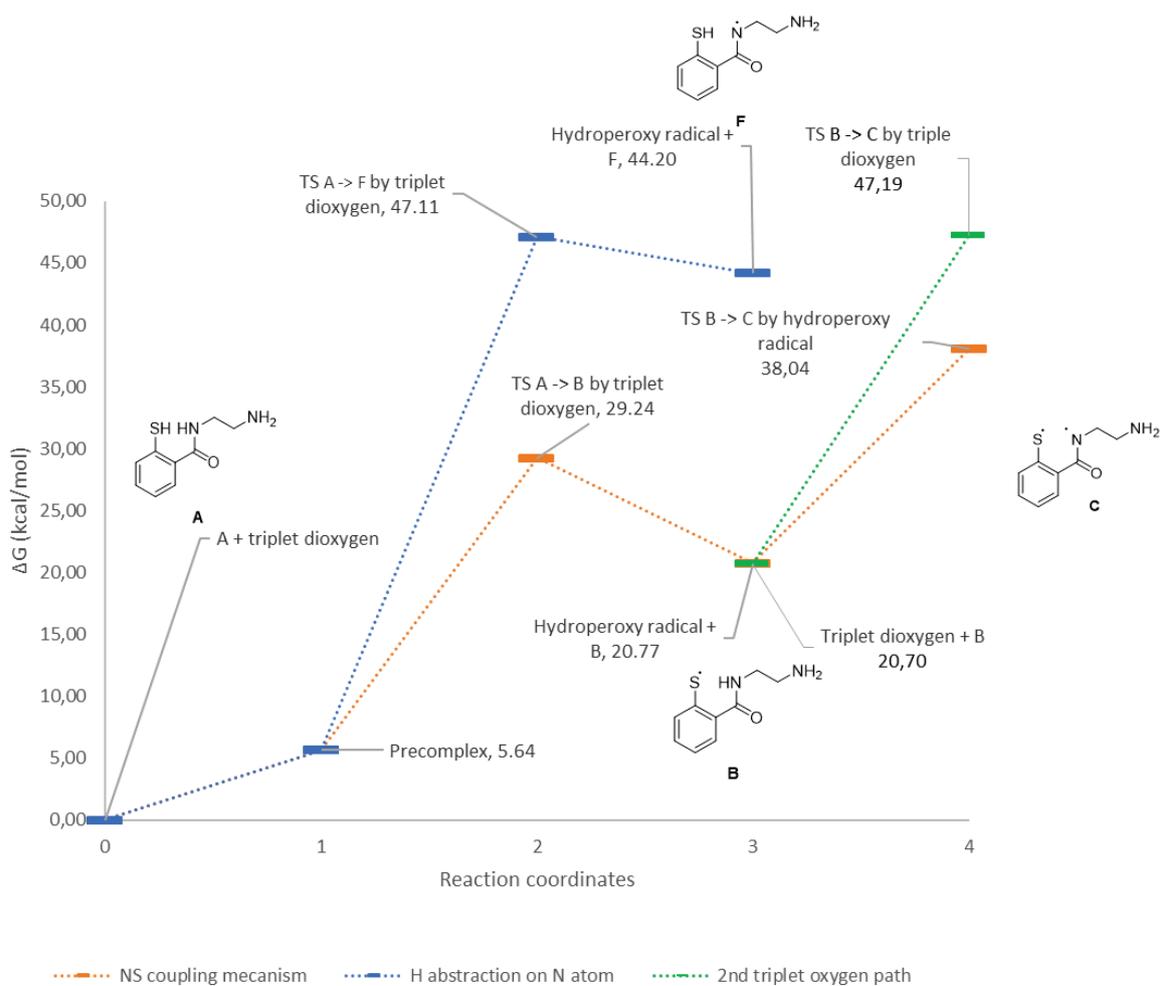
i. Investigation of a radical mechanism by DFT

The oxidative radical mechanism is similar to the mechanism of disulfide bond formation. Even though radical processes could be challenging to model^{70,372}, they could be well investigated by DFT³⁷³. Even diradical species were reported to be accurately described in DFT³⁷⁴. The mechanism of the disulfide bond formation in aqueous solutions is still discussed^{85,128,371a,375}. Some studies involved the superoxide radical whereas others reported the triplet dioxygen 3O_2 as oxidant. Triplet dioxygen was first chosen as abstractor of hydrogen. Indeed, it is the most probable oxidant in our system³⁷⁶. **A** was considered as a model of **II-69** (Scheme 188, Scheme 189). The thiophenol **A** had a similar amine than **II-69** in an ortho position, presenting the same arrangement of atoms for a NS coupling. **A** contained less atoms than **II-69**, facilitating the calculations, which were performed with the Austin-Frisch-Petersson (APFD) functional²¹⁷ with a 6-311G** basis set for C, H, O, N atoms and 6-311G++** for S atom. A Polarizable

Continuum Model (PCM) using the SMD variation³⁷⁷ was used for the description of water. Spin contamination³⁷⁸ was checked for all calculations, given the radicalar nature of the investigated mechanism.

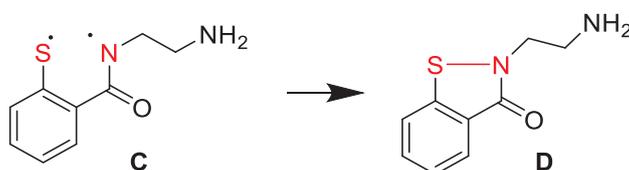
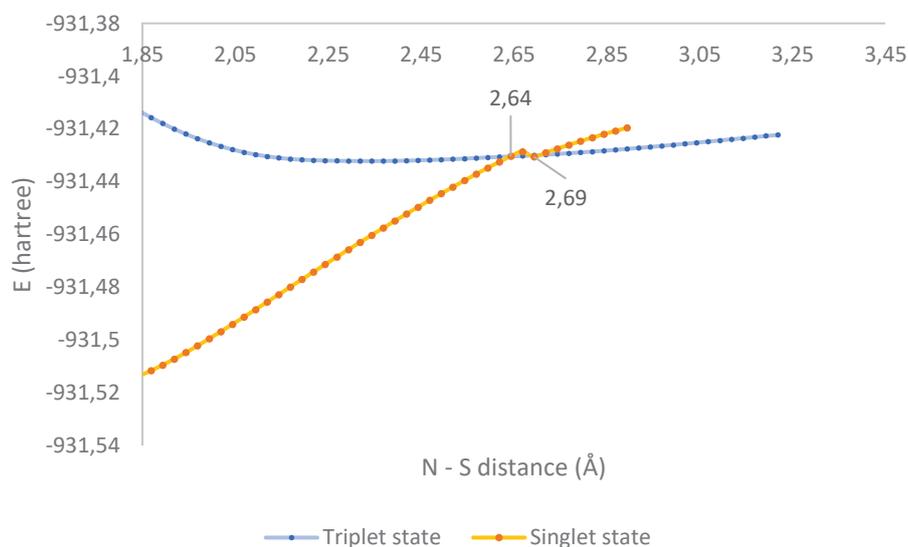


Scheme 188: Investigated radical mechanism for the NS coupling and disulfide bond formation by computational chemistry

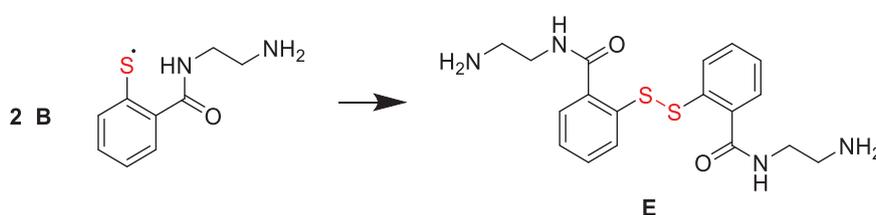
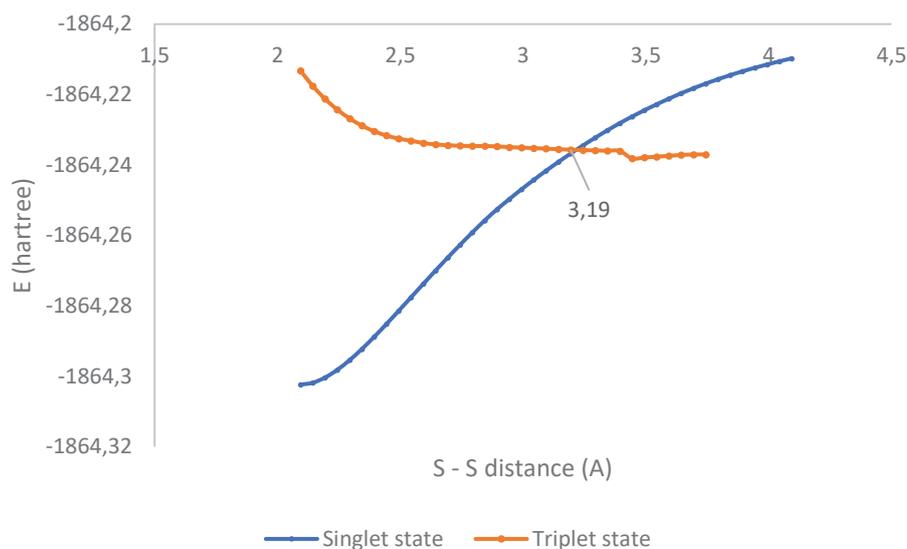


Scheme 189: Energy landscape of the NS and SS coupling on model compound A

The abstraction of the hydrogen atom can occur either on the S (thiol) or the N (amide) atom. DFT clearly showed that the abstraction was easier for the S atom (**A** → **B**) than for the N atom, from 17.87 kcal.mol⁻¹. Then, the formation of a second radical centre on the remaining heteroatom was found to be easier with triplet oxygen (**B** → **C**) than with hydroperoxyl radical, from 9.15 kcal.mol⁻¹. The last step consisted in evaluating the energetic cost of the recombination of the diradical intermediate **C** into **D**. However, DFT is a mono-determinant method, which is not suitable when a spin transition triplet → singlet is involved in the TS. TD-DFT could be a solution and was widely employed in photochemistry modelling³⁷⁹, as well as other approaches such as CAS-DFT³⁸⁰. But this method could not be tested here. Therefore, a simpler approach was followed. A relaxed scan of the N – S distance gave an approximative representation of the two triplet and singlet surfaces (Scheme 190). An intersection of the potential curves will be an estimation of the conical intersection, based on the Marcus theory³⁸¹. The scans were performed in restricted and unrestricted DFT. On our system, the curves did not form parabola. The intersection indicated a distance between the N and S atoms of 2.6 Å but the absence of parabola regarding the two curves avoided us to estimate a free energy of reaction. The combination of the two radicals was postulated to be fast and to cost few energies, involving no rearrangement nor distance variation.



*Scheme 190: Relaxed and unrestricted scan of the N – S distance for diradical **C**. Energy in hartree vs distance in Angstroms between the two radicals. At each step, the geometry is optimised, and a single point (SP) calculation is performed in both singlet and triplet states.*



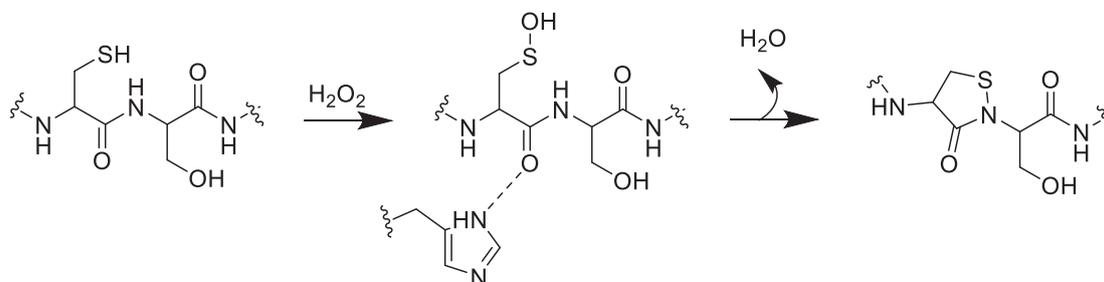
*Scheme 191: Relaxed and unrestricted scan of the S – S distance between two radical **B** Energy in hartree vs distance between the two radicals. At each step, the geometry is optimised, and a SP calculation is performed in both singlet and triplet states.*

The same approach was followed for the recombination of two thiol radicals **B** (Scheme 191). The two curves were not parabola once again. That suggested that the scans were not to evaluate a free energy of reaction regarding the recombination of diradical species

Simple DFT calculations was not adapted to study the proposed radicalar mechanism.

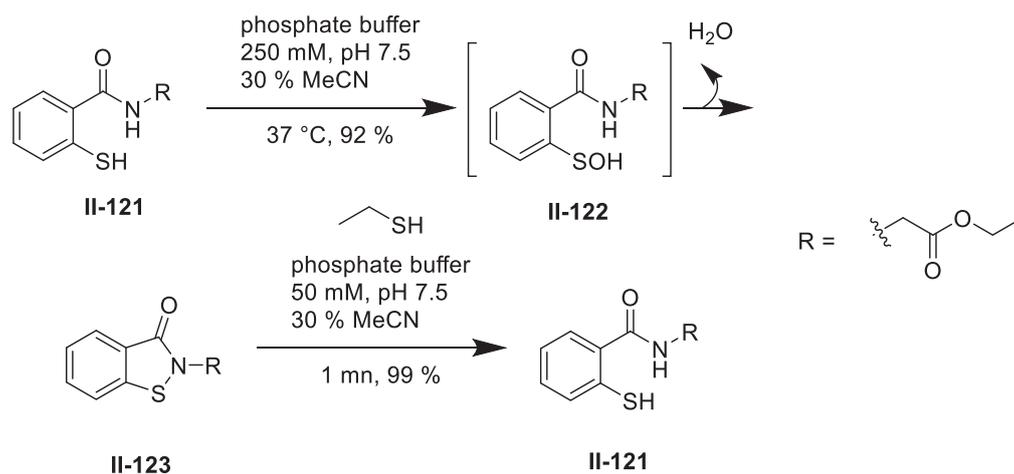
- ii. Description of a recently proposed electrophilic mechanism regarding the formation of sulfenyl amides

The formation of sulfenyl amide between a cysteine and a serine assisted by an histidine was reported in the literature on the protein tyrosine phosphatase 1B (Scheme 192)³⁸².



Scheme 192: Formation of a sulfenyl amide between cysteine 215 and serine 216, assisted by histidine 214 of the tyrosine phosphatase 1B

Later, a study modelled this formation of sulfenyl amide thanks to thiophenol bearing amide bond and amines side chain in ortho position (*Scheme 193*). In an aqueous buffer, thiophenol **II-121** spontaneously underwent overoxidation into a non isolated and unstable sulfenic acid **II-122**. **II-122** formed spontaneously a cycle by a sulfenyl amide bond formation and a release of water in a very high yield of 92 % in **II-65**. The authors suggested that the sulfenic acid was sufficiently electrophilic to undergo cyclisation with the amide. Interestingly, the use of ethanethiol enabled to reopen the cycle quickly and completely, releasing the thiol **II-121**. Ethanethiol acted as a reductor of the inactive benzoisothiazolidinone. A subsequent theoretical study by DFT revealed that the presence of a substituent in second ortho position of the thiophenol **II-121** favoured the cyclisation compared to **II-121** by facilitating a closer approach of the amides (NH) and sulfenic acid (OH) moieties but also by increasing the electrophilicity of the sulfur atom in the sulfenic acid³⁸³. Another computational study confirmed the previous affirmations and highlighted the important role of nonbounded S – N interactions³⁸⁴ as well as the influence of the substituent R of the amide facilitating the cyclisation by electronic assistance. Finally, a study of electrophilic sulfenic acids derived from cysteine showed that the 5-membered cycle could be opened in presence of C-nucleophile, such as piperidinedione or barbituric acids³⁸⁵. Our 1,4-bisthiophenols with side chain connections by amide bond in the 2,5 positions of the aromatic core corresponded to those studies highlighting the elements facilitating the formation of cyclic sulfenyl amides.



Scheme 193: Modelling of the sulfenyl amide formation^{382,384-385}. *R = various groups*

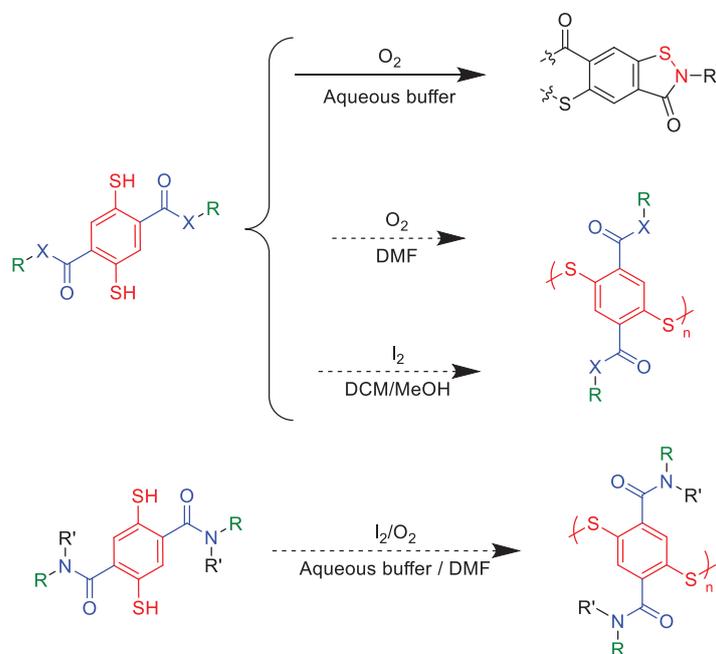
Based on these studies, the previous radicalar mechanism should be corrected. The basic functional groups such as amines bore by the amide of our 1,4-bisthiophenols in the 2,5 positions probably trigger the formation of cyclic sulfenyl amide. Given that the sulfenyl amide **II-123** could be opened, a control the redox potential during the DCLs of our 1,4-bisthiophenols could help to overcome the NS coupling.

The litterature explained the formation of sulfenyl amides by an ionic mecanism. It proceeded probably by an oxidation of thiols into sulfenyl acids, then a nucleophilic attack of the NH moieties on the electrophilic sulfur atom with a release of water. A reduction could open the cyclic and give the starting compounds.

c. Overcoming the NS coupling on N₂,N₅-bis(aminoalkyl)-1,4-bisthiophenols

The aim of this subpart was to find experimental conditions in order to favor the cyclo-oligomerisation of 1,4-bisthiophenols by disulfide bond formation, minimising the formation of cyclic sulfenyl amides.

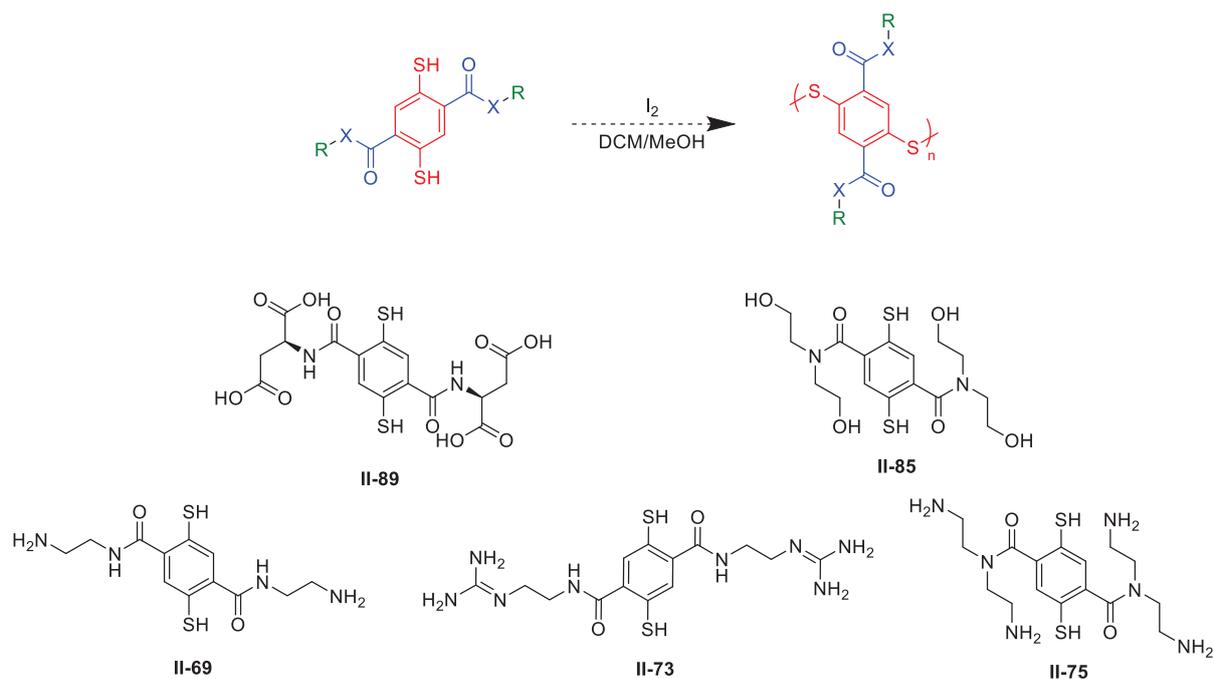
Different solutions were alternatively explored (Scheme 194). A quick oxidation of thiols into disulfide bond by iodine was investigated even though the NS coupling was an oxidation and could also be accelerated by iodine. Alternatively, tertiary amides were used instead of secondary amides, blocking the cyclisation. Finally DMF was evaluated as solvent. Indeed, the possible interaction between thiolate by the neighbouring aminonium through a stable salt bridge may be disrupted by a strong hydrogen bond donating/accepting solvent such as DMF compared to water³⁸⁶.



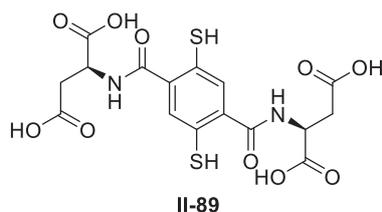
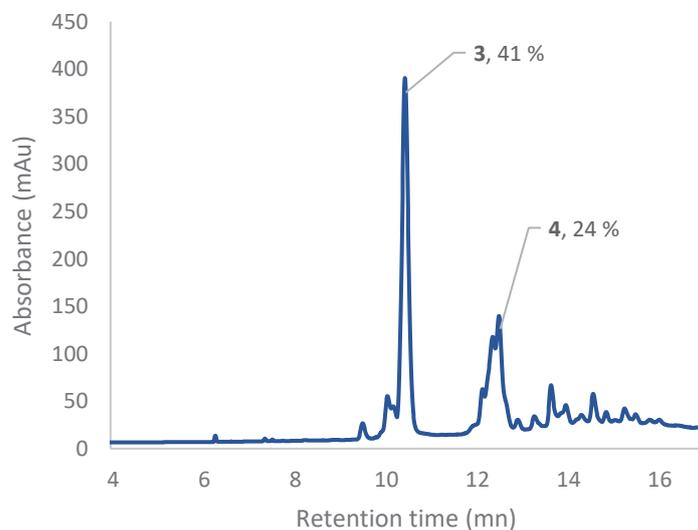
Scheme 194 : Strategies to avoid the formation of cyclic sulfenyl amides during the connection of 1,4-bisthiophenol monomers by disulfide bond formation. X: O, NH. Connection between the monomeric units forming the principal chain (red) and connection between the principal chain and the side chain bearing the functional groups (blue). R: functional groups dedicated to the recognition of modified proteins, R' alkyl groups.

- Oxidation with iodine

The iodine oxidation was tested in aqueous buffers and in organic media (DCM, DCM/MeOH, MeOH) for 5 functionalised 1,4-bisthiophenol monomers. The building blocks were dissolved to reach a concentration of 4 mM and one equivalent of iodine was added dropwise over 5 mn¹¹². The resulting mixtures were analysed by HPLC and ESI-MS analysis. The addition of iodine aqueous media (milliQ water, pH adjusted at 7.5) only resulted on overoxidation of thiols into sulfenic acids (ESI-MS analysis). In the case of **II-65**, **II-85**, **II-73**, **II-69** dissolved in a mixture of DCM/MeOH 3:1, precipitates were formed upon iodine addition. ESI-MS analysis of both supernatant and precipitates redissolved in DMSO revealed the presence of species containing sulfenic acids and sulfenyl amides. Regarding **II-89** dissolved in pure methanol, discrete cyclo-oligomers were formed in 5 mn without overoxidation of thiols, which was in contrast with the DCL in aqueous buffers and oxygen. The resulting mixture was analysed by LC-MS and HPLC and was mainly composed of cyclic trimers (41 %) and tetramers (24 %) (Scheme 196). Furthermore, these cyclo-oligomers were different from the one of the same size obtained in the DCL of **II-89** with oxygen and aqueous buffers given their different retention time on the same HPLC method. This suggested that iodine oxidation would differ in terms of diastereoselectivity than oxidation with air. It could be interesting to isolate these cyclo-trimer and -tetramer by precipitation or preparative HPLC.

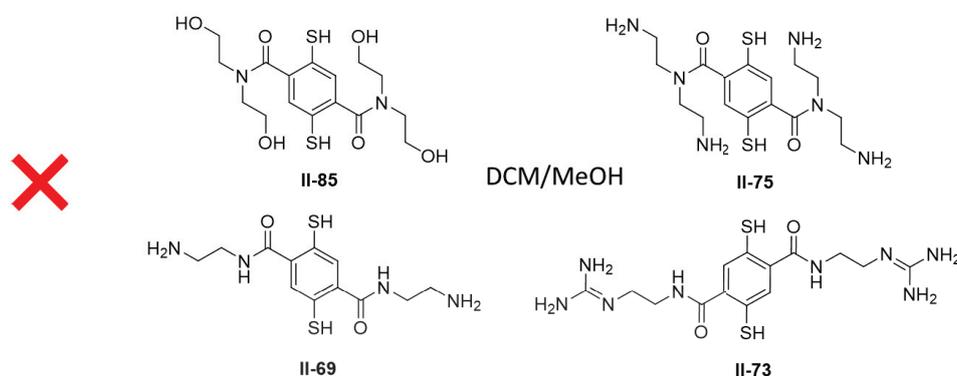
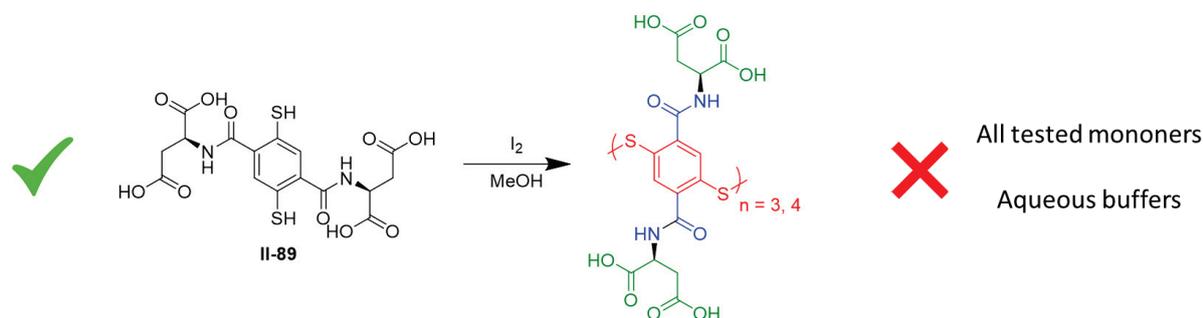


Scheme 195: Investigating iodine as an alternative oxidant that oxygen. The aim was to avoid the NS coupling observed in aqueous buffers. The iodine oxidation was tested in aqueous buffers and in organic media for 5 functionalised 1,4-bisthiophenol monomers. X: O, NH. Connection between the monomeric units forming the principal chain (red) and the connection between the principal chain and the side chain bearing the functional groups (blue). R: functional groups dedicated to the recognition of modified proteins.



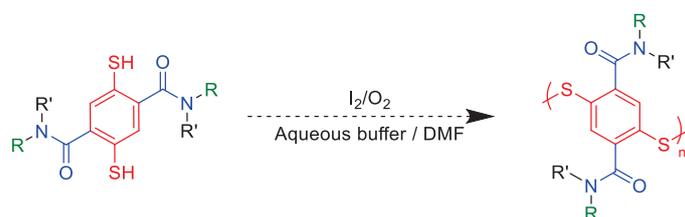
Scheme 196: Composition of the crude mixture of II-89 with iodine in MeOH. n: cyclic oligomer composed of n units

Nevertheless, the success of iodine to avoid NS coupling and overoxidation was limited (Scheme 197). It failed for 1,4-bisthiophenols bearing basic groups such as amines. Fortunately, this approach solved a part of the problem for 1,4-bisthiophenol **II-89** bearing acidic groups. Iodine was useful for all the dithiol building blocks bearing carboxylic acids that were studied in DCC in this thesis.



Scheme 197: Summarisation of the iodine alternative for an oxidation of thiols into disulfide bond without NS coupling. The carboxylate derivatives gave discrete disulfide macrocycles in methanol, while other monomers formed precipitates. The iodine addition in aqueous buffers resulted in precipitation for all monomers. X: O, NH. Connection between the monomeric units forming the principal chain (red) and the connection between the principal chain and the side chain bearing the functional groups (blue). R: functional groups dedicated to the recognition of modified proteins.

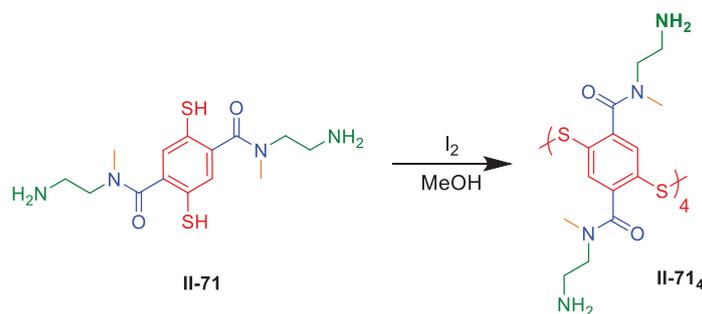
- Tertiary amides in order to avoid the cyclisation during the formation of the sulfenyl amide



Scheme 198: Using tertiary amides in order to avoid the NS coupling observed in aqueous buffers. Connection between the monomeric units forming the principal chain (red) and the connection between the principal chain and the side chain bearing the functional groups (blue). R: functional groups dedicated to the recognition of modified proteins. R' alkyl groups.

The secondary amides of our functionalised 1,4-bisthiophenols could suffer from NS coupling because of their available hydrogen and poor steric hindrance, at the opposite tertiary amides (Scheme 198). In aqueous buffers, the building blocks **II-71** quickly underwent overoxidation of thiols into sulfenic acids without formation of cyclic sulfenyl amide as expected. In addition, a cyclo-tetramer was

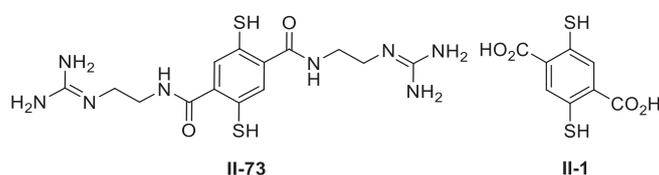
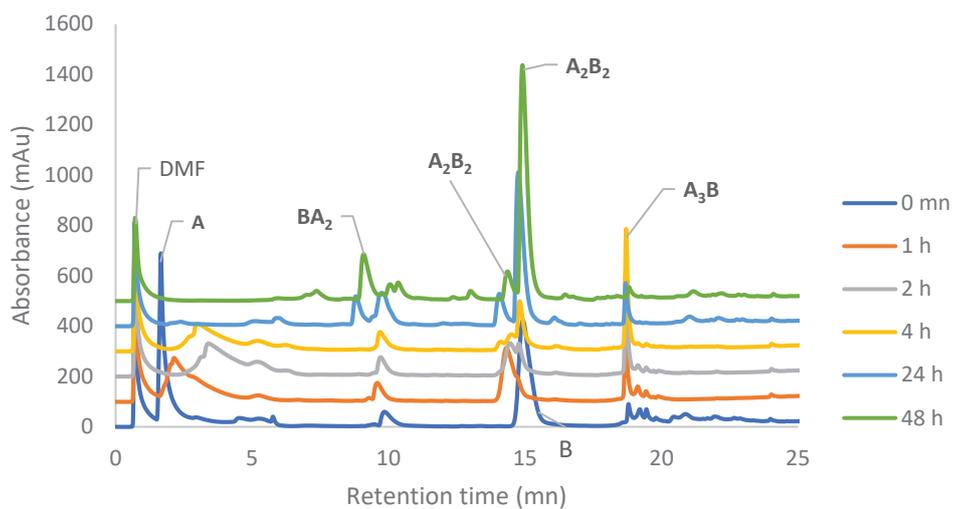
detected after addition of iodine into a solution of **II-71** in methanol (same protocol as above). The object was identified by HRMS analysis but no further isolation or characterisation was made.



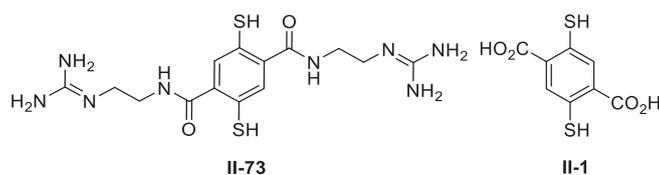
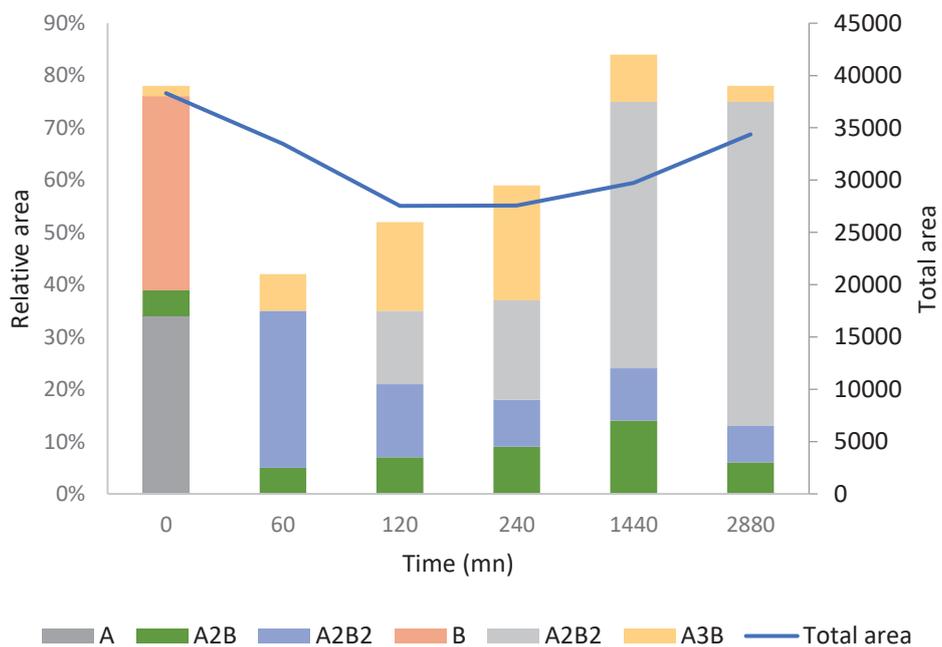
Scheme 199 : Tertiary amide blocking the NS coupling. Connection between the monomeric units forming the principal chain (red) and the connection between the principal chain and the side chain bearing the functional groups (blue). Green: functional groups dedicated to the recognition of modified proteins, orange: alkyl groups blocking the NS coupling.

- Oxidation with oxygen in DMF instead of aqueous buffers

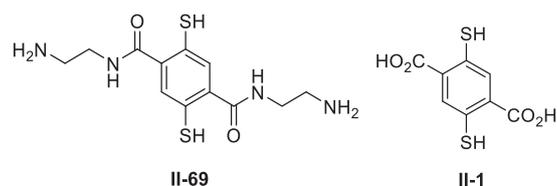
The possible interaction between thiolate by the neighbouring aminonium through a stable salt bridge may be disrupted by a strong hydrogen bond donating/accepting solvent such as DMF compared to aqueous buffers³⁸⁷. The same methodology as in aqueous media was followed with basic N_2,N_5 -bis(aminoalkyl)-1,4-bisthiophenols. On the one hand, regarding DCL of these dithiol alone, without a combination with other building blocks, ESI-MS analysis showed overoxidation of thiols into sulfenyl acids and sulfenyl amides, as well as linear dimers by disulfide bond formation, similarly to the analysis recorded in aqueous buffers. On the other hand, cyclohetero-oligomers of **II-1** with **II-73**, **II-65** and **II-69** were detected (*Scheme 200*, *Scheme 201*, *Scheme 202*, *Scheme 203*, *Scheme 204*, *Scheme 205*). Indeed, heterotetrameric macrocycles of formula $(\text{II-1})_3\text{X}$, $(\text{II-1})\text{X}(\text{II-1})\text{X}$ and $(\text{II-1})_2\text{X}_2$ were observed for $\text{X} = \text{II-69}$ and **II-73**, without overoxidation of thiols and NS coupling. The three libraries behaved similarly. The library based on **II-73** and **II-69** are presented below. The one generated from **II-65** had a similar profile but the LC/MS analysis could not be performed for technical reasons. ESI-MS analysis confirmed the presence of similar macrocycles.



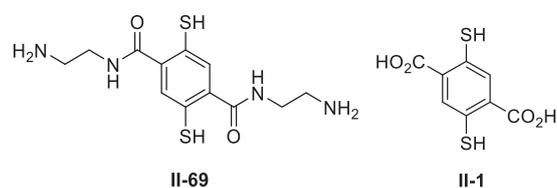
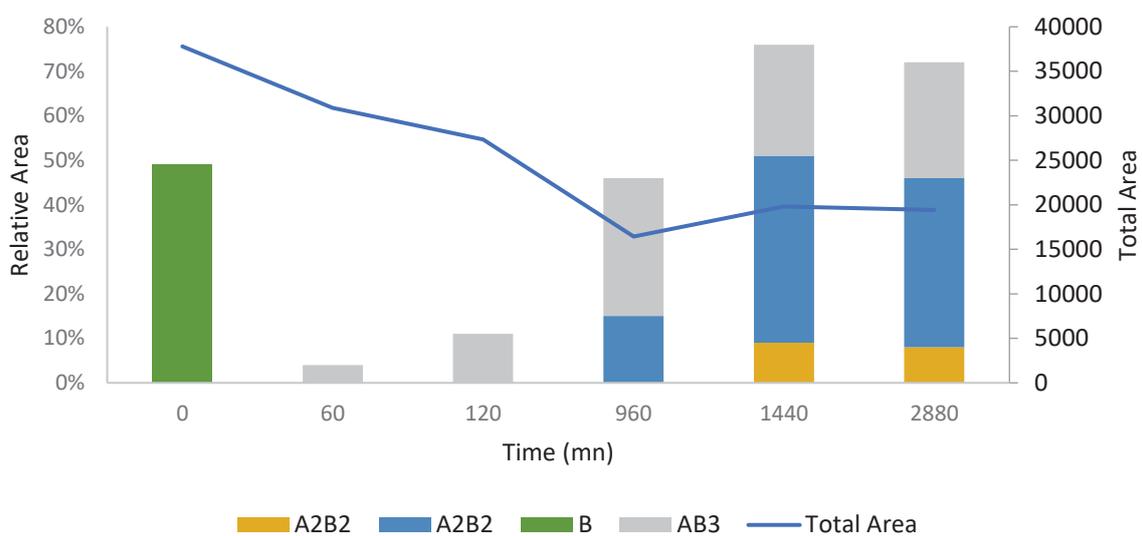
Scheme 200: Evolution of **II-73** = **A** and **II-1** = **B** (4 + 4 mM) in DMF, method D. X_nY_m cyclic disulfide oligomer composed of n X units and m Y units



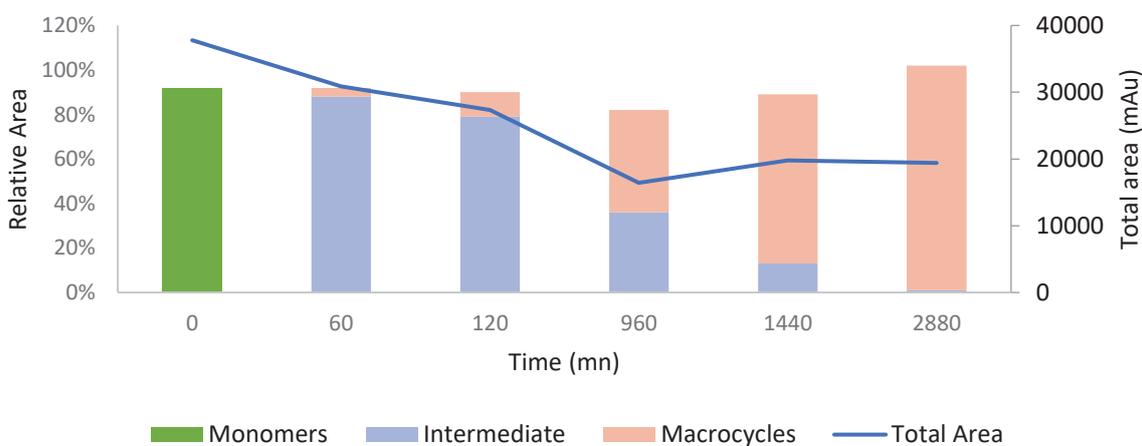
Scheme 201: Building block and macrocycle population evolution between **II-73** = **A** and **II-1** = **B** (4 + 4 mM) in DMF (1) X_nY_m cyclic disulfide oligomer composed of n X units and m Y units



Scheme 203: Evolution of **II-69 = A** and **II-1 = B** in DMF (4 mM + 4 mM), method D. X_nY_m cyclic disulfide oligomer composed of n X units and m Y units.



Scheme 204: Building block and macrocycle population evolution between **II-69 = A** and **II-1 = B** (4 + 4 mM) in DMF (1) X_nY_m cyclic disulfide oligomer composed of n X units and m Y units.



Scheme 205: Population evolution between **II-69** and **II-1** (4 + 4 mM) in DMF (2)

The influence of various templates at 2 mM (polyamines, aromatic amines, amino acids, carboxylic acids and peptides **II-125-158**) on the repartition of the cyclo-oligomers obtained from the cyclo-oligomerisation **II-69** and **II-1** at 4 mM respectively in DMF was studied (Appendix II-VII: Set of amplification data for the DCLs between **II-69** and **II-1** (4 + 4 mM) in DMF). The libraries were analysed after 12 h by HPLC. The same study was done on **II-69** alone at 8 mM with 2 mM of templates. Amines and carboxylic acids were chosen because of the possible electrostatic interaction that they could develop with the monomeric units **II-1** and **II-69** respectively. Amino acids and peptides could develop electrostatic interactions with both monomeric units. Concerning the DCLs of **II-69** alone, the templates did not help to avoid the overoxidation of thiols into sulfe/inyl acids and cyclic sulfenyl amides. Regarding the DCLs combining **II-69** with **II-4**, the repartition of the different cyclohetero-oligomers evolved slightly without a dominant amplification, compared to the untemplated library. The details of the quantitative are available in appendix (Appendix II-VII: Set of amplification data for the DCLs between **II-69** and **II-1** (4 + 4 mM) in DMF). The DCL with template **II-126** led to the best amplification of the cyclohetero-tetramers. A separation of the crude mixture on a semi-preparative C₈ reverse phase failed to isolate (**II-1**)₃(**II-69**) and two (**II-1**)₂(**II-69**)₂, which did not precipitate under acidification. This work could be done again with an optimisation of the separation and a direct lyophilisation of the fractions.

Finally, the addition of iodine in a solution of **II-1** dissolved with alternatively **II-69** and **II-73** in DMF failed to produce discrete cyclo-oligomers. The method should be applied in methanol since this media was successful regarding **II-89**. This study proved that the DMF was a good media to make cyclohetero-oligomers by oxygen oxidation between **II-1** and 1,4-bisthiophenol bearing basic amino side chains. However, the proportions of cyclohetero-tetramers were not amplified in presence of templates.

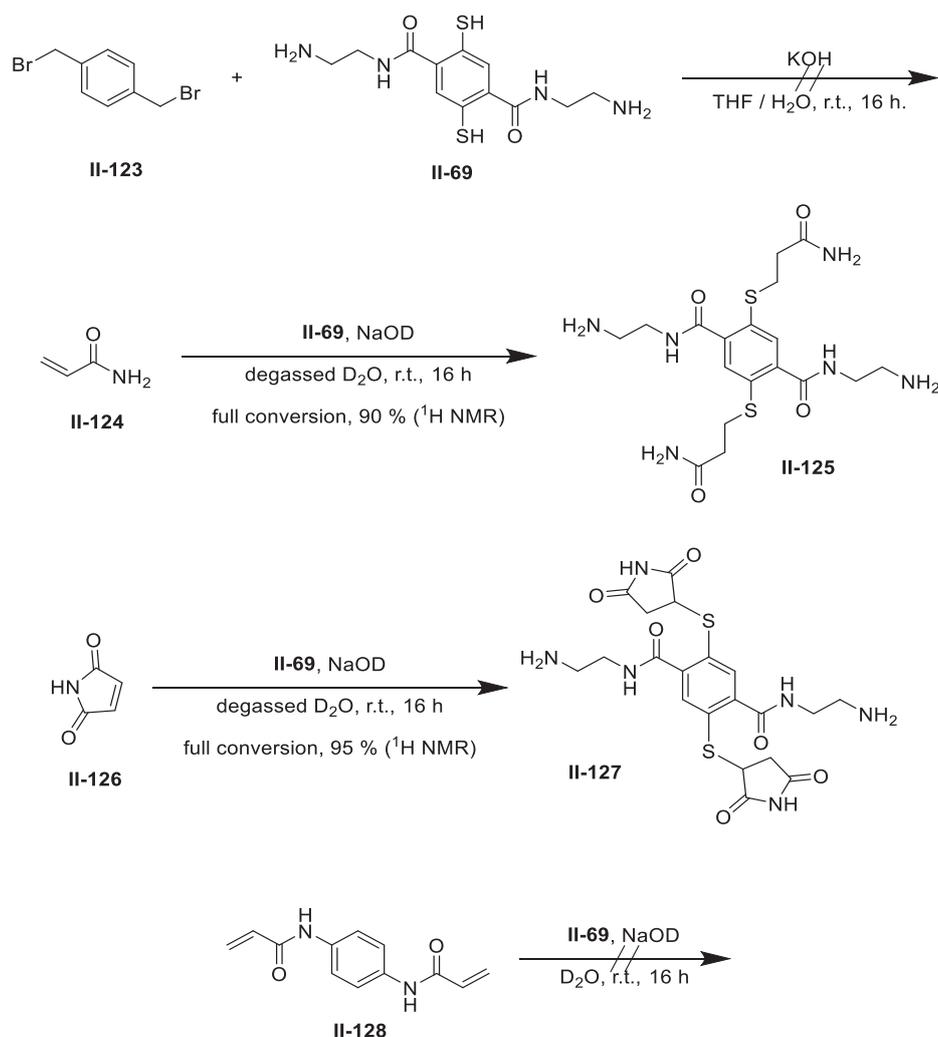
*Cyclohetero-tetramers between **II-1** and 1,4-bisthiophenols bearing basic amino side chains such as **II-73/69** were formed in DCL that are conducted in DMF with oxygen as oxidant.*

d. Exploration of the nucleophilicity of 1,4-bisthiophenols

The aim of this subpart was to connect the monomers following the strategy **A** by irreversible bonds thanks to the nucleophilicity of the thiols of the synthesised N₂,N₅-1,4-bisthiophenols. Macrocycles can be synthesised by virtue of synthetic chemistry rather than self-assembly³⁸⁸, especially by nucleophilic substitution^{278,389} between thiols and bromide derivative, forming thioether-bridge cyclophanes. The corresponding yields are sometimes good³⁹⁰ but usually poor³⁹¹ due to the poor selectivity between oligomerisation and polymerisation. The nucleophilicity of the thiols of **II-69** were benchmarked as an alternative for the formation of static macrocycles with dielectrophiles such as 1,4-bis(bromomethyl)benzene **II-123** and with model compounds such as acrylamide (*Scheme 206*). The concentrations were fixed at the millimolar range and the addition rates were low in order to favour cyclisation rather than polymerisation. Degassed THF/water or water were used. Firstly, a nucleophilic substitution of **II-69** with **II-123** α,α' -dibromo-*p*-xylene³⁹² was attempted and failed to create a new cyclophanes. ESI-MS and ¹H NMR analysis showed the presence of the starting α,α' -dibromo-*p*-xylene **II-123** and 1,4-bisthiophenols, suggesting a poor reactivity in the tested concentrations.

Acrylamide **II-124** and maleimide **II-126** were chosen as model compounds for a nucleophilic attack of the thiols given their electrophilicity, even though the thiol/ene addition was a reversible reaction. **II-69** was dissolved in the millimolar range in presence of acrylamide **II-124** in degassed water basified with sodium hydroxide (pH 8)³⁹³. ¹H NMR analysis revealed a total disappearance of the alkene signals

and ESI-MS analysis confirmed the presence of the desired adducts of the nucleophilic addition (**II-125**). Moreover, **II-126** maleimide³⁹⁴ was also successfully tested as Michael acceptor (ESI-MS analysis + ¹H NMR). The conversion was total (**II-127**). Given that these two acceptors reacted with **II-69**, an electrophile containing two electrophiles sites was employed in order to directly form cyclophanes. **II-128** was obtained in one step from 1,4-phenylenediamine³⁹⁵ with a yield of 62 %. A formation of an alternate macrocycle was targeted. Yet, despite several attempts, ¹H NMR and ESI-MS analysis indicated that no hetero-oligomers were formed. A polymerisation might be responsible and the addition rate of **II-128** should be decreased.



Scheme 206 : Nucleophilic reactivity of bromide derivative and Michael's acceptors with II-69

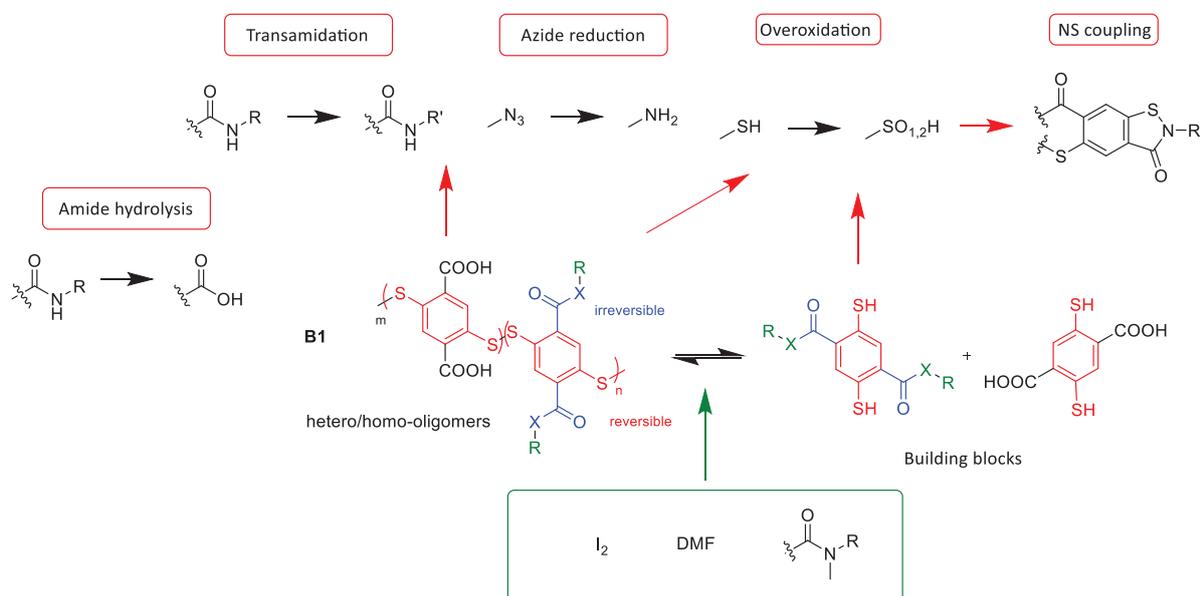
The nucleophilic attack of the thiols of 1,4-bisthiophenol **II-69** on bromide derivatives and Michael's acceptors was validated on model compounds, but the methodology was not optimised to form hetero-oligomers.

The thiols of functionalised 1,4-bisthiophenols such as II-69 could be used in nucleophilic attacks on bromide derivatives and Michael's acceptors.

e. Conclusion

The 1,4-bisthiophenols bearing side chains (mainly basic amino side chains) on the 2,5 position of the aromatic core showed an unexpected behaviour in aqueous solution (Scheme 207). Instead of forming

disulfide bond by a slow oxidation with oxygen in DCLs, the thiols were quickly oxidised into sulfenic acids and into cyclic sulfenyl amides. This oxidation was first supposed radicaler, but based on the litterature, the process was a sequence of oxidation of the thiols into sulfenic acids and then a nucleophilic attack of the NH moieties of the amide bond to the electrophilic sulfur of the sulfenyl acids. The cyclisation was probably favoured by the assistance of amines of the side chains. The use of DMF as solvent for the DCL and the use of 1,4-bisthiophenols with tertiary amides in the 2,5 positions avoided the cyclisation into sulfenyl amides. Regarding an example of tertiary amide, addition of iodine into a solution of 1,4-bisthiophenols led to a cyclic-tetramers. Cyclic hetero-oligomers between **II-1** and 1,4-bisthiophenols bearing amino side chains were observed, as well as homoligomers of **II-89**, without successful isolation. Nevertheless, overoxidation of thiols into sulfenic acids remained. In addition, the thiols of functionalised 1,4-bisthiophenols showed nucleophilic properties with bromide derivatives and Michael's acceptors. It could be used for the synthesis of static cyclophanes following the strategy **A** of cavitand design. A new design of aromatic building blocks could also be a solution, with more distant thiols and amide moiety, avoiding the cyclisation into sulfenyl amides.



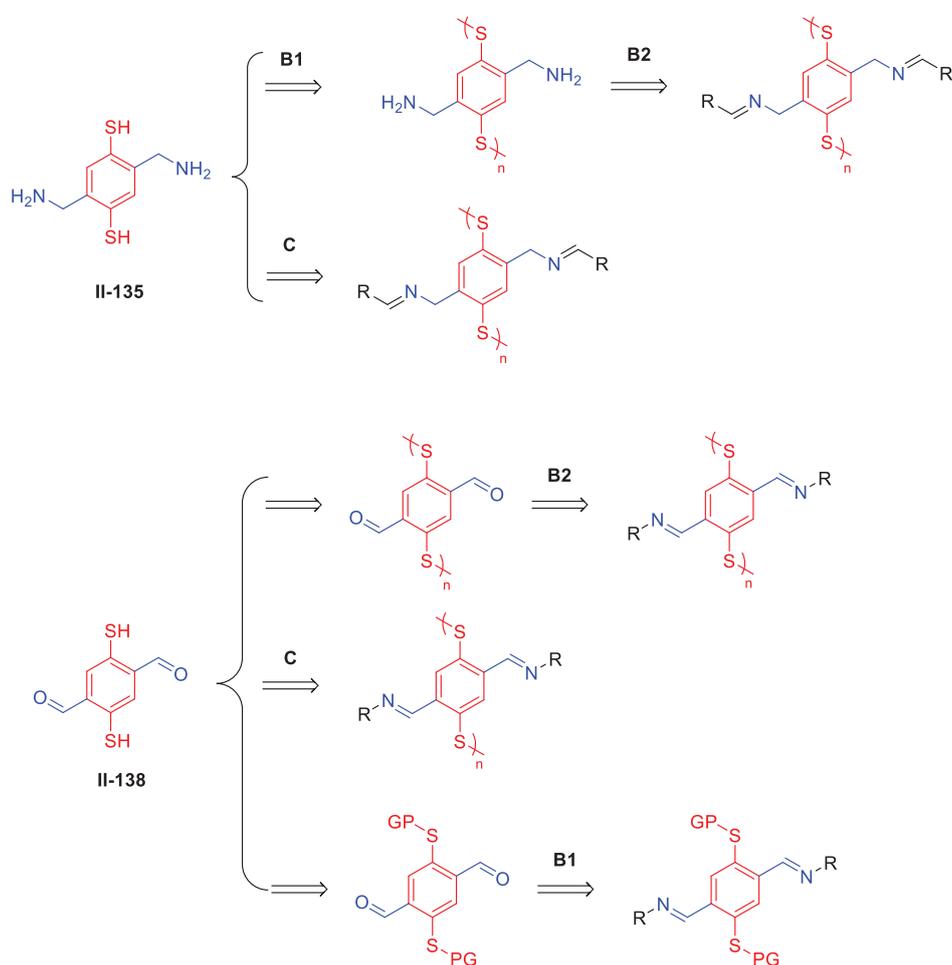
*Scheme 207: Summary of the activity of 1,4-bisthiophenols in DCLs. Following the strategy **B1**, the connection of monomers into the principle chain (red) will be obtained by disulfide bond formation, while the side chain bearing the functional groups dedicated to the recognition of modified proteins (green) will be connected by irreversible amide bond (blue) to the principle chain before the disulfide bond formation by DCC. The DCLs suffered from many undesired reactivities, highlighted in red. The use of iodine, DMF and disubstituted primary amide tackled a part of the issues.*

The series of functionalised 1,4-bisthiophenols underwent undesired reactivities such as overoxidation of thiols into sulfenic acids and sulfenyl amides, preventing their connection by self-assembly by disulfide bond formation. These undesired side reactivity was partially rationalised and avoided.

6. Towards combination of disulfide and hydrazone/imines exchanges for the synthesis of cavitants

In this subchapter, the strategy **C**, which consist in the cyclo-oligomerization and monomers functionalization using two orthogonal reversible linkages, was explored. 1,4-bisthiophenol bearing

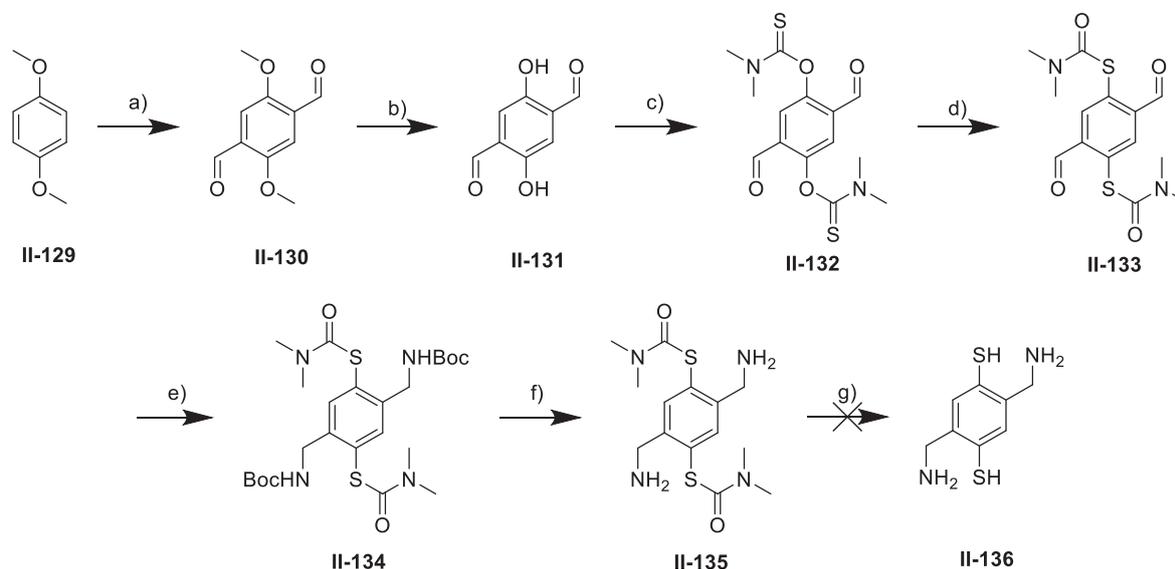
either aldehydes, acylhydrazone, and amines in the 2,5 positions of the aromatic core were targeted. Those functions are precursors of disulfides, acylhydrazones and imines. Disulfide exchange and acylhydrazone or imine exchange can occur simultaneously in basic conditions. In acidic conditions, the disulfide exchange is frozen while the other exchanges take place. In addition, 1,4-bisthiophenols bearing amines in 2,5 positions of the aromatic core are interesting to develop salt bridges with the side chains of the targeted proteins, following their protonation state in aqueous solutions. Besides, Consequently, an amino building block could be used for the strategy **C**, **B1** and **B2** given that an amide bond will be used to connect the side chain to the dithiophenol monomer. Regarding the aldehyde moieties, they could be used for post-functionalisation if the corresponding building block forms the principal chain of the cavitand. In addition, the thiols could be protected, enabling the construction of the side chain before deprotection and connection of the monomer into the principal chain of the cavitand. Such a building block corresponds to the strategy **C**, **B1** and **B2**. Based on the structure of **I-1**, the targets are presented below (Scheme 208).



*Scheme 208: Targeted building blocks for the synthesis of cavitands following the strategies **B1**, **B2** and **C**. Connection between the monomeric units forming the principal chain (red) and the connection between the principal chain and the side chain bearing the functional groups (blue). The strategies differ from their level of dynamic connections. R: functional groups dedicated to the recognition of modified proteins.*

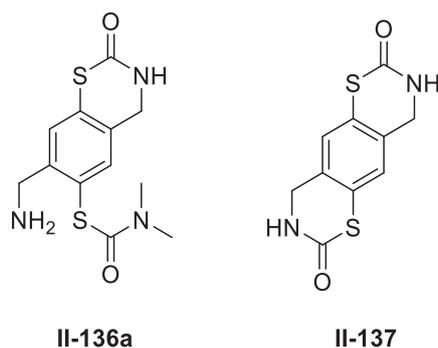
- Synthesis of 1,4-bisthiophenol bearing amines in the 2,5 positions of the aromatic core

This subpart develops the synthesis of the 2,5-bis(aminomethyl)benzene-1,4-dithiol (Scheme 208), beginning with the synthesis of **II-131** in a large scale following known procedures (Scheme 209). Formylation of 1,4-dimethoxybenzene **II-129** gave **II-130** with a yield of 51 %³⁹⁶. Deprotection of the methoxy groups by BBr_3 ³⁹⁶ with a yield of 68 %, followed by O-thiocarbamation¹⁶ of free alcohols led to **II-132** with a yield of a 71 %. A Newman-Kwart rearrangement changed O-thiocarbamate (**II-132**) into S-thiocarbamate (**II-133**) in NMP with a yield of 55 %.



*Scheme 209: synthesis of **II-136**; a) 1. $n\text{-BuLi}$, TMEDA, Et_2O , reflux, 16 h. 2. 1-formylpiperidine, r.t., 1 h. 51 %; b) 1. BBr_3 , DCM, r.t., 3 h. 2. H_2O , r.t., 1 h. 68 %; c) dimethylthiocarbamoyl chloride, DABCO, DMA, r.t., 48 h. 71 %; d) 210°C , 9 mn, NMP, 55 %; e) BocNH_2 , TFA, Et_3SiH , CHCl_3 , 40°C , 5 d., 67 %; f) TFA, tetrachloroethane, r.t., 20 mn, 81 %; g) see supplementary material*

Direct one pot reductive amination³⁹⁷ of **II-133** involving NH_4OAc , NaCNBH_3 in ethanol and 30 % aqueous ammonia failed to provide the desired diamine **II-135**. The crude mixture was positive to the Ellman's reagent³⁹⁸. Thiols react with 5,5'-dithiobis-(2-nitrobenzoic acid), cleaving its disulfide bond to give 2-nitro-5-thiobenzoate (TNB^-), and then yellow TNB^{2-} dianion in neutral pH. The observed yellow colour corresponded to the presence of thiolate generated from the S-thiocarbamates of **II-133**. The conversion of aldehydes was total under these conditions, attested by ESI-MS and ^1H NMR analysis. Thus, a two-step approach was preferred. A reductive amination³⁹⁹ of **II-133** with tert-butyl carbamate provided **II-134** with a yield of 67 %. ^1H NMR analysis of the reaction mixture showed that the kinetics of both imine formation and imine reduction were slow and required several days to complete. The electrophilicity of the aldehydes were decreased by the aromatic core of **II-133**. Deprotection of the Boc protecting group of amines of **II-134** into **II-135** was achieved using TFA³⁹⁹ in tetrachloroethane with a yield of 81 %. the choice of the solvent was guided by a preliminary screening. Indeed, the product was found to be degraded rapidly (> 1h) in DCM. The degradation process was not entirely elucidated. ESI-MS analysis combined with ^1H NMR analysis of the crude mixture showed that the primary amine moiety was reacting with the S-thiocarbamate in ortho position, leading to a bicyclic **II-136a** and tricyclic adduct **II-137** (Scheme 210). This behaviour was previously observed on similar substrates (6-membered aromatic rings with S-thiocarbamate in the 1,4 positions, and other functionalities in the 2,5 positions) by Dr. Dumartin and Dr. Pascal, former PhD students of the group^{82a,c}. Other conditions, involving trimethylsilane iodide failed to deprotect the S-thiocarbamate⁴⁰⁰.



Scheme 210: Cyclic undesired product obtained from the deprotecting reaction of II-135

At this stage, the experimental yields of the sequence were appreciable, reproducible, scalable and in good agreement with the literature. However, the final deprotection of S-thiocarbamates of **II-135** in basic conditions failed. ESI-MS and ^1H NMR analysis of the reaction were employed to monitor the composition of the crude mixture. The conditions of reaction involved sodium hydroxide (at various equivalent between 4 and 32) in alcoholic/aqueous media (temperature ranging from different temperatures and reaction times between 1 and 4 hours (Appendix II-II: Reaction conditions). Hydrides such as DIBAL-H^{352b,401} were also tested and proved to be inefficient to give the desired product. Indeed ESI-MS analysis showed the presence of species resulting from the attack of the amine on the S-thiocarbamate, leading to the previously mentioned bi- and tricyclic structure. The presence of an amine and thiophenol groups in ortho position is a real challenge to the synthesis of new building blocks, even though the undesired cyclisation depends on many parameters such as the conformational flexibility.

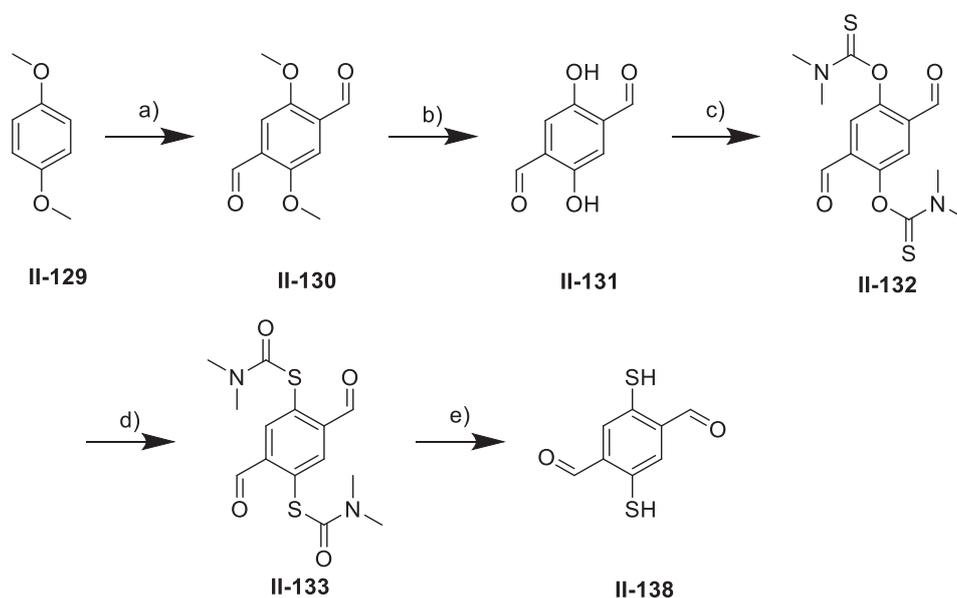
The synthesis of a 1,4-bisthiophenol bearing benzylic amines encountered synthetic issue. The nucleophilic properties of both amines and thiols prevented a deprotection without an undesired cyclisation.

- b. Synthesis of 1,4-bisthiophenol with aldehyde moieties in the ortho positions and subsequent formation of hydrazones

This subpart develops the synthesis of a 1,4-bisthiophenol of interest, bearing aldehydes in 2,5 positions of the aromatic core (Scheme 208). The synthesis relied on the previous synthetic sequence until compound **II-133** (Scheme 211). Then, a deprotection¹⁶ in basic conditions cleaved the S-thiocarbamate of **II-133** and released the thiol groups in **II-138** with a yield of 71%. The product was isolated by acidic precipitation. This step did not prove to be very reproducible and could be done at a scale of 200 mg. The conversion should be carefully monitored by ^1H NMR. Indeed, after the disappearance of the protons signals of the methyl bore by the S-thiocarbamates, the reaction should be immediately stopped by an acidification with HCl. Longer reaction time tended to complexify the crude mixture and gave less pure product. The side-products were not clearly identified.

This building block **II-138** was then study in aqueous solution, regarding its combination with acylhydrazides, its cyclo-oligomerisation and its protection of thiols (Scheme 213). Its solubility was limited to mixtures between DMSO or DMF and aqueous buffers. Sodium acetate 100 mM at pH 7 was the best buffer to dissolve **II-138**. In a DCL of **II-138** at 5 mM with DMF / sodium acetate 100 mM pH 7 (1:1 in volume), a precipitate appeared after 1 hour. The composition of the solid could not be determined because of its insolubility in organic and aqueous media. Neither linear nor cyclic oligomers were detected by ESI-MS analysis in the supernatant and in samples of the DCL before the

appearance of the precipitate. Despite the conditions supposed to be in favour of the cyclo-oligomerisation (millimolar concentration, slightly basic pH) we could not conclude on the properties of self-assembly by disulfide bond formation of **II-138**.

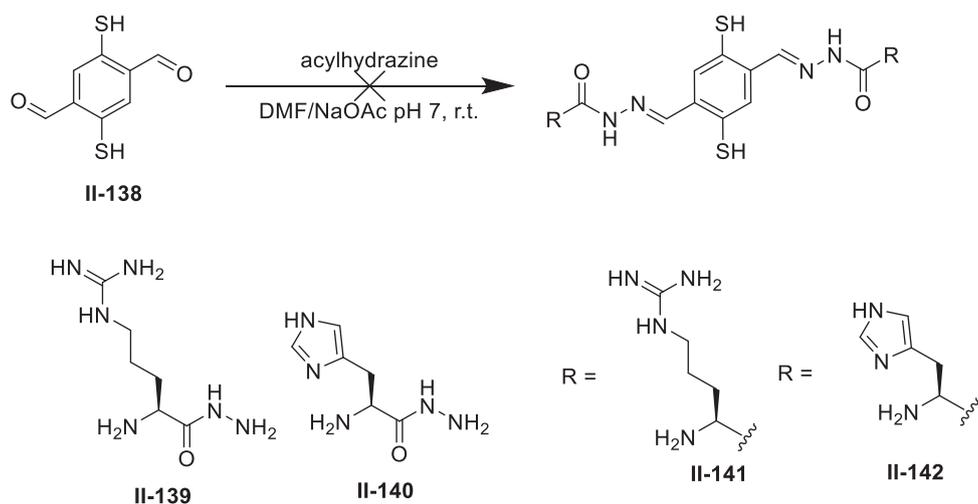


*Scheme 211: Synthesis of II-138 a) 1. *n*-BuLi, TMEDA, Et₂O, reflux, 16 h. 2. 1-formylpiperidine, r.t., 1 h. 51 %; b) 1. BBr₃, DCM, r.t., 3 h. 2. H₂O, r.t., 1 h. 68 %; c) dimethylthiocarbamoyl chloride, DABCO, DMA, r.t., 48 h. 71 %; d) 210°C, 9 mn, NMP, 55 %; e) NaOH, H₂O / isopropanol, 60°C, 6h., 71 %*

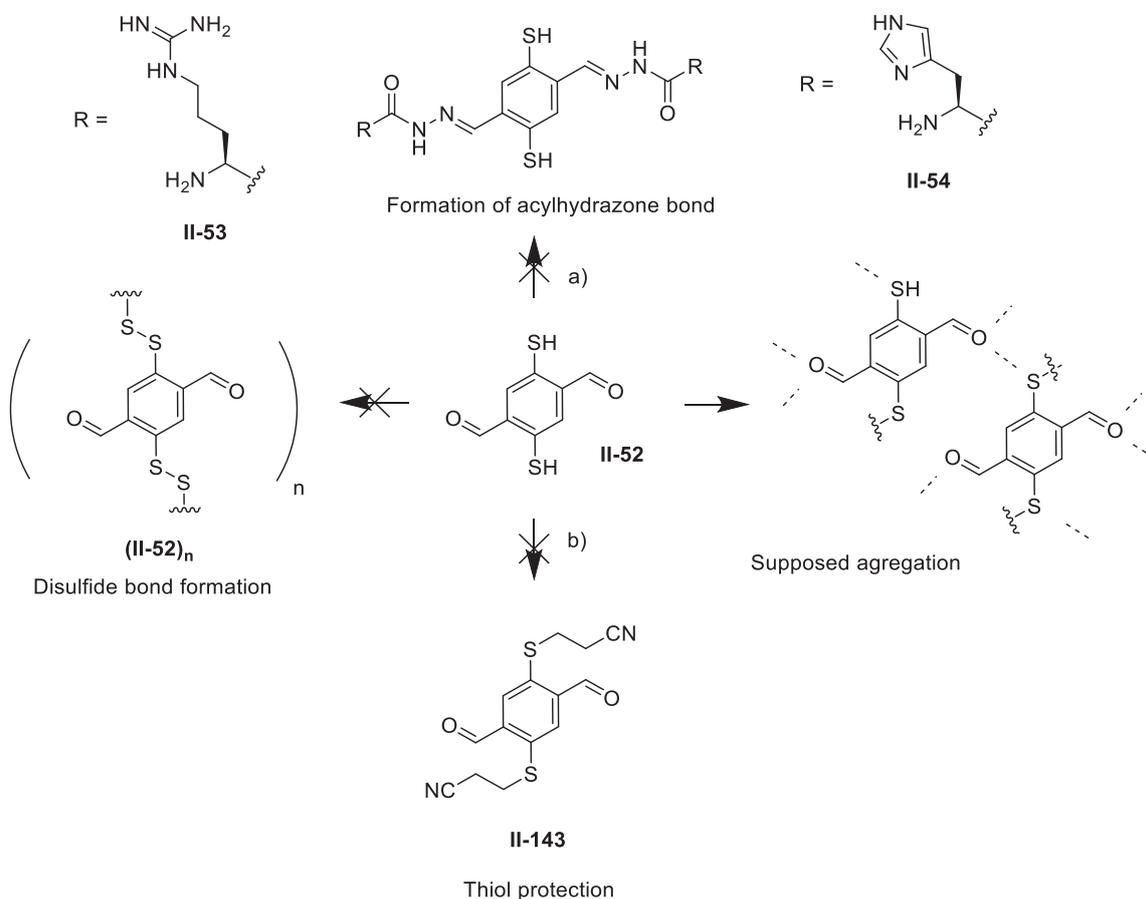
Then, **II-138** was combined in the same experimental conditions (apart from the solvent which was degassed) with hydrazides (**II-139** and **II-140** at 2 equivalents) to form the acylhydrazones **II-141** and **II-142** (Scheme 212). The compounds were soluble, but similarly to the previous experiments, a precipitate appeared after one hour. ESI-MS analysis of the supernatant and samples before the precipitations only indicated the presence of the starting acylhydrazides. A solution might be the addition of solid **II-52** form directly in a more concentrated solution of hydrazide. ESI-MS analysis was probably not the most appropriate analysis and should be combined with ¹H NMR analysis to control the composition of the crude mixtures.

In order to enable the condensation of the acylhydrazines on the aldehydes moieties of **II-138**, a protection of its thiols with cyanopropyl group (**II-143**) was investigated³⁴⁵. The end of the 48 hours of reaction between cyanopropyl bromide, sodium hydride in THF at reflux, was an insoluble precipitate. ESI-MS analysis of the supernatant did not show the presence of the desired product.

The presence of 2 aldehydes in ortho positions of a 1,4-bisthiophenol was probably the reasons many unsolved side reactivities were observed instead of the desired cyclo-oligomerisation by disulfide bond formation or the connection of side chain by acylhydrazone bond formation. A different backbone, spacing the chemical functions might solve this problem. The construction of a 1,4-bisthiophenol bearing hydrazides instead of aldehydes end groups in ortho positions is an alternative presented in the next part.



Scheme 212 : Formation of acylhydrazones from acylhydrazides and 1,4-bisthiophenol **II-138**



Scheme 213: Reactivity of **II-138**. a) DMF/NaOAc 100 mM pH 7. Hydrazide, r.t., b) NaH, THF, cyanopropyl bromide, reflux, 48 h.

The synthesis of a 1,4-bisthiophenol bearing aldehyde functions in the 2,5 positions of the aromatic core succeeded. Unsolved side reactivities were observed instead of the desired cyclo-oligomerisation by disulfide bond formation or the connection of side chain by acylhydrazone bond formation.

c. Conclusion

1,4-bisthiophenol bearing amines moieties in the 2,5 positions of the aromatic core were not obtained because of an unexpected cyclisation resulting from a nucleophilic attack of amines on the S-thiocarbamate during the last steps of the synthetic sequence. Nevertheless, a 1,4-bisthiophenol bearing aldehydes moieties in the 2,5 positions of the aromatic core was successfully synthesised. However, the cyclo-oligomerisation of this monomer by disulfide bond as well as the connection of side chain bearing acylhydrazides to the aldehydes of the monomer by acylhydrazone bond formation failed.

The formation of cavitands using simultaneously or one after the other exchange reactions based on disulfide bonds and complementary hydrazone or acylhydrazones bonds could not be evaluated. Synthetic issue and undesired reactivity highlighted the risks of neighbouring and reactive chemical functions on 1,4-bisthiophenols.

7. Conclusion

The intention of this chapter was to explore three strategies in order to synthesised tailored cavitands for the recognition of therapeutic proteins that are affected by PTM. The synthesis of 1,4-bisthiophenols bearing side chains on their 2,5 positions with functional groups such as alcohols and amines was successfully realised from **II-1** thanks to trityl and cyanopropyl protecting groups. Those monomers showed an unexpected behaviour in aqueous solution. Instead of forming disulfide bonds by a slow oxidation with air in DCLs, the thiols were quickly oxidised into sulfenic acids and sulfenyl amides. Specific conditions of DCLs and the use of 1,4-bisthiophenols bearing tertiary amides instead of secondary amides partially avoided NS coupling and other undesired reactivities. Regarding the synthesised 1,4-bisthiophenols bearing aldehyde in the 2,5 positions, the connection of monomers by disulfide bonds and the connection of the side chains to the principle chain by acylhydrazone bonds failed. The high reactivity of the aldehydic monomers was a recurrent issue during the synthetic experiments. Turning to corona[*n*]arenes did not help to obtain cavitands for the desired purpose because of a poor reproducibility of the reported protocols. The synthesis of novel corona[*n*]arenes failed. At this stage, no recognition of modified proteins was possible, but optimisations on several synthetic path could help to get tailored cavitands. The use of acylhydrazone instead of disulfide bonds as reversible bonds could be an alternative in order to connect the monomers. These compounds are usually very stable in aqueous media and do not suffer from side reactions.

Experimental section

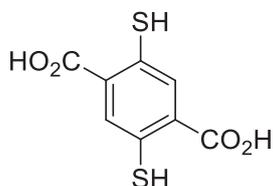
a. Synthetic procedures and characterisations

Only new compounds are fully characterised. Otherwise, only ^1H and ^{13}C NMR were recorded. "Match with literature" means a reference for the compound spectra. The protocol references are given in the main text. Adaptations and optimisations were done without specific mention. The new compounds were fully characterised by HRMS, 1D, 2D NMR, and mp analysis. The spectral data are described after their synthetic procedure. The corresponding spectra are not given in appendices for conciseness. They were compiled on previous reports and are electronically available. If a molecule was not fully characterised, for instance with macrocycles, spectral data could be delivered in appendices to support the discussion.

iii. Materials and methods

All commercial reagents were used as received. Solvents were dried using conventional methods. All reactions were carried out under nitrogen or argon. Column chromatography was performed using silica gel (0.040–0.063 nm). Reactions were monitored via TLC on a silica gel plate and visualised under UV light. ^1H NMR and ^{13}C NMR spectra were recorded on a Bruker DRX at 400 MHz or on a Bruker ALS at 300 MHz. J values are given in Hz. Mass spectra were acquired on an LCQ Advantage ion trap instrument, detecting positive or negative ions in the ESI mode. Melting points were determined by a Büchi Melting point B-540.

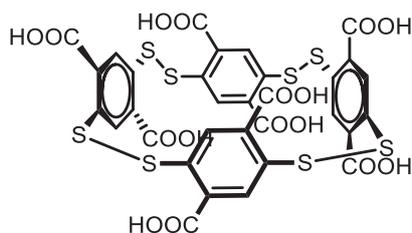
iv. Preparative procedures and characterisations



II-1 2,5-dimercaptoterephthalic acid

A solution of diethyl 2,5-bis((dimethylcarbamoyl)thio)terephthalate (2.6 g, 6.06 mmol, 1 eq) in degassed 1.3 M KOH (5.8 g, 104 mmol, 17 eq) in EtOH/H₂O (1:1, 80 mL) was refluxed under an inert atmosphere for 3 h. The reaction mixture was cooled in ice, and concentrated HCl (15 mL) was added until pH 1. A bright yellow precipitate was formed, filtered, and washed extensively with water, yielding compound **II-1** as a yellow solid (1.215 g, 95 %).

^1H NMR (300 MHz, MeOD) δ = 8.06 (s, 2 H, Ar) data matched with literature reference¹⁶

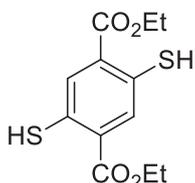


II-1₄

A solution of **II-1** 2,5-dimercaptoterephthalic acid (230 mg, 1 mmol, 1 eq) and spermine (50.6 mg, 0.25 mmol, 0.25 eq) in Tris [200 mM] buffer pH 7.4 (3.05 g in 100 mL of milliQ water) was stirred under air

for 24 h. Then, TFA was added until pH 1. The resulting precipitate was filtrated and dissolved in a minimum amount of borate buffer [200 mM] pH 8.3 and stirred for 5 mn. Then, TFA was added until pH 1. The sequence was done three times. Then, the solid was washed with milliQ water three times, and dried under vacuum to yield yellow powder (185 mg, 80 %).

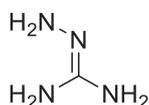
^1H NMR (300 MHz, D_2O , NaOD, pH 7.4) δ = 7.98 (s, 8 H, Ar) data matched with literature reference¹⁶



II-21 diethyl 2,5-dimercaptoterephthalate

A 0.75 M solution of freshly prepared EtONa in distilled EtOH (642 mg of Na, 27.96 mmol, 12 eq, in 37 mL of EtOH) was added to diethyl 2,5-bis((dimethylcarbamoyl)thio)terephthalate (1 g, 2.33 mmol, 1 eq. The mixture was heated under reflux for 1 h 30) under N_2 , and then cooled to 0°C . A 1 M solution of HCl was added until pH 1. The resulting precipitate was filtrated and washed with chilled water. The solid was purified by column chromatography (SiO_2 , cyclohexane: ethyl acetate = 7:3) to give a yellow solid (147 mg, 22 %).

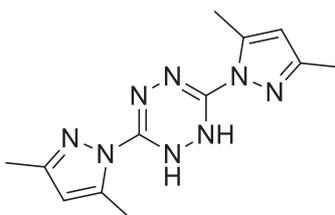
Yellow solid; mp: $132.7\text{-}134.8^\circ\text{C}$; HRMS (ESI) $[\text{M} + \text{Na}]^+$ found 309.0222, calculated 309.0226 for $[\text{C}_{12}\text{H}_{14}\text{NaO}_4\text{S}_2]^+$; ^1H NMR (400 MHz, CDCl_3) δ ppm = 7.96 (s, 2 H, Ar), 4.69 (s, 2 H, SH), 4.42 (q, $J = 7.15$ Hz, 4 H, CH_2), 1.44 (t, $J = 7.15$ Hz, 6 H, CH_3); ^{13}C NMR (400 MHz, CDCl_3) δ ppm = 165.5 (CO), 133.5 ($\text{C}_{\text{Ar-CO}}$), 133.3 (C_{Ar}), 129.4 ($\text{C}_{\text{Ar-S}}$), 61.9 (CH_2), 14.2 (CH_3)



II-40 carbonohydrazoneic diamide

Guanidine hydrochloride (19.1 g, 0.20 mol, 1 eq) was suspended in 1,4-dioxane (100 mL). Hydrazine monohydride (34.0 g, 0.68 mol 3.4 eq) was added dropwise under stirring and the mixture was refluxed for 2 h. A white precipitate formed after cooling to room temperature. The product was collected by filtration and washed with cold 1,4-dioxane. The solid was recrystallised from water if necessary to give a white powder (30 g, 99 %).

White powder; ^{13}C NMR (300 MHz, D_2O) δ ppm = 159.4 (C_{IV}), data matched with literature reference²⁰⁷.

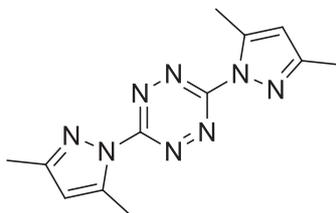


II-41 3,6-bis(3,5-dimethyl-1H-pyrazol-1-yl)-1,2-dihydro-1,2,4,5-tetrazine

To a solution of **II-42** carbonohydrazoneic diamide (30 g, 0.20 mol, 1 eq) in 400 mL of water, 2,4-pentanedione (40.9 mL, 0.40 mol, 2 eq) was added dropwise. The mixture was stirred overnight at

70°C. The precipitate was collected after filtration and washed with cold water to a yellow solid (43,2 g, 80 %).

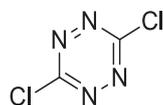
Yellow powder; ^1H NMR (300 MHz, CDCl_3) δ ppm = 8.09 (s, 2 H, NH), 5.96 (s, 2 H, H_{Ar}), 2.47 (s, 6 H, CH_3), 2.21 (s, 6 H, CH_3); ^{13}C NMR (300 MHz, CDCl_3) δ ppm = 149.9 (C_{IV}), 145.7 (C_{IV}), 142.2 (C_{IV}), 109.8 (C_{ArH}), 13.7 (CH_3), 13.4 (CH_3), data matched with literature reference²⁰⁷.



II-42 3,6-bis(3,5-dimethyl-1H-pyrazol-1-yl)-1,2,4,5-tetrazine

II-41 3,6-bis(3,5-dimethyl-1H-pyrazol-1-yl)-1,2-dihydro-1,2,4,5-tetrazine (43.0 g, 0.158 mol, 1 eq) was suspended in 180 mL of DCM. A NaNO_2 solution (29.5 g, 0.428 mol, 2.7 eq, in 35 mL of water) was added dropwise at 0°C under stirring. To this mixture, acetic acid (21.6 mL, 0.36 mol, 2.28 eq) was added dropwise. The solution turns to clear pink with gas evolution and was stirred for 4h. Then, the red organic layer was collected. The upper aqueous layer was washed with 2 x 300 mL DCM until the organic layer remained clear colourless. The organic layer was combined and washed with 5% K_2CO_3 , dried with MgSO_4 , filtered and evaporated to give a dark red solid (31.0 g, 74 %).

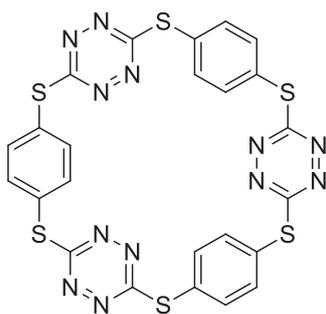
Dark red powder; ^1H NMR (300 MHz, CDCl_3) δ ppm = 6.13 (s, 2 H, Ar), 2.64 (s, 6 H, CH_3), 2.32 (s, 6 H, CH_3); ^{13}C NMR (300 MHz, CDCl_3) δ ppm = 159.0 (C_{IV}), 154.2 (C_{IV}), 143.5 (C_{IV}), 111.7 (C_{ArH}), 14.5, 13.6 (CH_3), data matched with literature reference²⁰⁷.



II-20 3,6-dichloro-1,2,4,5-tetrazine

II-42 3,6-bis(3,5-dimethyl-1H-pyrazol-1-yl)-1,2,4,5-tetrazine (43.0 g, 0.158 mol, 1 eq) was suspended in 20 mL of acetonitrile. 0.40 mL of hydrazine monohydrate (8.20 mmol) was added dropwise at room temperature. A precipitate starts forming during this procedure. Continue refluxing the mixture for 20 min. The dark red precipitate was collected after cooling to room temperature and washed with CH_3CN . The precipitate was re-suspended in 5 mL of anhydrous CH_3CN and cooled to 0 °C in ice bath. 1.9 g of trichloroisocyanuric acid (0.82 mmol) was dissolved in 5 mL of anhydrous CH_3CN and slowly added to the first slurry. Solution color turned from dark red to orange and a white precipitate formed during the addition of trichloroisocyanuric acid. The solution was stirred at room temperature for additional 30 min. The clear orange clear layer was passed through a short silicone gel plug and the orange band was eluted with anhydrous CH_3CN . The solvent was removed in vacuum to give an orange solid (4.94 g, 29 %).

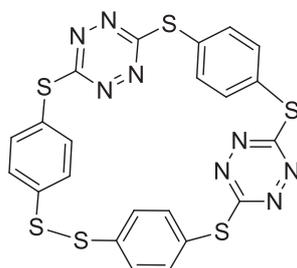
Orange powder; ^{13}C NMR (400 MHz, CDCl_3) δ ppm = 168.0 (C_{IV}), data matched with literature reference²⁰⁷.



II-50 A name could not be generated for this structure

Under argon protection, to a solution of DIPEA (436 μL , 2.50 mmol, 2.1 eq) in acetonitrile (176 mL), which was degassed and pre-heated to 75 $^{\circ}\text{C}$, was added dropwise a solution of 1,4-benzenedithiol (169 mg, 1.19 mmol, 1 eq) and 3,6-dichlorotetrazine (180 mg, 1.19 mmol, 1 eq) in a degassed solution of acetonitrile (12 mL) and dichloromethane (47 mL) during 30 min. After the mixture was cooled gradually to room temperature, 0.1 M hydrochloric acid solution (300 mL) was added. The resulting mixture was extracted with dichloromethane (4 \times 150 mL). The combined organic phase was washed with brine (3 \times 300 mL) and dried over anhydrous MgSO_4 . The resulting residue was purified by column chromatography (SiO_2 , cyclohexane : dichloromethane = 1:3) to give a red powder (79 mg, 29 %).

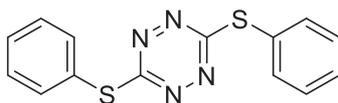
Red powder, mp: decomposition 234.6 $^{\circ}\text{C}$; HRMS (ESI) $[\text{M} + \text{Na}]^+$ found 682.9554, calculated 682.9524 for $[\text{C}_{24}\text{H}_{12}\text{N}_{12}\text{NaS}_6]^+$; ^1H NMR (400 MHz, CDCl_3) δ ppm = 7.68 (s, 12 H, Ar); ^{13}C NMR (400 MHz, CDCl_3) δ ppm = 173.5 (C_{IVS}), 135.7 (C_{Ar}), 128.8 ($\text{C}_{\text{Ar-S}}$). Data matched with literature reference^{320c}.



II-51 A name could not be generated for this structure

Under argon protection, to a solution of DIPEA (380 μL , 2.20 mmol, 2.1 eq) in acetonitrile (150 mL), which was degassed and pre-heated to 60 $^{\circ}\text{C}$, was added dropwise a solution of 1,4-benzenedithiol (149 mg, 1.05 mmol, 1 eq) and 3,6-dichlorotetrazine (160 mg, 1.05 mmol, 1 eq) in a degassed solution of acetonitrile (40 mL) and dichloromethane (10 mL) during 30 min. After the mixture was cooled gradually to room temperature, 0.1 M hydrochloric acid solution (300 mL) was added. The resulting mixture was extracted with dichloromethane (4 \times 150 mL). The combined organic phase was washed with brine (3 \times 300 mL) and dried over anhydrous MgSO_4 . The resulting residue was purified by column chromatography (SiO_2 , cyclohexane : ethyl acetate = 8:2) to give a red powder (100 mg, 49 %).

Red powder, mp: decomposition 70.6 $^{\circ}\text{C}$; HRMS (ESI) $[\text{M} + \text{Na}]^+$ found 602.9861, calculated 602.9401 for $[\text{C}_{22}\text{H}_{12}\text{N}_8\text{NaS}_6]^+$; ^1H NMR (400 MHz, CDCl_3) δ ppm = 7.68 (s, 4 H, Ar), 7.64 - 7.50 (m, 8 H, Ar); ^{13}C NMR (400 MHz, CDCl_3) δ ppm = 173.9 (C_{IVS}), 173.3 (C_{IVS}), 138.8 (C_{Ar}), 137.0 (C_{ArH}), 136.1 (C_{ArH}), 128.2 (C_{Ar}), 126.7 (C_{ArH})



II-57 3,6-bis(phenylthio)-1,2,4,5-tetrazine

To a stirred solution of DIPEA (143 μ L, 0.825 mmol, 2.5 eq.) in acetonitrile (10 mL) was added dropwise at the same speed through two separate addition funnels a solution of thiophenol (72.5 mg, 0.66 mmol, 2 eq.) in acetonitrile (10 mL) and a solution of 3,6-dichlorotetrazine (100 mg, 0.33 mmol, 1 eq.) 330 mg (1 mmol) in acetonitrile (10 mL). After completion of the addition, the solvent was evaporated. The remaining residue was purified by column chromatography (SiO_2 , cyclohexane : dichloromethane = 1:1) to give a red powder (58 mg, 59 %).

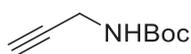
Red solid; mp: 130.8 - 132.6 $^{\circ}\text{C}$; HRMS (ESI) $[\text{M} + \text{H}]^+$ found 299.0420, calculated 299.0420 for $[\text{C}_{14}\text{H}_{11}\text{N}_4\text{S}_2]^+$; ^1H NMR (400 MHz, D_2O , NaOD) δ ppm = 7.68 - 7.61 (m, 2 H, Ar), 7.52 - 7.42 (m, 3 H, Ar); ^{13}C NMR (400 MHz, D_2O , NaOD) δ ppm = 173.9 (C_{IV} tetrazine), 135.6 (C_{Ar}), 130.4 (C_{Ar}), 129.8 (C_{Ar}), 125.6 ($\text{C}_{\text{Ar-S}}$)



II-61 di-tert-butyl but-2-ene-1,4-diyl dicarbamate

A solution of **II-60** *tert*-butyl-allylcarbamate (1.00 g, 6.35 mmol, 1 eq) and Grubb's I catalyst (261 mg, 0.318 mmol, 0.05 eq) in dry DCM (25 mL) under argon was refluxed overnight. Then, solvent was evaporated. The remaining residue was purified by column chromatography (SiO_2 , cyclohexane : ethyl acetate = 8:2) to give a grey powder (444 mg, 49 %).

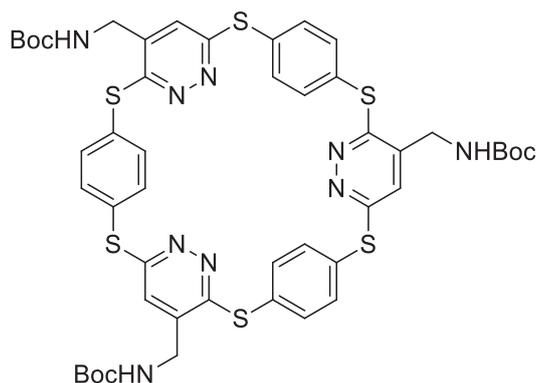
Grey powder; ^1H NMR (300 MHz, CDCl_3) δ ppm = 5.62 (t, $J = 2.8$ Hz, 2 H, $\text{H}_{\text{C}=\text{C}}$), 4.59 (br. s., 2 H, NH), 3.72 (br. s., 4 H, CH_2), 1.44 (s, 18 H, CH_3), data matched with literature reference³²²ⁱ.



II-63 tert-butyl prop-2-yn-1-yl carbamate

Di-*tert*-butyl dicarbonate (2.18 mL, 10.0 mmol, 1.0 equiv) was added dropwise at 0 $^{\circ}\text{C}$ to a solution of prop-2-yn-1-amine (0.686 mL, 10.0 mmol, 1.0 equiv) in CH_2Cl_2 (20 mL). After 1 h of stirring, the solvent was removed in vacuo and the resulting colourless oil was dried under high vacuum overnight to yield a white solid (1.502 g, 97 %), which was used as such without further purification.

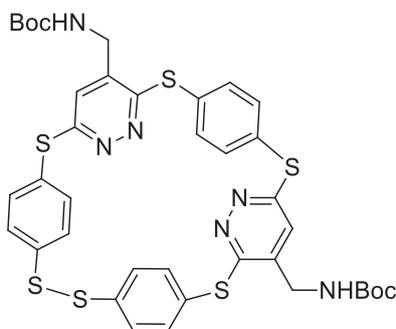
White solid ; ^1H NMR (300 MHz, CDCl_3) δ ppm = 4.85 (br. s., 1 H, NH), 3.96 - 3.83 (m, 2 H, CH_2), 2.20 (t, $J = 2.5$ Hz, 1 H, $\text{H-C}_{\text{alkyne}}$), 1.43 (s, 9 H, CH_3), matched with literature reference.^{322a}



II-59

In a sealed tube, a solution of **II-50** (20.0 mg, 0.030 mmol, 1 eq) with **II-60** *tert*-butyl-*N*-allylcarbamate (47 mg, 0.30 mmol, 10 eq) in dry chloroform (1.0 mL) was stirred for 24 hours at 130°C, until red colour disappeared. The crude product was analysed by mass spectrometry.

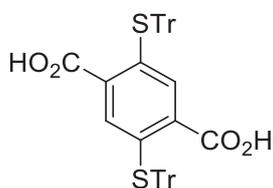
HRMS (ESI) [M + H]⁺ found 1042.2340, calculated 1042.2359 for [C₄₈H₅₁N₉O₆S₆]⁺



II-58

In a sealed tube, a solution of **II-51** (20.0 mg, 0.034 mmol, 1 eq) with *tert*-butyl-*N*-allylcarbamate (54 mg, 0.34 mmol, 10 eq) in dry chloroform (1.1 mL) was stirred for 24 hours at 130°C, until red colour disappeared. The crude product was analysed by mass spectrometry.

HRMS (ESI) [M + H]⁺ found 835.1360, calculated 815.1352 for [C₃₈H₃₉N₆O₄S₆]⁺

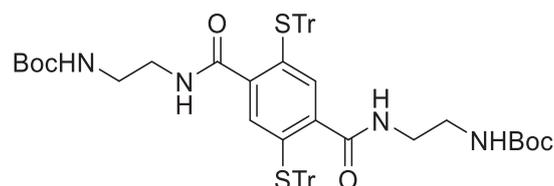


II-67 2,5-bis(tritylthio)terephthalic acid

A solution of **II-1** (500 mg, 2.17 mmol, 1 eq) and trityl chloride (1.575 g, 5.65 mmol, 2.6 eq) in degassed and dry DMF (15 mL) was stirred at room temperature under an inert atmosphere for 48 h. The resulting precipitate was filtrated, washed with water (3*20 mL), chloroform (3*20 mL) and cyclohexane (3*20 mL). The product was dried under vacuum, yielding compound as a yellow solid (1.507 g, 97 %).

Yellow solid; mp: degradation 405.0 °C; HRMS (ESI) $[M + Na]^+$ found 737.1791, calculated 737.1791 for $[C_{46}H_{34}NaO_4S_2]^+$; 1H NMR (400 MHz, DMSO- d_6) δ ppm = 7.31 - 7.21 (m, 30 H, Ar_{trityl}), 7.14 (s, 2 H, Ar); ^{13}C NMR (400 MHz, DMSO- d_6) δ ppm = 166.2 (CO), 147.7 (C_{Ar-trityl}-C), 143.0 (C_{trityl}-S), 129.6 (C_{Ar-trityl}), 127.9 (C_{Ar}-S), 127.7 (C_{Ar-trityl}), 127.5 (C_{Ar-trityl}), 126.6 (C_{Ar}-CO)

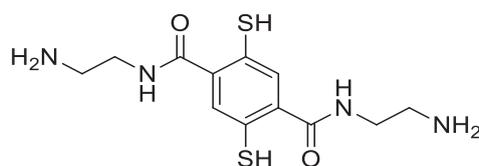
1H NMR (300 MHz, CDCl₃) δ ppm = 5.17 (br. s, 1 H, NHCO₂), 3.64 - 3.59 (m, 4 H, CH₂OCH₂CH₂OCH₂), 3.53 - 3.49 (m, 4 H, OCH₂CH₂O), 3.36 - 3.25 (m, 2 H, CH₂-CH₂-NHBoc), 2.87 (t, J = 5.1 Hz, 2 H, CH₂-NH₂), 1.71 (br. s., 2 H, NH₂), 1.43 (s, 9 H, CH₃) matched with literature reference⁴⁰²



II-68 di-tert-butyl (((2,5-bis(tritylthio)terephthaloyl)bis(azanediyl))bis(ethane-2,1-diyl))dicarbamate

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (393 mg, 0.55 mmol, 1 eq) with HOBT (194 mg, 1.43 mmol, 2.6 eq) in dry DCM (26 mL) was added at 0°C EDC (254 μ L, 1.43 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of the **II-144** tert-butyl (2-aminoethyl)carbamate (240 mg, 1.43 mmol, 2.6 eq) in dry DCM (12 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 1:1) to give a white solid (401 mg, 77 %).

White solid; mp: degradation 128.6 °C; HRMS (ESI) $[M + Na]^+$ found 1021.4014, calculated 1021.4003 for $[C_{60}H_{62}N_4NaO_6S_2]^+$; 1H NMR (400 MHz, CDCl₃) δ ppm = 7.34 - 7.21 (m, 30 H, Ar_{trityl}), 6.92 (s, 2 H, Ar), 4.93 (br. s, 2 H, NH), 4.76 (br. s, 2 H, NH), 3.20 - 3.12 (m, 4 H, CH₂-NHBoc), 3.11 - 3.04 (m, 4 H, CH₂-NHCO), 1.49 (s, 18 H, *t*Bu); ^{13}C NMR (400 MHz, CDCl₃) δ ppm = 167.1 (C_{Boc}O), 156.0 (CONH), 143.6 (C_{Ar}trityl), 139.3 (C_{Ar}-CONH), 133.0 (C_{Ar}-S), 132.2 (C_{Ar} trityl), 130.0 (C_{Ar} trityl), 127.9 (C_{Ar} trityl), 127.1 (C_{Ar}H), 79.4 (C_{tBu}), 70.9 (C_{trityl}-S), 40.5 (CONH-CH₂), 39.8 (CH₂-NHBoc), 28.4 (CH₃)

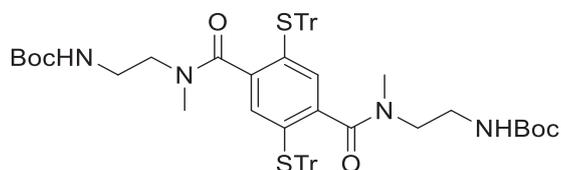


II-69 N¹,N⁴-bis(2-aminoethyl)-2,5-dimercaptoterephthalamide

A solution of **II-68** di-tert-butyl (((2,5-bis(tritylthio)terephthaloyl)bis(azanediyl))bis(ethane-2,1-diyl))dicarbamate (270 mg, 0.270 mmol, 1 eq) in a mixture of degassed DCM/TFA/Et₃SiH 45/55/5 (6.22 mL / 8.50 mL / 0.78 mL) was stirred for 1 h under inert atmosphere. Solvent and volatile compounds were evaporated. The remaining residue was dissolved in MeOH (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness, affording an orange powder as a TFA salt (183 mg, 99 %).

Orange solid; mp: 104.6 – 106.7 °C; HRMS (ESI) $[M + H]^+$ found 315.0928, calculated 315.0944 for $[C_{12}H_{19}N_4O_2S_2]^+$; 1H NMR (400 MHz, D₂O) δ ppm = 7.57 (s, 2 H, Ar), 3.67 (t, J = 6.0 Hz, 4 H, CH₂-NCO),

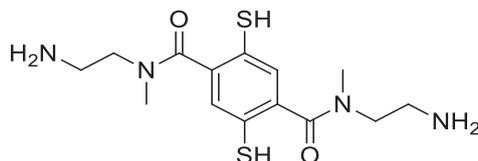
3.24 (t, J = 6.0 Hz, 4 H, CH₂-NH₂); ¹³C NMR (400 MHz, D₂O) δ ppm = 171.0 (CONH), 163.7 (C_{TFA}), 163.4 (C_{TFA}), 135.7 (C_{Ar}-CONH), 131.2 (C_{Ar}H), 129.3 (C_{Ar}-S), 118.1 (C_{TFA}F₃), 115.8 (C_{TFA}F₃), 39.5 (CH₂-NH₂), 37.9 (CH₂-NCO)



II-70 di-tert-butyl (((2,5-bis(tritylthio)terephthaloyl)bis(methylazanediy))bis(ethane-2,1-diyl))dicarbamate

To a solution of **II-61** 2,5-bis(tritylthio)terephthalic acid (400 mg, 0.56 mmol, 1 eq) with HOBT (197 mg, 1.46 mmol, 2.6 eq) in dry DCM (20 mL) was added at 0°C EDC (258 μL, 1.46 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of *tert*-butyl(2-(methylamino)ethyl)carbamate (253 mg, 1.46 mmol, 2.6 eq) in dry DCM (5 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 1:1) to give a white solid (444 mg, 77 %).

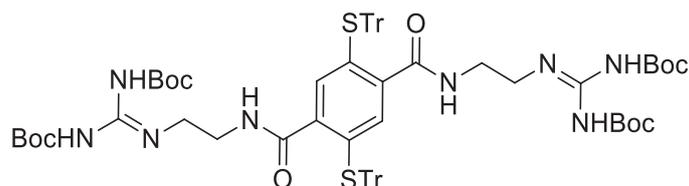
White solid; mp: degradation 117.9 – 120.0 °C; HRMS (ESI) [M + H]⁺ found 1027.4493, calculated 1027.4497 for [C₆₂H₆₇N₄O₆S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.35 - 7.19 (m, 30 H, Ar_{trityl}), 6.33 (s, 2 H, Ar), 4.81 (br. s., 2 H, NH), 3.50 - 3.38 (m, 4 H, CH₂-NHBoc), 3.25 - 3.15 (m, 4 H, CH₂-NMeCO), 2.36 (s, 6 H, CH₃), 1.42 (s, 18 H, *t*Bu); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 168.5 (C_{Boc}O), 156.0 (CONH), 143.6 (C_{Ar} trityl), 137.4 (C_{Ar}-CONH), 131.3 (C_{Ar}-S), 130.0 (C_{Ar} trityl), 129.9 (C_{Ar} trityl), 127.8 (C_{Ar} trityl), 127.1 (C_{Ar}H), 79.2 (C_{tBu}), 70.9 (C_{trityl}-S), 47.0 (CONMe-CH₂), 38.4 (CH₂-NHBoc), 37.6 (CH₃-N), 28.4 (CH₃ Boc)



II-71 N₁,N₄-bis(2-aminoethyl)-2,5-dimercapto-N1,N4-dimethylterephthalamide

A solution of **II-70** di-*tert*-butyl (((2,5-bis(tritylthio)terephthaloyl)bis(methylazanediy))bis(ethane-2,1-diyl))dicarbamate (400 mg, 0.389 mmol, 1 eq) in a mixture of degassed DCM/TFA/Et₃SiH 45/55/5 (6.22 mL / 8.50 mL / 0.78 mL) was stirred for 30 mn under inert atmosphere. Solvent and volatile compounds were evaporated. The remaining residue was dissolved in MeOH (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness, affording a white powder as a TFA salt (230 mg, 99 %).

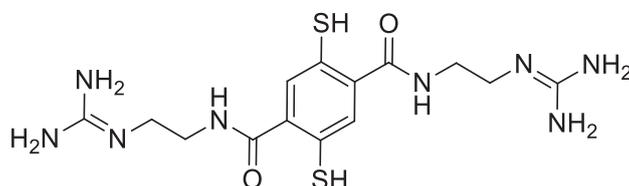
White solid; mp: 109.8 – 111.7 °C; HRMS (ESI) [M + H]⁺ found 343.1244, calculated 343.1257 for [C₁₄H₂₃N₄O₂S₂]⁺; ¹H NMR (400 MHz, D₂O) δ ppm = 7.43 (s, 2 H, Ar), 3.83 (t, J = 6.5 Hz, 4 H, CH₂-NCO), 3.56 (br. s., 2 H, NH), 3.33 (t, J = 6.4 Hz, 4 H, CH₂-NH₂), 3.17 (t, J = 6.8 Hz, 2 H, NH), 3.11 (s, 2H, SH); 2.93 (s, 6 H, CH₃); ¹³C NMR (400 MHz, D₂O) δ ppm = 171.5 (C_{TFA}F₃), 163.1 (CONH), 136.5 (C_{Ar}-CONH), 129.5 (C_{Ar}-S), 126.4 (C_{Ar}H), 117.7 (C_{TFA}F₃), 114.8 (C_{TFA}F₃), 45.0 (CH₂-NH₂), 36.8 (CH₂-NMeCO), 32.5 (CH₃)



II-72 (2,5-bis(tritylthio)terephthaloyl)bis(2-(2-aminoethyl)-1,3-di-Boc-guanidine)

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (500 mg, 0.7 mmol, 1 eq) with HOBT (247 mg, 1.82 mmol, 2.6 eq) in dry DCM (35 mL) was added at 0°C EDC (322 μ L, 1.82 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of the 2-(2-aminoethyl)-1,3-di-Boc-guanidine (550 mg, 1.82 mmol, 2.6 eq) in dry DCM (15 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 1:1) to give a white solid (529 mg, 58 %).

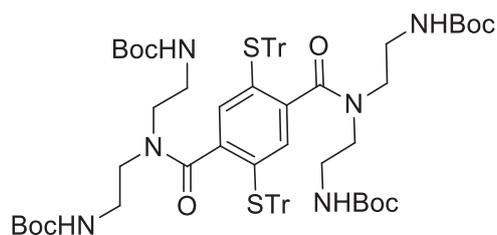
White solid; mp: degradation 115.0 °C; HRMS (ESI) [M + H]⁺ found 642.2853, calculated 642.2870 for [C₇₂H₈₄N₈O₁₀S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 11.45 (s, 2 H, NHBoc), 8.40 (t, J = 6.1 Hz, 2 H, NHCO), 7.27 - 7.25 (m, 30 H, Ar_{trityl}), 6.90 (s, 2 H, Ar), 5.07 (t, J = 5.7 Hz, 2 H, NHBoc), 3.35 (s, 4 H, N-CH₂), 3.24 - 3.17 (m, 4 H, CH₂-NHCO), 1.49 (s, 36 H, tBu); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 166.8 (C_{Boc}O), 163.4 (C_{Boc}O), 156.4 (CONH), 153.0 (NNC_{guanidine}N), 143.6 (C_{Ar} trityl), 139.5 (C_{Ar}-CONH), 132.8 (C_{Ar}-S), 132.5 (C_{Ar} trityl), 130.1 (C_{Ar} trityl), 127.9 (C_{Ar} trityl), 127.1 (C_{Ar}H), 83.3 (C_{tBu}), 79.2 (C_{tBu}), 77.2 (C_{tBu}), 71.0 (C_{trityl}-S), 39.9 (CONH-CH₂), 39.0 (CH₂-guanidine), 28.3 (CH₃), 28.2 (CH₃), 28.0 (CH₃), 26.9 (CH₃)



II-73 N₁,N₄-bis(2-((diaminomethylene)amino)ethyl)-2,5-dimercaptoterephthalamide

To a solution of **II-72** (2,5-bis(tritylthio)terephthaloyl)bis(2-(2-aminoethyl)-1,3-di-Boc-guanidine) (340 mg, 0.262 mmol, 1 eq) in a mixture of degassed DCM/TFA (6.1 mL / 8.4 mL) was added Et₃SiH (0.76 mL, 4.9 mmol, 19 eq) was added dropwise and the mixture was stirred for 3 h under inert atmosphere. Solvent and volatile compounds were evaporated. The remaining residue was dissolved in MeOH (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness, affording an orange powder as a TFA salt (136 mg, 83 %).

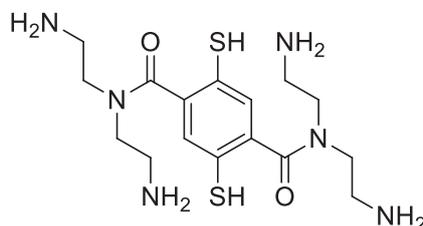
Orange solid; mp: degradation 226.4°C; HRMS (ESI) [M + H]⁺ found 399.1367, calculated 399.1380 for [C₁₄H₂₃N₈O₂S₂]⁺; ¹H NMR (400 MHz, D₂O) δ ppm = 7.48 (s, 2 H, Ar), 3.59 - 3.54 (m, 4 H, N-CH₂), 3.48 - 3.42 (m, 4 H, CH₂-NHCO); ¹³C NMR (400 MHz, D₂O) δ ppm = 170.2 (C_{guanidine}), 163.4 (C_{TFA}), 162.7 (C_{TFA}), 157.0 (CONH), 135.6 (C_{Ar}-CONH), 130.2 (C_{Ar}H), 128.1 (C_{Ar}-S), 117.7 (C_{TFA}F₃), 114.8 (C_{TFA}F₃), 40.4 (CH₂-N), 38.5 (CH₂-N)



II-74 tetra-tert-butyl (((2,5-bis(tritylthio)terephthaloyl)bis(azanetriyl))tetrakis(ethane-2,1-diyl))tetracarbamate

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (500 mg, 0.7 mmol, 1 eq) with HOBt (247 mg, 1.82 mmol, 2.6 eq) in dry DCM (35 mL) was added at 0°C EDC (322 μ L, 1.82 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of the amine **II-145** (552 mg, 1.82 mmol, 2.6 eq) in dry DCM (15 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 1:1) to give a white solid (598 mg, 66 %).

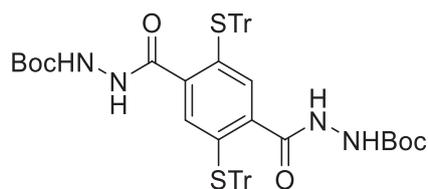
White solid; mp: degradation 225.9 °C; HRMS (ESI) [M + H]⁺ found 1285.6062, calculated 1285.6076 for [C₇₉H₈₉N₆O₁₀S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.35 - 7.22 (m, 32 H, Ar_{trityl}, Ar), 3.56 - 3.17 (m, 8 H, CH₂-NH(CO)), 2.93 - 2.59 (m, 8 H, CH₂-NH), 1.48 - 1.39 (m, 36 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 155.9 (Ar-CONH), 143.5 ((CO)_{Boc}), 129.9 (C_{Ar} trityl), 129.9 (C_{Ar} trityl), 127.9 (C_{Ar}-CONH), 128.1 (C_{Ar} trityl), 128.1 (C_{Ar} trityl), 127.5 (C_{Ar}-S), 127.4 (C_{Ar}-H), 77.2 (C_{Boc}), 45.8 (CH₂-CH₂-NHBoc), 39.2 (CH₂-NHBoc), 28.4 (CH₃)



II-75 N₁,N₁,N₄,N₄-tetrakis(2-aminoethyl)-2,5-dimercaptoterephthalamide

To a solution of **II-74** tetra-tert-butyl (((2,5-bis(tritylthio)terephthaloyl)bis(azanetriyl))tetrakis(ethane-2,1-diyl))tetracarbamate (200 mg, 0.155 mmol, 1 eq) in a mixture of degassed DCM/TFA (3.6 mL / 4.95 mL) was added Et₃SiH (0.45 mL, 2.9 mmol, 19 eq) was added dropwise and the mixture was stirred for 3 h. Solvent and volatile compounds were evaporated. The remaining residue was dissolved in MeOH (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness, affording a pale yellow powder as a TFA salt (133 mg, 99 %).

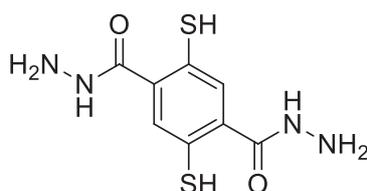
Pale yellow solid; mp: 86.6 – 89.1°C; HRMS (ESI) [M + H]⁺ found 401.1785, calculated 401.1788 for [C₁₆H₂₉N₆O₂S₂]⁺; ¹H NMR (400 MHz, D₂O) δ ppm = 7.49 (s, 2 H, Ar), 3.92 - 3.80 (m, 4 H, CH₂-NH₂), 3.67 - 3.50 (m, 4 H, CH₂-NH₂), 3.35 (t, J = 6.5 Hz, 4 H, CH₂-N), 3.15 (t, J = 7.1 Hz, 4 H, CH₂-N); ¹³C NMR (400 MHz, D₂O) δ ppm = 162.7 (CO), 129.6 (C_{Ar}-H), 117.6 (C_{Ar}-CO), 114.8 (C_{Ar}-S), 46.2 (CH₂-NH₂), 43.0 (CH₂-NH₂), 37.0 (CH₂-N), 36.6 (CH₂-N)



II-76 di-tert-butyl 2,2'-(2,5-bis(tritylthio)terephthaloyl)bis(hydrazinecarboxylate)

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (400 mg, 0.56 mmol, 1 eq) with HOBt (197 mg, 1.46 mmol, 2.6 eq) in dry DCM (20 mL) was added at 0°C EDC (258 μ L, 1.46 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of *tert*-butoxycarbonylhydrazine (193 mg, 1.46 mmol, 2.6 eq) in dry DCM (5 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 70:30) to give a white solid (351 mg, 65 %).

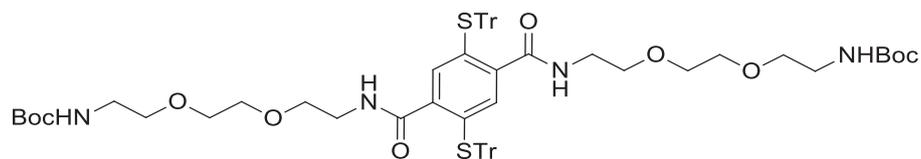
White solid; mp: degradation 145.9 – 148.0 °C; HRMS (ESI) [M + Na]⁺ found 965.3358, calculated 965.3377 for [C₅₆H₅₄N₄NaO₆S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.35 - 7.18 (m, 32 H, H_{Ar}, Trityl, H_{Ar}), 1.43 (s, 18 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 164.8 (C_{iv}-C_{Ar}), 154.4 (C_{iv}Boc), 143.3 (C_{Ar}H), 130.1 (C_{Ar}), 127.9 (C_{Ar}) 127.2 (C_{Ar}), 28.0 (C_{tBu})



II-77 2,5-dimercaptoterephthalohydrazide

A solution of **II-76** di-*tert*-butyl 2,2'-(2,5-bis(tritylthio)terephthaloyl)bis(hydrazinecarboxylate) (290 mg, 0.307 mmol, 1 eq) in a mixture of degassed DCM/TFA/Et₃SiH 45/55/5 (7.2 mL / 9.0 mL / 0.8 mL) was stirred for 45 mn under inert atmosphere. Solvent and volatile compounds were evaporated. The remaining residue was sonicated in degassed pentane (4*20 mL). The solid was filtrated and then evaporated to dryness, affording a yellow powder (79 mg, 99 %).

Yellow solid; mp: degradation 223.1 – 225.9 °C; HRMS (ESI) [M + H]⁺ found 259.0321, calculated 259.0318 for [C₈H₁₁N₄O₂S₂]⁺; ¹H NMR (400 MHz, D₂O) δ ppm = 7.62 (s, 2 H, Ar), 1.23 (s, 4 H, NH₂); ¹³C NMR (400 MHz, D₂O) δ ppm = 167.5 (C_{iv}-C_{Ar}), 138.6 (C_{iv}-C_{Ar}), 132.5 (C_{Ar}), 131.2 (C_{Ar}H), 130.0 (C_{Ar})

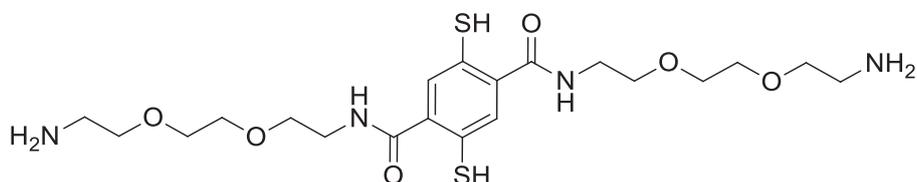


II-78 di-tert-butyl (((((((2,5-bis(tritylthio)terephthaloyl)bis(azanediy))bis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl))dicarbamate

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (674 mg, 0.94 mmol, 1 eq) with HOBt (108 mg, 0.804 mmol, 2.6 eq) in dry DCM (14 mL) was added EDC at 0°C (142 μ L, 0.804 mmol, 2.6 eq) and the

mixture was stirred for 1 hour under N₂. **II-146** tert-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate (200 mg, 0.804 mmol, 2.6 eq) in DCM (7 mL) was then added dropwise at 0°C. The mixture was stirred at room temperature for 48 h. Water was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 3:7) to give a white powder (238 mg, 36 %).

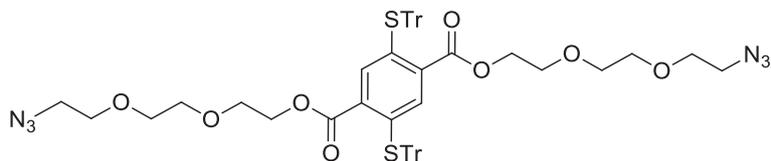
White powder, mp : 71.2 – 72.7 °C; HRMS (ESI) [M + Na]⁺ found 1175.5246, calculated 1175.5232 for [C₆₈H₇₉N₄NaO₁₀S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.31 - 7.18 (m, 30 H, H_{Ar}Trityl), 6.75 (s, 2 H, Ar), 5.37 (br. s., 2 H, NH), 4.94 (br. s., 2 H, NH), 3.54 - 3.48 (m, 8 H, HN-CH₂-CH₂-O, BocHN-CH₂-CH₂-O), 3.45 - 3.37 (m, 8 H, O-CH₂-CH₂-O), 3.36 - 3.26 (m, 8 H, BocHN-CH₂-CH₂-O, HN-CH₂-CH₂-O), 1.43 (s, 18 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 166.7 (NHCO), 155.9 (C_{IV}Trityl), 143.8 (C_{Ar}-C-S), 138.3 (C_{Ar}-CO), 132.7 (C_{Ar}-S), 130.9 (C_{Ar}H), 130.1 (C_{Ar} trityl), 127.8 (C_{Ar} trityl), 127.0 (C_{Ar} trityl), 79.3 (C_{IV} Boc), 71.1 (CH₂-O), 70.2 (CH₂-O), 70.2 (CH₂-O), 69.9 (CH₂-O), 40.3 (COHNCH₂), 39.5 (BocHNCH₂), 28.4 (CH₃)



II-79 N₁,N₄-bis(2-(2-(2-aminoethoxy)ethoxy)ethyl)-2,5-dimercaptoterephthalamide

To a solution of **II-78** di-tert-butyl (((((((2,5-bis(tritylthio)terephthaloyl)bis(azanediyl))bis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl))bis(oxy))bis(ethane-2,1-diyl))dicarbamate (190 mg, 0.162 mmol, 1 eq) in degassed and dry DCM (15mL) was added TFA (6 mL). Then, Et₃SiH (0.6 mL) was added dropwise and the mixture was stirred for 45 mn. Solvent and volatile compounds were evaporated. The remaining residue was dissolved in a 90/10 degassed mixture of MeOH/water (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness. The remaining residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 8:2) affording a yellow oil (80 mg, 99 %).

Yellow oil; HRMS (ESI) [M + Na]⁺ found 513.1803, calculated 513.1812 for [C₂₀H₃₄N₄NaO₆S₂]⁺; ¹H NMR (400 MHz, D₂O) δ ppm = 7.46 (s, 2 H, Ar), 3.73 - 3.68 (m, 16 H, CH₂-O), 3.55 (t, J = 5.3 Hz, 4 H, CH₂-NH₂), 3.14 (t, J = 4.8 Hz, 4 H, CH₂-NH); ¹³C NMR (400 MHz, D₂O) δ ppm = 169.9 (CO), 135.7 (C_{Ar}-CO), 130.0 (C_{Ar}H), 128.0 (C_{Ar}-S), 69.6 (CH₂), 69.6 (CH₂), 69.4 (CH₂), 68.7 (CH₂), 66.4 (CH₂), 39.3 (CH₂-NH₂), 39.0 (CH₂-NH)

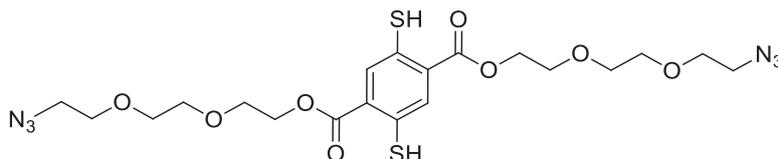


II-80 bis(2-(2-(2-azidoethoxy)ethoxy)ethyl) 2,5-bis(tritylthio)terephthalate

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (674 mg, 0.94 mmol, 1 eq) with HOBt (385 mg, 2.45 mmol, 2.6 eq) in dry DCM (50 mL) was added EDC at 0°C (452 μL, 2.45 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. **II-97** 2-(2-(2-azidoethoxy)ethoxy)ethanol (430 mg, 2.45 mmol, 2.6 eq) was then added dropwise at 0°C. The mixture was stirred at room temperature for 48 h. Brine

was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 8:2) to give a yellow oil (454 mg, 47 %).

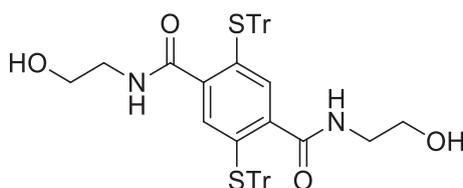
HRMS (ESI) [M + Na]⁺ found 1051.3443, calculated 1051.3493 for [C₅₈H₅₆N₆NaO₈S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.39 - 7.14 (m, 32 H, H_{Ar}, Ar_{Trityl}), 4.16 (t, J = 5.1 Hz, 4 H, CH₂-COO), 3.68 - 3.64 (m, 4 H, CH₂-CH₂-N₃), 3.62 - 3.58 (m, 4 H, CH₂-O), 3.52 - 3.48 (m, 4 H, CH₂-O), 3.47 (t, J = 4.9 Hz, 4 H, CH₂-CH₂-COO), 3.38 (t, J = 5.1 Hz, 4 H, CH₂-N₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.0 (Ar-COO), 143.3 (C_{Ar}-C-S), 136.0 (C_{Ar}-COO), 132.8 (C_{Ar} trityl), 130.7 (C_{Ar} trityl), 130.2 (C_{Ar} trityl), 127.7 (C_{Ar}-S), 126.9 (C_{Ar}H), 70.8 (C_{trityl}), 70.6 (CH₂-O), 70.4 (CH₂-O), 70.1 (CH₂-O), 68.7 (CH₂-C-COO), 64.1 (CH₂-COO), 50.6 (N₃-CH₂)



II-81 bis(2-(2-(2-azidoethoxy)ethoxy)ethyl) 2,5-dimercaptoterephthalate

To a solution of **II-80** bis(2-(2-(2-azidoethoxy)ethoxy)ethyl) 2,5-bis(tritylthio)terephthalate (100 mg, 0.097 mmol, 1 eq) in degassed and dry DCM (10 mL) was added TFA (0.4 mL, 5.22 mmol, 54 eq). The mixture was stirred at room temperature for 30 mn. Then, Et₃SiH (0.1 mL, 0.63 mmol, 6 eq) was added dropwise and the mixture was stirred for 15 mn. Solvent and volatile compounds were evaporated. The remaining residue was dissolved in a 90/10 degassed mixture of MeOH/water (20 mL). This methanolic phase was washed with degassed heptane (4*15 mL), and then evaporated to dryness, affording a yellow oil (38 mg, 72 %).

HRMS (ESI) [M + H]⁺ found 545.1480, calculated 545.1483 for [C₂₀H₂₉N₆O₈S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.98 (s, 2 H, Ar), 4.71 (br. s., 2 H, SH), 4.50 (t, J = 5.1 Hz, 4 H, CH₂-COO), 3.86 (t, J = 4.6 Hz, 4 H, CH₂-CH₂-N₃), 3.73 - 3.66 (m, 12 H, CH₂-O, CH₂-CH₂-COO), 3.38 (t, J = 4.9 Hz, 4 H, CH₂-N₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.4 (Ar-COO), 133.6 (C_{Ar}-C-S), 133.4 (C_{Ar}-COO), 129.2 (C_{Ar}-S), 70.7 (CH₂-O), 70.7 (CH₂-O), 70.1 (CH₂-O), 69.0 (CH₂-C-COO), 64.8 (CH₂-COO), 50.6 (N₃-CH₂)

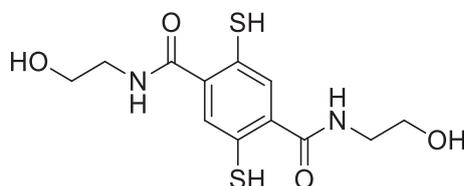


II-82 N₁,N₄-bis(2-hydroxyethyl)-2,5-bis(tritylthio)terephthalamide

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (400 mg, 0.56 mmol, 1 eq) with HOBt (198 mg, 1.44 mmol, 2.6 eq) in dry DCM (30 mL) was added at 0°C EDC (259 μL, 1.44 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. Ethanamine (87 μL, 1.44 mmol, 2.6 eq) was then added at 0°C. The mixture was stirred at room temperature for 48 h. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, dichloromethane: methanol = 95:5) to give a white solid (264 mg, 58 %).

White solid; mp: degradation 212.8 °C; HRMS (ESI) [M + Na]⁺ found 823.2622, calculated 823.2635 for [C₅₀H₄₄N₂NaO₄S₂]⁺; ¹H NMR (400 MHz, DMSO-d₆) δ ppm = 7.31 - 7.18 (m, 30 H, H_{trityl}), 6.86 (t, J = 5.7 Hz,

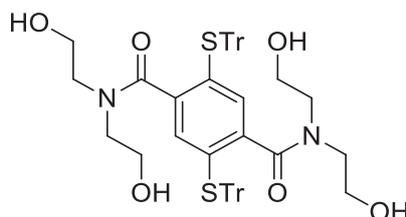
2 H, NH), 6.46 (s, 2 H, H_{Ar}), 4.68 (t, J = 5.4 Hz, 2 H, OH), 3.29 (q, J = 6.4 Hz, 4 H, CH₂-CH₂-OH), 3.08 (q, J = 6.3 Hz, 4 H, CH₂-NHCO); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 166.5 (CO), 143.9 (C_{Ar}-C-S), 139.8 (C_{Ar}-CONH), 131.7 (C_{Ar} trityl), 131.3 (C_{Ar}H), 130.0 (C_{Ar}-S), 128.4 (C_{Ar} trityl), 127.5 (C_{Ar} trityl), 71.0 (C_{trityl}), 60.0 (CH₂-OH), 42.0 (CH₂-N)



II-83 N₁,N₄-bis(2-hydroxyethyl)-2,5-dimercaptoterephthalamide

A solution of **II-82** N₁,N₄-bis(2-hydroxyethyl)-2,5-bis(tritylthio)terephthalamide (160 mg, 0.200 mmol, 1 eq) in a mixture of degassed DCM/TFA/Et₃SiH 45/55/5 (4.60 mL / 6.29 mL / 0.58 mL) was stirred for 45 mn under inert atmosphere. Solvent and volatile compounds were evaporated. The remaining residue was dissolved in MeOH (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness. The remaining residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give a white solid (20 mg, 31 %).

White powder; mp: degradation 120.2 °C; HRMS (ESI) [M + H]⁺ found 317.0619, calculated 317.0624 for [C₁₂H₁₇N₂O₄S₂]⁺; ¹H NMR (400 MHz, DMSO-d₆) δ ppm = 8.57 (s, 2 H, NH), 7.55 (s, 2 H, Ar), 5.27 (br. s., 1 H, OH), 4.74 (br. s., 1 H, OH), 3.50 (t, J = 6.1 Hz, 4 H, CH₂-OH), 3.29 (d, J = 6.1 Hz, 4 H, CH₂-NH); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 166.8 (CO), 134.7 (C_{Ar}-CONH), 130.1 (C_{Ar}-S), 130.0 (C_{Ar}H), 59.6 (CH₂-OH), 42.1 (CH₂-N)

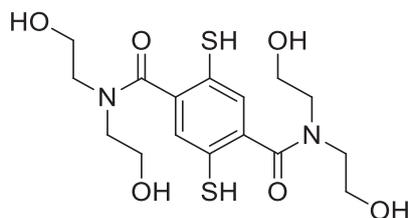


II-84 N₁,N₁,N₄,N₄-tetrakis(2-hydroxyethyl)-2,5-bis(tritylthio)terephthalamide

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (300 mg, 0.42 mmol, 1 eq) with HOBT (148 mg, 1.09 mmol, 2.6 eq) in dry DCM (29 mL) was added at 0°C EDC (194 μL, 1.09 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. Diethanolamine (114 mg, 1.09 mmol, 2.6 eq) was then added at 0°C. The mixture was stirred at room temperature for 48 h. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, dichloromethane: methanol = 95:5) to give a yellow solid (218 mg, 58 %).

Yellow solid; mp: 91.8 – 93.4 °C; HRMS (ESI) [M + H]⁺ found 889.3309, calculated 889.3340 for [C₅₄H₅₃N₂O₆S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.35 - 7.19 (m, 31 H, H_{trityl}, H_{Ar}), 6.58 (s, 1 H, H_{Ar}), 4.19 - 4.11 (m, 4 H, CH₂-OH), 3.56 - 3.51 (m, 4 H, CH₂-OH), 2.80 - 2.75 (m, 4 H, CH₂-N), 2.69 - 2.62 (m, 4 H, CH₂-N); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 169.2 (CO), 165.3 (CO), 143.5 (C_{Ar}-C-S), 143.3 (C_{Ar}-C-S), 139.7 (C_{Ar}-CONH), 138.1 (C_{Ar}-CONH), 135.9 (C_{Ar}-S), 133.1 (C_{Ar}H), 130.1 (C_{Ar} trityl), 129.9 (C_{Ar}-S), 128.3

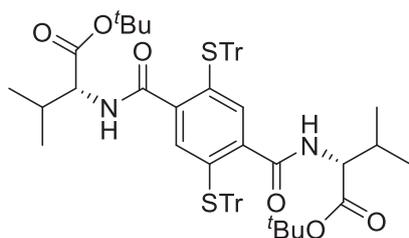
(C_{Ar}H), 127.8 (C_{Ar} trityl), 127.2 (C_{Ar} trityl), 71.5 (C_{trityl}), 70.6 (C_{trityl}), 64.6 (CH₂-OH), 60.6 (CH₂-OH), 52.8 (CH₂-N), 50.2 (CH₂-N)



II-85 N₁,N₁,N₄,N₄-tetrakis(2-hydroxyethyl)-2,5-dimercaptoterephthalamide

A solution of **II-84** N₁,N₁,N₄,N₄-tetrakis(2-hydroxyethyl)-2,5-bis(tritylthio)terephthalamide (200 mg, 0.225 mmol, 1 eq) in a mixture of degassed DCM/TFA/Et₃SiH 45/55/5 (5.18 mL / 7.12 mL / 0.65 mL) was stirred for 3 h under inert atmosphere. Solvent and volatile compounds were evaporated. The remaining residue was dissolved in MeOH (20 mL). This methanolic phase was washed with degassed heptane (4*20 mL), and then evaporated to dryness, affording a yellow oil (108 mg, 82 %).

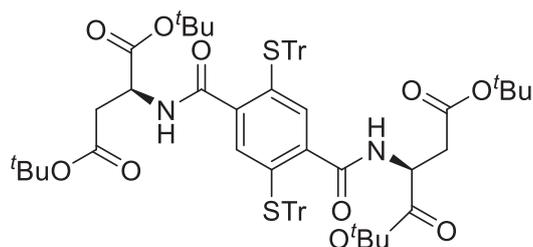
Yellow oil; HRMS (ESI) [M + H]⁺ found 405.1131, calculated 405.1149 for [C₁₆H₂₅N₂O₆S₂]⁺; ¹H NMR (400 MHz, D₂O) δ ppm = 7.37 (s, 2 H, Ar), 4.66 - 4.58 (m, 4 H, CH₂-OH), 3.89 - 3.80 (m, 4 H, CH₂-OH), 3.61 - 3.51 (m, 4 H, CH₂-NH), 3.32 - 3.22 (m, 4 H, CH₂-NH); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 165.9 (CO), 139.7 (C_{Ar}-CONH), 129.3 (C_{Ar}H), 125.9 (C_{Ar}-S), 60.8 (CH₂-OH), 58.5 (CH₂-OH), 56.4 (CH₂-N), 51.3 (CH₂-N)



II-86 (2R,2'R)-di-tert-butyl 2,2'-((2,5-bis(tritylthio)terephthaloyl)bis(azanediyl))bis(3-methylbutanoate)

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (500 mg, 0.7 mmol, 1 eq) with HOBT (247 mg, 1.82 mmol, 2.6 eq) in dry DCM (25 mL) was added at 0°C EDC (323 μL, 1.82 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of the amine hydrochloride (383 mg, 1.82 mmol, 2.6 eq) and DIPEA (0.935 mL, 5.49 mmol, 7.85 eq) in dry DCM (10 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 8:2) to give a white solid (328 mg, 46 %).

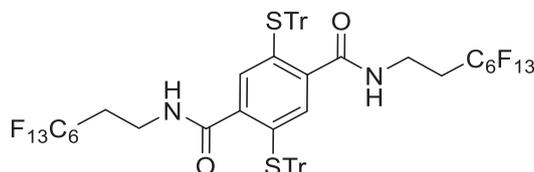
White solid; mp: 157.2-160.1 °C; HRMS (ESI) [M + H]⁺ found 1025.4578, calculated 1025.4592 for [C₆₄H₆₉N₂O₆S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.36 - 7.16 (m, 30 H), 6.80 (s, 2 H), 5.73 (d, J = 8.6 Hz, 2 H), 4.39 (dd, J = 5.0, 8.7 Hz, 2 H), 2.01 (dd, J = 6.8, 12.0 Hz, 2 H), 1.43 (s, 18 H), 0.77 (d, J = 6.8 Hz, 6 H), 0.67 (d, J = 6.8 Hz, 6 H); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 170.6 (CO-OtBu), 166.6 (Ar-CONH), 143.6 (C_{Ar}-C-S), 137.4 (C_{Ar}-CONH), 132.9 (C_{Ar} trityl), 130.0 (C_{Ar} trityl), 129.9 (C_{Ar} trityl), 127.8 (C_{Ar}-S), 126.9 (C_{Ar}H), 81.7 (C_{tBu}), 71.1 (C_{trityl}), 57.5 (N-CH), 31.7 (CH), 28.0 (CH₃ tBu), 18.8 (CH₃, isopropyl), 17.7 (CH₃, isopropyl)



II-88 (2S,2'S)-tetra-tert-butyl 2,2'-((2,5-bis(tritylthio)terephthaloyl)bis(azanediyl))disuccinate

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (200 mg, 0.28 mmol, 1 eq) with HOBT (110 mg, 0.73 mmol, 2.6 eq) in dry DCM (20 mL) was added at 0°C EDC (130 μ L, 0.73 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of the amine hydrochloride (206 mg, 0.73 mmol, 2.6 eq) and DIPEA (0.374 mL, 2.20 mmol, 7.85 eq) in dry DCM (10 mL) was then added at 0°C. The mixture was stirred at room temperature for 72 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 8:2) to give a white solid (214 mg, 65 %).

White solid; mp: 86.1-89.0 °C; HRMS (ESI) [M + H]⁺ found 1169.5004, calculated 1169.5014 for [C₇₀H₇₇N₂O₁₀S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.34 - 7.29 (m, 12 H, Ar), 7.25 - 7.17 (m, 18 H, Ar), 6.80 (s, 2 H, Ar), 6.28 - 6.22 (m, 2 H, NH), 4.60 (td, J = 4.6, 8.3 Hz, 2 H, CH-N), 2.71 (dd, J = 4.9, 17.1 Hz, 2 H, CH₂), 2.55 (dd, J = 4.8, 17.0 Hz, 2 H, CH₂), 1.41 (d, J = 2.2 Hz, 36 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 170.1 (CH₂-CO-OtBu), 169.1 (CH-CO-OtBu), 166.4 (Ar-CONH), 143.7 (C_{Ar}-C-S), 138.0 (C_{Ar}-CONH), 132.7 (C_{Ar} trityl), 130.8 (C_{Ar} trityl), 130.1 (C_{Ar} trityl), 127.8 (C_{Ar}-S), 126.9 (C_{Ar}-H), 82.0 (C_{tBu}, CH-COotBu), 81.1 (C_{tBu}, CH₂-COotBu), 71.4 (C_{trityl}), 49.4 (CH), 37.5 (CH₂), 28.1 (CH₃, CH₂-COotBu), 27.8 (CH₃ CH-COotBu).

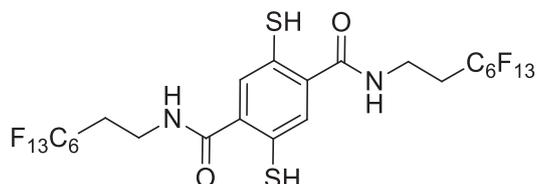


II-90 N₁,N₄-bis(3,3,4,4,5,5,6,6,7,7,8,8,8-tridecafluorooctyl)-2,5-bis(tritylthio)terephthalamide

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (78 mg, 0.109 mmol, 1 eq) with HOBT (40 mg, 0.24 mmol, 2.2 eq) in dry DCM (5.5 mL) was added EDC at 0°C (50 μ L, 0.24 mmol, 2.2 eq) and the mixture was stirred for 1 hour under N₂. 3,3,4,4,5,5,6,6,7,7,8,8,8-tridecafluorooctan-1-amine hydrochloride (96 mg, 0.24 mmol, 2.2 eq) was then added dropwise at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 8:2) to give a white powder (108 mg, 71 %).

White solid; mp: 199.4 degradation °C; HRMS (ESI) [M + Na]⁺ found 1427.2187, calculated 1427.2165 for [C₆₂H₄₂F₂₆N₂NaO₂S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.24 (s, 30 H, Ar_{Trityl}), 6.90 (s, 2 H, H_{Ar}), 4.92 (t, J = 6.0 Hz, 2 H, NH), 3.33 (q, J = 6.5 Hz, 4 H, N-CH₂), 2.29 - 2.06 (m, 4 H, CH₂-CF₂); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 166.7 (Ar-CON), 143.6 (C_{Ar}-C-S), 139.2 (C_{Ar}-CON), 133.1 (C_{Ar}-S), 132.1 (C_{Ar}-H), 130.0 (C_{Ar}

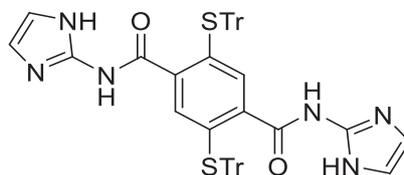
trityl), 127.9 (C_{Ar} trityl), 127.2 (C_{Ar} trityl), 71.0 (C_{trityl}), 32.1 (C-NH), 30.4 (C-CF₃); ¹⁹F NMR (400 MHz, CDCl₃) δ ppm = -126.18 - -126.00 (3 F, m), -123.64 (2 F, s), -122.83 (2 F, s), -121.81 (2 F, s), -114.16 (2 F, t, J=13.70 Hz), -80.73 (2 F, t, J=9.92 Hz)



II-91 N₁,N₄-bis(3,3,4,4,5,5,6,6,7,7,8,8,8-tridecafluorooctyl)-2,5-dimercaptoterephthalamide

II-90 N₁,N₄-bis(3,3,4,4,5,5,6,6,7,7,8,8,8-tridecafluorooctyl)-2,5-bis(tritylthio)terephthalamide (100 mg, 0.071 mmol, 1 eq) dissolved in a solution of degassed and dry DCM (1.64 mL, 45 % volume), TFA (2.25 mL, 55 % volume) and Et₃SiH (0.2 mL, 5 % volume). The mixture was stirred at room temperature for 45 mn. Then, solvent and volatile compounds were evaporated. The remaining residue was washed with degassed heptane (4*15 mL), and then evaporated to dryness, affording a white solid (40 mg, 61 %).

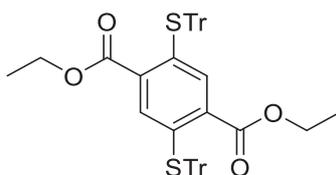
White solid; mp: 231.2 degradation °C; HRMS (ESI) [M - H]⁻ found 919.0015, calculated 919.0009 for [C₂₄H₁₃F₂₆N₂O₂S₂]⁻; ¹H NMR (400 MHz, (CD₃)₂CO + NaOD / D₂O) δ ppm = 8.10 (s, 2 H, H_{Ar}), 3.86 - 3.51 (m, 4 H, N-CH₂), 2.71 - 2.38 (m, 4 H, CH₂-CF₂); ¹³C NMR (400 MHz, (CD₃)₂CO + NaOD / D₂O) δ ppm = 170.6 (Ar-CON), 141.7 (C_{Ar}-CON), 139.0 (C_{Ar}H), 133.1 (C_{Ar}-S); 30.9 (C-NH), 30.4 (C-CF₃); ¹⁹F NMR (400 MHz, (CD₃)₂CO + NaOD / D₂O) δ ppm = -126.61 (3 F, s); -123.82 (2 F, s); -123.31 (2 F, s); -122.34 (2 F, s); -114.38 (2 F, s); -81.40 (2 F, s)



II-92 N₁,N₄-di(1H-imidazol-2-yl)-2,5-bis(tritylthio)terephthalamide

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (400 mg, 0.56 mmol, 1 eq) with HOBT (198 mg, 1.44 mmol, 2.6 eq) in dry DMF (30 mL) was added at 0°C EDC (259 μ L, 1.44 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. 2-amino-imidazole sulfate (190 mg, 1.44 mmol, 2.6 eq) was then added. The mixture was stirred at 80°C for 48 h. After cooling, water (200 mL) was then added. The resulting precipitate was filtrated and dried under vacuum at 50°C overnight. The crude product was purified by column chromatography (SiO₂, dichloromethane: methanol = 95:5) to give a yellow oil (88 mg, 19 %).

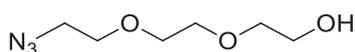
Yellow oil; HRMS (ESI) [M + H]⁺ found 845.2725, calculated 845.2727 for [C₅₂H₄₁N₆O₂S₂]⁺; ¹H NMR (400 MHz, DMSO-d₆) δ ppm = 7.76 - 7.67 (m, 1 H, NH), 7.57 (m, 1 H, NH), 7.32 - 7.18 (m, 32 H, H_{Trityl} + H_{Im}), 7.15 - 7.09 (m, 2 H, H_{Im}), 6.45 (s, 2 H, H_{Ar}); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 147.8 (CO), 143.8 (C_{Ar}-C-S), 132.7 (C_{Ar}-CONH), 129.6 (C_{Im_IV}), 129.0 (C_{ImH}), 128.6 (C_{ImH}), 128.3 (C_{Ar}-S), 127.8 (C_{Ar} trityl), 127.5 (C_{Ar} trityl), 126.6 (C_{Ar}H), 126.2 (C_{Ar} trityl), 80.6 (C_{trityl})



II-96 diethyl 2,5-bis(tritylthio)terephthalate

To a solution of **II-67** 2,5-bis(tritylthio)terephthalic acid (3.1 g, 4.34 mmol, 1 eq) with HOBt (1.53 g, 11.3 mmol, 2.6 eq) in dry DCM (200 mL) was added EDC (2.0 mL, 11.3 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. Ethanol (0.66 mL, 11.3 mmol, 2.6 eq) was then added dropwise at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, chloroform : methanol = 8:2) to give a yellow powder (2.56 g, 77,8 %).

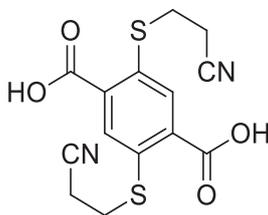
Yellow solid; mp: 230.6 – 232.1 °C; HRMS (ESI) [M + Na]⁺ found 793.2405, calculated 793.2417 for [C₅₀H₄₂NaO₄S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.38 - 7.31 (m, 15 H, H_{Ar}), 7.24 - 7.17 (m, 17 H, H_{Ar}), 4.04 (q, J = 7.2 Hz, 4 H, CH₂), 1.05 (t, J = 7.2 Hz, 6 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 165.2 (Ar-CON), 143.4 (C_{Ar}-C-S), 135.8 (C_{Ar}-CON), 133.1 (C_{Ar}H), 130.2 (C_{Ar} trityl), 127.7 (C_{Ar} trityl), 126.9 (C_{Ar} trityl), 70.8 (C_{trityl}), 60.9 (CH₂), 13.9 (CH₃)



II-97 2-(2-(2-azidoethoxy)ethoxy)ethanol

To a sodium azide solution (1.6 g, 24.6 mmol, 2 eq) in distilled water (100 mL) was added dropwise **5-1** 2-(2-(2-chloroethoxy)ethoxy)ethanol (2 g, 1.72 mL, 12.3 mmol, 1 eq). The mixture was heated at reflux and stirred for 24 h. After cooling, the aqueous phase was extracted with DCM (3*50 mL). The organic phase was dried over anhydrous Na₂SO₄, filtered, and dried under vacuum to yield a colourless oil (2.11 g, 98 %).

¹H NMR (300 MHz, CDCl₃) δ ppm = 3.73 - 3.68 (m, J = 3.7 Hz, 2 H, N₃-CH₂-CH₂-O), 3.66 - 3.62 (m, 6 H, CH₂-OC), 3.60 - 3.56 (m, 2 H, CH₂-CH₂-OH), 3.37 (t, J = 5.0 Hz, 2 H, CH₂-N₃), 2.60 (s, 1 H, OH) data matched with literature reference³⁴²

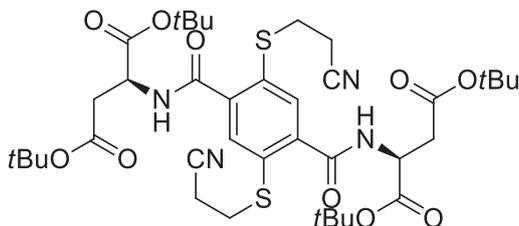


II-98 2,5-bis((2-cyanoethyl)thio)terephthalic acid

To a solution of NaH (313 mg, 13 mmol, 6 eq) in degassed and dry THF (5 mL) was added a solution of **II-1** (500 mg, 2.17 mmol, 1 eq) in degassed and dry THF (10 mL). The mixture was stirred at room temperature under an inert atmosphere for 1 h. Then, 3-bromopropanitrile (3.6 mL, 43.4 mmol, 20 eq) was added dropwise. The mixture was then heated at reflux under inert atmosphere for 48 h. The solvent was evaporated, and chloroform was added (50 mL). After sonication, the precipitate was

filtrated, and sonicate with 60 mL of HCl 1 M. The solid was then filtrated, and washed with distilled water (100 mL). The solid was dried under vacuum to yield a yellow powder (688 mg, 94 %).

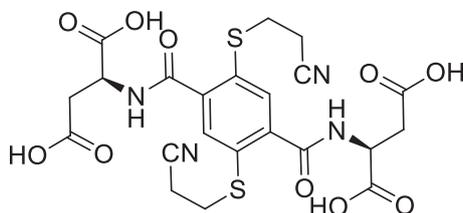
Yellow solid; mp: 234.6 - 237.4 °C; HRMS (ESI) $[M + Na]^+$ found 359.0133, calculated 359.0131 for $[C_{14}H_{12}NaO_4S_2]^+$; 1H NMR (400 MHz, DMSO- d_6) δ ppm = 13.59 (br. s., 2 H, COOH), 7.80 (s, 2 H, Ar), 3.26 (t, J = 6.8 Hz, 4 H, S-CH₂), 2.87 (t, J = 6.8 Hz, 4 H, CH₂-CN); ^{13}C NMR (400 MHz, DMSO- d_6) δ ppm = 166.7 (CO), 134.1 (C_{Ar}-CO), 132.9 (C_{Ar}-S), 119.2 (CN), 27.2 (CH₂-S), 16.6 (C-CN)



II-100 (2S,2'S)-tetra-tert-butyl 2,2'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(azanediyl))disuccinate

To a solution of **II-99** 2,5-bis((2-cyanoethyl)thio)terephthalic acid (100 mg, 0.297 mmol, 1 eq) with HOBT (104 mg, 0.773 mmol, 2.6 eq) in dry DCM (22 mL) was added at 0°C EDC (134 μ L, 0.773 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of the amine hydrochloride (218 mg, 0.773 mmol, 2.6 eq) and DIPEA (0.400 mL, 2.33 mmol, 7.85 eq) in dry DCM (11 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 1:1) to give a white solid (182 mg, 78 %).

White solid; mp: 134.3 – 137.2 °C; HRMS (ESI) $[M + H]^+$ found 791.3318, calculated 791.3354 for $[C_{38}H_{55}N_4O_{10}S_2]^+$; 1H NMR (400 MHz, CDCl₃) δ ppm = 7.79 (s, 2 H, Ar), 7.61 (d, J = 8.1 Hz, 2 H, NH), 4.86 (td, J = 4.2, 8.8 Hz, 2 H, CH-N), 3.26 - 3.12 (m, 4 H, S-CH₂), 2.99 (dd, J = 4.2, 17.4 Hz, 2 H, AB system, CH₂-CO₂tBu), 2.87 (dd, J = 4.4, 17.4 Hz, 2 H, AB system, CH₂-CO₂tBu), 2.67 (t, J = 7.3 Hz, 4 H, CH₂-CN), 1.49 (s, 18 H, CH₃), 1.45 - 1.44 (m, 18 H, CH₃); ^{13}C NMR (400 MHz, CDCl₃) δ ppm = 170.3 (CH₂-CO-OtBu), 169.3 (CH-CO-OtBu), 165.6 (Ar-CONH), 139.4 (C_{Ar}-CONH), 132.8 (C_{Ar}-S), 132.2 (C_{Ar}-H), 117.6 (CN), 82.7 (C_{tBu}, CH-COotBu), 81.6 (C_{tBu}, CH₂-COotBu), 49.7 (CH), 37.2 (CH₂-CH), 30.6 (CH₂-S), 28.0 (CH₃, CH₂-COotBu), 27.9 (CH₃ CH-COotBu), 18.0 (C-CN)

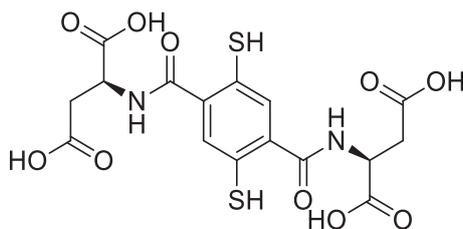


II-101 (2S,2'S)-2,2'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(azanediyl))disuccinic acid

To a solution of **II-100** (2S,2'S)-tetra-tert-butyl 2,2'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(azanediyl))disuccinate (100 mg, 0.126 mmol, 1 eq) in dry DCM (12 mL) was added TFA (8 mL) and triethylsilane (0.200 mL, 1.26 mmol, 10 eq) at 0°C. The mixture was stirred at room temperature for 4 h. Solvent and volatile compounds were evaporated. The remaining residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 258

0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give a white solid (39 mg, 55 %).

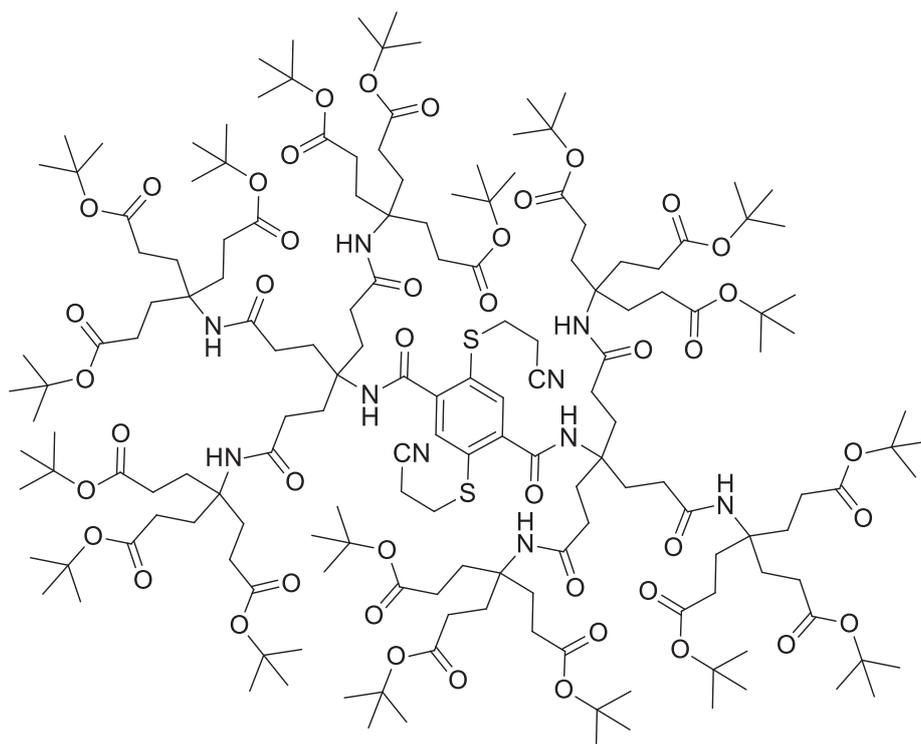
White solid, mp: 208.7 decomposition; HRMS (ESI) $[M + Na]^+$ found 589.0657, calculated 589.0670 for $[C_{22}H_{22}N_4NaO_{10}S_2]^+$; 1H NMR (400 MHz, MeOD) δ ppm = 7.66 (s, 2 H, Ar), 4.99 - 4.97 (m, 2 H, CH-N), 3.25 (dt, $J = 1.5, 7.0$ Hz, 4 H, S-CH₂), 3.05 (dd, $J = 5.3, 16.9$ Hz, 2 H, AB system, CH₂-CO₂H), 2.95 (dd, $J = 7.0, 16.7$ Hz, 2 H, AB system, CH₂-CO₂H), 2.78 (t, $J = 7.0$ Hz, 4 H, CH₂-CN); ^{13}C NMR (400 MHz, MeOD) δ ppm = 174.1 (CH₂-COOH), 173.8 (CH-COOH), 169.6 (Ar-CONH), 141.7 (C_{Ar}-CONH), 133.4 (C_{Ar}-S), 132.3 (C_{Ar}H), 119.8 (CN), 51.0 (CH), 36.9 (CH₂-CH), 31.2 (CH₂-S), 18.7 (C-CN)



II-101 (2S,2'S)-2,2'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(azanediyl))disuccinic acid

On **II-100** (2S,2'S)-2,2'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(azanediyl))disuccinic acid (165 mg, 0.232 mmol, 1 eq) was added under inert atmosphere at 0°C a degassed solution of CsOH.H₂O (780 mg, 4.65 mmol, 20 eq) in THF/MeOH (3.3 mL / 3.3 mL). The mixture was stirred at room temperature for 3 h. The resulting precipitate was filtrated and dried under vacuum. The solid was then stirred in a minimum amount of concentrate HCl to remove caesium ion. The precipitate formed was filter off and then dried under vacuum, affording a pale yellow powder. (106 mg, 99 %).

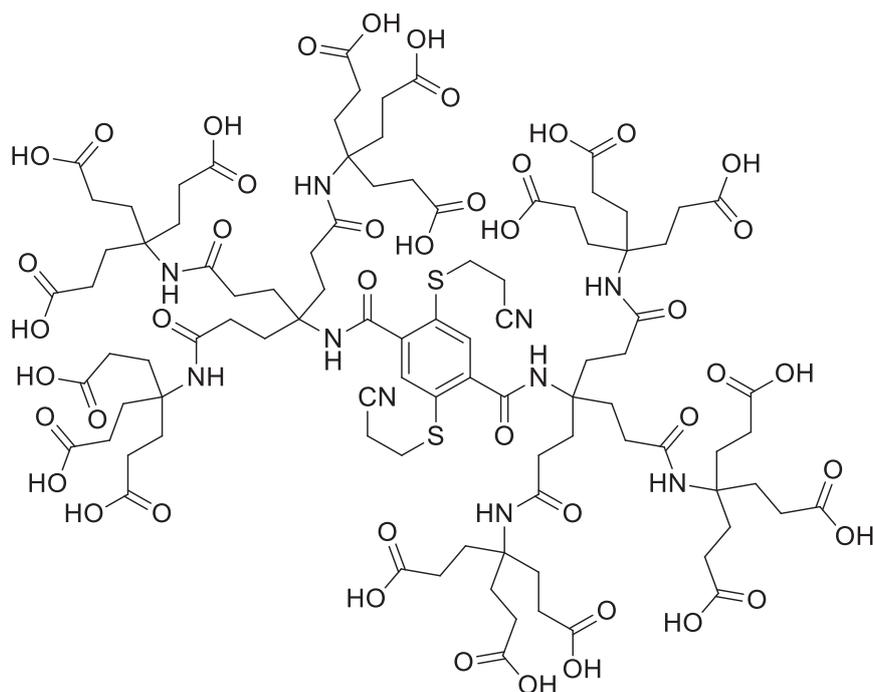
Pale yellow solid; mp: degradation 286 °C; HRMS (ESI) $[M - H]^-$ found 459.0158, calculated 459.0174 for $[C_{16}H_{15}N_2O_{10}S_2]^-$; 1H NMR (400 MHz, D₂O) δ ppm = 7.92 (s, 2 H, Ar), 4.55 (dd, $J = 4.3, 9.4$ Hz, 2 H, CH), 2.79 - 2.57 (m, AB system, 4 H, CH₂); ^{13}C NMR (400 MHz, D₂O) δ ppm = 179.1 (COOH), 178.9 (COOH), 169.6 (CONH), 139.1 (C_{Ar}-CONH), 137.5 (C_{Ar}H), 134.2 (C_{Ar}-S), 54.4 (CH), 40.0 (CH₂)



II-102 tetra-tert-butyl 4,4'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)(4-amino-4-(3-((1,7-di-tert-butoxy-4-(3-(tert-butoxy)-3-oxopropyl)-1,7-dioxoheptan-4-yl)amino)-3-oxopropyl)heptanedioyl)bis(azanediy))bis(4-(3-(tert-butoxy)-3-oxopropyl)heptanedioate)

To a solution of **II-98** 2,5-bis((2-cyanoethyl)thio)terephthalic acid (59 mg, 0.175 mmol, 1 eq) with HOBt (55 mg, 0.402 mmol, 2.3 eq) in dry DCM (10 mL) was added at 0°C EDC (71 μ L, 0.402 mmol, 2.3 eq) and the mixture was stirred for 1 hour under N₂. A solution of the **II-101** amine (500 mg, 0.350 mmol, 2.0 eq) in dry DCM (5 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane: ethyl acetate = 1:1) to give a white solid (213 mg, 38 %).

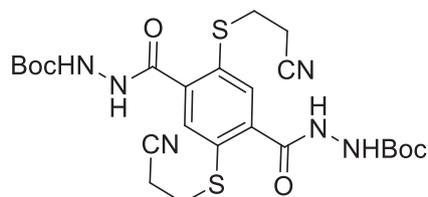
White solid; mp: 68.7 - 71.1 °C; HRMS (ESI) [M + 3 H]³⁺ found 1060.3148, calculated 1060.3109 for [C₁₆₆H₂₇₉N₁₀O₄₄S₂]³⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.66 (s, 2 H, Ar), 7.55 (br. s, 2 H, NH), 6.18 (br. s, 6 H, NH), 3.29 (t, J = 5.9 Hz, 4 H, CH-N), 2.78 (t, J = 6.0 Hz, 4 H, S-CH₂), 2.27 (s, 12 H, CH₂), 2.22 - 2.11 (m, 48 H, CH₂), 1.96 - 1.88 (m, 36 H, CH₂), 1.42 (s, 162 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 172.7 (CO-OtBu), 172.6 (CO-OtBu), 166.7 (Ar-CONH), 156.8, 141.1 (C_{Ar}-CONH), 131.4 (C_{Ar}-S), 130.6 (C_{Ar}H), 118.7 (CN), 80.5 (C_{IV}tBu), 58.7 (C_{IV}), 57.4 (C_{IV}), 49.1 (C_{IV}), 33.9 (CH₂-S), 31.6 (CH₂), 31.4, 29.7 (CH₂), 28.1 (CH₃), 26.9 (CH₂), 25.6 (CH₂), 24.9 (CH₂), 18.3 (CH₂-CN)



II-103 4,4',4'',4''',4''''-((4-(2,5-bis((2-cyanoethyl)thio)-4-((1,7-dioxo-4-(3-oxopropyl)heptan-4-yl)carbamoyl)benzamido)-4-(3-((1,5-dicarboxy-3-(2-carboxyethyl)pentan-3-yl)amino)-3-oxopropyl)heptanedioyl)pentakis(azanediyl))pentakis(4-(2-carboxyethyl)heptanedioic acid)

To a solution of **II-102** tetra-*tert*-butyl 4,4'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)(4-amino-4-(3-((1,7-di-*tert*-butoxy-4-(3-(*tert*-butoxy)-3-oxopropyl)-1,7-dioxoheptan-4-yl)amino)-3-oxopropyl)heptanedioyl)bis(azanediyl))bis(4-(3-(*tert*-butoxy)-3-oxopropyl)heptanedioate) (210 mg, 0.066 mmol, 1 eq) in dry DCM (20 mL) was added TFA (8 mL) at 0°C. The mixture was stirred at reflux overnight. Solvent and volatile compounds were evaporated. The remaining residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give a white solid (23 mg, 16 %).

White solid, mp: 156.3 °C decomposition; HRMS (ESI) [M + 2 H]²⁺ found 1085.4004, calculated 1085.3993 for [C₉₄H₁₃₄N₁₀O₄₄S₂]²⁺; ¹H NMR (400 MHz, MeOD) δ ppm = 7.63 (s, 2 H, Ar), 3.37 - 3.33 (m, 4 H, CH-N), 2.83 (t, J = 6.5 Hz, 4 H, S-CH₂), 2.40 - 2.22 (m, 48 H, CH₂), 2.10 (br. s., 12 H, CH₂), 2.07 - 1.97 (m, 36 H, CH₂); ¹³C NMR (400 MHz, MeOD) δ ppm = 177.2 (CO_{ext}), 175.6 (CO_{mid}), 169.9 (CO_{Ar}), 142.9 (C_{Ar}-CONH), 132.4 (C_{Ar}-S), 131.5 (C_{Ar}H), 120.4 (CN), 60.2 (C_{IV}), 58.8 (C_{IV}), 32.2, 31.0 (CH₂-CH-S), 30.6 (CH₂), 30.6 (CH₂), 30.5 (CH₂), 30.5 (CH₂-S), 29.4 (CH₂), 29.3 (CH₂), 18.9 (CH₂-CN)

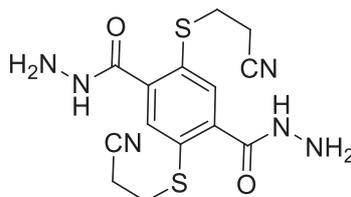


II-105 di-*tert*-butyl 2,2'-((2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(hydrazinecarboxylate)

To a solution of **II-98** 2,5-bis((2-cyanoethyl)thio)terephthalic acid (620 mg, 1.84 mmol, 1 eq) with HOBT (645 mg, 4.78 mmol, 2.6 eq) in dry DCM (40 mL) was added at 0°C EDC (845 μL, 4.78 mmol, 2.6 eq) and the mixture was stirred for 1 hour under N₂. A solution of *t*-butoxycarbonylhydrazine (632 mg, 4.78

mmol, 2.6 eq) in dry DCM (10 mL) was then added at 0°C. The mixture was stirred at room temperature for 48 h. Brine was added and the organic layer was washed twice with brine. The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, DCM) to give a white solid (437 mg, 42 %).

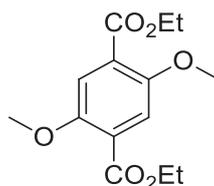
White solid, mp: 166.3 °C decomposition; HRMS (ESI) [M + Na]⁺ found 587.1693, calculated 587.1771 for [C₂₄H₃₂N₆NaO₆S₂]⁺; ¹H NMR (400 MHz, DMSO-d₆) δ ppm = 10.20 (br. s., 2 H, NH), 9.09 (br. s., 2 H, NH), 7.46 (s, 2 H, Ar), 3.24 (t, J = 7.3 Hz, 4 H, S-CH₂), 2.85 (t, J = 7.1 Hz, 4 H, CH₂-CN), 1.44 (s, 18 H, CH₃); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 166.0 (Ar-CONH), 155.2 (C_{Boc}=O), 136.5 (C_{Ar}-CONH), 132.4 (C_{Ar}-S), 128.5 (C_{Ar}H), 119.2 (CN), 79.4 (C_{iVBoc}), 28.1 (CH₃), 27.8 (C-S) 17.0 (CH₂-CN)



II-106 2,5-bis((2-cyanoethyl)thio)terephthalohydrazide

To a solution of **II-98** di-*tert*-butyl 2,2'-(2,5-bis((2-cyanoethyl)thio)terephthaloyl)bis(hydrazinecarboxylate) (437 mg, 0.77 mmol, 1 eq) in dry DCM (20 mL) was added TFA (9.2 mL) and triethylsilane (1.23 mL, 7.70 mmol, 10 eq) at 0°C. The mixture was stirred at reflux overnight. Solvent and volatile compounds were evaporated. The remaining residue was purified by column chromatography (SiO₂, DCM / MeOH 9:1) to give a white solid (100 mg, 43 %).

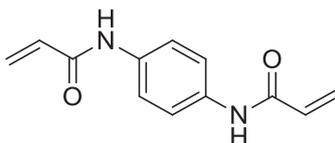
White solid, mp: 236.5 decomposition; HRMS (ESI) [M + H]⁺ found 365.0845, calculated 365.0849 for [C₁₄H₁₇N₆O₂S₂]⁺; ¹H NMR (400 MHz, DMSO-d₆) δ ppm = 9.83 (br. s., 2 H, NH), 7.39 (s, 2 H, Ar), 3.28 - 3.21 (t, J = 7.0 Hz, 4 H, S-CH₂), 2.83 (t, J = 6.9 Hz, 4 H, CH₂-CN); ¹³C NMR (400 MHz, DMSO-d₆) δ ppm = 165.8 (CO), 137.1 (C_{Ar}-CONH), 131.7 (C_{Ar}-S), 127.6 (C_{Ar}H), 119.3 (CN), 27.8 (CH₂-S), 17.1 (C-CN)



II-110 diethyl 2,5-dimethoxyterephthalate

To a solution of diethyl 2,5-dihydroxyterephthalate (500 mg, 1.97 mmol, 1 eq) and anhydrous K₂CO₃ (1.361 g, 9.85 mmol, 5 eq) in dry acetone (10 mL) was added MeI (0.6 mL, 9.85 mmol, 5 eq). The mixture was refluxed overnight. Then, solvent and volatile compounds were evaporated. The remaining residue was dissolved in chloroform (20 mL) and washed with water (2*30 mL). The organic phase was dried with MgSO₄, filtrated and evaporated to dryness, affording a white powder (482 mg, 87 %).

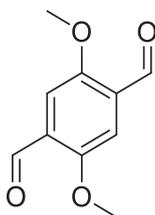
White powder; ¹H NMR (300 MHz, CDCl₃) δ ppm = 7.38 (s, 2 H, Ar), 4.39 (q, J = 7.2 Hz, 4 H, CH₂), 3.90 (s, 6 H, CH₃-O-), 1.40 (t, J = 7.2 Hz, 6 H, CH₃), data matched with literature reference^{351,403}



II-128 N,N'-(1,4-phenylene)bisacrylamide

1,4-phenylenediamine (5.6 mmol, 605.6 mg) was mixed with acetone (25 mL) and cooled to 0 °C in an ice bath. Acryloyl chloride (11 mmol, 0.90 mL) was mixed with acetone (15 mL) and added dropwise to the solution of 1,4-phenylenediamine. The mixture was stirred at room temperature overnight. A solution of saturated NaHCO₃ was then added, the resulting solid was filtered and washed with 100 mL of cold H₂O. The crude product was recrystallized from EtOH/MeOH, to afford white crystals that were collected by filtration (0.75 g, 62%).

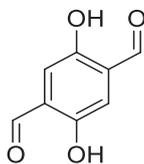
White solid; ¹H NMR (300 MHz, DMSO-d₆) δ ppm = 10.12 (br. s., 2 H, NH), 7.61 (s, 4 H, Ar), 6.50 - 6.33 (m, 2 H, =CH), 6.33 - 6.16 (m, 2 H, =CH₂), 5.73 (br. s., 2 H, =CH₂); data matched with literature³⁹⁵.



II-130 2,5-dimethoxyterephthalaldehyde

To a solution of **II-129**, 1,4-dimethoxybenzene (11.5 g, 83.26 mmol, 1 eq) in distilled Et₂O (276 mL) was added TMEDA (37.4 mL, 249.78 mmol, 3 eq). At 0 °C, *n*BuLi at 2.5 M in hexane (100 mL, 249.78 mmol, 3 eq) was added dropwise. The mixture was stirred at reflux overnight under inert atmosphere. At 0 °C, 1-formylpiperidine (27.74 mL, 249.78 mmol, 3 eq) was added dropwise. The mixture was stirred at room temperature for 1 h. Following the addition of distilled water (300 mL) and HCl 3 M (57.5 mL), the mixture was extracted with hot CHCl₃ (300 mL×4). The organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by recrystallisation in CHCl₃ to give an orange solid (9.51 g, 51 %).

¹H NMR (300 MHz, CDCl₃) δ ppm = 10.51 (s, 2 H, CHO), 7.46 (s, 2 H, Ar), 3.95 (s, 6 H, CH₃) data matched with literature reference⁴⁰⁴.

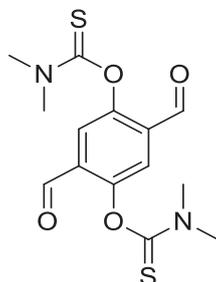


II-131 2,5-dihydroxyterephthalaldehyde

To a solution of **II-130** 2,5-dimethoxyterephthalaldehyde (6.59 g, 33.93 mmol, 1 eq) in distilled DCM (156 mL) was added at 0 °C BBr₃ at 1 M in distilled DCM (133.7 mL, 133.7 mmol, 4 eq). The mixture was stirred for 3 hours under N₂. Distilled water (300 mL) was then added at 0 °C. The mixture was extracted with hot CHCl₃ (300 mL×4) and washed with Rochelle's salt (300 mL) and water (2*300 mL). The organic

phase was dried over anhydrous Na_2SO_4 and filtered. After removing the solvent, the remaining residue was purified by recrystallisation in CHCl_3 to give a yellow solid (3.82 g, 68 %).

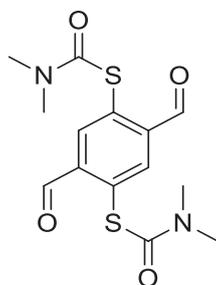
^1H NMR (300 MHz, CDCl_3) δ ppm = 10.23 (s, 2 H, CHO), 9.97 (s, 2 H, OH), 7.25 (s, 2 H, Ar) data matched with literature reference⁴⁰⁵



II-132 O,O'-(2,5-diformyl-1,4-phenylene) bis(dimethylcarbamothioate)

To a solution of **II-131** 2,5-dihydroxyterephthalaldehyde (423 mg, 2.546 mmol, 1 eq) and DABCO (1.142 g, 10.184 mol, 4 eq) in dry DMA (11.3 mL) was added at 0°C dimethylthiocarbamoyl chloride (1.252 g, 10.184 mmol, 4 eq) in 3.7 mL of dry DMA. The mixture was stirred at room temperature for 48 h under inert atmosphere. The white precipitate was filtrated and washed with water. Drying of the resulting solid gave a white powder (618 mg, 71 %).

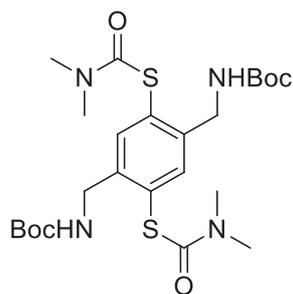
White solid; mp: $226.1 - 228.6^\circ\text{C}$; HRMS (ESI) $[\text{M} + \text{H}]^+$ found 341.0623, calculated 341.0624 for $[\text{C}_{14}\text{H}_{17}\text{N}_2\text{O}_4\text{S}_2]^+$; ^1H NMR (300 MHz, CDCl_3) δ ppm = 10.07 (s, 2 H, CHO), 7.67 (s, 2 H, Ar), 3.52 - 3.48 (m, 6 H, CH_3), 3.45 (s, 6 H, CH_3); ^1H NMR (400 MHz, DMSO-d_6) δ ppm = 10.03 (s, 2 H, CHO), 7.65 (s, 2 H, Ar), 3.43 (s, 6 H, CH_3), 3.41 (s, 6 H, CH_3); ^{13}C NMR (400 MHz, DMSO-d_6) δ ppm = 187.3 (CHO), 185.5 (CS), 151.8 (C_{ArS}), 133.1 ($\text{C}_{\text{Ar-CHO}}$), 123.9 (C_{ArH}), 42.9 (CH_3), 38.5 (CH_3)



II-133 S,S'-(2,5-diformyl-1,4-phenylene) bis(dimethylcarbamothioate)

In a sealed tube was charged **II-132** O,O'-(2,5-diformyl-1,4-phenylene) bis(dimethylcarbamothioate) (400 mg, 1.175 mmol, 1 eq) in dry NMP (10 mL). The mixture was heated at 210°C for 9 mn under microwave. After cooling at 0°C for 1 h, the pink precipitate formed was filtrated and washed with ethanol. Drying of the solid gave a pink powder (220 mg, 55 %).

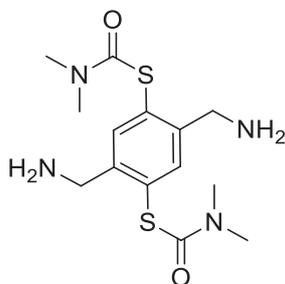
Pink solid; mp: $221.9 - 224.3^\circ\text{C}$; HRMS (ESI) $[\text{M} + \text{H}]^+$ found 341.0625, calculated 341.0624 for $[\text{C}_{14}\text{H}_{17}\text{N}_2\text{O}_4\text{S}_2]^+$; ^1H NMR (300 MHz, CDCl_3) δ ppm = 10.26 (s, 2 H, CHO), 8.19 (s, 2 H, Ar), 3.19 (s, 6 H, CH_3), 3.04 (s, 6 H, CH_3); ^1H NMR (400 MHz, DMSO-d_6) δ ppm = 10.18 (s, 2 H, CHO), 8.03 (s, 2 H, Ar), 3.06 (br. s., 12 H, CH_3); ^{13}C NMR (400 MHz, DMSO-d_6) δ ppm = 189.4 (CHO), 162.9 (CO), 139.9 ($\text{C}_{\text{Ar-CHO}}$), 136.0 (C_{Ar}), 133.5 (C_{ArS}), 36.5 (CH_3)



II-134 di-tert-butyl ((2,5-bis((dimethylcarbamoyl)thio)-1,4-phenylene)bis(methylene))dicarbamate

To the solution of compound **II-133** *S,S'*-(2,5-diformyl-1,4-phenylene) bis(dimethylcarbamothioate) (50 mg, 0.147 mmol, 1 eq) and *t*-butyl carbamate (100 mg, 0.852 mmol, 5.8 eq) in the mixture of MeCN (1.5 mL) and CHCl₃ (3 mL) was added triethylsilane (0.136 μL, 0.852 mmol, 5.8 eq) and TFA (43 μL, 0.56 mmol, 3.8 eq). The resulting mixture was stirred 72 hours at room temperature. Following the addition of saturated aqueous NaHCO₃ (1 mL), the mixture was extracted with CHCl₃ (20 mL×4). The combined organic phase was dried over anhydrous Na₂SO₄ and filtered. After removing the solvent, the remaining residue was purified by column chromatography (SiO₂, cyclohexane : ethyl acetate = 8:2) to give a white solid (57 mg, 71 %).

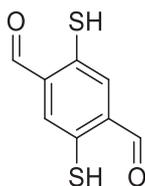
White solid; mp: 53.2-55.7°C; HRMS (ESI) [M + H]⁺ found 565.2121, calculated 565.2125 for [C₂₄H₃₈N₄NaO₆S₂]⁺; ¹H NMR (400 MHz, CDCl₃) δ ppm = 7.60 (s, 2 H, Ar), 5.20 (br. s., 2 H, NH), 4.40 (s, 4 H, CH₂), 3.11 (br. s., 6 H, NCH₃), 3.00 (br. s., 6 H, NCH₃), 1.44 (s, 18 H, CH₃); ¹³C NMR (400 MHz, CDCl₃) δ ppm = 166.1 (N-CO-S), 155.9 (N-CO-O), 142.9 (C_{Ar}-CH₂), 138.2 (C_{Ar}-S), 130.3 (C_{Ar}H), 79.3 (C_{Boc}), 42.9 (CH₂), 37.1 (NCH₃), 28.7 (NCH₃), 28.4 (CH₃ Boc).



II-135 *S,S'*-(2,5-bis(aminomethyl)-1,4-phenylene) bis(dimethylcarbamothioate) (di TFA salt)

To a solution of **II-134** di-*tert*-butyl ((2,5-bis((dimethylcarbamoyl)thio)-1,4-phenylene)bis(methylene))dicarbamate (200 mg, 0.586 mmol, 1 eq) in tetrachloroethane (16.5 mL) was added TFA (5.50 mL). The mixture was stirred at room temperature for 20 mn. The mixture was then evaporated to dryness, yielding a white powder (468 mg, 81 %).

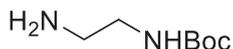
White solid; mp: 204.9-206.1 °C; HRMS (ESI) [M + H]⁺ found 343.1252, calculated 343.1257 for [C₁₄H₂₃N₄O₂S₂]⁺; ¹H NMR (400 MHz, DMSO-*d*₆) δ ppm = 8.36 (br. s., 2 H, H_{TFA}), 7.83 (s, 2 H, Ar), 7.20 (t, J = 51.1 Hz, 2 H, NH), 4.13 (s, 4 H, CH₂), 3.11 (br. s., 6 H, CH₃), 2.94 (br. s., 6 H, CH₃); ¹³C NMR (400 MHz, DMSO-*d*₆) δ ppm = 164.1 (SCON), 139.5 (C_{Ar}-CH₂), 137.9 (C_{Ar}-S), 131.2 (C_{Ar}), 40.5 (CH₂), 37.2 (CH₃), 37.2 (CH₃)



II-138 2,5-dimercaptoterephthalaldehyde

A solution of **II-133** di-*tert*-butyl ((2,5-bis((dimethylcarbamoyl)thio)-1,4-phenylene)bis(methylene))dicarbamate (200 mg, 0.59 mmol, 1 eq) in degassed NaOH (143 mg, 3.67 mmol, 6.5 eq) in isopropanol/H₂O (8:2, 10 mL) was stirred at 60°C under an inert atmosphere for 5 h. The reaction mixture was cooled in ice, and concentrated HCl (3 mL) was added until pH 1. A precipitate was formed, filtered, and washed with water and DCM. The product was dried under vacuum, yielding compound a green brown solid (82 mg, 71 %).

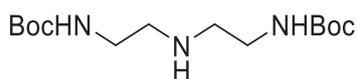
Green brown solid; mp: 147.4°C degradation; HRMS (ESI) [M - H]⁻ found 196.9737, calculated 196.9736 for [C₈H₅O₂S₂]⁻; ¹H NMR (300 MHz, D₂O, NaOD) δ ppm = 10.57 (br. s., 2 H), 7.56 (s, 2 H); ¹³C NMR (400 MHz, D₂O, NaOD) δ ppm = 197.9 (CHO), 146.1 (C_{Ar}-CHO), 139.4 (C_{Ar}S), 135.1 (C_{Ar}H)



II-144 tert-butyl (2-aminoethyl)carbamate

To a solution of 1,2-diaminoethane (1.4 mL, 20 mmol) in chloroform (20 mL) was added dropwise at 0°C a solution of di-*tert*butyldicarbonate (4.37 g, 20 mmol) in chloroform (10 mL) over a period of 3 h. After stirring at room temperature for 16 h, the mixture was washed with brine (3 * 20 mL) and water (20 mL), dried over MgSO₄ and concentrated in vacuo to afford a colourless oil (3.2 g, 99 %).

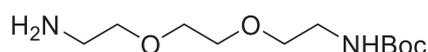
¹H NMR (300 MHz, CDCl₃) δ ppm = 4.89 (br. s., 1 H, NH), 3.17 (q, J = 5.9 Hz, 2 H, CH₂-NH), 2.79 (t, J = 6.0 Hz, 2 H, CH₂-NH₂), 1.48 - 1.42 (m, 9 H, CH₃), 1.36 (br. s., 2 H, NH₂); data matched with literature reference^{201b}.



II-145 di-tert-butyl (azanediylbis(ethane-2,1-diyl))dicarbamate

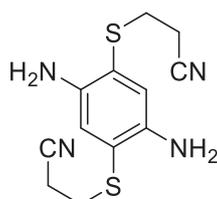
To a solution of imidazole (1.566 g, 23 mmol, 1 eq) in DCM (10 mL) was added di-*tert*-butyl dicarbonate (5.24 g, 24 mmol, 1.05 eq) portion wise. The reaction mixture was stirred for one hour at room temperature. The reaction mixture was washed with 3*20 mL water, dried over Na₂SO₄, filtered and the volatiles were removed under reduced pressure. The residue was dissolved in 15 mL toluene and diethylenetriamine (1.2 mL, 11 mmol, 0.5 eq) was added. The reaction mixture was stirred for two hours at 60°C. Then, 20 mL DCM was added, and the organic phase was washed with 2*20 mL water. The organic phase was dried over Na₂SO₄, filtered and reduced under reduced pressure. The remaining residue was purified by column chromatography (SiO₂, DCM : MeOH : NH₃ = 10:1:0.1) to give a colourless oil (2.889 g, 86 %).

¹H NMR (300 MHz, CDCl₃) δ ppm = 5.04 (br. s., 2 H, (CO)NH), 3.24 - 3.09 (m, 4 H, CH₂-NH(CO)), 2.69 (t, J = 5.9 Hz, 4 H, CH₂-NH), 1.41 (s, 18 H, CH₃); ¹³C NMR (300 MHz, CDCl₃) δ ppm = 156.1 (CO), 79.0 (C_{tBu}), 48.7 (CH₂-N), 40.3 (CH₂-NH(CO)), 28.3 (CH₃), data matched with literature reference²¹⁰.



II-146 tert-butyl (2-(2-(2-aminoethoxy)ethoxy)ethyl)carbamate

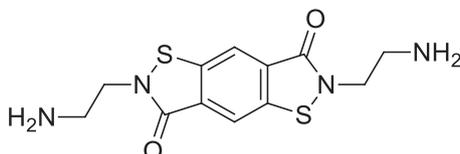
To a stirred solution of 1,8-diamino-3,6-dioxaoctane (2.0 g, 13.5 mmol) in DCM (100 mL) was added at 0°C a solution of di-*t*-butyl dicarbonate (440 mg, 13.5 mmol) in DCM (35 mL) over 30 min. Stirring was continued overnight, after which the solvent was evaporated under vacuum. The resulting oil was subjected to separation by flash column chromatography on silica gel using DCM/MeOH/ammoniac (95:5/2) as eluent to yield a clear oil (1,82 g, 91 %).



II-150 3,3'-((2,5-diamino-1,4-phenylene)bis(sulfaneyl))dipropanenitrile

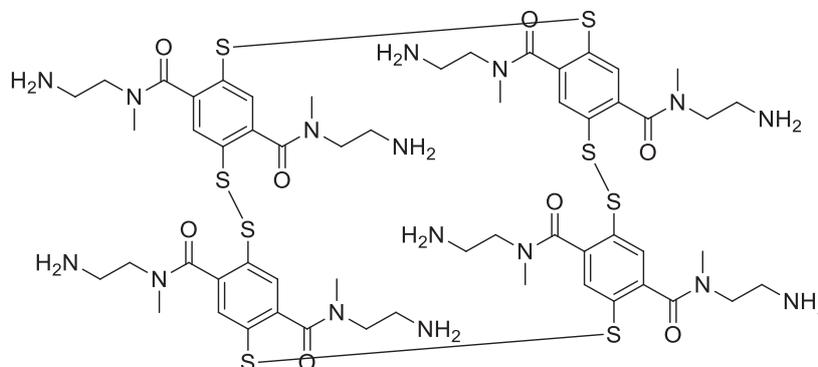
To a solution of **II-36** 2,5-diamino-1,4-benzedithiol dihydrochloride (2.0 g, 8.16 mmol, 1 eq.) in a 1.8 M NaOH (1.44 g, 36 mmol, 4.4 eq.) degassed aqueous solution (20 mL) was added dropwise at 0°C 3-bromopropionitrile (1.5 mL, 18 mmol, 2.2 eq.). Then, cetyltrimethylammonium bromide (74 mg, 0.204 mmol, 0.025 eq.) was added under argon. The mixture was stirred overnight. The resulting precipitate was then filtrated, washed with water and dried under vacuum to yield a yellow solid (2.201 g, 98 %).

Yellow solid; ¹H NMR (300 MHz, CDCl₃) δ ppm = 6.87 (s, 2 H, Ar), 3.94 (br. s., 4 H, NH₂), 2.99 (t, *J* = 7.0 Hz, 4 H, S-CH₂), 2.57 (t, *J* = 7.0 Hz, 4 H, CH₂-CN), ¹³C NMR (300 MHz, CDCl₃) δ ppm = 140.9 (C-N), 122.5 (C_{Ar}), 118.6 (C_{Ar}), 118.1 (CN), 29.6 (C-S), 18.4 (C-CN); matched with literature reference⁴⁰⁶.



II-151 2,6-bis(2-aminoethyl)benzo[1,2-d:4,5-d']diisothiazole-3,7(2H,6H)-dione

HRMS (ESI) [M + H]⁺ found 311.0633, calculated 311.0631 for [C₁₂H₁₄N₄O₂S₂]⁺



(II-71)₄

To a stirred solution of **II-71** N₁,N₄-bis(2-aminoethyl)-2,5-dimercapto-N₁,N₄-dimethylterephthalamide (1.2 mg, 0.004 mmol, 1 eq) in methanol (0.5 mL) was added dropwise over 5 mn [50 mM] iodine solution in MeOH until a yellow colour was persistent. The mixture was analysed by mass spectrometry.

HRMS (ESI) [M + 3 H]³⁺ found 454.4781, calculated 454.4776 for [C₅₆H₈₃N₁₆O₈S₈]³⁺

b. DCC, HPLC and UPLC/HRMS analysis

vii. Library preparation

In a typical experiment, the building blocks were individually dissolved in an aqueous buffer (Tris pH 7.4 200 mM for example) at a given pH or in organic media to obtain building block stock solutions at [2-10] mM. The stock solutions were freshly prepared and quickly engaged in the DCLs: in 1.5 mL vials equipped with a turbulent were placed 500 µL of the building block 1 and 500 µL of building block 2 solution stocks. The vials were left open to the air and stirred at room temperature for 48 hours or more. Templates were freshly prepared at 100 mM solution stocks in the corresponding media. If required, 10 µL were added in the library.

viii. HPLC analysis

Analysis are performed using a 1260 infinity Agilent Technologies system at 25 °C.

Method A: column Agilent Eclipse Plus C₈, 3.5 µm, 4.6 x 150 mm, V_{inj} = 3 µL, flow 1 mL/mn. Gradient water + 0.05 % TFA / MeCN + 0.05 % TFA in 25 mn method (t₀ 80/20, t₅ 60/40, t₈ 0/100).

Method B: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 µL, 2.7 µm, 2.1 x 50 mm flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 22 mn method (t₀ 100/0, t₁₅ 70/30, t₂₂ 0/100).

Method C: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 µL, 2.7 µm, 2.1 x 50 mm flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 28 mn method (t₀ 100/0, t₂ 100/00, t₂₀ 80/20, t₂₅ 0/100, t₂₈ 0/100)

Method D: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 µL, 2.7 µm, 2.1 x 50 mm flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 25 mn method (t₀ 100/0, t_{1.5} 100/00, t₁₅ 85/15, t₂₄ 0/100, t₂₅ 0/100)

Method E: Agilent Nucleodur HILIC EC 250 4.6 V_{inj} = 5 µL, 5 µm, 4.6 x 250 mm flow 0.8 mL/mn. Gradient AcONH₄ 50 mM / MeCN in 25 mn method (t₀ 10/90, t₂₅ 60/40)

Method F: Restek Raptor Biphenyl, 2.7 µm, 3.00 x 50 mm, V_{inj} = 3 µL, flow 0.6 mL/mn. Gradient water + 0.1 % formic acid / MeOH + 0.1 % formic acid in 17 mn method (t₀ 99/1, t_{0.5} 99/1, t₁₅ 50/50, t₁₆ 1/99, t₁₇ 1/99)

Method G: Phenomenex Kinetex C₁₈ 100 A 2.6 µm 100 x 2.1 mm V_{inj} = 10 µL, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t₀ 90/10, t₁₀ 50/50, t₁₅ 40/60, t₁₇ 0/100, t₂₀ 0/100)

Method H: Phenomenex Kinetex C₁₈ 100 A 2.6 µm 100 x 2.1 mm V_{inj} = 10 µL, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t₀ 100/0, t₁ 100/0, t₁₅ 40/60, t₁₇ 0/100, t₁₉ 0/100)

Method I: Restek Raptor Biphenyl, 2.7 μm , 3.00 x 50 mm, $V_{\text{inj}} = 5 \mu\text{L}$, flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeOH + 0.1 % formic acid in 17 mn method (t_0 70/30, t_{15} 0/100, t_{17} 0/100)

Method PREP₁: Zorbax Eclipse XDB C₈ 5.0 μm 250 x 9.4 mm $V_{\text{inj}} = 100 \mu\text{L}$, flow 5.0 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t_0 92/8, t_2 92/8, t_{14} 50/50)

ix. UPLC analysis

Analysis are performed using a DIONEX ultimate 3000 Thermo Scientific system.

Method J: 30 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm $V_{\text{inj}} = 4 \mu\text{L}$, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 20 mn method (t_0 90/10, t_{20} 50/50)

Method K: 45 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm $V_{\text{inj}} = 4 \mu\text{L}$, flow 0.4 mL/mn. Gradient water + 0.1 % TFA / MeCN + 0.1 % TFA in 8 mn method (t_0 85/15, t_8 58/42)

Method L: 45 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm $V_{\text{inj}} = 4 \mu\text{L}$, flow 0.4 mL/mn. Gradient water + 0.1 % TFA / MeCN + 0.1 % TFA in 8 mn method (t_0 85/15, t_8 40/60)

x. UPLC/HRMS analysis

Analysis are performed using a U3000 Thermo Fisher Scientific / QTOF Impact II Bruker system. The column used was an Agilent Poroshell 120 EC C₈, 2.7 μm , 5 x 2.1 mn, $V_{\text{inj}} = 3 \mu\text{L}$, flow 0.8 mL/mn, Gradient water + 0.1 % formic acid / MeCN:MeOH (50/50) (or MeCN) + 0.1 % formic acid for typical run, other methods have been used. Mass spectrometry parameters depend on the analysis, UV detection performed at 250 and 350 nm.

c. Crystallography

Crystals were obtained by slow diffusion of hexane into a solution of the compound of interest in DCM. Suitable crystals from II-51 (not deposited on CCDC) were selected and mounted on a Gemini kappa-geometry diffractometer (Agilent Technologies UK Ltd) equipped with an Atlas CCD detector using Mo radiation ($\lambda = 0.71073 \text{ \AA}$). Intensities were collected at 150 K by means of the CrysAlisPro software³³. Reflection indexing, unit-cell parameter refinement, Lorentz-polarisation correction, peak integration and background determination were carried out using CrysAlisPro software. An analytical absorption correction was applied using the modelled faces of the crystal²¹¹. The resulting sets of hkl were used for structure solution and refinement. The structures were solved using direct methods with SIR97²¹² and the least-squares refinement on F_2 was achieved using CRYSTALS software²¹³. OLEX2²¹⁴ and SHELXT²¹⁵ software were also used for this purpose. All non-hydrogen atoms were refined anisotropically. The hydrogen atoms were initially positioned geometrically and initially refined with soft restraints on the bond lengths and angles to regularise their geometry (C–H in the range 0.93–0.98 \AA) and $U_{\text{iso}}(\text{H})$ (in the range 1.2–1.5 times U_{eq} of the parent atom), after which the positions were refined with riding constraints.

d. DFT details

33 CrysAlisPro, Agilent Technologies, Version, 1.171.34.49, (release 20–01–2011 CrysAlis171.NET), compiled Jan 20 2011, 15:58:25.

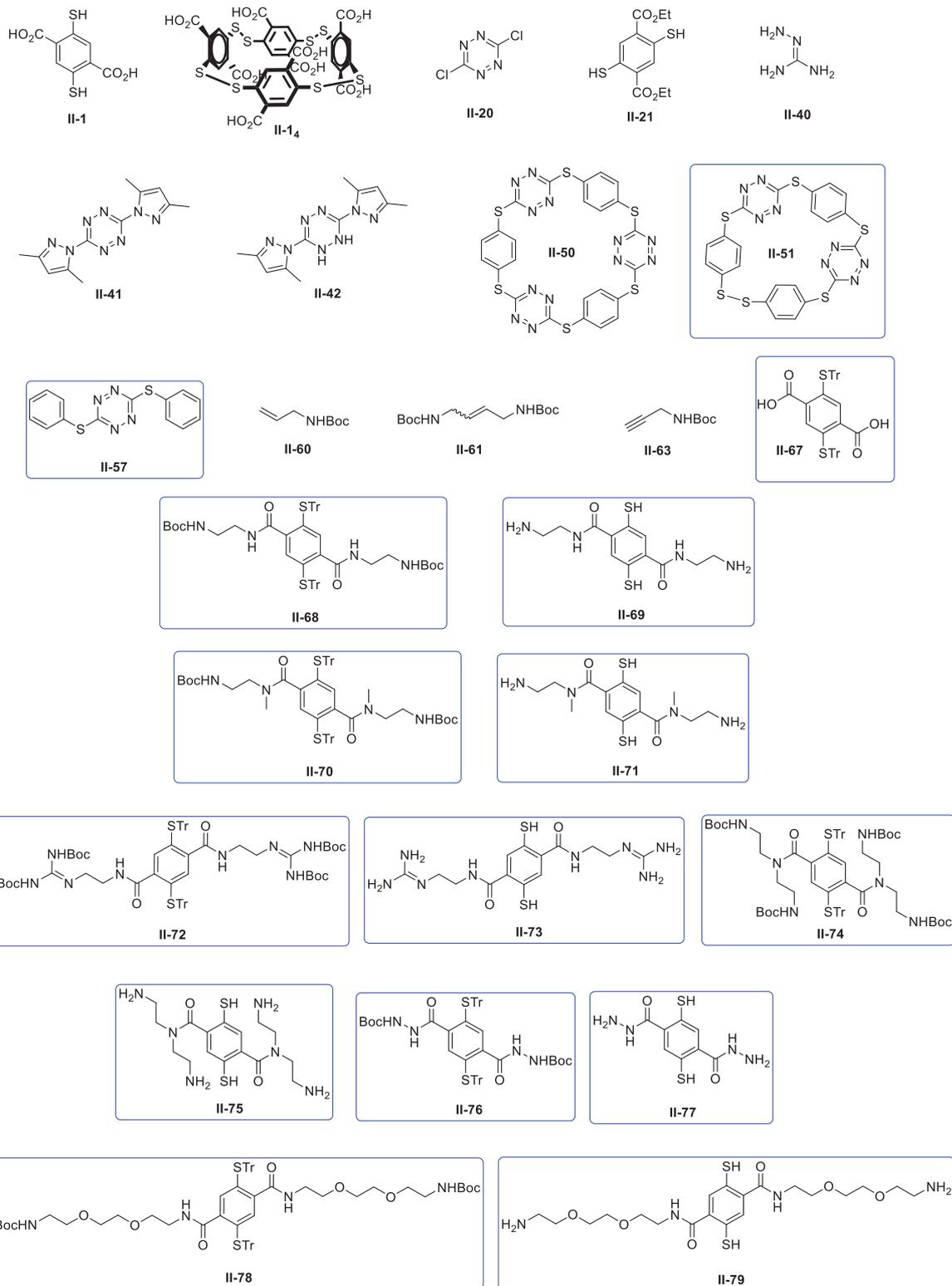
All the molecules have been fully optimized at B3LYP⁷² or APFD²¹⁷ functional /6-311G** (C, H, O, N) 6-311++G** (S) level of theory with Gaussian 09 Rev B.02³⁴. Vibrational frequencies have been computed to characterize transition states and stationary points. IRC were employed to check transition states.

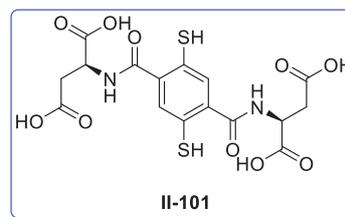
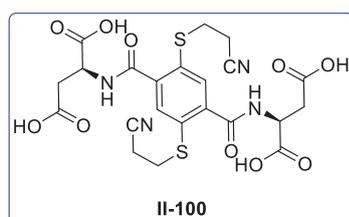
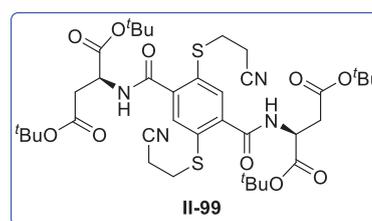
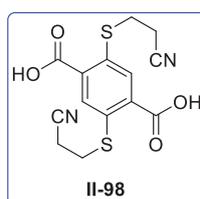
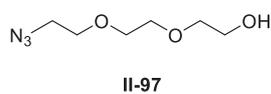
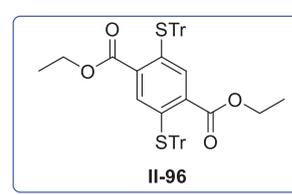
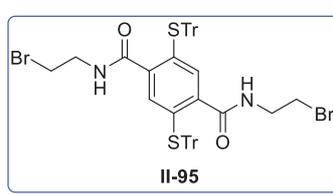
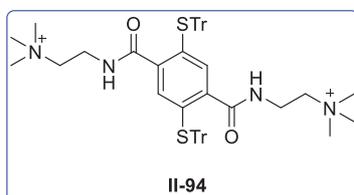
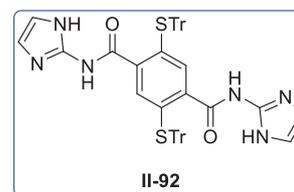
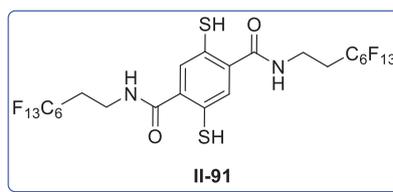
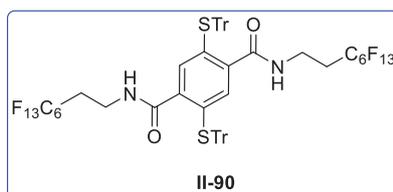
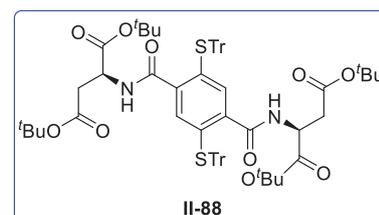
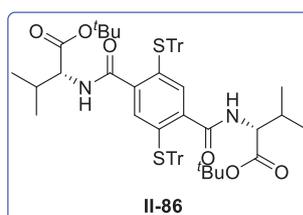
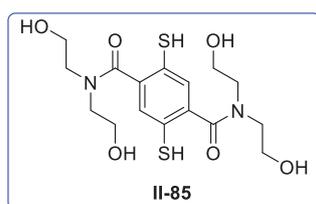
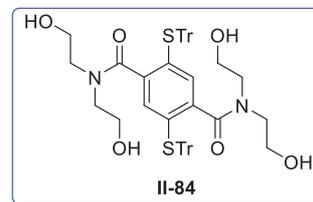
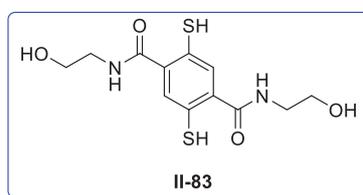
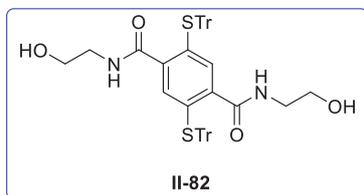
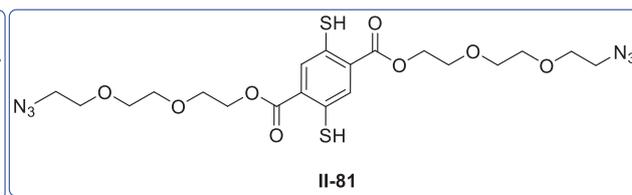
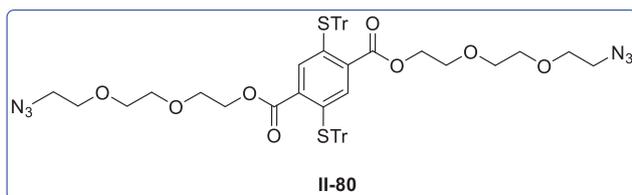
34 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino and G. Zhe, G. 09, Revision D.01.

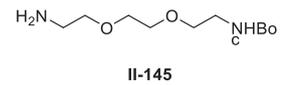
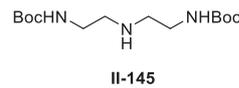
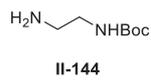
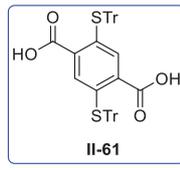
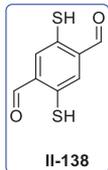
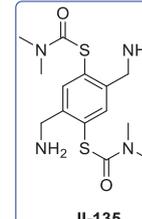
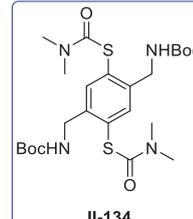
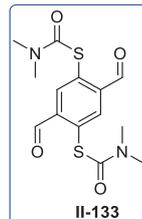
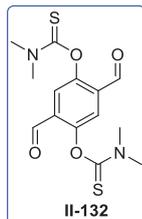
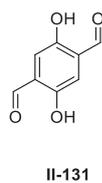
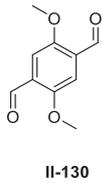
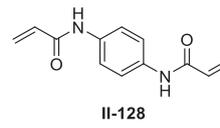
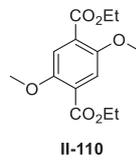
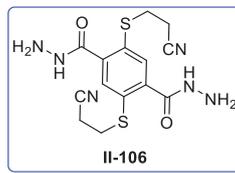
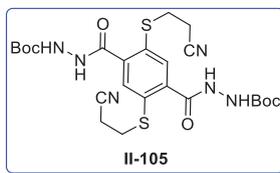
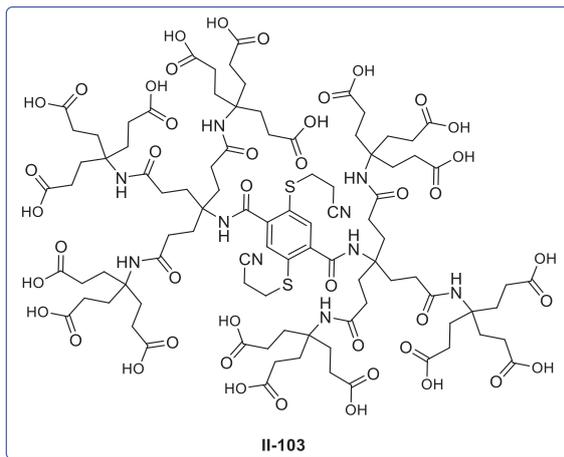
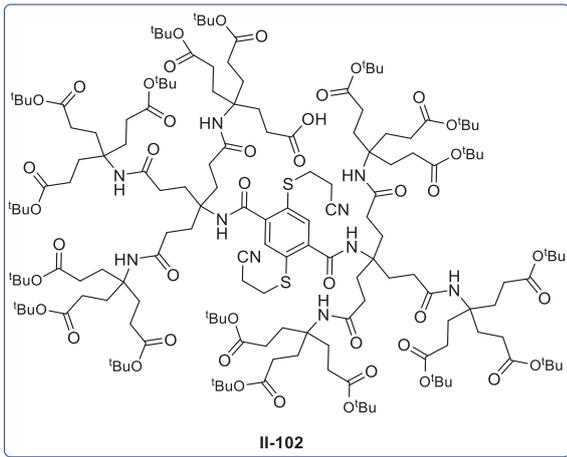
Appendices

a. Appendix II-I: List of molecules

Synthesised molecules of chapter 2, for isolated and ESI-MS analysis + ^1H NMR identification on pure product. New and fully characterised molecules in a box.







b. Appendix II-II: Reaction conditions

• II-135 -> II-136

Test number	Conditions	Observations (qualitative)
1	NaOH _(aq) (4 eq) /isopropanol, 60°C, 1 h ¹⁶	No conversion
2	NaOH _(aq) (16 eq) /isopropanol, reflux, 4 h	No conversion
3	NaOH _(aq) (16) H ₂ O /ethanol, reflux, 4 h	Conversion, II-136, II-137
4	NaOH _(aq) (32 eq) /isopropanol, reflux, 1 h	Conversion, II-136, II-137
5	LiAlH ₄ , DME, 100 °C, 3 h	Conversion + degradation (mass ESI (+) 399.1 439.1 492.6 268.3 238.1 (-) 276.8)
6	DIBAL-H, 80°C, 16 h	Conversion, degradation (mass ESI (+) 399.1 439.1 492.6 268.3 238.1 (-) 276.8)

Scheme 214: Deprotecting conditions of II-134 S-thiocarbamate for the II-135 obtention

• Synthesis of protected amines

Entry	Starting amine	Protected amine	Yield
1 ^{201b}	II-149 <chem>NCCN</chem>	II-144 <chem>CC(C)(C)NCN</chem>	99 %
2 ⁴⁰⁷	II-150 <chem>NCCNCN</chem>	II-145 <chem>CC(C)(C)NCCNC(C)(C)C</chem>	51 %
3 ⁴⁰²	II-151 <chem>NCCOCCOCCN</chem>	II-146 <chem>CC(C)(C)NCCOCCOCCN</chem>	291 %

Scheme 215: Protection of amines with Boc protecting group for further peptide-like coupling with II-61

• II-133 -> II-107

Test number	Conditions	Observations
1	KMnO ₄ , 1,4-dioxane/water, r.t., 4 h. ^{350d}	Insoluble, no conversion
2	NaClO ₂ , NaH ₂ PO ₄ , DMSO/water, 65°C, 6 h. ^{350c}	Insoluble, no conversion
3	KMnO ₄ , pyridine, r.t., 4 h. ^{350a}	Insoluble, degradation
4	NaClO ₂ , CHCl ₃ /TFA 5/1, r.t., 16 h. ^(no reference)	No conversion, partial degradation
5	K ₂ Cr ₂ O ₇ , CHCl ₃ /TFA 5/1, r.t., 16 h. ^(no reference)	No conversion, partial degradation
6	PB-SiO ₂ -KMnO ₄ , CHCl ₃ , r.t., 16 h ^{350e}	No conversion

7	<i>m</i> -CPBA, CHCl ₃ , reflux, 16 h. ^{350b}	No conversion
---	---	---------------

Scheme 216: Reaction conditions for the **II-133** oxidation into **II-107**

• **II-130 -> II-108**

Test number	Conditions	Observations
1	PB-SiO ₂ -KMnO ₄ , CHCl ₃ , r.t., 16 h ^{350e}	Formation of carboxylic acids
2	<i>m</i> -CPBA, CHCl ₃ , reflux, 16 h. ^{350b}	No conversion

Scheme 217: Reaction conditions for the **II-130** oxidation into **II-108**

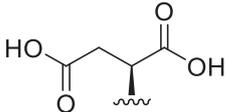
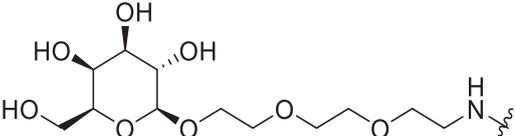
• **II-110 -> II-111**

Test number	Conditions	Observations
1	DIBAL-H, DCM, -78°C -> r.t., 16 h ^{352a}	No conversion
2	LiAlH ₄ , Et ₂ O, -78°C -> r.t., 16 h ^{352b}	No conversion

Scheme 218: Reaction conditions for the **II-110** oxidation into **II-111**

• DCLs conditions involving functionalised 1,4-bis thiophenols with **II-1**

Variation on the R group		Buffers	Building block concentration
II-69		Tris 200 mM pH 7.4 Ammonium acetate 200 mM pH 6.7	[2-20] mM
II-71			
II-73		Phosphate buffer 200 mM 7.4, CaCl ₂ 100 μM MnCl ₂ 100 μM	
II-65			
II-85		Citric acid buffer 200 mM pH 7.4	
II-77		NaCl 100 μM pH 7.4	
II-79		MilliQ water, pH adjusted at 7.4 with NaOH	
II-81			

II-89			
II-113 ³⁵			

Scheme 219: DCL conditions involving functionalised 1,4-bis(2-hydroxyethyl)phenols with II-1

35 II-113 is a building block applied on lectin recognition(c) Pascal, Y., Université Claude Bernard Lyon 1, 2018.. Sample provided by Dr. Pascal. Should be considered only for combination with II-1, not for the whole study.

c. Appendix II-III: Data of the binding assays between II-1₄ and GFPs

Sequences of the GFPs: (one letter code)

GFP MUT2 A206K STSX2 HIS6:

GSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWPTLVTTTFAYGLQCFARYPDH
MKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYIM
ADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGI
THGMDELYKSTSSTSLEHHHHHH

GFP MUT2 A206K STS HIS6:

GSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWPTLVTTTFAYGLQCFARYPDH
MKQHDF
KSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYIMADKQKNGI
KVNFKIR
NIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGITHGMDELYKSTSLEHH
HHHH GFP MUT2 A206K HIS6:

GSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWPTLVTTTFAYGLQCFARYPDH
MKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYIM
ADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAAGI
THGMDELYKLEHHHHHH

GFP MUT2 MAA A206K HIS6:

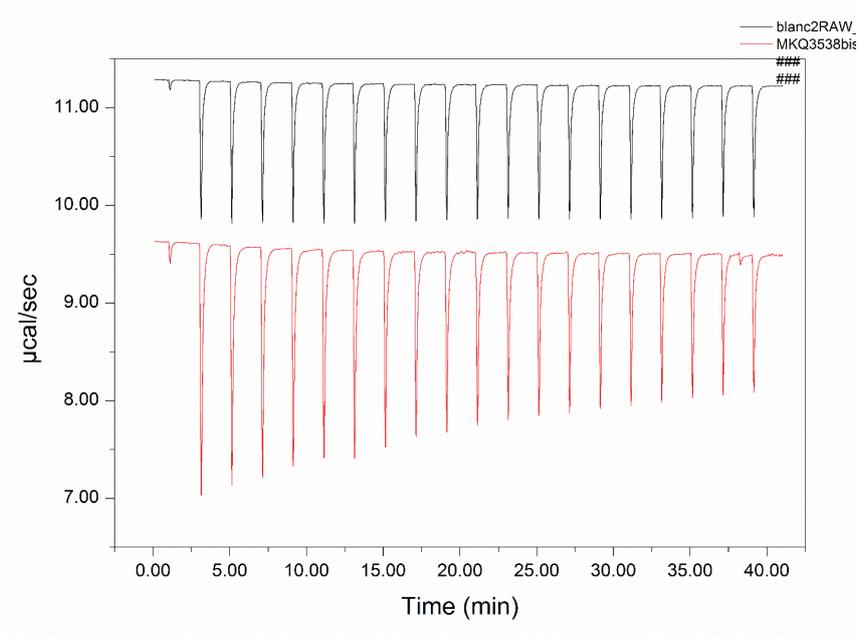
MAASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWPTLVTTTFAYGLQCFARYP
DHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYI
MADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAA
GITHGMDELYKLEHHHHHH

GFP MUT2 MKQ A206K HIS6:

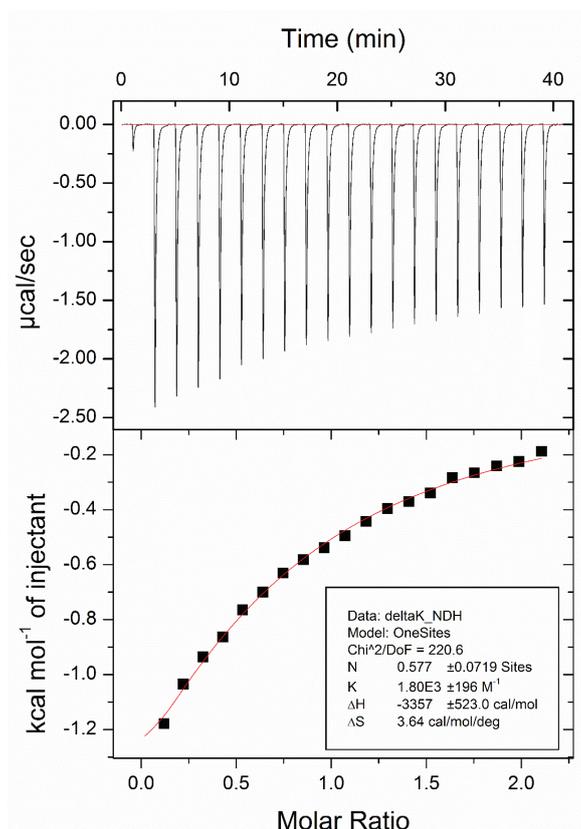
MKQSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWPTLVTTTFAYGLQCFARYP
DHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYI
MADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAA
GITHGMDELYKLEHHHHHH

GFP MUT2 MKQ A206K NOK :

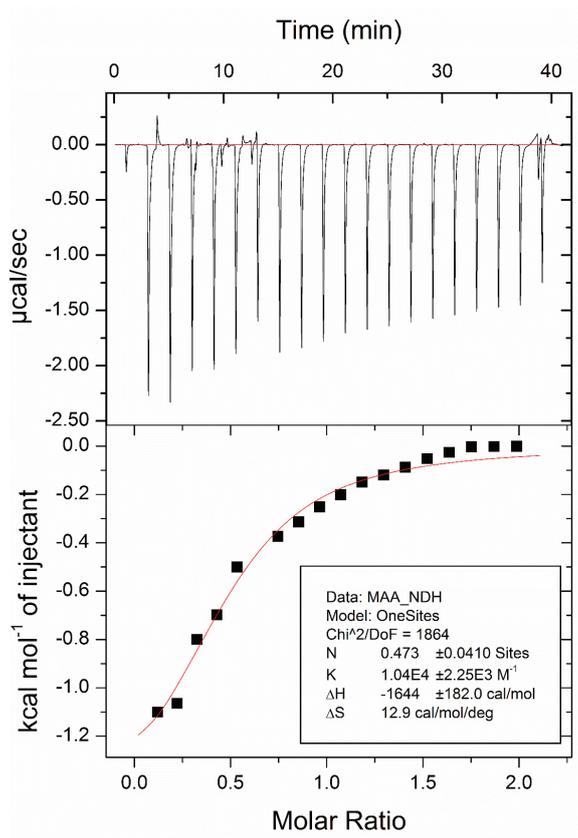
MAASKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLT LKFICTTGKLPVPWPTLVTTTFAYGLQCFARYP
DHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYN SHNVYI
MADKQKNGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDH MVLLEFVTAA
GITHGMDELY



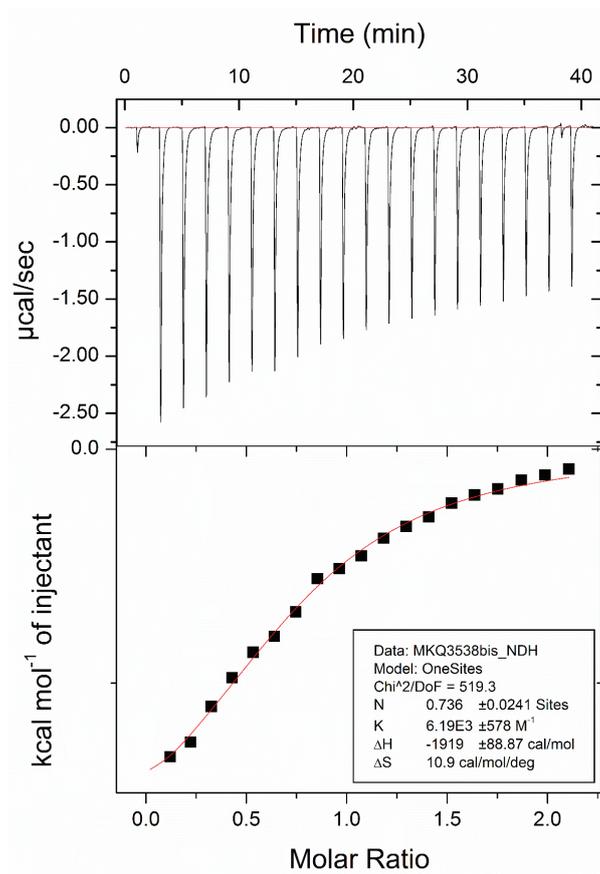
Scheme 220: ITC thermograms of **II-1₄** (5.6 mM) in Tris 100 mM pH 8 compared to MKQ (0.56 mM) and **II-1₄** (5.6 mM) in Tris 100 mM pH 8



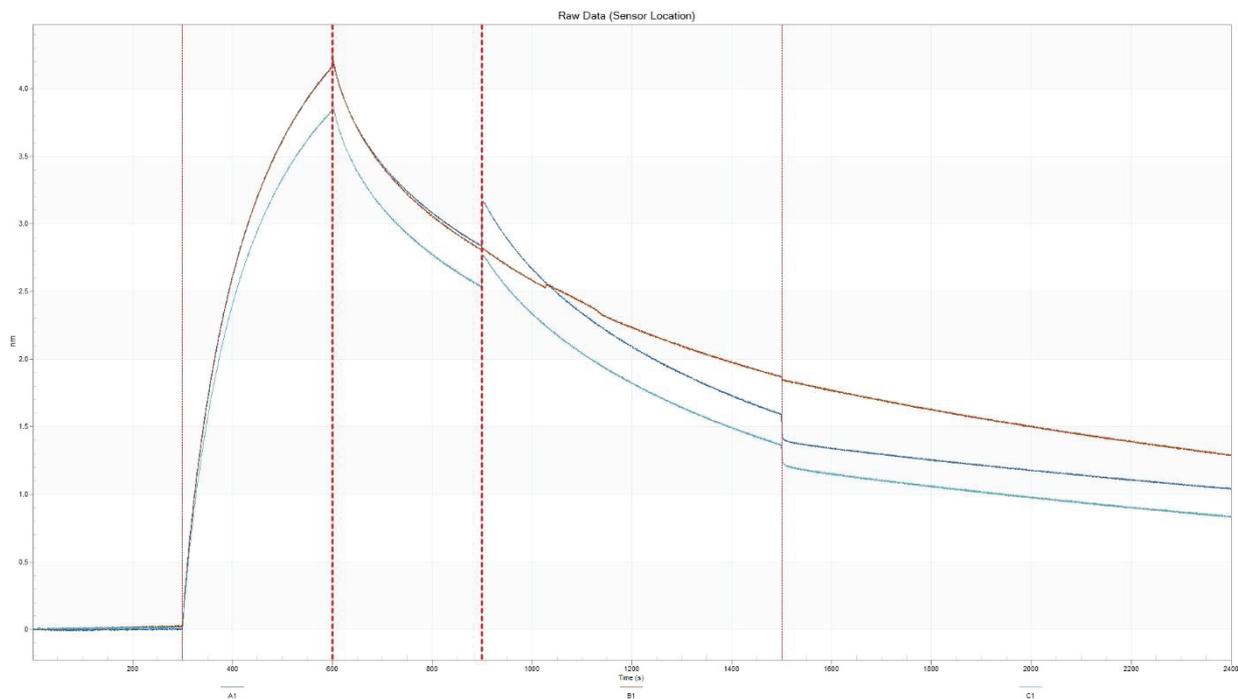
Scheme 221 : ITC thermogram of ΔK (0.56 mM) and **II-1₄** (5.6 mM) in Tris 100 mM pH 8



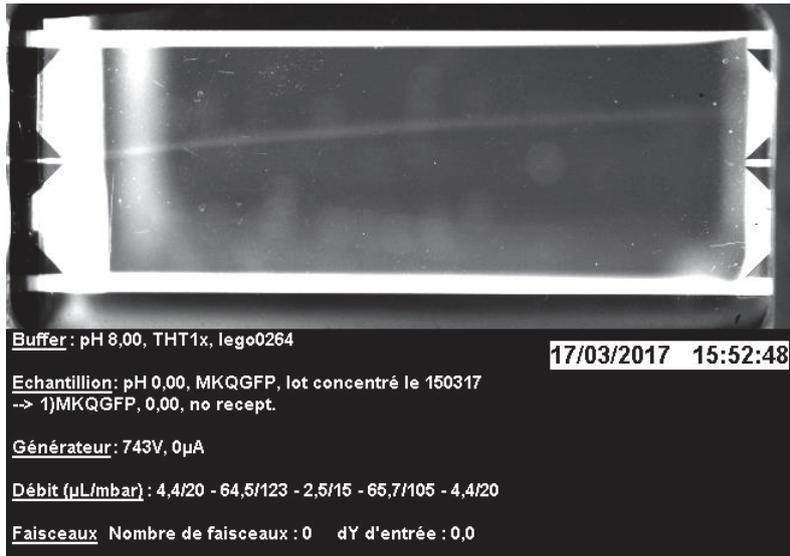
Scheme 222: ITC thermogram of MAA (0.56 mM) and **II-1₄** (5.6 mM) in Tris 100 mM pH 8



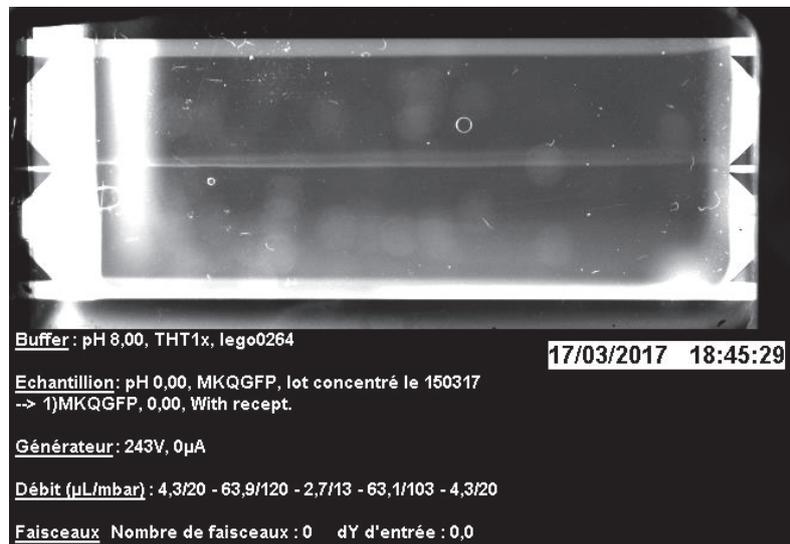
Scheme 223: ITC thermogram of MKQ (0.56 mM) and **II-1₄** (5.6 mM) in Tris 100 mM pH 8



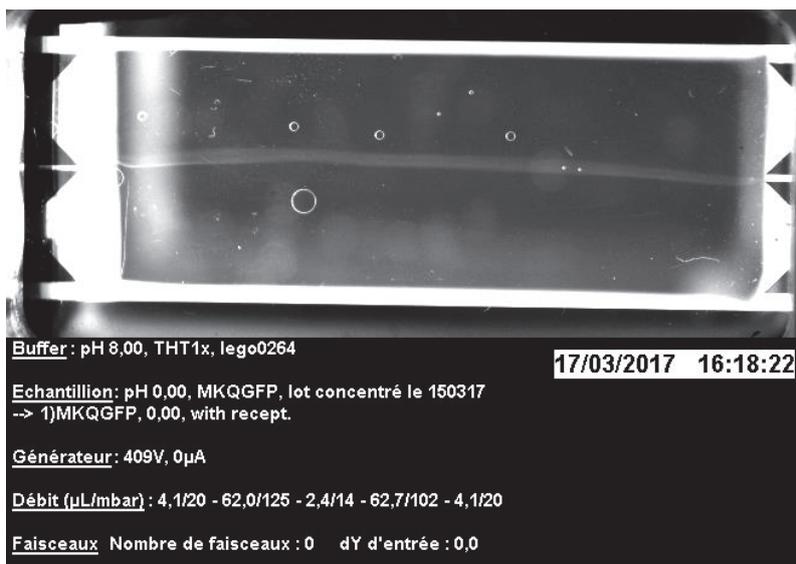
Scheme 224: Octet[®] thermograms, binding (nm) vs time (s), pale blue MKQ + **II-1₄**, blue MKQ + buffer, red MAA + **II-1₄**



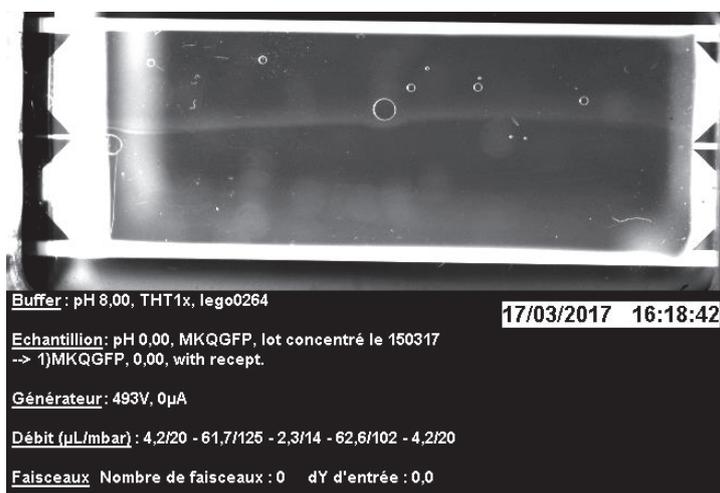
Scheme 225: μ -FFE with MKQ



Scheme 226: μ -FFE with MKQ + II-1₄ (1)



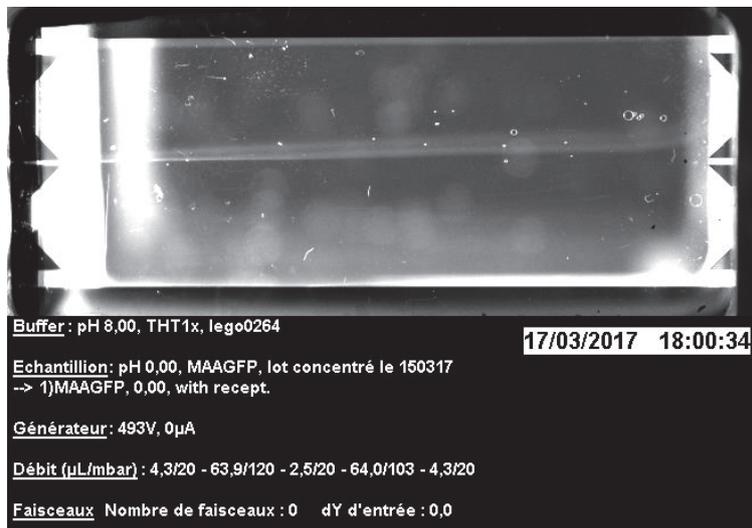
Scheme 227: μ-FFE with MKQ + II-1₄ (2)



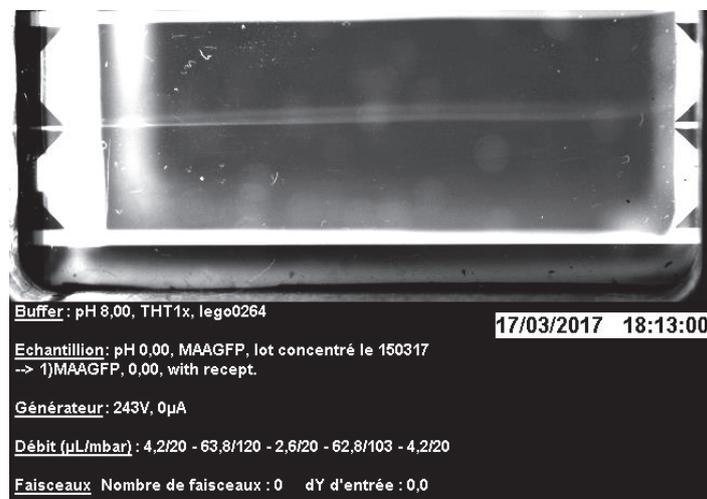
Scheme 228: μ-FFE with MKQ + II-1₄ (3)



Scheme 229: μ -FFE with MAA



Scheme 230: μ -FFE with MAA + II-1₄ (1)



Scheme 231: μ -FFE with MAA + II-1₄ (2)

d. Appendix II-IV: Additional HRMS or LC/MS data

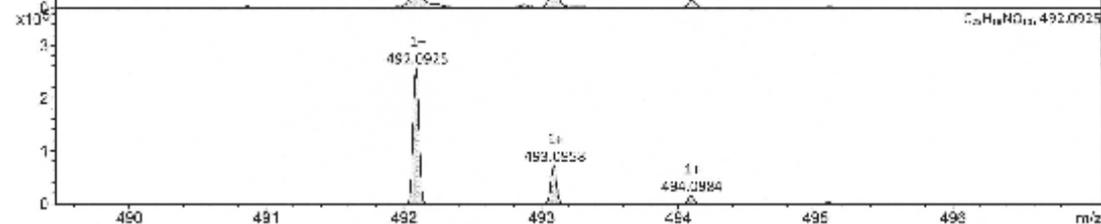
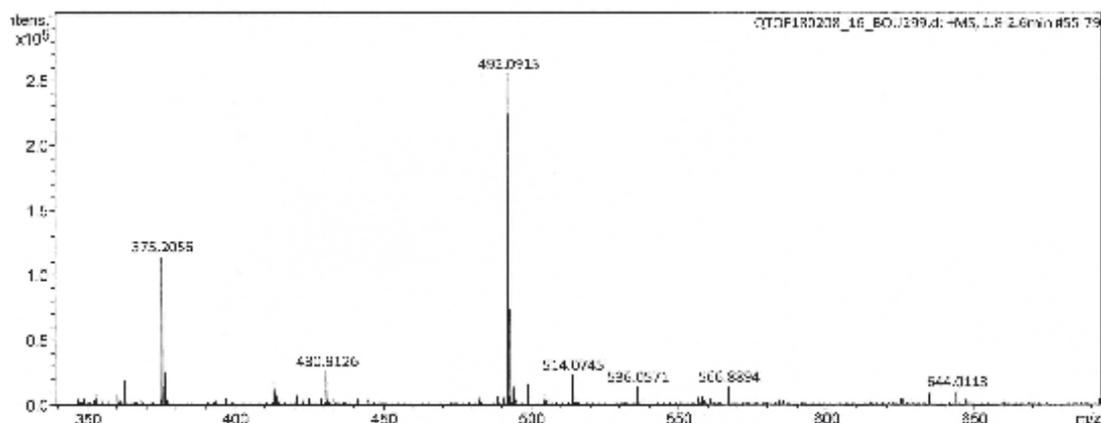
CENTRE COMMUN DE SPECTROMETRIE DE MASSE

Analysis Info

Analysis Name QTOF180208_16_BOU299.d
 Method 2016_03_17_Infusion_50-1000_pos.m
 Comment
 Acquisition Date 2/8/2018 4:30:48 PM
 Instrument / Ser# mic-QTOF-Q 228888.10
 231

Acquisition Parameter

Source Type	ESI	Ion Polarity	Positive	Set Nebulizer	0.4 Bar
Focus	Active	Set Capillary	3500 V	Set Dry Heater	200 °C
Scan Begin	50 m/z	Set End Plate Offset	-500 V	Set Dry Gas	4.0 l/min
Scan End	1000 m/z	Set Collision Cell RF	400.0 Vpp	Set Divert Valve	Waste



Meas. m/z	Ion Formula	Sum Formula	m/z	err [ppm]	mSigma	z	Adduct
492.0915	C25H18NO10	C25H17NO10	492.0925	2.2	4.3	1+	M+H
514.0745	C25H17NNaC10		514.0745	-0.1	9.3	1+	M+Na

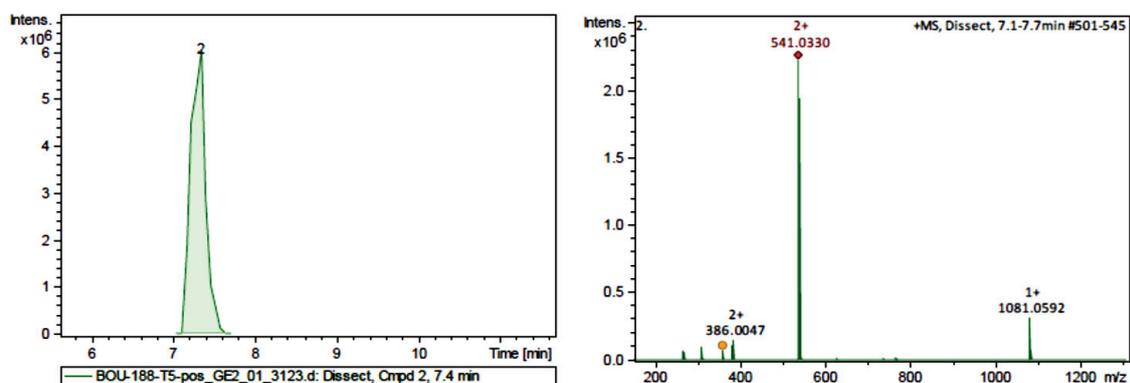
Scheme 232: HRMS spectrum of compound (II-71)₄

Meas. m/z	Ion Formula	Score	m/z	err [ppm]	mSigma	Adduct
519.155582	C20H31N4O8S2	50.16	519.157782	4.2	10.5	M+H

Scheme 233: HRMS spectrum of compound (II-120)

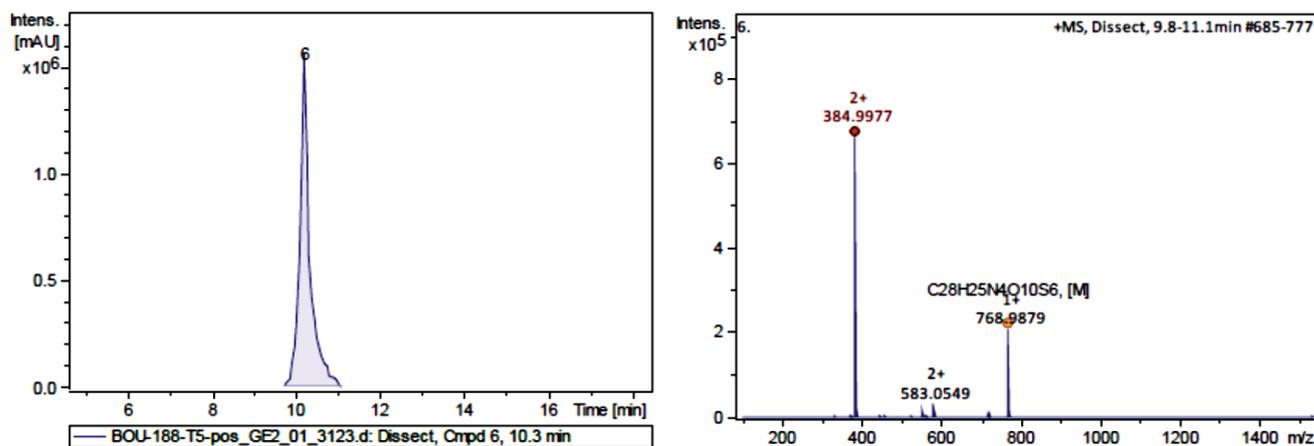
Meas. m/z	Ion Formula	Score	m/z	err [ppm]	mSigma	Adduct
1087.27421	C40H55N12O16S4	100	1087.27363	-0.5	81.2	M+H
1104.30077	C40H58N13O16S4	100	1104.30018	-0.5	20.1	M+NH4
1125.22897	C40H54KN12O16S4	100	1125.22952	0.5	36	M+K

Scheme 234 : HRMS spectrum of compound (II-81)₂



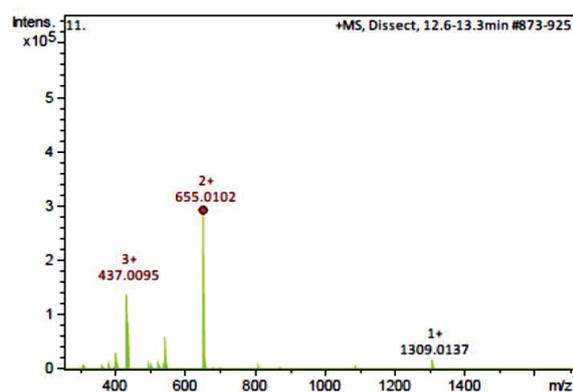
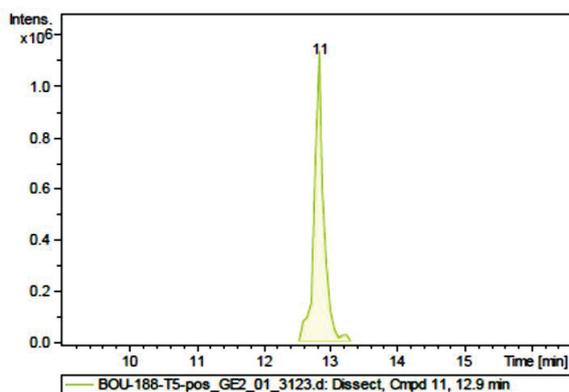
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	z	Adduct
361.0246	C40H43N8O12S8	361.025	1	18.2	100	3+	M+H

Scheme 235: HRMS from LC/MS, corresponding to (II-1)₂(II-69)₂



Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	z	Adduct
768.9879	C28H25N4O10S6	768.9889	1.4	22.9	100	1+	M+H
384.9977	C28H26N4O10S6	384.9981	1	15.5	100	2+	M+H

Scheme 236: HRMS from LC/MS, corresponding to (II-1)₂(II-69)₂

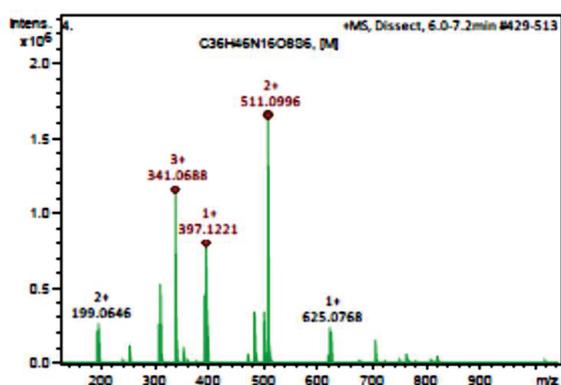
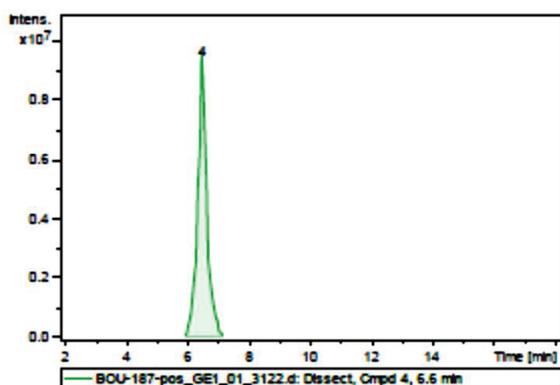


Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	z	Adduct
655.0102	C48H46N8O16S10	655.0114	1.9	28.4	100	2+	M+H
437.0095	C48H47N8O16S10	437.01	1.1	20.5	100	3+	M+H

Scheme 237 : HRMS from LC/MS, corresponding to (II-1)₃(II-69)₂

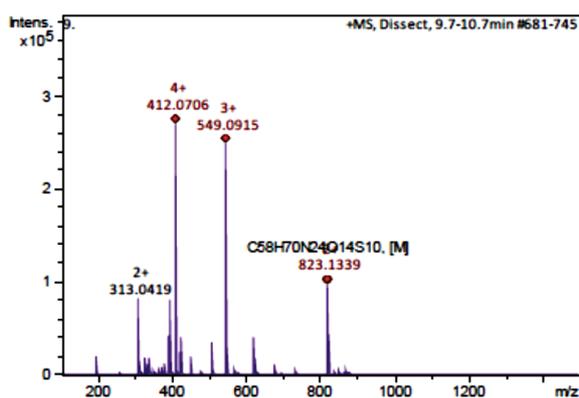
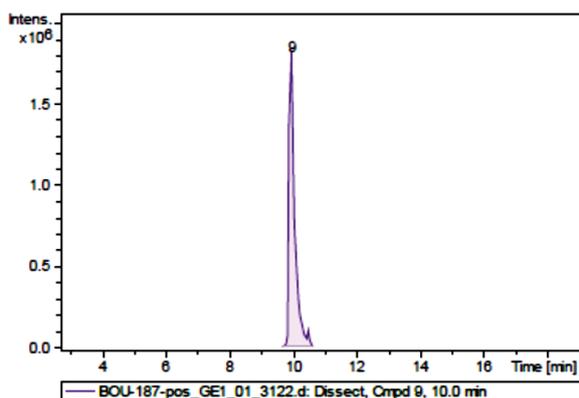
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	z	Adduct
538.9801	C38H30N8O14S8	538.9829	5.2	13.5	17.45	2-	M-H

Scheme 238: HRMS from LC/MS, corresponding to (II-1)₃(II-73)



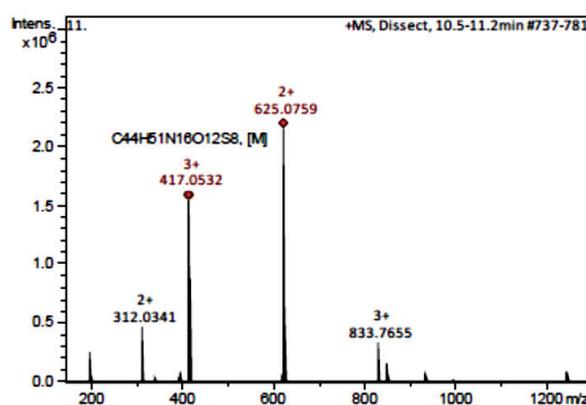
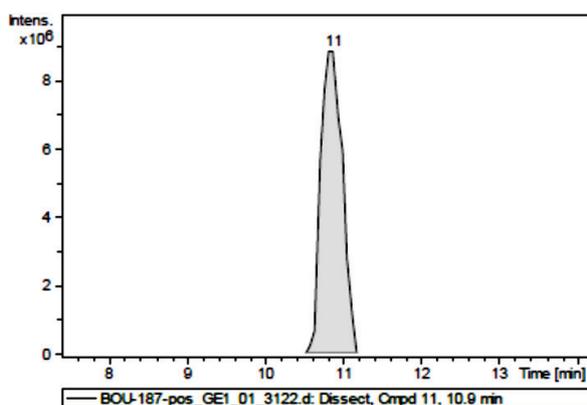
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
511.0996	C36H46N16O8S6	511.0999	0.6	9.8	100	M+H	2+

Scheme 239: HRMS from LC/MS, corresponding to (II-1)(II-73)₂



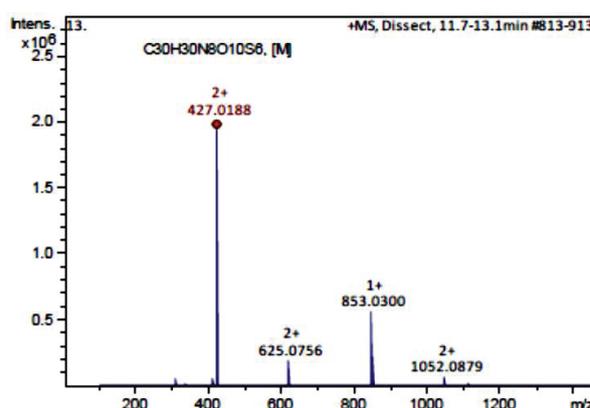
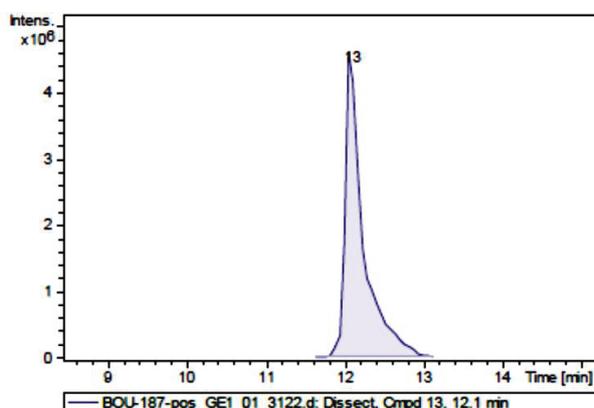
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
412.0706	C58H72N24O14S10	412.0711	1.4	11.7	100	M	4+
549.0915	C58H71N24O14S10	549.0924	1.7	18	100	M	3+
823.1339	C58H70N24O14S10	823.135	1.3	25.5	59.05	M	2+

Scheme 240: HRMS from LC/MS, corresponding to $(II-1)_2(II-73)_3$



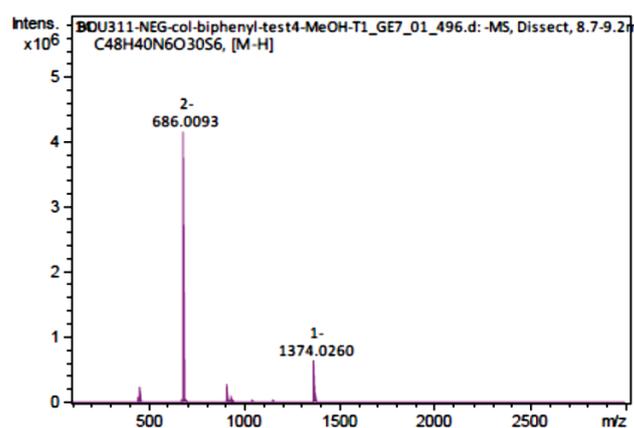
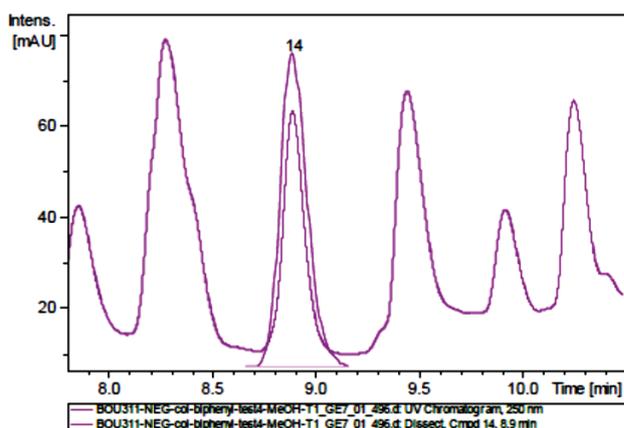
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
417.0532	C44H51N16O12S8	417.0541	2.1	4.7	100	M+H	3+

Scheme 241: HRMS from LC/MS, corresponding to $(II-1)_2(II-73)_2$



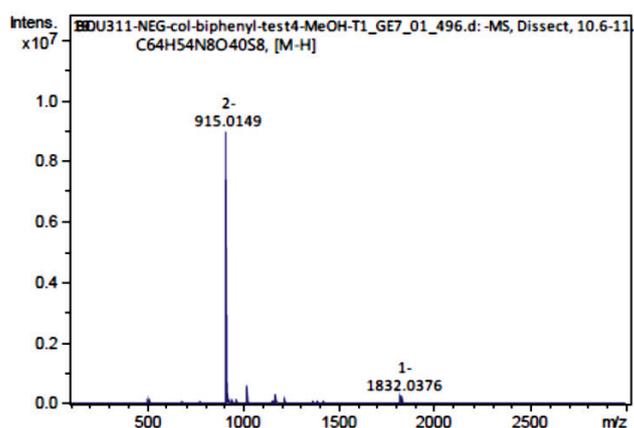
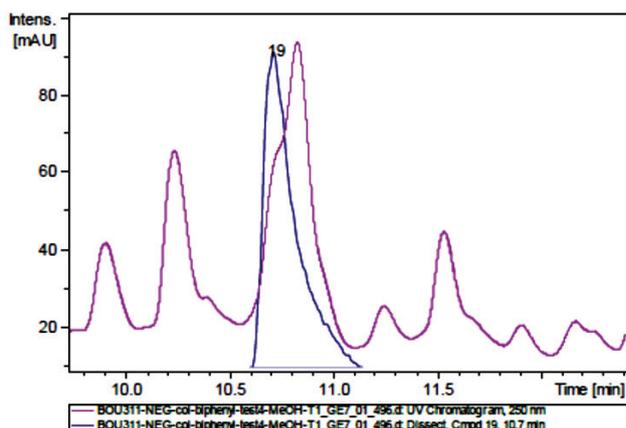
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
427.0188	C30H30N8O10S6	427.0199	2.6	9.7	100	M+H	2+

Scheme 242: HRMS from LC/MS, corresponding to $(II-1)_2(II-73)$



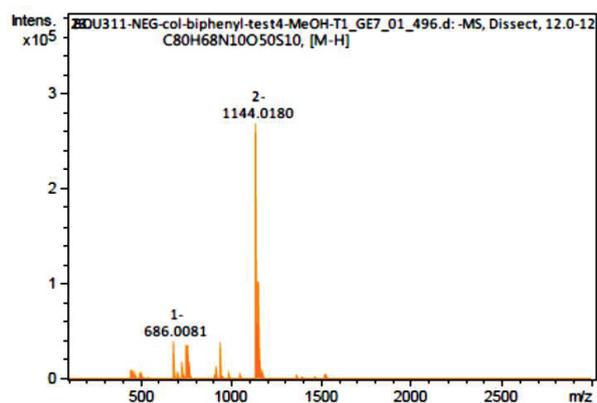
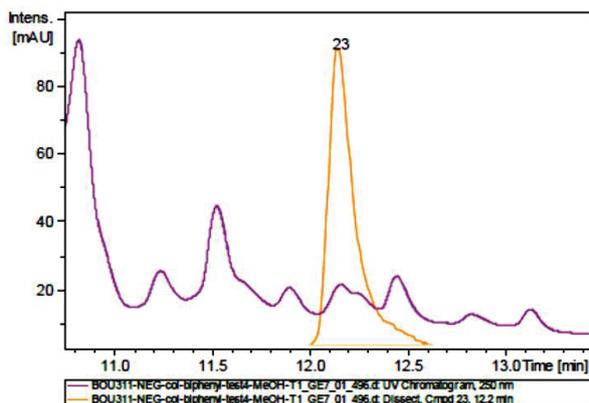
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
686.0093	C48H40N6O30S6	686.0062	-4.6	27.2	100	M-H	2-

Scheme 243: HRMS from LC/MS, corresponding to (II-101)₃



Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
915.0149	C64H54N8O40S8	915.0107	-4.5	15.8	100	M-H	2-

Scheme 244: HRMS from LC/MS, corresponding to (II-101)₄



Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
1144.018	C80H68N10O50S10	1144.0152	-2.4	23.3	100	M-H	2-

Scheme 245: HRMS from LC/MS, corresponding to (II-101)₅

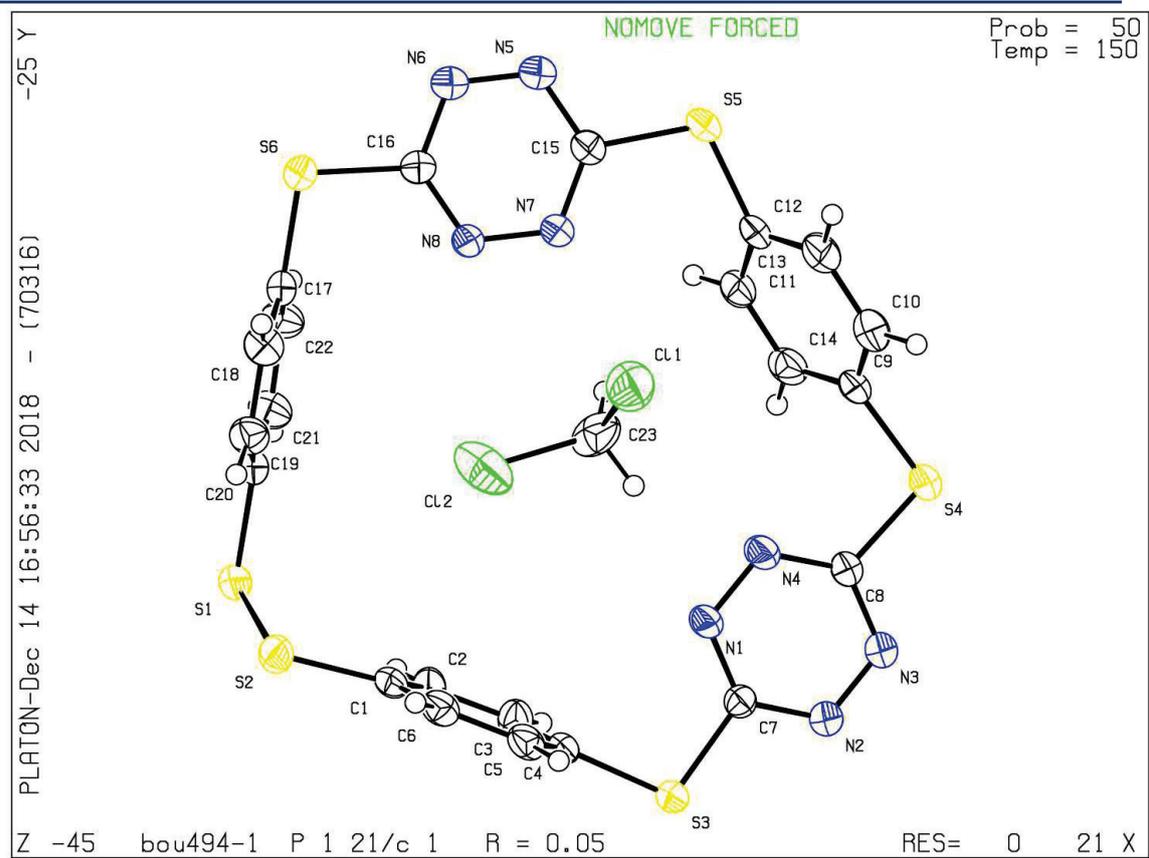
Meas. m/z	Ion Formula	m/z	err [ppm]	mSigma	Score	Adduct	z
445.0005	C15H13N2O10S2	445.0017	2.7	13.2	100	M-H	1-

Scheme 246: HRMS from LC/MS, corresponding to (II-118)₂

e. Appendix II-V: CheckCIFiles and crystallographic data for II-51

Olex report electronically available.

Bond precision: C-C = 0.0042 Å Wavelength=0.71073
Cell: a=11.3263(11) b=18.4484(19) c=13.7289(17)
alpha=90 beta=100.515(10) gamma=90
Temperature: 150 K
Calculated Reported
Volume 2820.5(5) 2820.5(5)
Space group P 21/c P 1 21/c 1
Hall group -P 2ybc -P 2ybc
Moiety formula C22 H12 N8 S6, C H2 Cl2 C H2 Cl2, C22 H12 N8 S6
Sum formula C23 H14 Cl2 N8 S6 C23 H14 Cl2 N8 S6
Mr 665.68 665.68
Dx, g cm⁻³ 1.568 1.557
Z 4 4
Mu (mm⁻¹) 0.706 0.706
F000 1352.0 1352.0
F000' 1356.43
h, k, lmax 15, 25, 19 15, 24, 18
Nref 8007 7235
Tmin, Tmax 0.870, 0.942 0.696, 0.906
Tmin' 0.731
Correction method= # Reported T Limits: Tmin=0.696 Tmax=0.906
AbsCorr = ANALYTICAL
Data completeness= 0.904 Theta(max)= 29.712
R(reflections)= 0.0516(5307) wR2(reflections)= 0.1376(7235)
S = 1.071 Npar= 352



OLEX2 report for 7-1'

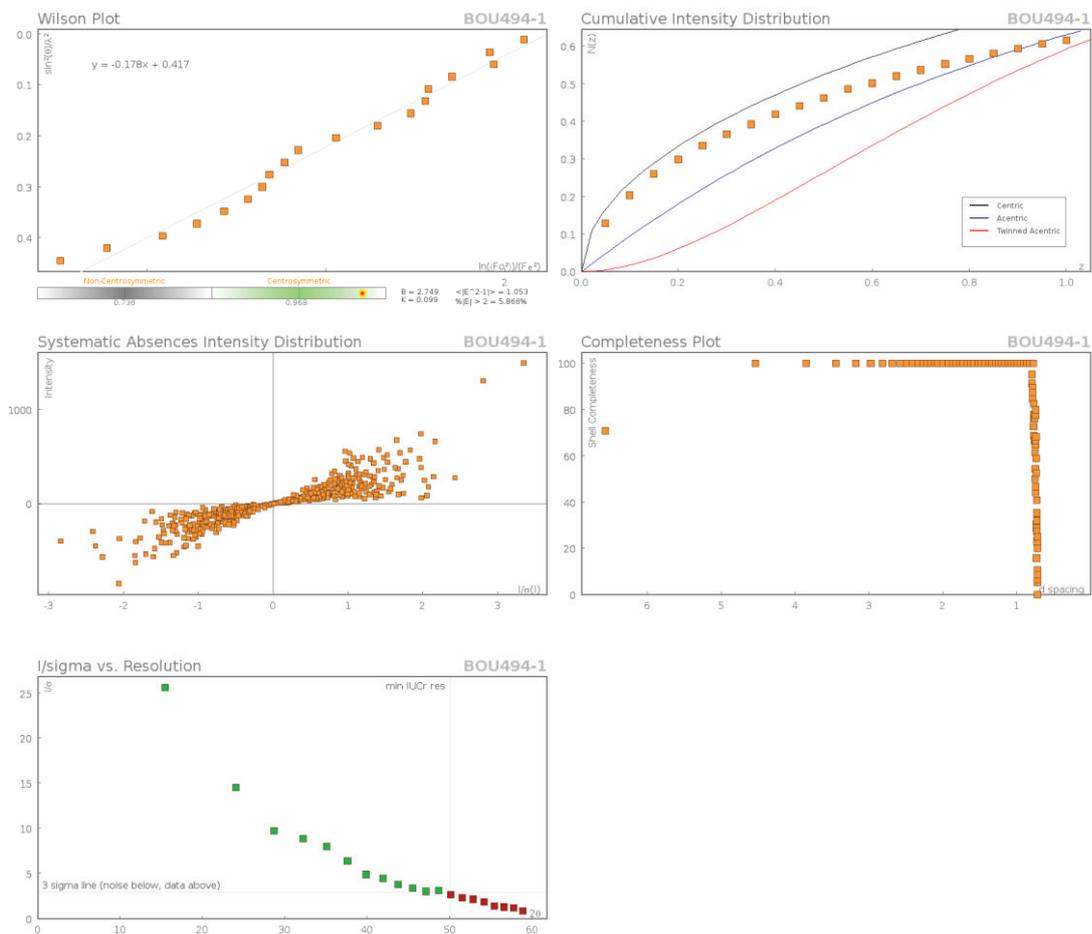
Submitted by: **Benjamin Ourri**
 Université de Lyon
 Solved by: **Erwann Jeanneau**
 Sample ID: **BOU494-1**

Crystal Data and Experimental

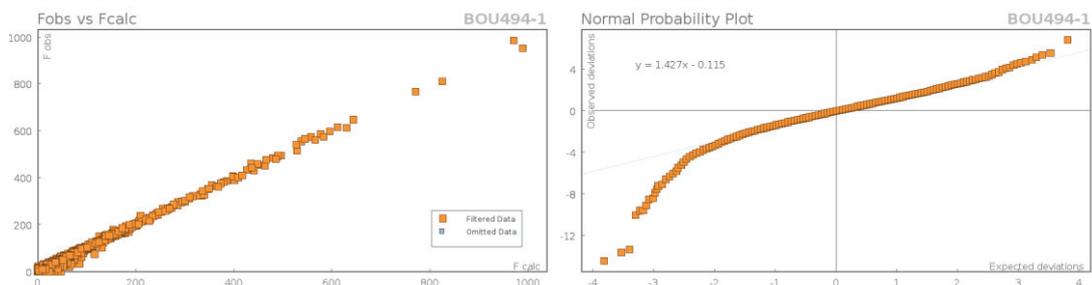
Structure Quality Indicators

Reflections:	d min (Å)	0.72	I/σ	23.1	R _{int}	5.74%	complete 100% (IUCr)	100%
Refinement:	Shift	-0.001	Max Peak	1.0	Min Peak	-0.9	Goof	1.071

Data Plots: Diffraction Data



Data Plots: Refinement and Data



Reflection Statistics

Total reflections (after filtering)	38753	Unique reflections	7235
Completeness	0.904	Mean I/σ	13.57

hkl _{max} collected	(15, 23, 18)	hkl _{min} collected	(-14, -24, -18)
hkl _{max} used	(14, 24, 18)	hkl _{min} used	(-15, 0, 0)
Lim d _{max} collected	100.0	Lim d _{min} collected	0.36
d _{max} used	7.1	d _{min} used	0.72
Friedel pairs	10481	Friedel pairs merged	1
Inconsistent equivalents	17	R _{int}	0.0574
R _{sigma}	0.0432	Intensity transformed	0
Omitted reflections	0	Omitted by user (OMIT hkl)	0
Multiplicity	(13889, 6693, 2519, 667, 146, 79, 7)	Maximum multiplicity	17
Removed systematic absences	764	Filtered off (Shel/OMIT)	0

Images of the Crystal on the Diffractometer



f. Appendix II-VI: DFT TS coordinates

TS coordinates for H-abstraction A -> B by molecular oxygen

C	-1.00159800	-0.89210400	-0.97842300
C	0.29935000	-1.29589800	-0.76683200
C	1.29691000	-0.38905000	-0.32877600
C	0.92124400	0.96756700	-0.06551700
C	-0.40360000	1.33743500	-0.27077300
C	-1.35563000	0.43672400	-0.73375700
H	-1.74249600	-1.60489300	-1.32556500
H	0.58871200	-2.32721700	-0.93750300
H	-0.67953300	2.36088000	-0.04455000
H	-2.37696100	0.76789900	-0.89289100
C	1.79242700	2.07876300	0.48022200
O	1.26695900	3.02559600	1.05438200
N	3.12001500	1.98987600	0.29243100
H	3.49061400	1.17579100	-0.18009000
C	4.02819600	2.99582600	0.80272900
C	4.13120400	4.20988400	-0.11625300
H	5.00890800	2.52598400	0.92545100
H	3.68314900	3.32008400	1.78768000
H	4.39983500	3.86880600	-1.12970700
H	3.14175200	4.67150700	-0.18220400
N	5.05949500	5.17626100	0.45858200
H	6.01330500	4.84278700	0.35535700
H	5.01187000	6.05078000	-0.05325000
S	2.89603500	-1.04690200	-0.17797400
H	2.90628800	-0.63302800	1.55023500
O	2.66491400	-0.39094000	2.57886700
O	1.46802700	0.02535200	2.62209800

TS coordinates for H-abstraction A on S atom by molecular oxygen

C	-1.11825800	-1.16177800	-0.16195900
C	0.12935600	-1.56587300	0.28695300
C	1.02272400	-0.64788300	0.85116600

C	0.63375800	0.70315000	0.95094600
C	-0.64226600	1.08225100	0.51887000
C	-1.51566100	0.16921800	-0.04247700
H	-1.78595800	-1.89506400	-0.60415800
H	0.41309800	-2.61032300	0.20116600
H	-0.91914000	2.12577200	0.62419000
H	-2.49317500	0.48652800	-0.39009000
C	1.49180700	1.78791300	1.49145000
O	1.01882500	2.78637400	2.02255900
N	2.84876400	1.59530400	1.29607200
H	3.20296400	1.12820400	-0.00681100
C	3.70040800	2.58388100	1.81049300
C	3.75561900	3.84720100	0.80194400
H	4.71971400	2.19889200	1.88663500
H	3.34925300	2.98563600	2.76483900
H	4.04672800	3.46790500	-0.17890200
H	2.73565600	4.23387800	0.76136300
N	4.65460200	4.85040000	1.22541600
H	5.63276900	4.59170500	1.19153800
H	4.42460900	5.28185000	2.11196700
S	2.57317800	-1.22692700	1.49164100
O	2.34249000	0.99929200	-1.72750500
O	3.44615600	0.92015500	-1.05102600
H	2.30704100	-2.53030200	1.26026000

TS coordinates for B -> C with hydroperoxyl radical

C	-1.00929900	-1.27473500	-0.14336200
C	0.32386300	-1.50452600	0.16174500
C	1.12163300	-0.45326000	0.64127800
C	0.56030300	0.83204500	0.78748300
C	-0.78155700	1.04017300	0.50099000
C	-1.56706900	-0.00680000	0.03068400
H	-1.62217600	-2.09150100	-0.51185200
H	0.75277000	-2.49475500	0.04726300

H	-1.19538700	2.03301100	0.64407400
H	-2.61310400	0.16269700	-0.20334400
C	1.40564400	1.95840600	1.26825000
O	0.93542900	2.93454200	1.83832500
N	2.73730600	1.79186200	0.99393700
H	3.10041600	1.26858000	-0.18102100
C	3.68989900	2.63875400	1.66420200
C	3.89630800	3.96212000	0.89482600
H	4.64550500	2.10571200	1.70969500
H	3.36123400	2.85738500	2.68611800
H	4.19938200	3.73050400	-0.13156400
H	2.93763000	4.48482500	0.84021500
N	4.88068300	4.84256000	1.49452300
H	5.78027400	4.37615400	1.56293200
H	4.60660200	5.08325900	2.44229200
S	2.77119900	-0.73363100	1.08406600
O	2.16526500	0.82155800	-1.83437200
O	3.35648300	0.67697500	-1.12932700
H	2.22189700	0.10379800	-2.48183100

TS coordinates for B -> C with molecular oxygen

C	-1.11825800	-1.16177800	-0.16195900
C	0.12935600	-1.56587300	0.28695300
C	1.02272400	-0.64788300	0.85116600
C	0.63375800	0.70315000	0.95094600
C	-0.64226600	1.08225100	0.51887000
C	-1.51566100	0.16921800	-0.04247700
H	-1.78595800	-1.89506400	-0.60415800
H	0.41309800	-2.61032300	0.20116600
H	-0.91914000	2.12577200	0.62419000
H	-2.49317500	0.48652800	-0.39009000
C	1.49180700	1.78791300	1.49145000
O	1.01882500	2.78637400	2.02255900
N	2.84876400	1.59530400	1.29607200
H	3.20296400	1.12820400	-0.00681100

C	3.70040800	2.58388100	1.81049300
C	3.75561900	3.84720100	0.80194400
H	4.71971400	2.19889200	1.88663500
H	3.34925300	2.98563600	2.76483900
H	4.04672800	3.46790500	-0.17890200
H	2.73565600	4.23387800	0.76136300
N	4.65460200	4.85040000	1.22541600
H	5.63276900	4.59170500	1.19153800
H	4.42460900	5.28185000	2.11196700
S	2.57317800	-1.22692700	1.49164100
O	2.34249000	0.99929200	-1.72750500
O	3.44615600	0.92015500	-1.05102600
H	2.30704100	-2.53030200	1.26026000

TS coordinates for II-51 with *tert*-butyl allyl carbamate
iEDDA reaction TS:

C	1.68583700	4.63704200	-0.46083700
C	3.79188400	3.79175800	0.67699800
C	-0.79847300	3.94938500	-1.44736400
C	-1.99054800	4.00278700	-0.73080000
H	-2.19575100	4.84708300	-0.08148100
C	-0.52144200	2.86856400	-2.28305300
H	0.40058900	2.84717600	-2.85304100
C	-1.42704100	1.82314700	-2.38491000
H	-1.22430000	0.97567500	-3.02889100
C	-2.60651500	1.86016600	-1.63633100
C	-2.89920300	2.95573900	-0.82397000
H	-3.81090500	2.95652400	-0.23332400
C	-1.73624500	-3.01035700	-0.73166400
C	-2.88403000	-0.88995000	-1.32065200
C	0.71694400	-4.24295200	-0.46068400
C	1.50561600	-4.19193500	0.68913100
H	1.04379600	-4.32625000	1.66185700
C	2.87312600	-3.97565200	0.59564100
H	3.48415200	-3.93916500	1.49079600

C 3.45606000 -3.79738300 -0.65649600
 C 2.66962200 -3.83785400 -1.80916500
 H 3.12159100 -3.68558200 -2.78482700
 C 1.30580900 -4.07489100 -1.71387000
 H 0.69154800 -4.10502000 -2.60501700
 C 5.61429800 -1.20156100 0.94200300
 C 5.51766200 -0.39522600 -0.19313900
 H 5.56836000 -0.84053500 -1.18074700
 C 5.37839400 0.97489700 -0.05296800
 H 5.33221200 1.60595200 -0.93404200
 C 5.31991700 1.54794100 1.21779300
 C 5.39113500 0.74233600 2.34908300
 H 5.32667900 1.18712000 3.33626900
 C 5.53913800 -0.63364500 2.21183000
 H 5.58996900 -1.26160500 3.09615200
 S 5.31964800 3.31386900 1.41750100
 S 0.31738800 5.33028000 -1.32979400
 S -3.82751000 0.57107300 -1.69029400
 S -1.01898700 -4.59129100 -0.29347500
 S 5.21031600 -3.58717900 -0.92486000
 S 5.95776100 -2.95497500 0.88845700
 N 2.62386900 5.54770500 -0.13447500
 N 3.70158900 5.11494400 0.44140200
 N 2.83419000 2.90128700 0.42424600
 N 1.74692100 3.33773200 -0.17097300
 N -3.67089100 -1.87352100 -0.78210100
 N -3.08994600 -2.97267700 -0.49239600
 N -1.85188900 -1.26483600 -2.14380900
 N -1.26524000 -2.36151000 -1.85224500
 C -7.27467800 -0.30581400 1.48139500
 C -8.00269600 -1.62282100 1.70324500
 H -9.07492500 -1.49177300 1.53682100
 H -7.63190800 -2.38465900 1.01333300
 H -7.85003500 -1.97756200 2.72572500

C -7.43003700 0.15349400 0.03508900
 H -8.49235100 0.27271700 -0.19533200
 H -6.92441800 1.10388600 -0.13310700
 H -7.00995600 -0.59425000 -0.64270800
 C -7.75200400 0.74935400 2.47354500
 H -7.58733900 0.40668300 3.49899400
 H -7.22489400 1.69133900 2.32601900
 H -8.82428300 0.91638500 2.33751300
 O -5.88296900 -0.66780000 1.74715000
 C -4.90347000 0.23307100 1.58806300
 O -5.04098200 1.42025200 1.37020800
 N -3.69944900 -0.39332700 1.73428500
 H -3.72222100 -1.39736300 1.80763900
 C -2.48418400 0.30010800 1.41210800
 H -1.90870000 0.53216900 2.32160000
 H -2.77221600 1.25832500 0.98556100
 C -1.53271400 -0.42303500 0.48966200
 H -0.92305900 0.23249200 -0.12883600
 C -1.04731300 -1.68708800 0.74797700
 H -0.03363600 -1.93813400 0.45567600
 H -1.44739900 -2.26803500 1.57446200

TS coordinates for II-51 with *tert*-butyl allyl carbamate retro-DA (N₂ release) reaction TS:

C 1.71738500 4.80782200 -0.44830200
 C 3.95948300 3.85648900 0.26704900
 C -0.87511800 4.18392100 -1.12849800
 C -1.92678800 4.05546100 -0.22736400
 H -2.03542800 4.76864700 0.58271100
 C -0.71941700 3.27318000 -2.17352400
 H 0.09256300 3.39593300 -2.88198500
 C -1.61014900 2.22082300 -2.30592100
 H -1.50954500 1.50631400 -3.11604200
 C -2.65274100 2.07809800 -1.38680200
 C -2.82152100 2.99717300 -0.35344100

H	-3.62547500	2.85631300	0.36235800	N	-3.23657200	-3.04280200	-1.10620000
C	-1.44038300	-2.85164200	-1.36575600	N	-2.04080000	-1.01141600	-2.54260200
C	-2.65855300	-0.63512700	-1.38075000	N	-1.39109200	-2.13170900	-2.52665700
C	0.85131100	-4.33726700	-1.05453600	C	-7.12705200	-0.24046100	1.89529900
C	1.42274400	-5.25199500	-0.16900100	C	-7.99598200	-1.48521400	1.99174600
H	0.81373900	-6.04151200	0.25900200	H	-9.04938400	-1.20252100	2.06017200
C	2.76383000	-5.14246500	0.16906700	H	-7.86106200	-2.11564100	1.10947100
H	3.21562900	-5.85148600	0.85464600	H	-7.73481600	-2.06893600	2.87814500
C	3.53002900	-4.09798800	-0.34676100	C	-7.44925400	0.54261100	0.62676000
C	2.95989500	-3.18577800	-1.23601500	H	-8.50020900	0.84405900	0.64452300
H	3.56241700	-2.38286800	-1.64646300	H	-6.82758000	1.43364200	0.54681000
C	1.62799400	-3.31473700	-1.60511800	H	-7.28548400	-0.08212500	-0.25538100
H	1.18701800	-2.62983100	-2.32238500	C	-7.27114000	0.61507900	3.14939500
C	5.18646100	-1.17908300	1.41181200	H	-6.98281400	0.04144600	4.03476500
C	5.59549400	-0.69657900	0.16949800	H	-6.65086700	1.50856600	3.08859400
H	5.87599400	-1.39101300	-0.61512200	H	-8.31595600	0.91677600	3.26426300
C	5.66285400	0.67247100	-0.05024200	O	-5.77826100	-0.80393700	1.81102600
H	5.99081000	1.05238200	-1.01204300	C	-4.71254200	-0.00065300	1.67156800
C	5.33203400	1.56338600	0.96728400	O	-4.71785500	1.21265300	1.68506400
C	4.91307300	1.08191400	2.20554100	N	-3.59765800	-0.77325500	1.52092100
H	4.65246900	1.77769400	2.99569800	H	-3.73096400	-1.77031500	1.55693300
C	4.82787700	-0.28451500	2.42388300	C	-2.32673600	-0.20022600	1.17643200
H	4.49335500	-0.65698800	3.38759300	H	-1.58464500	-0.39231200	1.96244200
S	5.57207800	3.30839200	0.72344300	H	-2.47392000	0.87846500	1.13789500
S	0.22255000	5.57702500	-0.98242100	C	-1.74158500	-0.70466000	-0.14318700
S	-3.81235600	0.73687700	-1.54660400	H	-0.91546300	-0.03073100	-0.40185200
S	-0.85811800	-4.55083900	-1.49377400	C	-1.18948800	-2.13239700	-0.04859600
S	5.24978700	-3.96693700	0.09454400	H	-0.11307200	-2.11993000	0.13621500
S	5.17619300	-2.89969700	1.87754000	H	-1.64706500	-2.70857300	0.76299600
N	2.74674500	5.66885700	-0.32588000	TS coordinates for II-51 with <i>tert</i> -butyl allyl carbamate			
N	3.89036500	5.18353900	0.04292000	rearomatisation reaction TS:			
N	2.93701400	3.00801100	0.17559600	C	0.89551500	4.67519700	-0.55647800
N	1.77503900	3.50200800	-0.19136000	C	3.20769000	4.13402500	0.34195300
N	-3.77812100	-2.02154000	-1.09180700	C	-1.56520300	3.66699100	-1.32259100

C	-2.70998600	3.51745100	-0.54797800	S	-0.37126900	-4.80497000	-0.75776100
H	-2.98508900	4.28716400	0.16530300	S	5.75448000	-3.24013500	-0.33472600
C	-1.19455200	2.67714800	-2.23483900	S	5.93375200	-2.30507700	1.50848200
H	-0.31661400	2.81382300	-2.85647300	N	1.74145500	5.70606500	-0.35284200
C	-1.96357200	1.53241700	-2.34982700	N	2.92019000	5.42981700	0.10487200
H	-1.70343500	0.76434900	-3.07241000	N	2.37333100	3.11507500	0.15234300
C	-3.10645400	1.37101400	-1.55766600	N	1.17473300	3.39592500	-0.31453000
C	-3.48705300	2.36786100	-0.66534000	N	-2.60711000	-2.13511500	-2.53571600
H	-4.34751300	2.22005600	-0.02395200	N	-1.88328500	-3.23605000	-2.27218500
C	-1.05072800	-3.27210100	-1.28544400	C	-7.23816900	-0.58673700	1.85077500
C	-2.78373400	-1.29329600	-1.51627100	C	-8.16030700	-1.79659500	1.84284400
C	1.33252600	-4.33140200	-0.63856000	H	-9.19746600	-1.47873500	1.97553700
C	2.10613800	-4.84943000	0.40447300	H	-8.07907600	-2.33424800	0.89497500
H	1.65391800	-5.51641600	1.13160900	H	-7.89968700	-2.48187400	2.65327400
C	3.44765500	-4.52109900	0.50146400	C	-7.55715600	0.33502900	0.67779800
H	4.05426600	-4.93697600	1.29896600	H	-8.60301300	0.64872100	0.73888800
C	4.01938800	-3.63322700	-0.41225100	H	-6.92369500	1.22110200	0.68872000
C	3.24822100	-3.10302800	-1.44404100	H	-7.41092400	-0.19177500	-0.26878900
H	3.69652500	-2.42063600	-2.15850100	C	-7.31841100	0.14112100	3.18864800
C	1.91208100	-3.46891100	-1.57214600	H	-7.05033700	-0.53601800	4.00456100
H	1.32425400	-3.11341400	-2.41415000	H	-6.65059200	1.00183700	3.20821700
C	5.48668500	-0.60114500	1.22974500	H	-8.34322000	0.48526200	3.35378600
C	5.48918300	0.02020700	-0.01936100	O	-5.91917200	-1.19509900	1.67849900
H	5.70882200	-0.55785300	-0.91035400	C	-4.81714600	-0.43060500	1.61542400
C	5.24336700	1.38150000	-0.11126400	O	-4.77581300	0.77920000	1.71015600
H	5.27128800	1.87372200	-1.07759200	N	-3.72755600	-1.23590000	1.45498700
C	4.98441800	2.12695700	1.03740100	H	-3.92150200	-2.20584400	1.25783800
C	4.95649300	1.50236600	2.28102700	C	-2.46432700	-0.70136700	0.99437100
H	4.74422400	2.08271500	3.17241700	H	-1.67641700	-0.89509500	1.73396400
C	5.20220300	0.13925100	2.37793500	H	-2.58406400	0.38508900	0.94267100
H	5.17898600	-0.34546000	3.34940600	C	-2.06367900	-1.30876800	-0.31895300
S	4.84473200	3.89675000	0.94886200	H	-0.61796900	-0.95141500	-0.91579400
S	-0.65760700	5.19068100	-1.20554300	C	-0.77048800	-2.06823800	-0.43818100
S	-4.09071700	-0.10161600	-1.73579700	H	0.24064900	-1.46832500	-1.03595000

H -0.24227600 -2.20691600 0.50380100

TS coordinates for II-51 with *tert*-butyl prop-2-yn-1-ylcarbamate iEDDA reaction TS:

C 1.38353400 4.81482800 -0.71248700
C 3.69156000 4.05745200 0.02330400
C -1.16301000 3.95244100 -1.23179100
C -1.98297800 3.65963400 -0.14796500
H -2.00273700 4.32061300 0.71158700
C -1.13839500 3.11899500 -2.34838200
H -0.50990100 3.37007600 -3.19575100
C -1.92008300 1.97487700 -2.37417000
H -1.90731600 1.32298700 -3.23722200
C -2.71166000 1.65919800 -1.26825800
C -2.76040400 2.50694100 -0.16193300
H -3.39432400 2.25886200 0.68423600
C -1.26183400 -3.14265200 -1.68781200
C -2.69952000 -1.12859000 -1.47169300
C 1.25267200 -4.28689300 -1.25428200
C 1.85821600 -5.30074600 -0.50704700
H 1.33366100 -6.23622400 -0.33993400
C 3.11693800 -5.10215000 0.04350400
H 3.58443300 -5.88213300 0.63492700
C 3.76784700 -3.88209500 -0.13137000
C 3.16808200 -2.87304200 -0.88588800
H 3.67871400 -1.92741800 -1.03084300
C 1.92134300 -3.07489900 -1.45984700
H 1.46304100 -2.30131600 -2.06381300
C 5.06524700 -0.77993100 1.61717500
C 5.64276900 -0.38852600 0.41080500
H 6.06175700 -1.13392200 -0.25686400
C 5.68195700 0.95724600 0.06920700
H 6.12416700 1.26529600 -0.87230300
C 5.16682000 1.91574800 0.93677900
C 4.59880900 1.52563700 2.14788300

H 4.19809300 2.27365800 2.82331700

C 4.53347100 0.18215700 2.48038800
H 4.07194400 -0.12013300 3.41573700
S 5.32928100 3.64039600 0.53396300
S -0.18672800 5.44307600 -1.21972200
S -3.79804500 0.24598000 -1.26181800
S -0.34831600 -4.66565000 -1.93246400
S 5.38876900 -3.65919700 0.57220900
S 5.00987900 -2.45646100 2.22318300
N 2.35664800 5.74447800 -0.66979300
N 3.53567400 5.35787000 -0.28996800
N 2.73235300 3.13412400 -0.02989800
N 1.53744900 3.52681200 -0.40656800
N -3.35614200 -2.33142200 -1.58111400
N -2.62545100 -3.36535500 -1.71814600
N -1.52793900 -0.96112200 -2.14997600
N -0.79929600 -1.99875800 -2.30154400
C -7.20712200 -0.01305600 1.77332100
C -8.29616000 -1.05150100 1.55376500
H -9.27910500 -0.57511500 1.58604800
H -8.17473700 -1.53506800 0.58146600
H -8.25594300 -1.81941100 2.33018900
C -7.21961300 1.02190900 0.65300200
H -8.19680200 1.51181700 0.62522000
H -6.45263000 1.78036400 0.80657500
H -7.05188600 0.53739200 -0.31248700
C -7.34206600 0.62942100 3.14935800
H -7.28725700 -0.13340700 3.93101000
H -6.55842200 1.36748000 3.31767500
H -8.31458500 1.12381100 3.22372300
O -5.98516500 -0.82030000 1.70459800
C -4.78318800 -0.24675800 1.82657000
O -4.55640600 0.92658700 2.04804700
N -3.81362100 -1.19248900 1.67404300

H	-4.09633900	-2.11554500	1.38556600	C	5.59270300	-0.44368200	0.54536000
C	-2.43496600	-0.79282100	1.52262200	H	6.03311900	-1.19610800	-0.10028200
H	-1.90052300	-0.81411300	2.48220800	C	5.65580500	0.89960400	0.19822400
H	-2.40291300	0.23960000	1.16652100	H	6.13991100	1.19920000	-0.72522200
C	-1.72571200	-1.69351900	0.61782300	C	5.11054000	1.86664000	1.03756500
C	-0.99418200	-2.62551000	0.27697100	C	4.48862700	1.48764400	2.22541500
H	-0.24323100	-3.36049900	0.48986900	H	4.06455800	2.24251200	2.87861000

TS coordinates for II-51 with *tert*-butyl prop-2-yn-1-ylcarbamate retro-DA (N₂ release) reaction TS:

C	1.41675200	4.78611200	-0.77216900	H	3.89647200	-0.14739400	3.47872800
C	3.69145200	4.01100400	0.04493200	S	5.30165400	3.58766000	0.63236000
C	-1.13508200	3.96072100	-1.34474200	S	-0.13215200	5.43378000	-1.31602400
C	-1.95451600	3.66608300	-0.26074100	S	-3.84103800	0.29626300	-1.46250700
H	-1.95214800	4.31086300	0.61139800	S	-0.36568100	-4.61210600	-2.02623400
C	-1.13671800	3.14858800	-2.47679800	S	5.28356400	-3.70758100	0.70126300
H	-0.50739900	3.40039700	-3.32341500	S	4.86871800	-2.49759600	2.33832000
C	-1.95091100	2.02696100	-2.51892000	N	2.39571200	5.70804600	-0.69544200
H	-1.96641200	1.39106400	-3.39453800	N	3.55726900	5.31232700	-0.27466400
C	-2.74692800	1.71005600	-1.41585100	N	2.73321900	3.09153400	-0.06045000
C	-2.76236000	2.53451400	-0.29175300	N	1.55508900	3.49308000	-0.47990000
H	-3.39051700	2.28190800	0.55624100	N	-3.51565400	-2.41632200	-1.82094600
C	-1.21508100	-3.06743800	-1.66857500	N	-2.84460200	-3.37578000	-1.95762200
C	-2.63888900	-1.05295000	-1.40985800	N	-1.67546000	-1.00130000	-2.49556700
C	1.20712700	-4.29296200	-1.26004000	N	-0.96487800	-2.00838600	-2.63276100
C	1.79693800	-5.32929300	-0.53710900	C	-6.94849200	0.00483500	2.25529100
H	1.27619100	-6.27445500	-0.42235600	C	-8.04058200	-1.05338800	2.28867400
C	3.03846100	-5.14043400	0.05713600	H	-9.00990300	-0.58564000	2.47802300
H	3.49549700	-5.93765900	0.63367900	H	-8.09097400	-1.58353800	1.33441100
C	3.68425400	-3.91108700	-0.05442100	H	-7.84567300	-1.78163700	3.07994400
C	3.09787300	-2.87708800	-0.78756400	C	-7.18396500	0.98297200	1.10891600
H	3.60537900	-1.92360400	-0.88420000	H	-8.15733500	1.46404200	1.23860200
C	1.86793700	-3.06863900	-1.39724300	H	-6.41290400	1.75211800	1.08203300
H	1.42539000	-2.27498200	-1.98904500	H	-7.18782700	0.45253400	0.15291100
C	4.96130700	-0.82408600	1.72821700	C	-6.84273300	0.71315300	3.60137300
				H	-6.62502700	-0.00873800	4.39349200

H	-6.05896300	1.46964600	3.58583000	C	4.33089900	1.81948100	-0.55574600
H	-7.79635000	1.19576000	3.83239700	C	3.23675800	1.21641900	-1.17678600
O	-5.74597900	-0.79302700	1.99992400	H	2.41285300	1.82526500	-1.53245800
C	-4.54938200	-0.20059200	1.88071300	C	3.19965700	-0.15998800	-1.34871100
O	-4.31606000	0.98379000	2.01231600	H	2.36138400	-0.62965800	-1.84884800
N	-3.60955000	-1.14803200	1.60303000	C	1.59910400	4.08417500	0.79967100
H	-3.93897300	-2.09031500	1.46574100	C	1.21950300	4.37300400	-0.50983200
C	-2.27748600	-0.80048300	1.16684000	H	1.97166400	4.45152600	-1.28750300
H	-1.52571500	-1.12526300	1.89418700	C	-0.12192400	4.56807600	-0.81228700
H	-2.22739200	0.28753600	1.10611300	H	-0.42130000	4.79203300	-1.83072000
C	-1.97657400	-1.44597800	-0.15069100	C	-1.08474000	4.49378400	0.19021600
C	-1.20834800	-2.53696200	-0.30679400	C	-0.70573500	4.20790200	1.50017000
H	-0.65538900	-3.03898800	0.47694000	H	-1.45693200	4.14863600	2.28016900
TS coordinates for II-51 with propargyl amine iEDDA				C	0.62887800	3.98871400	1.80156900
reaction TS:				H	0.91934000	3.74976600	2.82030500
C	-4.72680900	1.05091500	-0.42384200	S	-2.77813600	4.87655100	-0.19518500
C	-3.52137400	3.27779300	-0.25538100	S	-5.64911900	-0.44577100	-0.59774400
C	-4.36421900	-1.66065800	-0.38753000	S	-1.19003100	-4.86561500	0.20716300
C	-4.26297700	-2.34777900	0.81756800	S	4.32034900	-2.71789000	-1.01247000
H	-4.94307600	-2.12088200	1.63108300	S	4.40966700	3.58794600	-0.36364300
C	-3.49943500	-1.94929600	-1.44077500	S	3.28410200	3.90665500	1.35461300
H	-3.60235500	-1.42209200	-2.38291900	N	-5.45646500	2.16597500	-0.61484700
C	-2.51240200	-2.91086700	-1.28761700	N	-4.84121200	3.30445800	-0.52469300
H	-1.83565000	-3.13702500	-2.10012900	N	-2.80622800	2.17609800	-0.03020400
C	-2.38284900	-3.57159300	-0.06443200	N	-3.43163200	1.02429200	-0.11338800
C	-3.26864300	-3.30398500	0.98048500	N	1.42067300	-4.98605000	0.01700400
H	-3.16311000	-3.82514900	1.92715300	N	2.58554500	-4.53889100	-0.23745400
C	2.63966800	-3.22053600	-0.64628900	N	0.46468600	-3.22025700	-1.20755500
C	0.37664500	-4.12270100	-0.19350800	N	1.63453500	-2.77551500	-1.47442500
C	4.25201800	-0.94565800	-0.86648800	N	-0.15886400	-3.90059500	3.28060500
C	5.35301000	-0.34009300	-0.25453600	H	0.72709900	-4.03806800	3.75340300
H	6.16454900	-0.95212900	0.12626400	C	-0.12005200	-2.70047000	2.46897300
C	5.39721900	1.03937300	-0.11117000	H	-0.04985500	-1.82531600	3.12709800
H	6.24552300	1.51230600	0.37192900	H	-1.06336300	-2.58909800	1.93295100

C	1.01526600	-2.58927600	1.53470100	H	2.02575500	4.08562400	-1.88451700
C	2.12350200	-2.25705600	1.10685900	C	1.59340300	4.30746300	0.20799500
H	3.02457300	-1.68162800	1.17417000	C	1.94314300	3.93821300	1.50476700
H	-0.32372800	-4.71318500	2.69730500	H	1.38986100	4.33719200	2.34795900

TS coordinates for II-51 with propargyl amine retro-DA (N₂ release) reaction TS:

C	-3.35775700	3.39724600	-0.35235800	S	0.34920700	5.54831800	-0.06454200
C	-1.13936800	4.60708100	-0.11516900	S	-4.93457500	2.65941000	-0.63835700
C	-4.56853300	0.92878500	-0.45455100	S	-3.73407200	-3.50627800	0.08147000
C	-5.10960300	0.23779700	0.62528600	S	2.15596500	-4.67146400	-0.38301800
H	-5.70628000	0.76694900	1.36072600	S	5.67828900	0.58459300	-0.69913200
C	-3.80322200	0.26172900	-1.41075100	S	5.14906500	1.58273100	1.04828900
H	-3.40843200	0.80478400	-2.26244900	N	-3.36269500	4.73688100	-0.49814000
C	-3.55367000	-1.09368000	-1.27045100	N	-2.23074700	5.35459100	-0.37349200
H	-2.96329300	-1.62630000	-2.00793800	N	-1.14744700	3.29607400	0.11866600
C	-4.06188400	-1.78015300	-0.16472100	N	-2.29846800	2.67053900	-0.00116200
C	-4.85364300	-1.12050200	0.77250000	N	-1.55154400	-5.07433100	0.52367700
H	-5.22525200	-1.66372400	1.63363300	N	-0.41097700	-5.33282400	0.41219800
C	0.48684600	-4.03094400	-0.19805600	N	-1.41277000	-3.52306800	-1.33440800
C	-1.93358200	-3.50755900	0.02618600	N	-0.20503300	-3.79132900	-1.45013900
C	3.08744800	-3.16461300	-0.50184500	N	-2.49671100	-2.05340500	2.79025400
C	4.35092300	-3.15199100	0.09240700	H	-2.01302000	-2.65891300	3.44340100
H	4.70636000	-4.02744100	0.62638500	C	-1.58383200	-1.55104200	1.77747900
C	5.14169200	-2.01518800	0.01619500	H	-0.69185000	-1.04177200	2.17471500
H	6.12207500	-1.99863800	0.48018300	H	-2.11722800	-0.80911200	1.17505700
C	4.66248500	-0.87418100	-0.62600100	C	-1.08722300	-2.64720600	0.87980000
C	3.40444000	-0.89014800	-1.22937600	C	0.21472600	-2.96114200	0.75471300
H	3.03874900	-0.00569600	-1.73931100	H	1.02311300	-2.49200700	1.30075800
C	2.62382200	-2.03623400	-1.18367200	H	-2.89590000	-1.28581000	3.31678100
H	1.65864900	-2.05646700	-1.67763000				
C	3.72495800	2.56827600	0.62790400				
C	3.36482400	2.92695800	-0.67050400				
H	3.92978900	2.54477300	-1.51383600				
C	2.29895600	3.79188100	-0.87649700				

TS coordinates for II-20 and thiophenol S_NAr step 1 TS:

C	2.27273000	0.91940500	0.49256700				
C	3.41520000	0.75945200	-0.28444800				
C	3.76940700	-0.49800200	-0.76975800				
C	2.96485700	-1.59472100	-0.46517100				

C 1.82257300 -1.43676200 0.31166700
 C 1.45324300 -0.17638400 0.80934600
 H 1.99888600 1.90258700 0.86202900
 H 4.03244800 1.62398700 -0.51346300
 H 3.22741300 -2.58125600 -0.83750700
 H 1.19485500 -2.29328100 0.53721900
 S -0.02473700 0.03030400 1.73497400
 C -1.24913900 1.37618300 -0.39156500
 C -1.71055600 -1.00870500 -0.36049300
 N -2.73157400 -0.22854100 0.03815300
 N -2.49209100 1.04450100 0.01218000
 N -0.86113800 -0.61290300 -1.31723300
 N -0.61730400 0.66629800 -1.33979700
 Cl -1.84018600 -2.69948200 -0.06517000
 H 4.66068100 -0.62175600 -1.37716800
 Cl -0.76179300 3.01457500 -0.18404700

TS coordinates for II-20 and thiophenol SNAr step 2 TS:

C 3.03081900 0.42244900 0.39530300
 C 4.10505200 -0.21521900 -0.21664100
 C 4.00783600 -1.55700600 -0.57788200
 C 2.83633100 -2.26620900 -0.32407800
 C 1.76003600 -1.63663700 0.29227900
 C 1.85926100 -0.29112300 0.64762700
 H 3.09233100 1.46893900 0.67226100
 H 5.01859700 0.33733300 -0.41222000
 H 2.75976400 -3.31190000 -0.60462200
 H 0.84421500 -2.18016100 0.49700700
 S 0.49684100 0.51247600 1.45948100
 C -0.58383100 0.93459100 0.06243900
 C -2.69429600 -0.45258600 -0.14620600
 N -2.82355200 0.72597300 0.45144900
 N -1.77995300 1.49356900 0.52526400
 N -1.71904900 -0.72482100 -1.01239600
 N -0.65008500 -0.00001900 -0.98166300

Cl -4.03914400 -1.55545000 -0.07065800
 H 4.84762800 -2.05144300 -1.05601700
 Cl 0.41380500 2.58163400 -0.87779600

TS coordinates for II-20 and salicylic acid SNAr step 1 TS:

C -2.25214100 -0.43034600 0.29741000
 C -3.14459600 -1.10676800 -0.53660600
 C -2.76067500 -2.21704300 -1.28178100
 C -1.44231200 -2.65993300 -1.21020900
 C -0.53045600 -1.98293800 -0.40889500
 C -0.91004400 -0.86172000 0.34540800
 H -4.16455200 -0.73668500 -0.59029100
 H -1.12041900 -3.52767700 -1.77937900
 H 0.50192900 -2.31401300 -0.37138600
 S 0.32915900 -0.00175900 1.26175000
 C 2.57045400 -0.19298400 0.08173300
 C 1.03060200 1.35303500 -0.83699300
 N 2.07281700 1.95024200 -0.20312300
 N 2.91507000 1.10971900 0.30009700
 N 1.30094700 0.30969600 -1.65731000
 N 2.16338100 -0.52599600 -1.16845900
 Cl -0.32330300 2.34333500 -1.25286000
 H -3.48187000 -2.73111900 -1.91093700
 Cl 3.50043500 -1.38673800 0.93556700
 C -2.81407200 0.70340100 1.14894600
 O -3.49259300 1.56393500 0.54167000
 O -2.58977500 0.62877200 2.37901300

TS coordinates for II-20 and salicylic acid SNAr step 2 TS:

C 1.49468600 1.40163100 0.89971000
 C 2.53721800 1.59417200 1.79821700
 C 3.60872600 0.70680800 1.83481100
 C 3.62403600 -0.36618300 0.95470000
 C 2.59720200 -0.57940300 0.03407500
 C 1.51083100 0.31239500 0.02117600
 H 2.50581900 2.44695600 2.47009700

H	4.43920800	-1.08228900	0.94692400
S	0.17539000	0.02916700	-1.13829200
C	-1.23881700	0.89616500	-0.45658400
C	-2.83802900	-0.80397300	0.47804400
N	-2.21644800	-0.03199500	1.36993100
N	-1.46045600	0.91973300	0.93041900
N	-3.11505000	-0.36608200	-0.74719900
N	-2.36817100	0.57029500	-1.23951200
Cl	-3.69980600	-2.20013100	1.06546900
H	4.42303300	0.85403000	2.53783200
Cl	-0.96613500	2.99935700	-0.92882000
H	0.66926800	2.09851100	0.88542300
C	2.73294800	-1.74565900	-0.93909200
O	1.95515700	-1.70806100	-1.92853300
O	3.59803100	-2.60570100	-0.68352500

TS coordinates for II-20 and amino-thiophenol SNAr step 1 TS:

C	1.46061700	-1.55387000	0.05352900
C	2.56758900	-1.85864900	-0.75250300
C	3.65920900	-1.00343400	-0.82453000
C	3.66813800	0.18037100	-0.08759800
C	2.57206900	0.49430800	0.70892500
C	1.45067800	-0.34141200	0.79694500
H	2.56187200	-2.78440000	-1.32324400
H	4.51952700	0.85283900	-0.13260800
H	2.56703900	1.41376200	1.28618000
S	0.05195500	0.08814300	1.76833500
C	-0.77943000	1.63132500	-0.41419700
C	-2.21176600	-0.35531800	-0.34580000
N	-2.76529800	0.76773900	0.12876900
N	-2.00840000	1.82668600	0.07726500
N	-1.27077300	-0.36426100	-1.28764900
N	-0.51531500	0.69660800	-1.35052100
Cl	-3.04463600	-1.83020700	-0.03006500

H	4.50463500	-1.26530200	-1.45442200
Cl	0.32456900	2.94895800	-0.31586500
N	0.36958300	-2.38746300	0.09953600
H	0.50996500	-3.35138200	-0.15723700
H	-0.24437400	-2.22966900	0.88836000

TS coordinates for II-20 and amino-thiophenol SNAr step 2 TS:

C	2.89460000	0.54373500	0.05065500
C	3.87431700	-0.18372300	-0.60763100
C	3.69936400	-1.56132800	-0.76012300
C	2.57217700	-2.19608100	-0.26803000
C	1.56585000	-1.47470000	0.40113200
C	1.75206800	-0.08160700	0.55321400
H	4.76027900	0.30926900	-0.99247600
H	2.45037500	-3.26784600	-0.39765100
S	0.54862900	0.87692700	1.42892900
C	-0.65512600	1.22226900	0.09871700
C	-2.42494100	-0.53935900	-0.26894700
N	-2.78805800	0.51722900	0.44927700
N	-1.92990900	1.48150000	0.59741300
N	-1.44340100	-0.49917500	-1.16706300
N	-0.54714100	0.42435200	-1.04588900
Cl	-3.49064000	-1.91436100	-0.29503000
H	4.45650300	-2.15004800	-1.26973000
Cl	-0.01154300	3.15430800	-0.63253600
H	2.99933800	1.61526300	0.18625400
N	0.43289000	-2.10624100	0.83742600
H	0.48390000	-3.10275500	0.97462800
H	-0.13187200	-1.61270400	1.51267700

TS coordinates for II-20 and zwitterion amino-thiophenol SNAr step 1 TS:

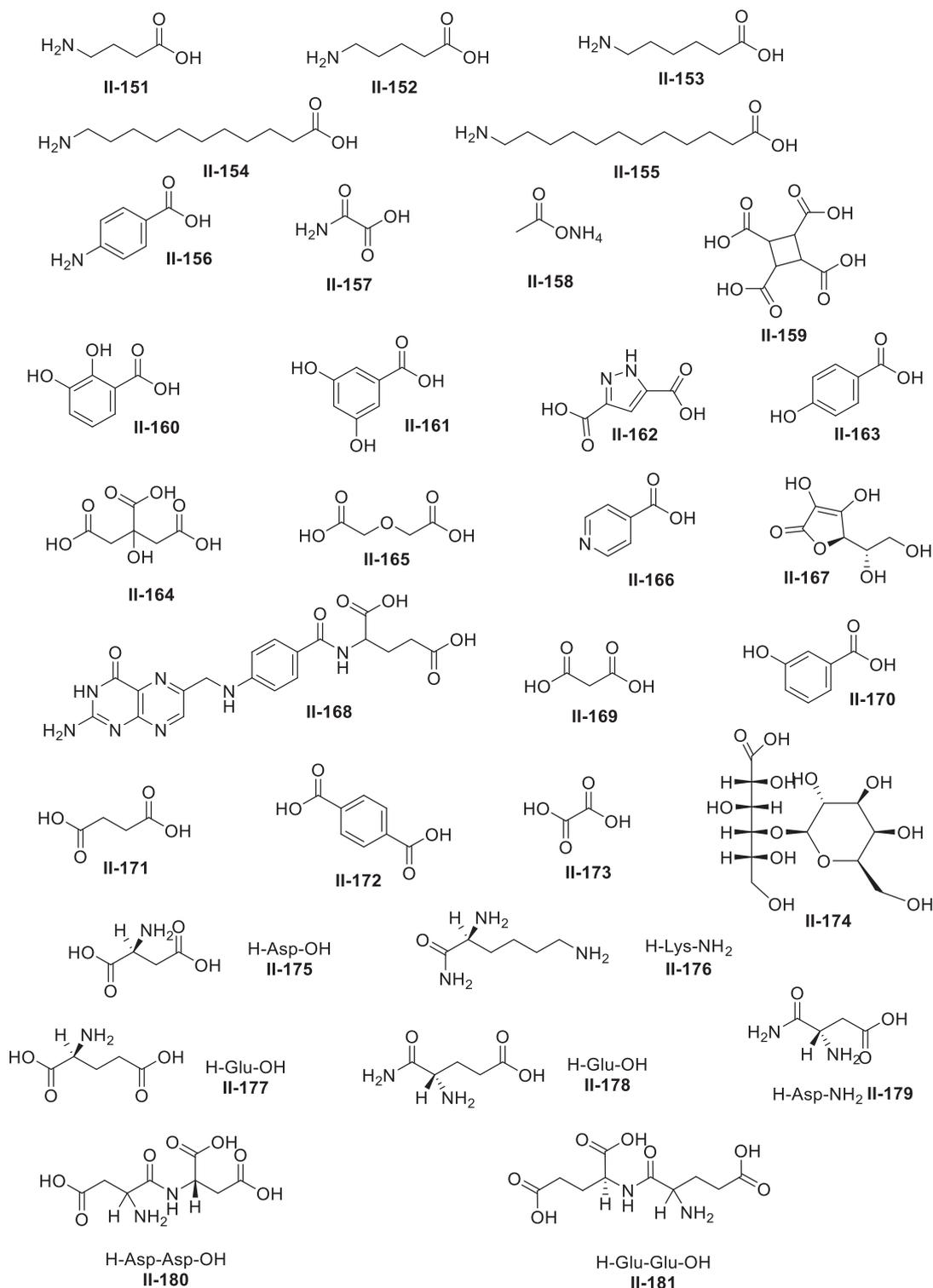
C	2.61628100	-0.02191800	0.58518000
C	3.71599600	-0.57299900	-0.04657900
C	3.54877500	-1.74706900	-0.77361100
C	2.29276400	-2.34407600	-0.84958300

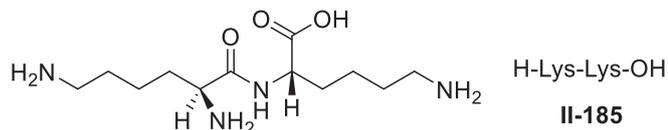
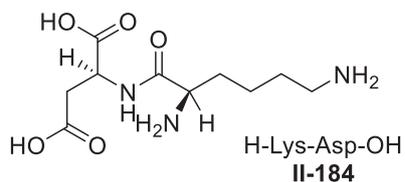
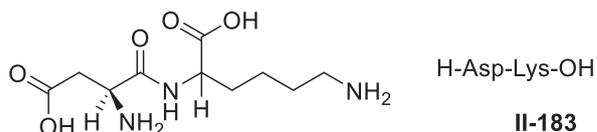
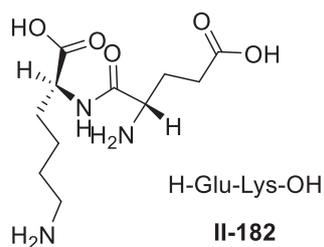
C	1.20101600	-1.77922100	-0.20197700	N	1.86195400	-1.59978600	0.40860700
C	1.34162200	-0.59920100	0.53596900	N	1.89654800	0.66338300	-1.04383900
H	4.68767700	-0.09513700	0.02193000	N	0.81085200	-0.04037300	-1.05743200
H	2.16272600	-3.25954200	-1.41729000	Cl	4.19167300	1.42431400	0.00872000
H	0.22434800	-2.24477900	-0.26160300	H	-4.25117500	2.72134600	-1.12201500
S	-0.01748700	0.17333700	1.33405500	Cl	-0.44107500	-2.51853200	-0.89547300
C	-0.77991400	1.20809800	-0.49869000	N	-3.00012400	-1.42250000	0.42796400
C	-2.60075800	-0.36094200	-0.12262200	H	-2.97968400	-1.66352300	1.42218000
N	-2.86743900	0.90660400	0.22436000	H	-3.90265500	-1.71910600	0.05451400
N	-1.96050400	1.78239100	-0.06790500	H	-2.24675500	-1.98255600	-0.04126100
N	-1.83486000	-0.64069000	-1.17784000				
N	-0.91739900	0.23243700	-1.46636300				
Cl	-3.62556700	-1.60425900	0.49784200				
H	4.39961300	-2.19145500	-1.27775900				
Cl	0.46795900	2.38531000	-0.90190400				
N	2.75294000	1.21708300	1.35193500				
H	3.34459400	1.10514600	2.17786700				
H	3.12601800	1.98779600	0.79465400				
H	1.79029700	1.46591200	1.66855200				

TS coordinates for II-20 and zwitterion amino-thiophenol
SNAr step 2:

C	-2.76252400	0.00242300	0.20661500
C	-3.73365600	0.76406700	-0.41685700
C	-3.49313800	2.11781900	-0.63565900
C	-2.29174000	2.68881400	-0.23299600
C	-1.32199800	1.90976100	0.38880000
C	-1.54665800	0.55380100	0.61435200
H	-4.67046300	0.31480400	-0.73036100
H	-2.10492000	3.74316100	-0.40453200
H	-0.37946700	2.34410200	0.70168300
S	-0.31337000	-0.43094200	1.42549400
C	0.69802600	-0.96447300	-0.01689200
C	2.83669500	0.35169700	-0.15346400
N	2.92774700	-0.85856700	0.38802600

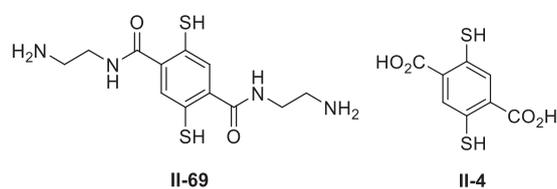
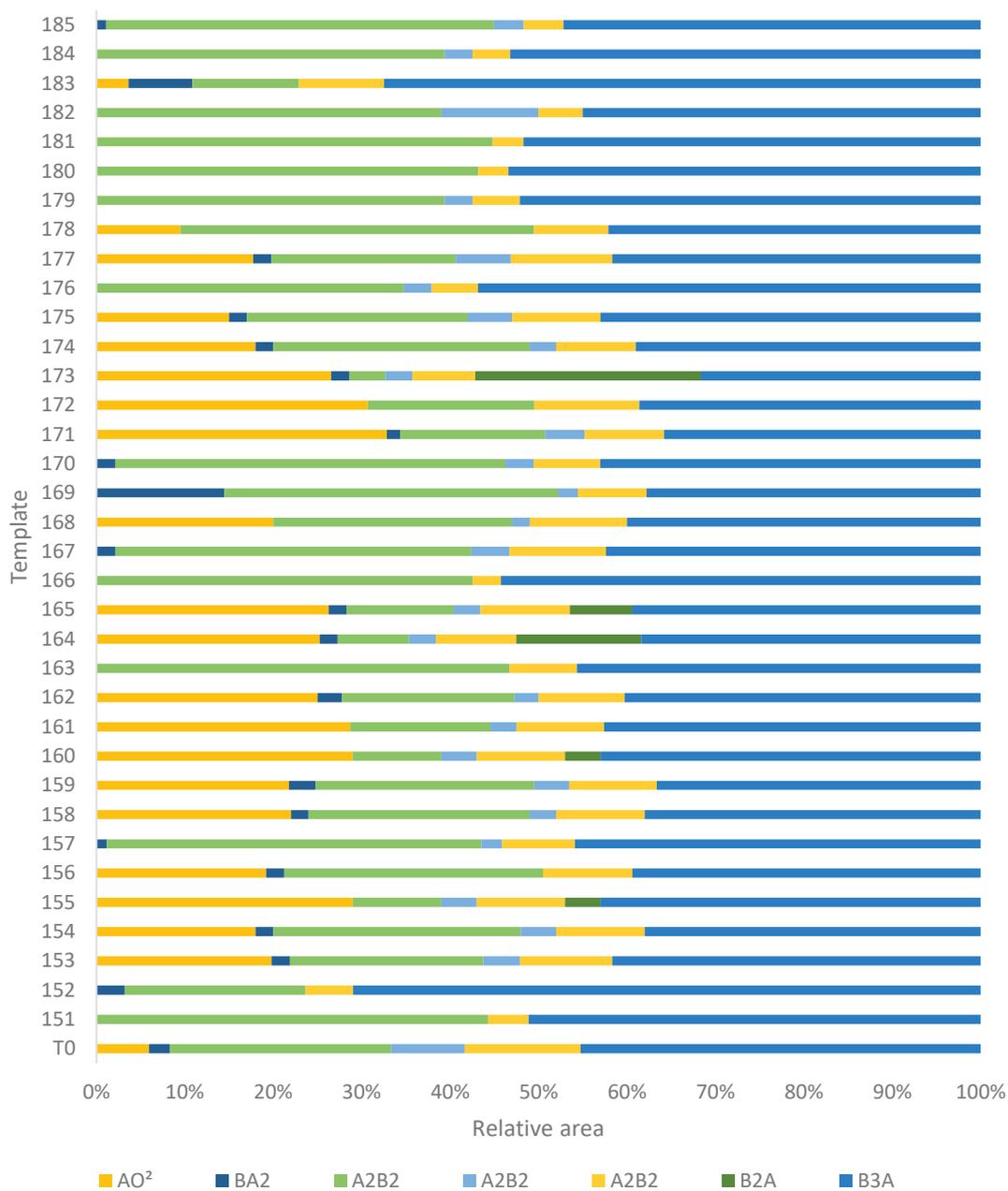
g. Appendix II-VII: Set of amplification data for the DCLs between **II-69** and **II-1** (4 + 4 mM) in DMF





Scheme 247: Templates used for DCLs of II-1 and II-69.

Templat e	AO ²	BA2	A2B2	A2B2	A2B2	AB2	AB3	Total area
T0	5%	2%	21%	7%	11%	0%	38%	1641
151	0%	0%	39%	0%	4%	0%	45%	2423
152	0%	3%	19%	0%	5%	0%	66%	1128
153	19%	2%	21%	4%	10%	0%	40%	1687
154	18%	2%	28%	4%	10%	0%	38%	1870
155	29%	0%	10%	4%	10%	4%	43%	1446
156	19%	2%	29%	0%	10%	0%	39%	1676
157	0%	1%	36%	2%	7%	0%	39%	1946
158	22%	2%	25%	3%	10%	0%	38%	1717
159	22%	3%	25%	4%	10%	0%	37%	1667
160	29%	0%	10%	4%	10%	4%	43%	1446
161	29%	0%	16%	3%	10%	0%	43%	1430
162	18%	2%	14%	2%	7%	0%	29%	2148
163	0%	0%	43%	0%	7%	0%	42%	1912
164	25%	2%	8%	3%	9%	14%	38%	1597
165	26%	2%	12%	3%	10%	7%	39%	1548
166	0%	0%	40%	0%	3%	0%	51%	2177
167	0%	2%	37%	4%	10%	0%	39%	1768
168	20%	0%	27%	2%	11%	0%	40%	1627
169	0%	13%	34%	2%	7%	0%	34%	2203
170	0%	2%	41%	3%	7%	0%	40%	2162
171	22%	1%	11%	3%	6%	0%	24%	1594
172	31%	0%	19%	0%	12%	0%	39%	1334
173	26%	2%	4%	3%	7%	25%	31%	1802
174	18%	2%	29%	3%	9%	0%	39%	1790
175	15%	2%	25%	5%	10%	0%	43%	1621
176	0%	0%	33%	3%	5%	0%	54%	2483



Scheme 249: Evolution of the macrocycle proportions by templating the DCLs of II-1 = B and II-69 = A (4 + 4 mM, DMF). Oⁿ: n overoxidation. X_nY_m cyclic disulfide oligomer composed of n X units and m Y units.

Chapter 3: Poly-dendrigrfts of lysine: flexible objects for various chemistry

Abstract

Dendrigrafts of lysine (DGL) are tree-like polymers of lysine. These recent molecules were first reported in the late 80's⁴⁰⁸ and are now widely employed in many fields, from imaging, catalysis to biomaterials engineering⁴⁰⁹. DGLs present an important density of positive charges resulting from the protonation of the lysine residues at physiological conditions. These polycations are adapted to develop strong electrostatic interactions with polyanions, such as GAGs, a family of linear polysaccharides. The most negatively charged GAG is heparin, a linear sulfated polysaccharide that consists of repeating subunits of α -1,4 linked uronic acid and D-glucosamine. This polymer, which is extracted from porcine liver, is an anticoagulant drug that disrupts the coagulation cascade leading ultimately to blood clots. The only FDA-approved drug in case of heparin overdose is protamine, a small arginine-rich peptides with important side effects. A second generation of DGL inhibited the anticoagulant activity of unfractionated heparin (UFH), low molecular-weight heparin (LMWH) and fondaparinux more efficiently than protamine did in human plasma. A computational study by molecular dynamic simulations revealed that this superior effect of DGL was due to its conformational flexibility. In fact, the branched architectures enabled the dendrimer to marshal more efficiently its charges in the association with heparin compared to protamine. The simulations revealed that mid size and long heparin formed a more stiffened complex with DGL, which explains its better neutralisation of UFH and LMWH compared to protamine. Regarding the binding of heparin with protamine, the study showed that the relative size of the partners has a substantial influence on their corresponding binding affinities, where the number of direct contacts (i.e., electrostatic or hydrogen bonds) is not the only factor guiding the observed selectivity. As a matter of fact, non-contact interactions were identified as an important contributor to the binding affinities. The simulations explained the experimental neutralisation of UFH and LMWH by protamine and the inactivity of protamine against Fondaparinux. Finally, the positive surface of DGLs was used to quench the fluorescence of homopolymers of acid aspartic and fluorescein upon aggregation. The disruption of the complex by an external partner restored the fluorescence and was used as a strategy of detection, resulting in various limits of detection for spermine and bacteria. The DGLs' ability to mimic esterases was also validated.

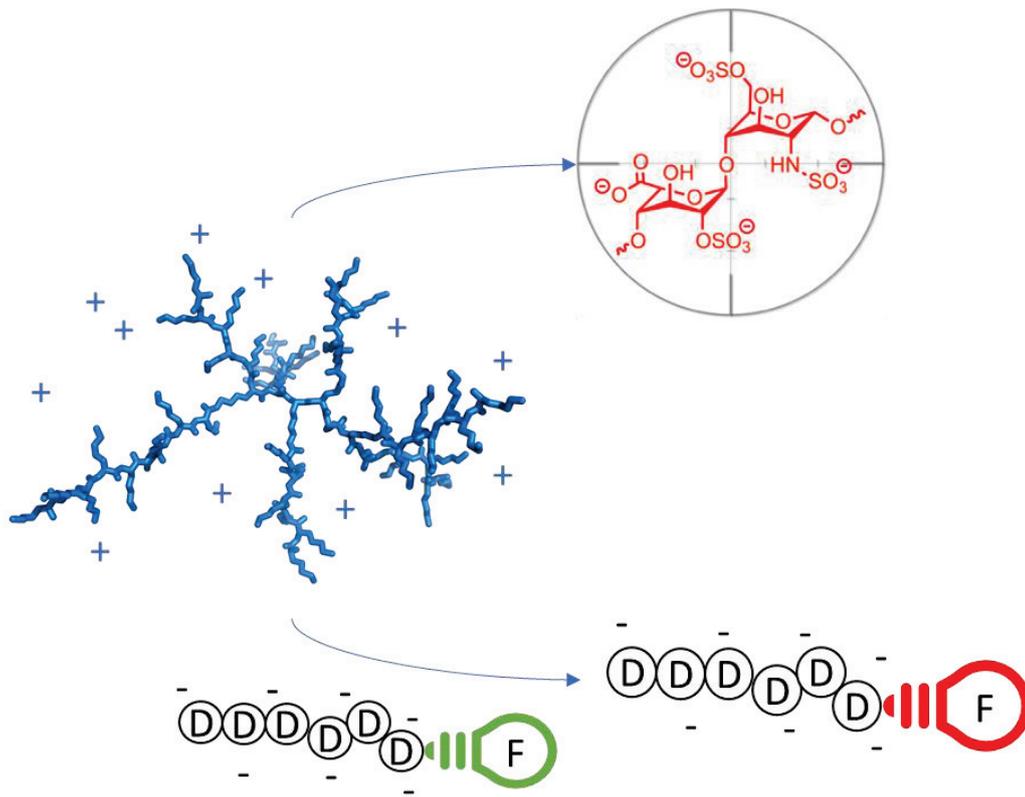


Figure 24 : DGLs' cationic surface are used to inhibit heparin's anticoagulant properties and to quench fluorescence of carboxyfluorescein derivatives upon aggregation

1. Introduction

The objective of this chapter was to study the recognition in physiological conditions of biomacromolecules by dendrigrafts of lysine (DGLs).

In the last two chapters, the aim was to synthesise disulfide bridge cyclophanes through dynamic covalent chemistry. These cyclophanes were dedicated to the recognition of biomolecules in water. Regarding this strategy of design, systematic changes can be readily made but the modelling is imperfect and can lead to ineffective systems. In this chapter, a different approach was followed: rather than synthesising complex molecular architectures, we directly selected and used commercially available scaffolds among a large number of structures, such as DGLs, and benchmarked them for various applications linked to their specific properties, for instance for the neutralisation of heparin, a drug injected in case of thrombosis. Regarding this strategy, the main drawback is the usually the difficult modification of the commercial scaffolds.

Thrombosis refers to the formation of blood clots. The venous thrombosis is a common complication occurring after major surgery such as pulmonary or cardiovascular interventions⁴¹⁰. Coagulation is a complex process resulting from the activation of a cascade of blood factors. A fibrin is then created, fixing the damage. Genetic diseases are often the cause of coagulation trouble. In this case, anticoagulants regulate the coagulation cascade and are also used to control the blood fluidity during surgery. Anticoagulants are usually administrated by intravenous way to prevent blood clotting⁴¹¹. Among them are heparins and vitamin K antagonists, direct inhibitors of factor Xa and thrombin⁴¹².

Heparin is a family of polysulfated glycosaminoglycans, composed of a pentasaccharide Fondaparinux, fractionated (LMWH) and unfractionated heparin (UFH)⁴¹³. The last two bind with factor Xa and thrombin, disrupting the coagulation cascade (Figure 25). Native heparin is extracted from porcine liver. Then, short polymeric forms are obtained thanks to a depolymerisation with heparinases which are enzymes lysing the glycosidic linkage between hexosamines and uronic acids. LMWHs are extensively used as anti-coagulant⁴¹⁴ and show less accidents than UFH. In fact, contaminants can be found in UFH because of its animal origin, and its purity can vary from batch to batch⁴¹⁵. Therefore, overdose of heparin can arise, causing bleeding. LMWHs showed 1.4 % of bleeding incidents compared to 2.3 % for UFH⁴¹⁵. In this case, protamine, a poly-arginine, is the only FDA-approved medicine neutralising the anticoagulant activity of heparin. However, protamine, which is produced by extraction of salmon sperm, has several drawbacks such as anaphylactic complications, variable disponibility and a reduced spectrum of activity on different classes of heparin since it is ineffective on Fondaparinux⁴¹⁶.

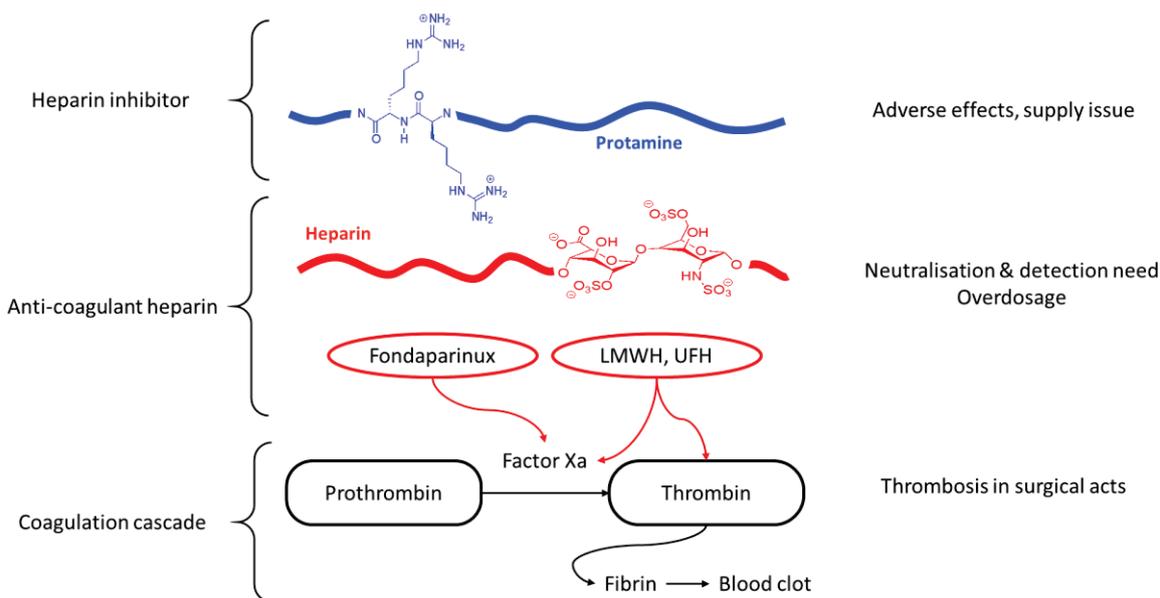


Figure 25: Protamine and heparin in the coagulation cascade. Protamine inhibits heparin which is an anticoagulant used to stop the coagulation cascade. Heparin is a family of GAGs divided into Fondaparinux, LMWHs and UFHs.

Over the last decade, many alternatives of protamine were proposed, based on polycationic molecules developing strong electrostatic interactions with heparin, the most negatively charged biopolymers³⁶. However, few candidates are currently in a late-stage development. In particular, a dendrimeric compounds UHRA (universal heparin reversal agent) proved to be a universal reversal of heparin in aqueous buffers⁴¹⁷. Dendritic polymers are tree-like molecules of high size with layered architecture⁴¹⁸. They are commercially available, fully synthetic and have been widely employed for biomedical applications⁴¹⁹. A variety of dendritic polymers exists, each having biological properties such as multivalency, self-assembling, electrostatic interactions, chemical stability, low cytotoxicity, and solubility⁴²⁰. Given that UHRA neutralised each class of heparin, we chose this scaffold and the commercial available DGLs as a possible inhibitor of heparin.

In a nutshell, the objectives of this chapter were to study the recognition of bio(macro)molecules in water such as heparin with DGLs by a combination of experimental and computational experiments.

2. Recent advances in the neutralisation of heparin

This bibliographic subchapter gives elements on the current state-of-the-art research on the neutralisation of heparin. Analogues of protamine are needed to tackle its disadvantages, such as reduced activity on LMWH and anaphylactic complications. The mechanism of action of protamine is mainly based on electrostatic interactions and H-bonds, due to its peptide backbone and its cationic amino acids such as arginine and lysine. These features have inspired the scientists designing molecules

³⁶ Ourri, B., Vial L., Lost in (Clinical) Translation: Recent Advances in Heparin Neutralisation and Monitoring, ACS Chemical Biology, Accepted manuscript
315

for heparin neutralisation. The future inhibitors should be active against all heparin sizes, with high and specific binding properties without toxicity. Currently, the newly developed molecules in this domain could be classified in three categories, which have not been evolving since the last decade: poly-peptides / bio-inspired molecules, auto-assembling systems, and dendrimers / polymers (Figure 26). Even though the main point is always to reproduce the electrostatic interaction of the protamine, many ways of presenting a positive surface to heparin are now explored.

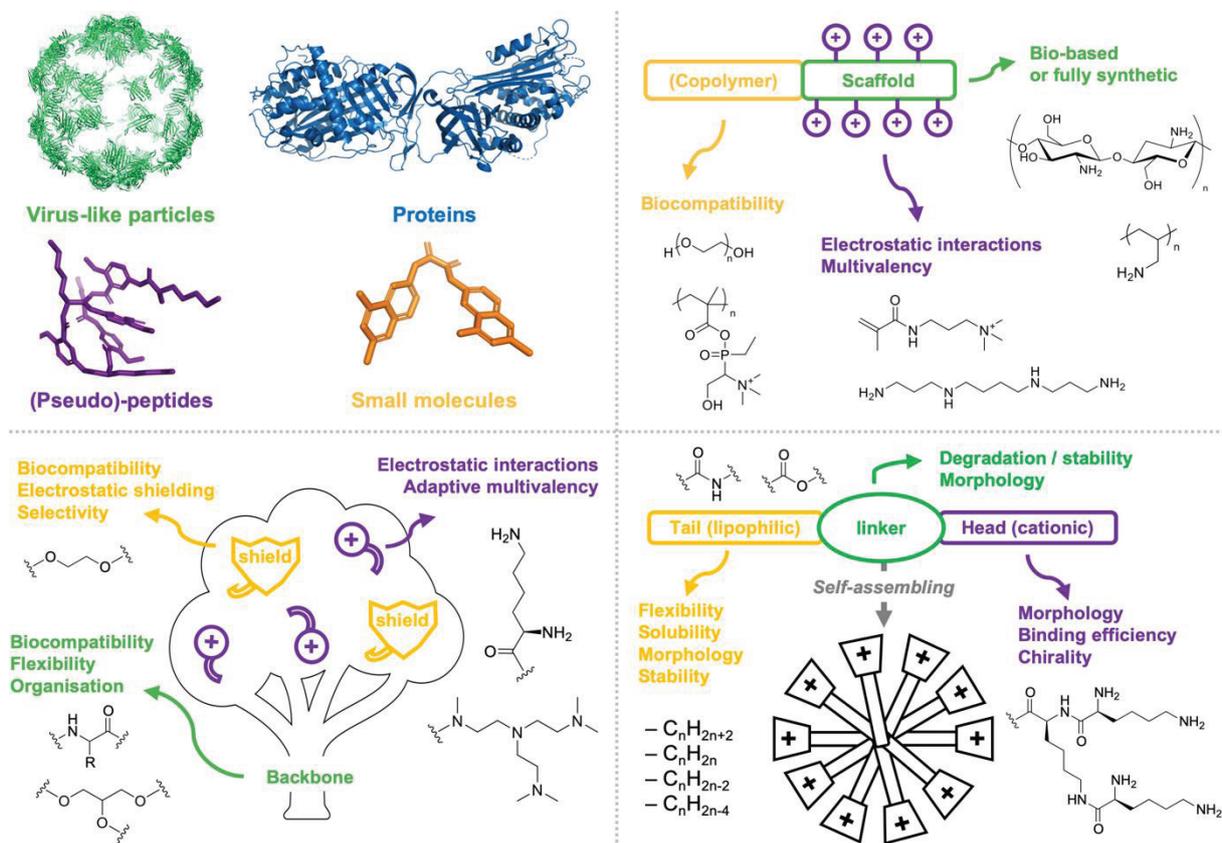


Figure 26 : Recent approaches toward heparin neutralisation involving bio-inspired binders (upper left), linear polymers (upper right), dendritic polymers (lower left), and self-assembled micelles (lower right).

a. Proteic/peptidic/pseudo-peptidic binders

Numerous natural proteins are known to bind heparin⁴²¹. Some of these proteins and their inactivated variants have been considered as potential heparin antidotes. For instance, a modified recombinant-inactivated antithrombin (riAT) neutralised efficiently all heparin sizes in a model rat cardiopulmonary bypass study (CPB)⁴²², without side effects⁴²³. Given that AT is involved in the coagulation cascade as an irreversible inhibitor of thrombin, its inactivation neutralised its anticoagulation properties and its heparin affinities. The half-time of UFH neutralisation proved to be higher than with protamine, without the side effects on hemodynamic parameters and plasma histamine levels generally observed with the FDA-approved antidote. riAT It paves the way to explore other inactivate molecules involved

in the coagulation cascade. Andenaxet, currently in phase III, is an inactive recombinant factor Xa. The molecule bound tightly to all heparin sizes in presence of AT, with nanomolar affinities⁴²⁴.

Native (unmodified) proteins, which are not involved in the coagulation cascade can also be considered. Histones are alkaline proteins wrapping DNA within the cell nuclei, known to bind with heparinoids and were thus investigated as protamine analogues through activated partial thromboplastin time tests (aPTT) in plasma⁴²⁵. The anticoagulating activity induced by therapeutic doses of heparin required low concentrations of histones to be inhibited by 50% in plasma ($IC_{50} = 1.8$ and $6.6 \mu\text{g/mL}$ for the neutralisation of 0.2 and 0.6 IU/mL UFH, respectively). Nevertheless, histones presented a high risk of side effects because of its involvement in inflammation processes.

Engineering of proteins that are not known as heparin binders has also been investigated. For instance, envelop and capsid proteins can self-assemble into virus nanoparticles. A single amino acid mutation on a bacteriophage Q β virus-like particle (VLPs) yielded a large cationic region that was replicated 180 times over the capsid⁴²⁶. They contain no viral materials, hence induce no immune response and are well tolerated in the human body. Their size and their surface are an interesting platform to functionalise. The neutralisation of UFH was complete in human plasma by APTT assays, at 1 mg/mL for $1.3 \mu\text{g/mL}$ of heparin, while protamine started to display aberrant behavior starting from 0.07 mg/mL with increased coagulation times.

Poly-peptides benefit from being based on the same structure as protamine, including H-bonds sites and cationic amino acids residues. Compared to proteins, peptides have a higher activity per unit mass and their manufacturing costs is lower (synthesis for small peptides vs. recombinant engineering for proteins). One of the most important advantages relies on the ease of variation allowed by this kind of synthesis. Consequently, a high screening of similar molecules leads to optimised systems faster than other approaches. In addition, long molecules prevent immunogenicity and are usually more specific, minimising side effects.

Semaphorin3a is known to bind to glycosaminoglycans (GAGs) by its C-terminal region. Hence, two derived peptides from this region, FS2 and NFS3 were found to bind heparin in low micromolar range but their ability to restore clotting has not been investigated so far⁴²⁷. Basically, protamine is active thanks to its arginine (67 % content). Thus, protamine imitators, which contained up to 3 and 20 arginine residues were synthesised⁴²⁸. It has been found that a minimum of 8 arginine is required to have anti-UFH activity. A peptide composed of 15 arginines neutralised UFH *in vivo* on rats by aPTT and anti-factor Xa assays similarly with protamine. However, the alkaline nature of both antidotes caused increased histamine release with rapid infusion. Having the same side effects as protamine is probably the main drawback of a strategy miming its action. Escaping from the protamine backbone avoids the same side effects.

One of the main attractive features of pseudo-peptides is their low susceptibility to proteolytic degradation *in vivo*. For instance, PER977 is a pseudo-peptide di-arginine piperazine, interacting with heparin through electrostatic interactions and H-bonds^{417b}. The literature concerning PER977 displays contradictory results: some authors claimed that this species did not bind to LMWH,^{417b} while others demonstrated that PER977 neutralised LMWH efficiently⁴²⁹, with limited side effects⁴²⁹⁻⁴³⁰. While

PER977 does not exhibit any secondary structure, one could also envision pseudo-peptides adopting precise conformations. In this context, peptide-like tetrameric foldamer $\text{H}_2\text{N}-(\text{Lys-Sal})_4-\text{CONH}_2$, made of *L*-lysine and a salicylic acid derivative (5-amino-2-methoxy-benzoic acid, Sal) was explicitly designed to form two-stranded β -sheet-like structures in which the basic lysine side chains project from one face of the structure in order to bind to fondaparinux⁴³¹. $\text{H}_2\text{N}-(\text{Lys-Sal})_4-\text{CONH}_2$ self-assembled into hexameric aggregate upon association with the pentasaccharide, inhibiting its anti-factor Xa activity in an aqueous buffer ($\text{EC}_{50} = 3,6 \mu\text{M}$). The foldamer was also active versus UFH in human plasma ($\text{IC}_{50} = 1,1 \mu\text{M}$ for the neutralisation of 0.5 U/mL from aPTT assay). However, the stability of the foldamers in a complex biological media such as diluted blood should be assessed. The presentation to a surface of interactions by the β -sheets and the lysine cations was an original approach, breaking with the paradigm of designing inhibitors of heparin only by means of electrostatic interactions. Following the same vision, members of the peptide-like surfen (Bis-2-methyl-4-amino-quinoyl-6-carbamide) family completely neutralised UFH (IC_{50} low- μM range), LMWH (μM range) and Fondaparinux (10-40 μM) in aqueous buffers and UFH and Fondaparinux (IC_{50} in the μM range for the neutralisation of 0.2 IU/mL) in mice plasma at therapeutic dose⁴³². It was suggested that the surfen molecules aligned themselves by H-bonds in a cooperative manner along the oligosaccharide, blocking the formation of AT-heparin complexes.

The coagulation cascade itself is a large source of inspiration for heparin inhibitors, promising efficient drug with high specificity, like inactivated protein. Even small peptide-likes have significant advantages due to their tailored design, including cationic sites, H-bonds, and tertiary structures. The main rewards of this category of antagonists are the ease of derivatisation. Nevertheless, polypeptides based on the same mode of action as protamine will probably meet the same side effects and will not overcome the current problems of the approved drug. Another strategy bet on fully artificial self-assembling systems.

b. Self-assembled binders

Self-assembling is the spontaneous and reversible organisation of multiple simple building blocks into a single complex superstructure. Among self-assembled superstructures, micelles are formed by the aggregation of surfactant molecules, which are made of a hydrophobic tail, a linker, and a polar head. Aggregates that bear cationic polar heads present a large positively charged and organised surface to heparin. Over the last decade, Smith *et al.* developed a new family of self-assembled multivalent micelles (SAMul) toward heparin binding⁴³³. The concept relied on the presentation of a large positively charge and organised surface to heparin, hence being well adapted to long heparin. Each part of the building blocks plays a key role in both the self-assembling process and the subsequent heparin binding event.

Even though the aliphatic tails are buried inside the micelles, their nature has a major influence on the micellar properties. A too long aliphatic chain disfavoured solubility, yielding large and uncontrolled aggregates that should be avoided⁴³⁴. For instance, a decrease of the chain length from a C_{18} to a C_{14} tail increased the solubility and decreased the CMC and aggregation size of the micelles⁴³⁴. In addition, excessive lipophilicity led to disaggregation over time upon binding to plasma proteins such as albumin.⁴³⁵ Preventing rigidity in the design of the tail was also recommended since the packed

surfactants were destabilised as the degree of unsaturation increased⁴³⁶. For a C₁₈ chain containing from 1 to 3 double bonds, critical micelle concentration increased from 42 to 78 μM in aqueous buffer, while the cation/anion charge excess at which 50% of a dye is displaced (Mallard Blue) from the anticoagulant's surface (CE₅₀) for UFH increased from 0.80 to 2.30 in the same medium.

The linkages between the tails and the heads of the building blocks were mostly amide and esters bonds, sometimes triazole⁴³⁵. Because esters could undergo hydrolysis over time, they avoided the accumulation of SAMul in excess (i.e., unbind to heparin), which is highly desirable in order to avoid side-effects. Free micelles fell out over a 24 hours timescale in physiological conditions, while micelles interacting with heparin seemed to be protected from degradation⁴³⁵. The incorporation of a spacer between the linker and the head could help to establish H-bonds with the targeted heparin, as protamine does. The spacer could improve the morphology of the assembly by the preorganisation of the head and by enhancing the potential expression of their chirality⁴³⁷. Modelisation showed that a spacer like a glycine also changed the polarity and the shape of the head, optimising the surface presented to heparin.

The last part of the surfactant is the head, which also controlled the morphology of the self-assemblies. First- and second-generation dendrons of lysine yielded worm-like micelles and spherical micelles, respectively.⁴³⁶ The worm-like assembly was better shape-matched to heparin and displayed a slightly better capacity to bind UFH from a simple Mallard Blue displacement assay in human serum (CE₅₀ ≈ 1 vs. CE₅₀ ≈ 25 for the spherical micelles). In addition, the worm-like assembly seemed to protect the lipophilic part of the micelle to be disrupted by interactions with proteins like albumin. The nature of the cationic motif also played a part. Two heads bearing an identical positive net charge of 2, N,N-di-(3-aminopropyl)-N-methylamine and spermidine, exhibited distinctive binding efficacy (CE₅₀ = 0.69 vs. 0.34 in aqueous buffer, respectively). Surprisingly, increasing the net charge of the head did not lead to improved affinities for heparin since electrostatic repulsion between the building blocks destabilised the micelles⁴³⁸. Indeed, a variation on the ligand charge^{438b} from + 1 to + 3 highlighted that, for this C₁₆ tail and polyamine category, the most effective heparin binder was a + 2 charged. The +1 charge was found to be poorly soluble, and the +3 did not outperform one of the +2 in terms of) displacement with MalB (CE₅₀ 0.34 vs 0,49 in aqueous buffer. Finally, introducing chirality in the head could also improve the affinity toward UFH. For instance, a variation with two successive chiral alanine and lysine residues showed that the self-assembly of SAMul was more stable with the LL/DD rather than DL/LD sequence, which was translated into EC₅₀ values (120 to 140 μM range in aqueous buffer, respectively)^{437b}.

The auto-assembling of surfactants was favoured in the presence of heparin and allowed the presentation of a positively surface to the polyanion. The electrostatic interactions were optimised by multivalence, not only because of the adaptive capacity of the micelles but also thanks to the flexible heparin. Auto assembling was straightforward synthetic approach to generate multivalent systems compared to classical covalent routes. The micelles formed highly stable complexes with heparin and unbound micelles were degraded after 24 h. However, every polyanions or anions could be trapped by these micelles, putting in light potential adverse effects which were not assessed so far. In addition, the risk of destabilisation in plasma should be studied. The intrinsic anticoagulant activity of the

micelles and surfactants should also be studied. Furthermore, the pharmacokinetic and the activity of the degraded micelles were not described. Finally, the strategy of presenting a large cationic surface to heparin was well adapted to long heparin, but it seemed to be less effective for LMHW and Fondaparinux. They did not benefit from effective flexibility to bind with the micelle by multivalence. The amphiphilic molecules self-assembling into SAMul could be straightforward modified, which was important for the search of a lead micelles. This ease of synthesis was a major advantage of self-assembling systems compared to other classes of antagonists. Small molecules, which further aggregate, could be easily manufactured in a large scale and purified. Self-assembling systems benefited from multivalence and their adaptive shape, which echoed to the flexible heparin, allowing an optimisation of electrostatic interaction between the two partners. This strategy moved away from only replicating the protamine mode of action and its associated side effects. The flexibility of heparin was also targeted by the last category of heparin antagonist.

c. Polymeric/dendritic binders

Heparin being a polyanion, polycations have been logically studied as protamine analogues. Polymers are high molecular weight molecules built by the repetition of one or several units. Some of these monomeric units can carry cationic functions at physiological pH, hence exposing a large positively charged molecular surface to heparin. The polymers can either be bio-based or fully synthetic. The monomer could be combined by polymerisation chemistry, or a pre-existent raw polymer could be functionalised. In all cases, the choice of the polymer's backbone was of crucial importance since functionalizing distinct materials with the same moieties did not yield identical reversal activity against heparin and the same source of contamination and side effects. Most of the polymers described here are derived from bio-source constituents. There are oligosaccharides, like Dextran, a dextrose ramified polymer made by enzymatic action, Chitosan, consisted of glucosamine, extracted from shellfish. It could also be a synthetic substance, such as polyglycerols, widely used for biomedical applications⁴³⁹.

14 polymers made from Dextran (Dex), γ -cyclodextrin (GCD), pullulan (Pul) and hydroxy-cellulose were grafted with cationic groups, mainly derived from GTMAC (glycidyltrimethyl-ammonium chloride), also from PAH (poly(allylamine hydrochloride)), spermine and from APTMAC (N-acrylamidopropyl-N,N,N,-trimethylammonium chloride)⁴⁴⁰. Polymers Dex-6-GTMAC, Pul-GTMAC and GCD-GTMAC1 had similar degree of substitutions functionalization but contrasted UFH binding abilities, and the mass ratio of polymers to UFH required to bind 90 % of UFH in aqueous buffer was 2.7, 1.9 and superior to 8, respectively. The degree of functionalization also plays a significant role since a minimum of *ca.* 25 positive charges on Dex-GTMAC were required to start observing UFH binding in solution. The most substituted Dex-GTMAC bearing *ca.* 260 cationic motifs turned out to be the most potent antidote to UFH (similar to protamine), fully restoring the parameters of the thrombosis development and typical measures of the coagulation cascade (i.e., aPTT and anti-factor Xa activity) in rats and mice. Neutralisation abilities of this candidate for protamine replacement was neither reported on LMWH nor on fondaparinux. Yet, its biocompatibility is undoubtedly superior to protamine as it did not induce any immune response, was normotensive and hemodynamically neutral in model animals. In addition, the polymer was rapidly eliminated (half-life of 12.5 ± 3.0 minutes) without any blood, accumulation in tissues and organ toxicity⁴⁴¹. Dextran was inexpensive and accessible, but the raw material was

highly dispersed. Another polymer made of chitosan, HTCC, was substituted by the same cationic group, GTMAC⁴⁴¹ with the same degree of substitutions compared to the previous dextran polymers and was active against LMHW in aqueous buffers. In addition, the substitution of raw chitosan annihilated its side effects on fibrinogen. This example highlights that polymer bearing the same cationic group but not on the same raw material had different heparin reversal activity.

The raw material could be fully synthetic, such as in PAH. Directly inspired by protamine, a PAH polymer composed of 9 % of amines substituted by arginine neutralised UFH similarly to protamine in vivo (APTT, rat model)⁴⁴². Functionalisation of raw PAH lowered his cytotoxicity, which contrasts with the general statement that arginine-rich peptides like protamine interact more strongly with lipidic bilayers than lysine-rich peptides⁴⁴³. Arginine-rich peptides are consequently toxic to cells from concentrations around 800 nM to 50 µM while lysine-bearing polymers are only cytotoxic at higher concentrations (>50 µM).⁴⁴⁴ Other studies are required to evaluate several toxicity parameters, as well as pharmacokinetic and adverse effects of the PAH polymer. Indeed, the amide links could be cleaved in physiological conditions, restoring the initial structure of PAH, which is known to be toxic for the human body. A PAH-LYS or a copolymer LYS-ARG could be an alternative solution.

Since a strong density of positive charge along the polymeric chains often leads to increased toxicity, a well-known strategy toward its reduction consists in copolymerizing it with nontoxic and biocompatible copolymer blocks such polyethylene glycol (PEG)⁴⁴⁵. For instance, a copolymer PEG₄₁-PMAPTAC₅₃ (PMAPTAC poly(3-methacraloylamino)propyl trimethylammonium chloride) showed limited side effects on blood count and pressure⁴⁴⁶. In addition, another similar copolymer of PEG and PMAPTAC displayed total fibroblast cell viability up to 150 µg/mL, while the corresponding non-PEGylated analogue was found to be strongly toxic (viability < 50%) at concentrations above 50 µg/mL⁴⁴⁷. The candidate restored coagulation and completely reversed both clotting time and anti-factor Xa activity caused by UFH, the LMWH enoxaparin or fondaparinux in Wistar rats. In addition, the copolymer did not modify the cardiorespiratory parameters of the animals, nor it damaged or significantly accumulated in their kidneys and liver. The inherent toxicity of vinyl polymers was also lowered by the incorporation of a PEG block. Copolymer made of PEG and PDMAEMA (poly(2-(dimethylamino)ethyl methacrylate) revealed that the length of the PEG co-block did not impact on the heparin affinity, assessed by methylene bleu (MB) displacement in aqueous buffer⁴⁴⁶. However, the length of PDMAEMA block was found crucial: a minimal length was required for the cationic co-block to achieve efficient neutralisation, but a longer chain was not useful. DLS, quartz crystal microbalance and surface plasmon resonance analysis proved that the complexes between the polymers and heparin gave more stable colloidal forms when the PEG block was longer. This indicated that PEG co-block not only increased biocompatibility, but also improved the pharmacokinetic. Indeed, the PEG reduced large aggregation of complexes in blood, which was one of the sources of protamine adverse effects.

Polymers are potential candidates thanks to their multivalent properties which can be optimised by copolymerisation. Artificial polymers had a low dispersity compared to bio-sourced raw materials. Most of the presented polymers could be easily manufactured, but their polycationic nature could lead to important side effects. Furthermore, their activity against LMWH and Fondaparinux were limited, because of a lack of organisation and spatial optimisation. Their degradation profiles were not clearly

elucidated. Variation on polymer design was quite easy but was limited by the 2 dimensions of these objects.

Contrary to linear (co)polymers, dendrimers are three-dimensional objects with layered architecture that benefit from their higher flexibility for heparin binding. In fact, when the core of 3 generations of polycationic poly(amidoamine) dendrimers (PAMAMs) was modified with a polyphenylenevinylene (PPV) motif, therefore rigidifying the whole structure, it turned out that their relative ability of binding UFH per cationic charge (CE_{50} independent of the generation (0.63 – 0.67)) in a buffered aqueous medium dropped in comparison with unmodified PAMAMs (CE_{50} decreased from 0.75 to 0.38 for G_1 to G_2 , and then increased to 0.45 for G_3). It demonstrated that the overall flexibility of dendrimers enables them to adapt their global shape more easily to maximise the total number and efficiency of contacts between the partners⁴⁴⁸. In addition, there was no interest in increasing the charge density for PAMAM family to improve the binding with heparin.

According to the literature, large cationic dendritic structures may disrupt key platelet functions⁴⁴⁹. As mentioned before, the introduction of PEG chains was an effective strategy to limit antidotes' drawbacks. A dendritic candidate reported by Kizhakkedathu *et al.* was named UHRA, standing for universal heparin reversal agent. Its design was based on a hyperbranched polyglycerol (HPG) scaffold bearing 20 trivalent ammonium methylated tris(2-aminoethyl)amine HBG (heparin binding group) motifs, capped with short methylated PEG chains as protective shell against non-specific interactions⁴⁵⁰. An ITC study in aqueous buffer of the association with UFH, enoxaparin and Fondaparinux proved that the dendrimeric architecture improved the binding compared to linear mPEG-HBG.

UHRA not only neutralised completely UFH and LMWH at 0,1 mg/mL *in vitro* assay in human plasma, but it antagonised Fondaparinux at 90 %, which was remarkable. This superior reversal activity of UHRA observed *in vitro* could be effectively translated *in vivo*, where the dendritic polymer completely and rapidly neutralised the activity of both UFH and LMWH enoxaparin, whereas protamine was only effective in neutralising UFH but could reverse only 60% enoxaparin activity in rats. UHRA exhibited high biocompatibility and limited side effects. Half-life time was 40 mn in rat, and no accumulation on organs or other adverse effects were observed^{417a}. Insights into the role of the PEG chains was provided by the authors⁴⁵¹. Isothermal titration calorimetry (ITC) experiments revealing that the enthalpic contribution to UFH binding was higher for UHRA in comparison with protamine, which could be related to its higher number of positive charge (60 vs. 22). However, an entropic penalty was observed for the dendritic polymer in comparison with the protein, which was due to the compression of the PEG chains upon interaction with UFH. This entropic penalty was the price to pay for the eradication of non-specific interactions with anionic blood components. The dendritic polymer lacking its PEG chains induced the aggregation of blood components, which was – as for protamine – at the origin its intrinsic anticoagulant property. It was also found that UHRA activity originated from the efficient disruption of the AT-UFH complex in the coagulation cascade. Such an understanding of the mechanism and the driven forces behind the binding event between the anticoagulant and its antidote is necessary for future optimisation and clinical development.

A comparison with Andexanet and PER977 on heparin UFH, enoxaparin and Fondaparinux was conducted by ITC^{417b}. The constants of association were in the range of micromolar and sub micromolar in aqueous buffers. The study confirmed that the candidate did not bind to clotting proteins such as fibrinogen, Xa and AT. UHRA appeared as a highly specific and safe candidate compared to the current drug. This highlighted the influence of methylated PEG brush, deserving a high affinity for all sizes of heparin and preventing undesired interactions. Indeed, the PEG layer acts as a shield, preventing interaction with clotting proteins and marshalling cationic charge by organisation and cooperativity. The PEG layer reduced the charge density of the object without compromising its affinity. Compared to copolymerisation with PEG co-blocks, the layered structure allowed a more diffuse positive charge density, whereas the co-block strategy only worked by segments, and did not benefit from the hyperbranched structure.

Polymers showed high potential as heparin inhibitors, based on their polycationic nature, which was also their Achille hell. Indeed, specificity and biocompatibility were problematic. One solution was copolymerisation with PEG co-blocks. Besides, polymers were rigid objects, limiting their adaptation to heparin. Their synthesis, based on cheap and available raw material, was simple and easily scalable. Turning to dendrimers presented the advantage of flexible multivalency, which enabled optimal electrostatic interactions. Furthermore, layer-structure allowed a shielding dedicated to biocompatibility and specificity, which was illustrated with UHRA and mPEG. The chemical nature of the cationic group dedicated to heparin binding was not essential, as well as his number, once a minimal amount of cationic charge was reached. The major drawback of this family was the expensive synthesis and purification, and the poor control of the dispersity, two major points for industrialisation. Nowadays, few dendrimeric systems had been investigated, paving the way for other studies and variation on dendrimer architecture.

d. Conclusion

Binding efficiently and specifically to a unique polyanion in the human body is a hard challenge. So far, current research for the design of new candidates are divided into three categories: poly-peptides and peptide mimic, auto-assembling systems and polymers and dendrimers. Nature is a large source of inspiration, and coagulation cascade has been explored as a pool of potential antagonists. Deactivation of native proteins, or proteomic for bio-inspired protein-like drug has the advantages of high specificity and low toxicity, if not a copy of protamine structure. At the opposite, auto-assembling systems are easy to synthesise and purify. These molecules aggregate in the presence of heparin and bind strongly thanks to their adaptive shape and their multivalency. Nevertheless, their side effects and behaviour in human body need more study to evaluate their full potential. Finally, polymers have shown interesting affinity and biocompatibility thanks to copolymerisation, but they have been limited by their specificity. This class of antagonists is inexpensive and easily accessible, whereas dendrimers are more difficult to synthesise and quite polydisperse. But these tree-like objects are very promising, given their layered structure, conferring them shielding effect, high affinity, and biocompatibility. Most of the candidates bind several classes of heparin, such as UFH, enoxaparin and Fondaparinux, and many of them avoid adverse effects of protamine as thrombosis generation, organ damage and

intrinsic activity in the blood. There is no denying that one of these molecules will replace one day the old protamine and will prevent surgical complications for a lot of patients.

3. Overview of dendrigrafts of lysine

This bibliographic subchapter provides elements on the objects used in this chapter: the dendrigraft of lysine. Dendritic polymers are a recently developed class of polymers. They were discovered in the late 1980s⁴⁰⁸ and have been a great success over the last decades. The family members are dendrimers, hyperbranched polymers and dendrigrafts⁴⁵². Similarly to polymers, they are composed of the repetition of a dendron with a tree-like structure (Figure 27). This structure contains a core, which is grafted with successive layers⁴⁵³. These layers consist of the repetition of the dendron and are capped by functionalised chemical groups. The arrangement of the surface is responsible for the presence of internal cavities. The dendritic polymers are composed of a succession of homogenous layers, starting from the core to the surface.

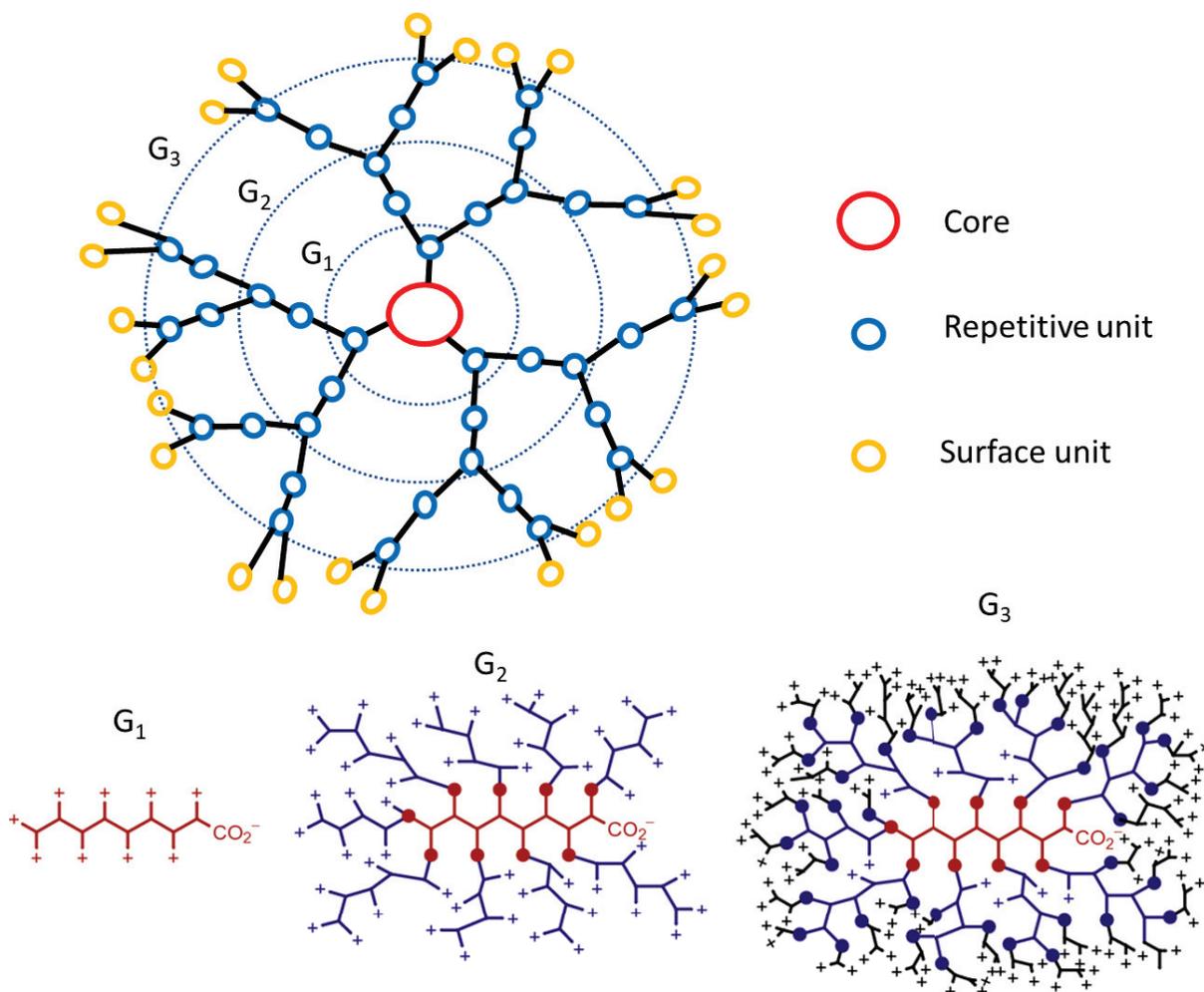


Figure 27: Structures of dendritic polymers: Top: structure of a dendrimer, bottom: structure of the dendrigrafts of lysine DGLs, positively charged at physiological pH. G_n : Dendrigraft of lysine of generation n .

DGLs are synthetic cationic dendrigrafts of lysines. Dendrigrafts have an intermediate controlled architecture, between dendrimers and hyperbranched polymers⁴⁵⁴. Their molecular weight increases quicker for each generation compared to dendrimers⁴⁵⁵. In addition, their dendrons are usually cheap and commercially available, which results in costless synthetic structures. Their synthesis are achieved by three different approaches. The divergent synthesis starts from the core to the peripheric surface. The first step consists in the synthesis of the first generation. The dendrons are grafted to the core and then they are functionalised with the precursor of the further generation⁴⁵⁶. Mistakes in the dendrimer synthesis come from uncompleted branching sequences. Purifications are difficult and are not employed at industrial scale. This plan is employed for the synthesis of the family of PAMAM dendrimers⁴⁵⁷.

The convergent approach builds the dendritic polymers from the external surface to the core. The dendrons are progressively synthesised and then grafted to the core⁴⁵⁸. By this method, the dendritic compounds are more monodisperse but less large because of steric hindrance⁴⁵⁹. Finally, the convergent and divergent synthesis combine in a third method. The dendrons are grafted on a functional core presenting the first layers⁴⁶⁰. This method is facilitated by click chemistry and particularly CuAAC. The synthetic method controls the architecture and the functionality of the dendrimeric polymers⁴⁵⁶.

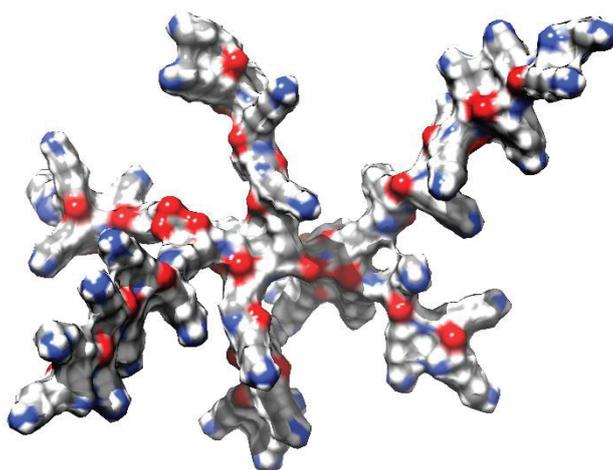


Figure 28: DGL G_2 surface representation from MD snapshot. Red negative electronic density, blue positive electronic density

DGLs are synthesised through a divergent method⁴⁶¹ (Figure 29). DGLs were reported for the first time in the late 1980's⁴⁶². The first generation is obtained by the condensation of amino acid **III-1** into *N*-carboxyanhydride (NCA) **III-2**, which led to G_1 . For the further generations, the previous G_{n-1} generation is used as a macroinitiator for the ring opening polymerisation of Lys-(TFA)-NCAs. After each polymerization step, the selective removal of the TFA ϵ -NH₂ protecting groups give new initiation sites⁴⁶³. The structure is thermodynamically controlled by the precipitation of the TFA-protected polymer and kinetically controlled by the steric hindrance. A TFA deprotection releases the G_n DGL at a multigram scale with a high-branching ratio. The generations 1 to 8 are characterised (illustration of G_2 in Figure 28).

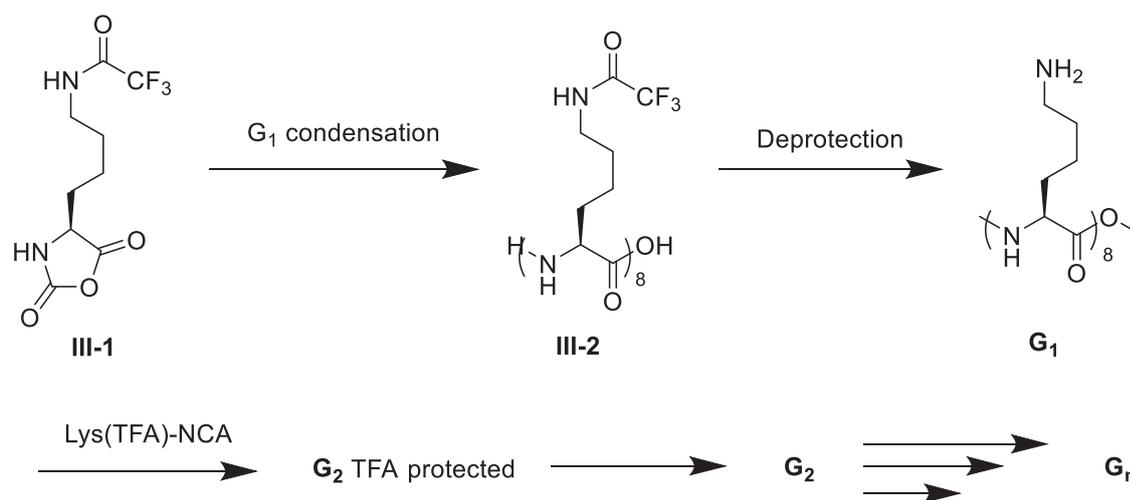


Figure 29: DGL divergent synthesis from Lys(TFA)-NCA

The DGLs G_1 to G_4 used in this study are provided by the COLCOM company³⁷. The amino surfaces are various, including lysinic, aspartic, succinamic, maleimidic, guanidinylated surfaces, with antibodies and fluo labelling. The fluorescent tag can also be labelled at the core. Their polydispersity is low. The main physical properties are presented in the next table (Table 17). The data are presented for a G_n TFA salt with a lysine surface. The protonation states and net charges were determined by potentiometric titrations⁴⁶⁴.

Generation	G_1	G_2	G_3	G_4	G_5
M	1950	10081	24480	70945	208994
Number of lysine groups	8	48	123	365	963
Protonation state at pH 7.8 (%) ¹	/	86	79	79	/
Net charge at pH 7.8 ¹	+7	+41	+92	+280	+727

Table 17 : Physical properties of DGLs with lysinic surface. ¹Accurate Mw and net charges at pH 7.4 for DGLs G_1 – G_5 were extracted from potentiometric titrations⁴⁶⁴.

Dendrimeric compounds are tunable and structured nano-objects. They have numerous functionalities and properties, among them multivalency. Multivalent binding is based on multiple and simultaneous molecular recognition processes. They can be cooperative and reinforce the bond formation compared to separate similar ligands⁴⁶⁵. Multivalency depends on the ligand architecture, particularly for biological systems⁴⁶⁶. Thanks to multivalency, dendrimeric polymers had applications in many fields, such as material sciences, electronic, catalysis and others⁴⁶⁷. For instance, the two components of this family, PAMAM and PEI, were employed in biomedical technologies⁴⁶⁸. The DGLs were mainly used for biological applications⁴⁶⁹. Nanoparticles of DGLs and γ -polyglutamic acid were efficient and biodegradable gene carrier vectors⁴⁷⁰. The DGLs toxicity was low when they were tuned with PEG⁴⁷⁰. Their native forms were nonimmunogenic^{4b}. These properties were indispensable for biological

³⁷ <http://www.colcom.eu/uk/index.php> collaborative contract with Dr. Laurent Vial.

applications. The dendrigrafts of lysine were recently applied in imaging^{409a}, biocatalysis^{409b}, biomaterials engineering^{409c,d}, drug and gene delivery^{4a,471} and finally sensing^{464a,472}. However, despite high expectations, there is currently few clinical translations of these dendrimeric systems⁴⁷³. The DGLs with lysinic surface are positively charged objects, hence interacting with negatively charged molecules such as GAGS thanks to strong electrostatic interactions. Multivalent complexes were formed between the two macromolecules, which was explored as a sensing strategy.

The current method monitoring heparin concentration is aPTT and has many drawbacks. Adapted to measure anti-thrombin III activity, aPTT are routinely used to monitor heparin levels in blood⁴⁷⁴. However, aPTT is affected by many analytical factors that are unrelated to the activity of heparin⁴⁷⁵. Variations from devices have been observed as well as the increased aPTT time due to biological factors such as inflammatory state⁴⁷⁶. Despite these limitations, aPTT is still trusted and employed to monitor heparin level, because of their availability and the familiarity with the tests in the hospital⁴⁷⁷.

An alternative based on a turn on platform was proposed by a previous PhD student of the group. DGLs quenched the fluorescence of a aspartic acid homo polymer labelled with fluorescein group D_nCF **III-3_n** (n number of aspartic acid monomer) by aggregation^{472c}. Then, the introduction of heparin displaced the reporter and restored the signal because of a higher interaction between the DGL and the GAG (Figure 30).

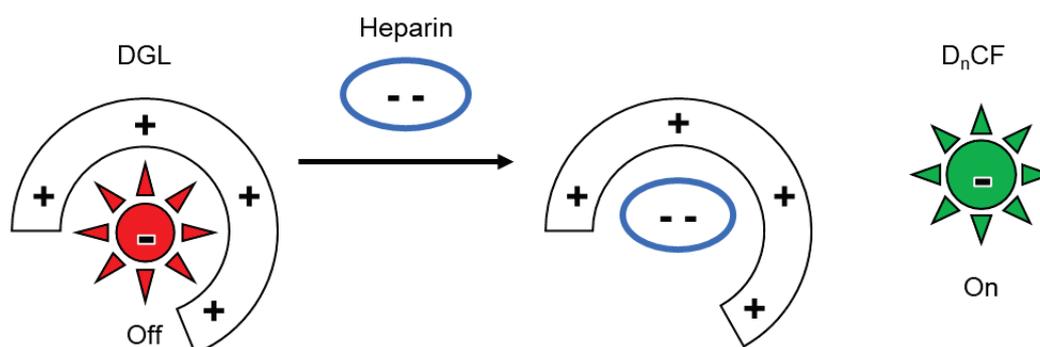


Figure 30: Turn on sensing strategy for heparin detection with DGL and D_nCF as reporter

The fluorescence of a **III-3₇** was completely quenched with a DGL of generation 4 (**G₄**). After that, it was restored by the competitive binding of heparin with **G₄**. This system was able to detect heparin in whole blood at clinical levels, with a LOD of 0.12 USP^{472c}. However, the sensor was not selective for chondroitin sulfate B, another GAG, which reflected the DGL high-binding capacity. Several variations on the size of the reporter and the DGL were possible. For instance, a **G₃** / **III-3₃** association discriminated GAGs in aqueous buffer^{472a}. UFH and chondroitin sulfate induced a positive or negative variation on the signal following the reporter loading on **G₃**. They had indeed two different negative electronic densities and hence two different interactions with **G₃**. One displaced the reporter while the other compacted the indicator on the **G₃** surface. The differentiation of mixtures was validated as well as the identification of pure GAGs thanks to a linear discriminant analysis. To sum up, DGLs are promising heparin binders.

By virtue of their architecture and multivalency, cationic DGLs are interesting dendrigrafts for heparin sensing. They could replace protamine as heparin inhibitors.

4. Exploration of DGL applications

The aim of this subchapter was to explore the possible medicinal applications of DGLs. Among them, the neutralisation of heparin and the development of a sensing platform for biomolecules were investigated.

a. Neutralisation of heparin

The aim of this subpart was to evaluate the neutralisation of heparin by DGLs. Heparins are the most negatively charged GAGs, while DGLs are large cationic macromolecules. Thus, strong electrostatic interactions reinforced by multivalency are expected between the two objects. The neutralisation of heparin by potent inhibitors are evaluated by dye displacement. The dyes are Mallard Blue⁴⁷⁸, Toluidine Blue and most frequently Azure A⁴⁷⁹. The protocol consists in the restoration of the blue absorbance upon addition of the inhibitor on the heparin/dye complex in buffered water. The neutralisers are compared by their EC_{50} , i.e. their required concentration to displace 50 % of Azure A from heparin. The heparin was UFH, which is commonly used for inhibitor benchmarking. The UFH mechanism of action is still poorly understood. This heparin has strong anticoagulant activity but also important haemorrhagic complications⁴⁸⁰. DGLs **G₁** to **G₅** were benchmarked with Azure A displacement assay^{464a,481}. The results are summed up in table³⁸ (Table 18). CE_{50} is the antidote charge excess required to displace 50 % of Azure A.

Generation	Protamine ²	G₁	G₂ ²	G₃ ²	G₄ ²	G₅ ²
Net charge at pH 7.8 ^{1, 2}	+22	+7	+41	+92	+280	+727
EC_{50} (μ M)	11.7	128	4.63	1.70	0.54	0.23
CE_{50}	0.79	2.79	0.61	0.50	0.48	0.51

*Table 18 : CE_{50} , EC_{50} data for UFH (Mw = 14.6 kDa) inhibitor from Azure A displacement assay, in PBS 1X at pH 7.4. ¹Net charges at pH 7.4 for DGLs **G₁**–**G₅** were extracted from potentiometric titrations⁴⁶⁴. ²experiment done by Dr. Francoia. The pH titration of **G₁** is available in appendix.*

A protamine concentration of 11.71 μ M displaced 50 % of the dye. Regarding **G₁**, the linear first generation, the same effect was reached with a concentration of 128 μ M (Appendix III-II: Fluorescence). **G₁** carried only 7 positive charges compared to protamine, explaining the difference for the two linear molecules. The electrostatic interaction between the inhibitor and UFH was indeed weaker. Then, the EC_{50} gradually decreased for the higher generations, from 4.63 μ M to 0.23 μ M for

³⁸ This work started with Dr. Francoia. He did the Azure A displacement assay for **G₂** to **G₅** and the anti-Xa assays. The results are presented for completeness. The UFH molecular weight of heparin is assumed as that of the sodiated analogue of the heparin repeat unit (i.e. 665.40 g.mol⁻¹). As supplied, UFH only contains ca. 30-40% of material with the active sequence of repeat units. However, the whole sample contains anionic saccharide units which can bind, even if they are in the wrong sequence. Hence the total concentration of the anionic disaccharide was reported irrespective of whether it is present in the active form of heparin or not.

G₂ to **G₅**. The CE₅₀ values described the binding ability per cationic charge. The CE₅₀ values decreased from protamine to **G_{2/5}** (0.79 to approximately 0.50). The architecture and the density of positive charge were similar from **G₂** to **G₅**^{464b}, which was coherent with the similar CE₅₀ values. The lower value compared to protamine highlighted the multivalency offered by the dendrimeric DGLs. There were no improvements in terms of CE₅₀ from **G₃** to **G₅**. Consequently, cheaper **G₂** and **G₃** were selected for a neutralisation of heparin in human plasma.

The reversal activity was characterised by antifactor Xa assay⁴⁸². The explored heparins were UFH (Mw = 14.6 kDa), enoxaparin (LMHW, Mw = 4.6 kDa) and fondaparinux (Mw = 1.7 kDa)^{464a}. Regarding UFH, DGLs were more efficient than protamine, which was in agreement with their respective CE₅₀ values. A weak paradoxical effect disqualified **G₃**, i.e. a restoration of the anti-Xa activity. Furthermore, **G₃** is more expensive than **G₂**. This DGL showed 90 % of enoxaparin neutralisation, which was higher than protamine (50 - 60 %). Regarding Fondaparinux, protamine was ineffective while **G₂** inhibited 60 % of anti-Xa activity. **G₂** showed an ability to neutralise the three heparin classes in vitro⁴⁸¹.

Besides, ITC was used to characterise the binding between **G₂** and heparins. The experiments in PBS 1 M were not adapted to detect associations between **G₂** and Fondaparinux. The buffer plays a major role in ITC experiments, particularly for this system⁴⁸³, suggesting that PBS was not the best buffer and was too concentrated. Therefore, Tris 200 mM was used. In this buffer, **G₂** showed high affinity with UFH. The association constant was in the range of 10⁷ M⁻¹ (Table 19), with only 19 and 17 % of absolute error ($\Delta K/K$, acceptance 20 %). The stoichiometry N was near 0.33, suggesting that one **G₂** could bind with 3 UFHs simultaneously. However, the experiment was only duplicated and the sigmoidal factor was too important (1100, upper limit 1000)¹³⁶. It was due to the upper limit of ITC measurement which relies in 10⁷ M⁻¹.

Measurement	1	2	average
N	0.348 +/- 0.001	0.298 +/- 0.001	0.323
K (M ⁻¹)	5.89 10 ⁷ +/- 1.10 10 ⁷	2.88 10 ⁷ +/- 4.48 10 ⁶	4.38 10 ⁷
ΔH (cal/mol)	8401 +/- 53	8392 +/- 62	8396
ΔS (cal/mol/°)	63.7	62.3	63.0

Table 19: Thermodynamic data from ITC for **G₂** (50 μ M) vs UFH (250 μ M) in Tris 200 mM pH 7.4

Only the heat of Fondaparinux dilution could be recorded in the case of the titration of **G₂** at 50 μ M with Fondaparinux (Tris 200 mM 7.4, Fondaparinux 771 μ M). The association constant was supposed to be weaker compared with UFH but given that the biological experiments showed a neutralisation of Fondaparinux, a formation of complex should be observed. The ITC experiments should be done with more concentrated samples. Yet, because of the expensive price of Fondaparinux, it was not possible to prepare a higher concentration (for instance 15 mM Fondaparinux, 1 mM DGL) in order to measure a heat of interaction. The ITC study was hence not complete and did not provide additional data on the binding event. This data could be helpful in order to understand the DGL mechanism of action. A future campaign of measurement could be done. ITC optimisations on biological objects such as heparin and macromolecules are difficult but other inhibitors of heparin were successfully studied by ITC^{417b}. In addition, the **G₂** toxicokinetic and toxicologic profile should be determined before further

clinical trials. The biocompatibility of such cationic macromolecules could be important. In vivo activity should also be evaluated in the future.

G₂ was an efficient and universal heparin reversal agent. The neutralisation was more effective compared to protamine. ITC failed to provide further insights into the binding mechanism.

b. Detection of biomolecules

Given our expertise in the use of the sensing platform composed of a DGL and a polymer of aspartic acid and fluorescein (**III-3_n**) for the detection of heparin, the aim of this subpart was to extend this sensor array to other bioobjects.

DGLs and **III-3_n** were dissolved in HEPES buffer (10 mM, pH 7.4) at a concentration corresponding to a full loading of the **III-3_n** on the DGL, with a complete extinction of the fluorescence of the carboxyfluorescein moiety. In addition, molecules were added at increasing concentrations in order to disrupt the complex between DGL and **III-3_n**, restoring the fluorescence.

Homopolymers of aspartic acid **III-3_n** were obtained thanks to a traditional Fmoc/tBu peptidic coupling strategy on a 2-chloro-trityl resin^{472a,c} (Figure 31) and then capped with a carboxyfluorescein moiety. The yields were satisfying.

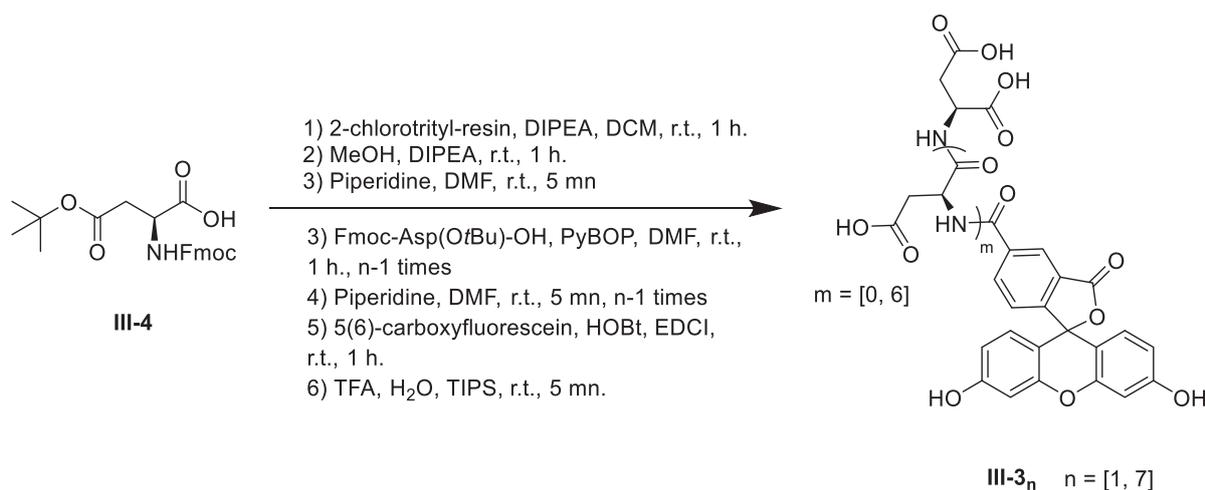


Figure 31: Iterative reporter synthesis. **III**: 66 %, **III-3₂**: 90 %, **III-3₃**: 86 %, **III-3₄**: 72 %, **III-3₅**: 56 %, **III-3₆**: 78 %, **III-3₇**: 90 %, (major 5-regioisomer of carboxyfluorescein was represented for clarity).

G₄ was arbitrarily chosen as a partner and was dissolved at 54 nM in a HEPES buffer (10 mM pH 7.4). There was a fluorescence increase upon an increasing concentration of carboxyfluorescein (0-100 μM) on **G₄**, showing that carboxyfluorescein did not form a complex with **G₄**. At the opposite, the fluorescence of the different homopolymers of aspartic acid and carboxyfluorescein **III-3_n** were efficiently quenched by **G₄** by the formation of complexes until full loading of the dendrigraft. After that, the fluorescence was restored and linearly increased, similarly to a calibration curve without **G₄** (Appendix III-II: Fluorescence). The concentration of the fluorescent polymers varied between 0 and 20 μM. The loadings of **III-3_n** were in the range of 2.0 – 5.5 μM (Table 20). The small polymers seemed

to have a higher loading (5.4 and 5.5 μM for **III-3₁** and **III-3₂**) than the longer polymers (2.1 and 2.0 μM for **III-3₆** and **III-3₇**). There was no proportionality between the size and the loading. The increase of negative charge by additional aspartic acid unit was supposed to saturate the DGL faster. However, the DGL surface topology is complex and could not be assimilated to a simple sphere. The D_nCF could induce a different folding and organisation at the surface and inside the cavities of the DGL from a size to the other one.

III-3_n	III-3₁	III-3₂	III-3₃	III-3₄	III-3₅	III-3₆	III-3₇
Loading (μM)	5.5	5.4	2.8	3.9	5.1	2.1	2.0

Table 20: III-3_n concentration required for G₄ saturation at 54 nM

A disruption of the complex between **G₄** and **III-3_n** by an additional molecule will displace the dye and can be used to find the concentration of this molecule thanks to a calibration curve. The dye displacement assay consisted in the preparation of a solution at the loading point. Then, the object of interest was introduced in the **G₄/ III-3_n** solution and the fluorescence was recorded. The aim was to detect fluorescence for the lowest concentration of the target, resulting in a limit of detection (LOD). The starting concentration was 54 μM of **G₄**. If required, it could be increased or decreased as long as the complex was formed and the fluorescence quenched. The choice of **III-3_n** was dictated by the desired **G₄/ III-3_n** association. A lower size generally eased the dye displacement, but the complex could be dissociated for lower starting concentrations of **G₄**. Different objects of interest were investigated (Table 21). The data are available in appendices (Appendix III-II: Fluorescence).

Firstly, spermine restored the fluorescence of the previous **III-3₄/G₄** system at 1 mM in HEPES buffer (10 mM, pH 7.4) (entry 1, table 5). Nevertheless, the restoration could not be performed at biological concentration (around 50 nM⁴⁸⁴) with **G₄** at 54 nM. Spermine was tetracationic at physiological pH and competed with **G₄** for binding with the anionic polymer of aspartic acid and fluorescein, restoring the fluorescence of the reporter initially aggregated on the DGL. A more diluted concentration of **G₄** was used (5.40 and 0.540 nM) in order to facilitate the disruption of the complex **G₄/III-3₂** and to lower the LOD. Nevertheless, the complex was completely dissociated at these concentrations.

The direct detection of deamidated therapeutic proteins was also explored (entries 5 – 10, table 5, see chapter 2 for the context). A restoration of the fluorescence was not observed for 2 antibodies **F_{ab}** and **Ig₁** and their deamidated analogues with a concentration of **G₄** at 54 nM and **III-3₄**. Consequently, **III-3₁** was chosen because this small polymer was more easily removed from the surface of **G₄** compared to **III-3₄**. However, no restoration of fluorescence was detected with concentration of therapeutic proteins up to 100 mM. The samples **F_{ab,de}** and **Ig_{1,de}** were deamidated at 5 % of their glutamine and asparagine residues, which poorly changed the sample electronic properties. This deamidation level could not provide enough anionic charges to these antibodies compared to their native counterparts to become a competitor of **III-3_n** for **G₄**.

Various assays of detection were carried with BL21(DE3) model³⁹, a no-pathogen and no-competent Escheria Coli. Detections were effective, but the lowest concentration of bacteria inducing fluorescence restoration was only 10⁵ bacteria/mL, which was not suitable for applications. Starting from a complex of **G₄** at 54 nM and **III-3₁**, the weakest binder from the **III-3_n** family, the bacteria were detected at a concentration of 10⁶ bacteria/mL. Decreasing the starting concentration of complex at 5.40 nM completely dissociated the complex, while increasing its concentration to 2.70 μM did not lower the LOD (10⁷ bacteria/mL). The use of strongest binder of **G₄** such as **III-3₄** and **III-3₇** prevented from a disruption of the complex upon increasing concentration of bacteria.

Entry	[III-3 _n]	[G ₄]	Object	Concentration range	Observations		
1	III-3₄ 3.86 μM	54.0 nM	Spermine	0 – 100 mM	Detection 1 mM		
2	III-3₄ 386 nM	5.40 nM		0 – 1.45 μM	Complex dissociation		
3	III-3₄ 38,6 nM	0.540 nM		0 – 100 mM			
4	III-3₄ 3,86 nM	540 pM					
5	III-3₁ 5.50 μM	54.0 nM	F _{ab}	0 – 100 μM		No displacement	
6			F _{ab,dea}				
7			Ig ₁				
8			Ig _{1,dea}				
9	III-3₄ 3.86 μM		F _{ab}				
10			F _{ab,dea}				
11	III-3₁ 4.22 μM	54.0 nM	BL21(DE3)		10 ³ / 10 ⁸ (bacteria / mL)		Detection 10 ⁶
12	III-3₁ 211 μM	2.70 μM					Detection 10 ⁷
13	III-3₁ 422 nM	5.40 nM		Complex dissociation			
14	III-3₁ 42.2 nM	0.540 nM		Complex dissociation			
15	III-3₄ 12.9 μM	54.0 nM		No displacement			
16	III-3₄ 1.29 μM	5.40 nM		Complex dissociation			
17	III-3₇ 8.59 μM	54.0 nM		No displacement			
18	III-3₇ 859 nM	5.40 nM		No displacement			

³⁹ <https://www.uniprot.org/proteomes/UP000001509>, 31/07/19

Table 21: Turn on sensing strategy involving **III-3_n** anionic reporter and cationic **G₄** with various objects of interest in HEPES buffer (10 mM, pH 7.4)

DGLs were known to have affinity for Gram-negative bacteria cells. A study by capillary electrophoresis reported an association constant of $1.1 \cdot 10^6$ for **G₄**⁴⁸⁵. The bacteria membranes presented negative charges on their surface, interacting with the positively charged DGLs, explaining the disruption of the complex **G₄/ III-3_n**. Another strategy for the detection of bacteria was applied. **G₂CF** (**G₂** with a terminal carboxyfluorescein) was known to pass through the membrane of a bacteria by phagocytosis⁴⁸⁵. **G₂CF** was commercially available. A calibration curve was determined and is available in appendix (Appendix III-II: Fluorescence). BL21(DLE) were incubated with a large excess of **G₂CF**, then centrifuged and washed to remove all the remaining **G₂CF** in excess. After lysis by ultrasound, fluorescence was measured and proved that **G₂CF** crossed the membrane and accumulated in the bacteria. Indeed, fluorescence was recorder from the mixture resulting from the lysis of the bacteria which was far higher than the lysis of bacteria without incubation with **G₂CF**. This system could be optimised and further used for biomedical applications.

DGLs proved their ability to be a part of sensing systems based on a turn on strategy or by direct detection with **G₂CF**. Nevertheless, no concrete applications could be currently envisioned before improvement of the LODs. The described sensing array for GAGs^{472a,c} could be optimised and proposed for a clinical translation. Indeed, operating systems in pure blood are rare. In addition, nanoparticles could be advantageously combined with DGLs in a new array strategy to reach sub nanomolar LOD and important linear range.

*DGLs effectively quenched the fluorescence of **III-3_n** by aggregation. The turn on sensing system was applied successfully for spermine and bacteria.*

c. **G₃** acting like an hydrolase

The aim of this subpart was to evaluate the ability of DGLs to perform enzymatic-like reactions. Hydrolases are enzymes that are responsible for hydrolysis reactions. Among them, peptidases and esterases cleave ester and amide bonds. They are usually found in cellular lysosomes⁴⁸⁶. Hydrolases are now used as catalysts for organic processes under mild conditions. They showed high stability, efficiency, and availability for a variety of bioprocesses⁴⁸⁷. Hydrolases are usually activated by cations. DGLs are polycationic molecules and can thus be hydrolase activators. In addition, DGLs whose size is superior to 2 exhibit free amines, which are responsible for the nucleophilic attack of the hydrolysis. Consequently, the DGLs could be both activators and hydrolases. **G₃** was the cheapest DGL bearing free amines and was dissolved with 8-octanoyloxy pyrene-1,3,6-trisulfonic acid **III-5** at 15 μ M in a HEPES buffer (100 mM pH 7.4).

G₃ immediately hydrolysed the ester bond of **III-5**, enhancing the fluorescence of the pyrene moiety (Figure 32, Figure 33). The graph was presented in free amine concentration (143 per **G₃**). 2 μ M of **G₃** restored half of the fluorescence of the pyrene moiety. Lysine amide NH₂ was assessed as a blank. The fluorescence was quenched for all lysine concentrations. Therefore, lysine amide did not hydrolyse the

ester bond of **III-5** at the same concentration of free lysine compared to **G₃**. That proved the influence of the DGL architecture in the hydrolysis of ester bond.

The literature reported that the enzyme substrate **III-5** was used for lipases and esterases assays⁴⁸⁸. Thus, **G₃** could be later investigated as a lipase and esterase. Lipase-catalysed reactions in organic solvents were widely studied and show interesting selectivity⁴⁸⁹. Esterase-catalysed reactions had also a variety of applications⁴⁹⁰.

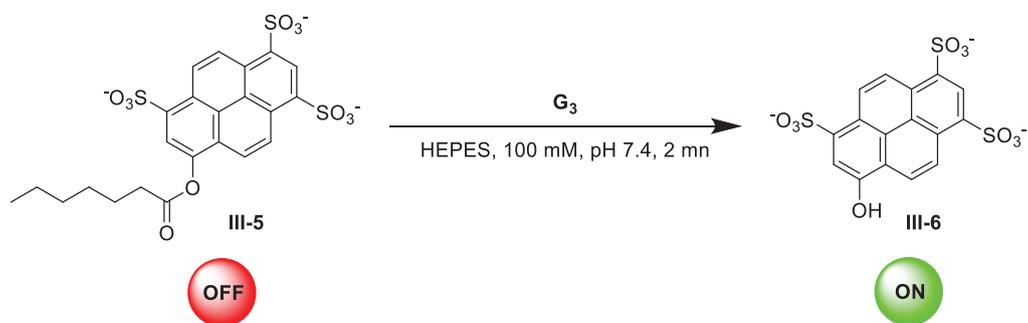


Figure 32: **G₃** hydrolysis of **III-5** in HEPES buffer

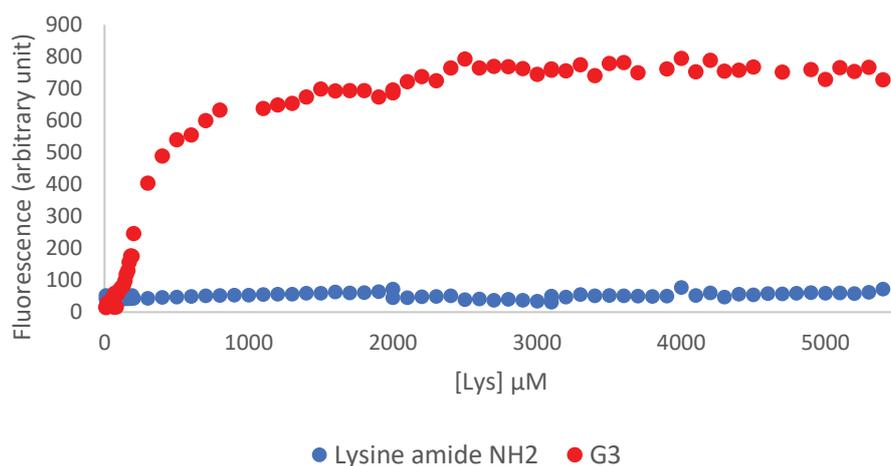


Figure 33: Hydrolysis of **III-5** (15 μM) by **G₃**, HEPES buffer 100 mM pH 7.4, 2 mn. Represented by effective lysine concentration

G₃ was successfully assessed as lipase and esterase, paving the way for the use of DGL for esterase/lipase-catalysed reactions in organic chemistry.

e. Conclusion

The dendrigrafts of lysine had a positive surface in physiological conditions and a particular shape, including a ramified architectures and internal cavities. These unique features led us to use DGLs in association with different objects that presents anionic charge in water. DGL quenched the fluorescence of anionic homopolymers of aspartic acid and carboxyfluorescein upon agregation. This fluorescence was turned on by the disruption of the complex following the introduction of spermine

and Gram-negative bacteria with a limit of detection. These limits are, however, not sufficiently low to enable concrete applications. Improvements and optimisations could be done by changing the reporter and by using nanoparticles. In addition, DGL showed tremendous neutralisation of heparin's anticoagulant properties with a superior effect than protamine. The molecule should be now tested on animal models and its toxicity should be evaluated. Finally, **G₂** mimicked the enzymatic activity of esterases and could be explored for catalysis.

5. Understanding the DGL – protamine / heparin binding by molecular dynamic simulations

The aim of this subchapter was to obtain structural and energetic information explaining the superior neutralisation activity of **G₂** for heparin compared to protamine, by means of computational chemistry. Molecular dynamic simulations were successfully applied in the literature to study the interaction between heparins and a potential inhibitor⁴⁹¹, as well for the study of the DGLs topology^{464b}. Consequently, it was particularly adapted for our system involving a DGL **G₂**, protamine and heparins of various sizes. This work was a collaboration with Professor Gerald Monard from Lorraine University⁴⁰. The computational resources were provided by the mesocentre EXPLOR from Lorraine University.

⁴⁰ Most of the simulations were prepared and run by Pr. Monard, analysis was done in tandem.

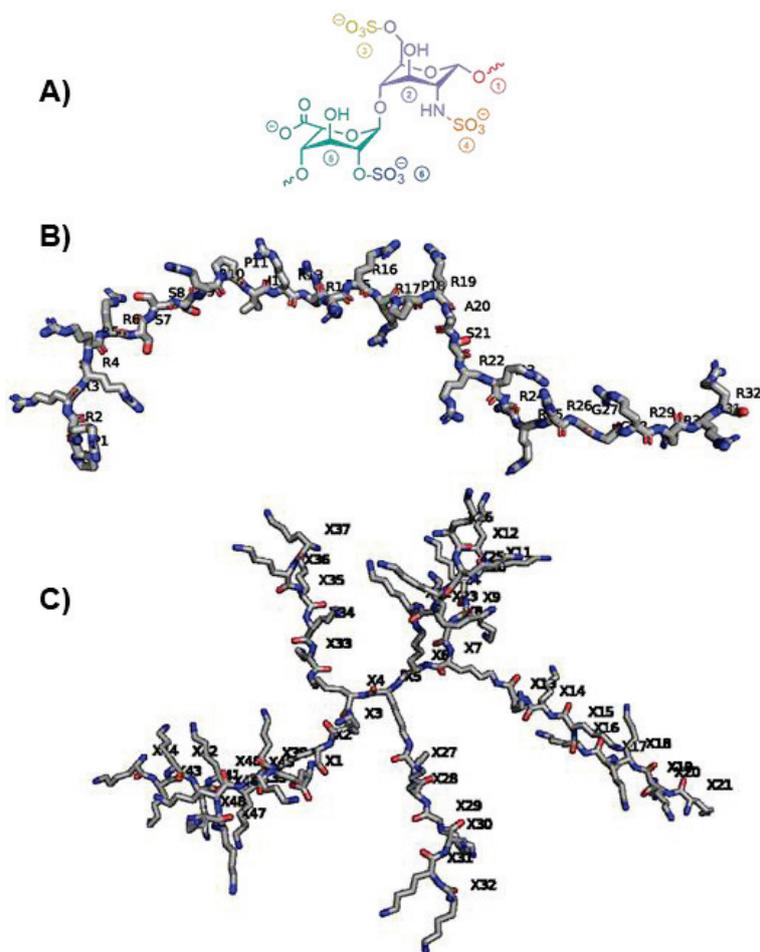


Figure 34 : Representation of some partners involved in the molecular dynamic simulations : A) repetition of the disaccharide units found in Fondaparinux and 2 medium-weight ideal heparins, named **dp14/Ex** and **dp44/UFH**. The labelling of the repetition unit is represented. B) Protamine **P₁** in sticks (Chimera software¹¹⁸) with one letter code labelling. C) DGL **G₂-1** in sticks (Chimera software¹¹⁸) with one letter code labelling (X because of the lysine residues were built manually see experimental section) (grey carbon, red oxygen, blue nitrogen)

8 different topologies of **G₂G₂-1/8**^{464b} were built in order to reproduce the dispersity of the dendrimer (Figure 34). 4 known protamines⁴¹⁶ structures **P_{1/4}** were selected. Their primary structures are developed in the experimental section (Molecular dynamic). The simulated heparins were **Fx** fondaparinux, and 2 medium-weight ideal heparins, named **dp14/Ex** (representative of real enoxaparin) and **dp44/UFH** (representative of real UFH). The sequences were built by the repetition of the two major disaccharides units (L-iduronic acid and D-glucosamine) found in active heparin (respectively 7 and 22 repetitions)⁴⁹².

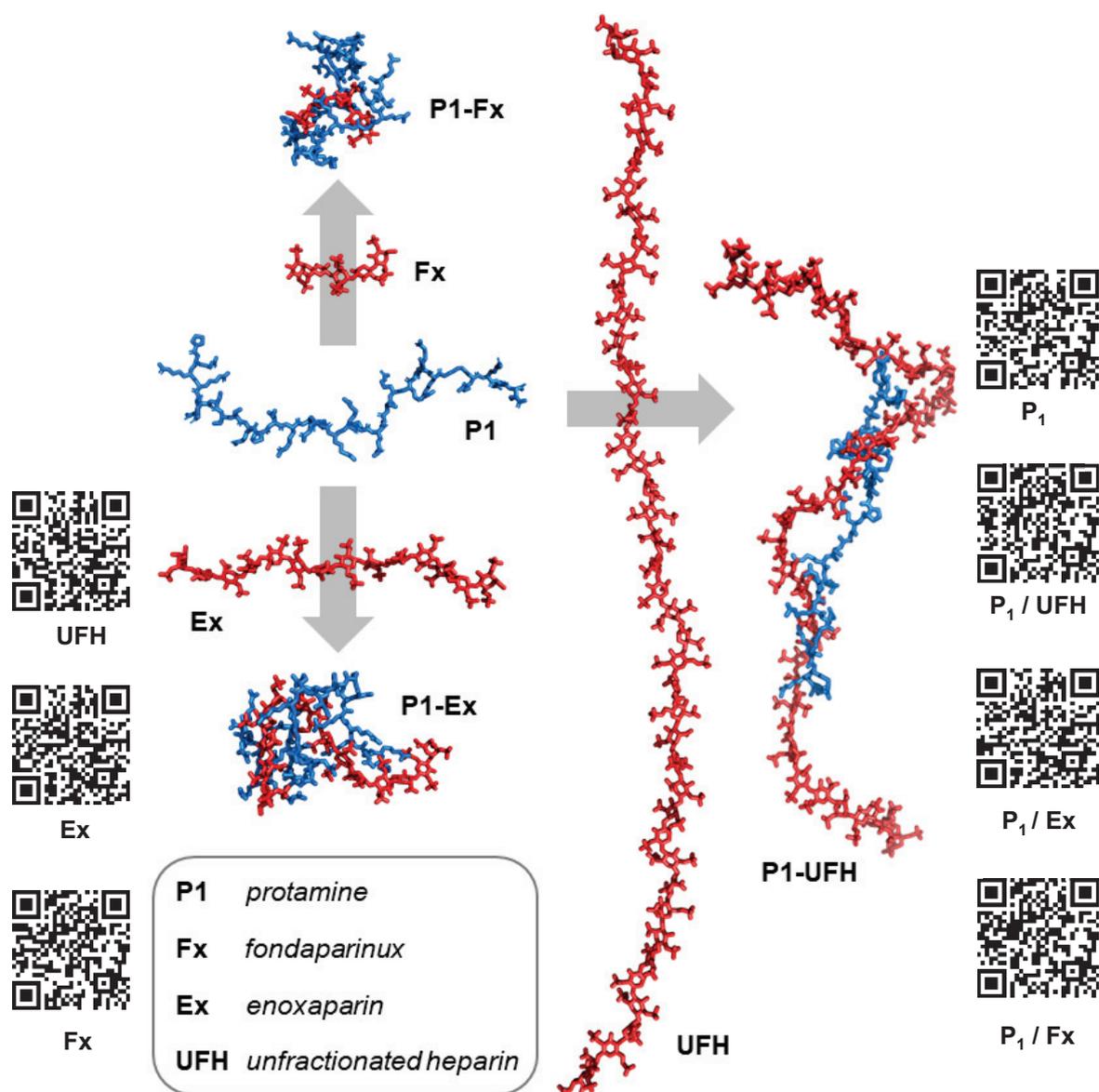
MD simulations on the stoichiometric complexes between the 8 **G₂** / 4 protamine **P₁₋₄** with **Fx**, **Ex** and **UFH** respectively were performed during 1 μ s in explicit TIP3P water²²⁰ thanks to the Amber18 package^{95b}. The computational details are given in the experimental section (Molecular dynamic). Each protamine **P_{1/4}** and DGL **G₂-1/8** displays a unique sequence. The trajectories of each sequence of protamine and DGL respectively showed similar behaviours. Therefore, the presented analysis involves

P₁ and **G₂₋₁**. The statistics are done over all the sequences. Cpptraj post-processing package was used^{95a}.

a. Study of the binding of protamine with heparin

As introduced, the evolution of complexes between 4 protamines **P_{1/4}** and 4 heparins **Fx**, **Ex**, and **UFH** were simulated. Representative snapshots at the end of the simulations are presented below (Scheme 250).

The formation of the complexes and their stability were evaluated thanks to the evolution of the contact surfaces between heparin and its antidote. The complexes were formed quickly (20 ns) and then the surface contacts remained unchanged, showing that the complexes were stable. Besides, the percentage of heparin's surface in contact with protamine decreased with the length of heparin, while the percentage of protamine's surface in contact with heparin increased with the length of heparin. It suggested a complementary of shape between **P₁** and **Ex**, which could be observed on the snapshots on the complexes (Scheme 250).



Scheme 250 : Representative snapshots taken at the end of the MD trajectories for protamine P_1 , and its stoichiometric complexes with fondaparinux F_x , enoxaparin E_x , and unfractionated heparin UFH . For the full set of computed complexes, see the appendix (Appendix III-V: Supporting information regarding the MD simulations of protamines with heparins)

The 2D RMS deviation maps of P_1 alone and within its complexes with heparins also illustrated the rigidification of the protamine's backbone upon binding and the stability of the complexes (Figure 36). Furthermore, middle size heparin E_x induced a higher stiffening of protamine's backbone compared to F_x and UFH . The complex of P_1 and E_x had the better complementarity of size and showed the greater stiffening upon binding. Therefore, the complex should involve optimisation contacts and interactions compared to the others with shorter F_x and longer UFH .

As a complement, average gyration radii of heparins were calculated to illustrate the antidote's ability to distort heparins (*Figure 37*). Average gyration radius was the evolution of the average distance within the molecule to its mass centre during the simulations. The radii indicated that shorter **Fx** was not deformed upon binding, at the opposite of longer **UFH**. **Ex** was also deformed by the association with protamine but to a lesser extent than **UFH**. That was coherent with the snapshots from the simulations (*Scheme 250*) and should be associated with an increasing entropic penalty. This penalty should be counterbalanced by the more numerous electrostatic interactions between the partners as the heparin length increased.

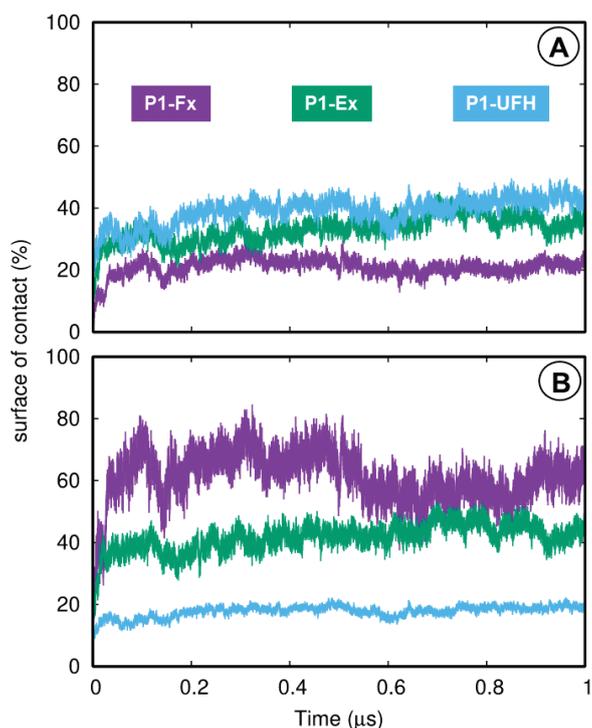
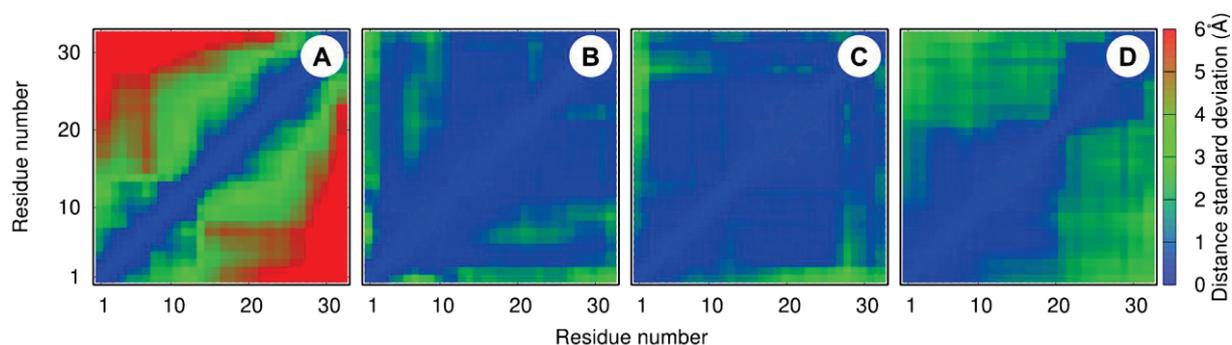


Figure 35: percentage of protamine P_1 's surface (A) or heparin's surface (B) in contact with its partner within their complexes. For the full set of computed complexes, see the appendix (Appendix III-V: Supporting information regarding the MD simulations of protamines with heparins)



*Figure 36: 2D RMS deviation maps of protamine P_1 alone (A), and within its complexes with **Fx** (B), **Ex** (C), and **UFH** (D). For the full set of computed complexes, see the appendix (Appendix III-V: Supporting information regarding the MD simulations of protamines with heparins)*

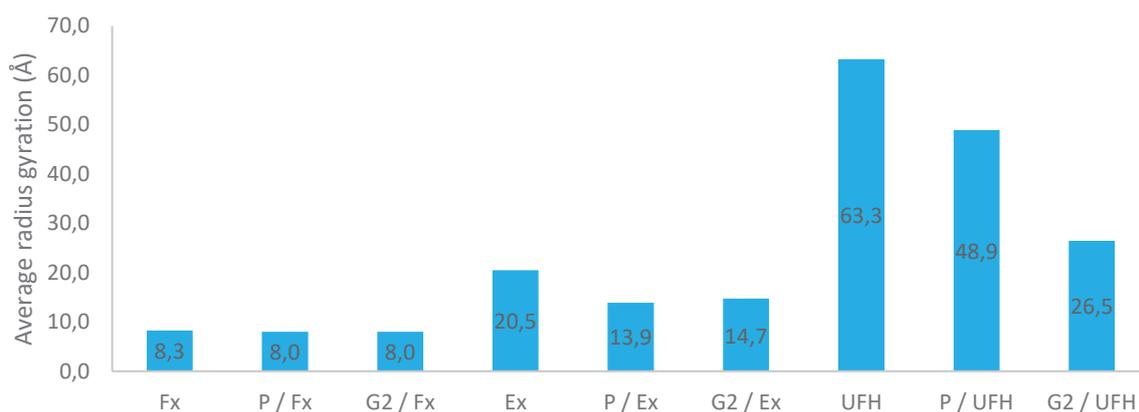


Figure 37: Average gyration radii of heparin alone and within its complexes with protamine.

The formation of the complexes resulted from the electrostatic interactions between the protamine's positive charge and the heparin's anionic charge (in which at least one pair of heavy atoms is within a hydrogen bonding distance (i.e., a salt bridge), or between a hydrogen atom from either protamine or heparin and any heavy atom from its partner (i.e., a hydrogen bond)). The average number per frame of salt bridges (N_E) and hydrogen bonds (N_H) engaged by either heparin (N_{hep}) or protamine (N_{pro}) in the complexes (from the simulations) revealed that salts bridges are not surprisingly the main contributors to the binding events for heparin of all sizes, accounting for 91 % to 96 % of the number of contacts between partners (Figure 38). Interestingly, the total number of contacts ($N_{\text{tot,hep/prot}} = N_{E,\text{hep/prot}} + N_{H,\text{hep/prot}}$) was the same regarding the complexes of protamine with **UFH** ($N_{E,\text{hep}} + N_{\text{tot,hep}} = 17.9$, $N_{\text{tot,prot}} = 19.5$) and **Ex** ($N_{\text{tot,hep}} = 16.9$, $N_{\text{tot,prot}} = 17.9$). The total number of contacts was significantly lower for complexes of protamine with **Fx** ($N_{\text{tot,hep}} = 10.4$, $N_{\text{tot,prot}} = 8.6$). Based on these values, protamine should form complexes: 1) as effectively with **UFH** as **Ex**, neutralising their anticoagulant activity similarly; 2) less effectively with **Fx** than with **UFH** and **Ex**, neutralising its anticoagulant activity less efficiently. However, experimentally, protamine fully neutralises **UFH**, partially **Ex** and is inactive against **Fx**. Therefore, the total number of contacts could not account directly for the ability of protamine to inhibit **UFH** and **Ex**'s anticoagulant activity.

MMPBSA method^{93a,221} provided the free energies of association ΔG_{bind} between heparin and protamine (Figure 39). ΔG_{bind} increased with the heparin length in agreement with the experimental ability of protamine to reverse heparin's anticoagulant properties of each size. The difference of tendencies between the total number of contacts N_{tot} and the free energies of association ΔG_{bind} suggested that additional non-contact interactions might be responsible for the better affinity of protamine for **UFH** compared to **Ex**.

A linear interaction energy (LIE) analysis (Figure 38) provided the average number of favourable electrostatic and Van der Waals interactions per frame engaged by either heparin ($N_{\text{LIE,hep}}$) or protamine ($N_{\text{LIE,prot}}$) within the complexes. The number of favourable interactions engaged by the protamine in the complexes with **UFH** ($N_{\text{LIE,prot}} = 20.0$) and **Ex** ($N_{\text{LIE,prot}} = 19.3$) were similar and lower regarding the complex with **Fx** ($N_{\text{LIE,prot}} = 14.0$). It did not explain the lower free energy of association

ΔG_{bind} of protamine for **UFH** compared to **Ex**. However, the number of favourable interactions engaged by heparin in the complexes with protamine decreased with the heparin length (**UFH**: $N_{\text{LIE,hep}} = 34.3$, **Ex**: $N_{\text{LIE,hep}} = 26.4$, **Fx**: $N_{\text{LIE,hep}} = 11.4$). This indicated that some residues of the longer **UFH**, which are not in direct contact with protamine, provided additional non-contact interactions $N_{\text{non-contact}}$ ($N_{\text{non-contact,hep}} = N_{\text{LIE,hep}} - N_{\text{H,hep}} - N_{\text{E,hep}}$, **UFH**: 7.7 ± 1.5 ; **Ex**, **Fx**: 0.6 ± 0.2) stabilising the complexes. Non-contact interactions therefore represented half, one quarter and one twentieth of all interactions engaged in complexes with protamine by **UFH**, **Ex**, and **Fx**, respectively. In comparison, the number of non-contact interactions provided by protamine are negligible in both complexes with **UFH** and **Ex**, and low in complex with **Fx** ($N_{\text{non-contact,prot}} = N_{\text{LIE,prot}} - N_{\text{H,prot}} - N_{\text{E,prot}}$; **UFH**: 0.5 ± 0.2 ; **Ex**: 1.4 ± 0.4 , **Fx**: 5.0 ± 1.0).

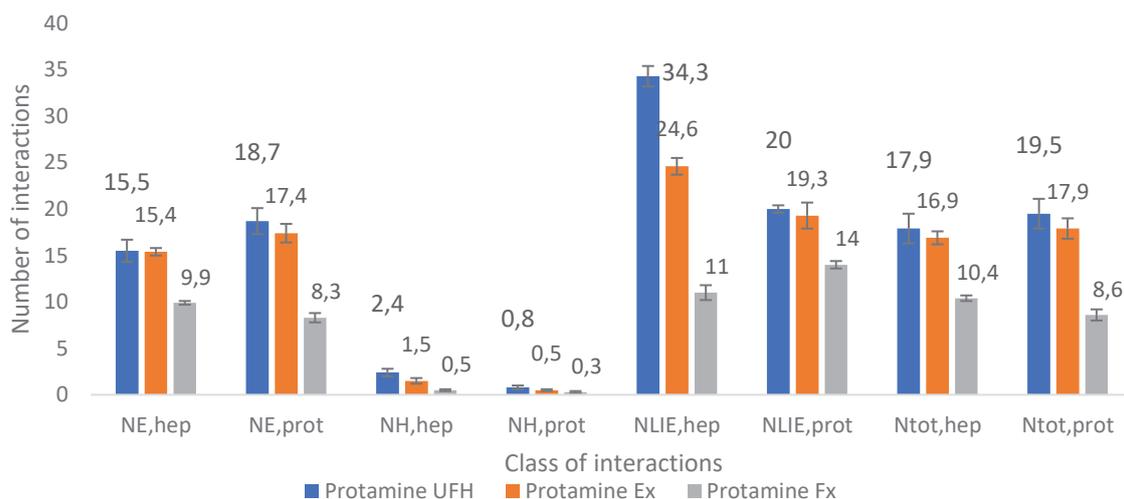


Figure 38 : Average number of salt bridges per frame engaged by heparin ($N_{\text{E,hep}}$) or protamine ($N_{\text{E,prot}}$), average number of hydrogen bonds (other than salt bridges) per frame engaged by heparin ($N_{\text{H,hep}}$) or protamine ($N_{\text{H,prot}}$), average number of favourable interactions per frame engaged by heparin ($N_{\text{LIE,hep}}$) or protamine ($N_{\text{LIE,prot}}$), total number of contacts engaged per frame by heparin ($N_{\text{tot,hep}}$) or protamine ($N_{\text{tot,prot}}$). For the full set of computed complexes, see the appendix (Appendix III-V: Supporting information regarding the MD simulations of protamines with heparins).

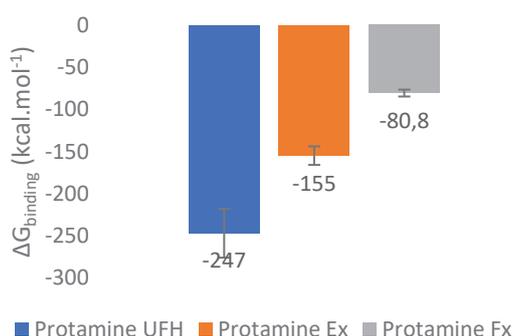


Figure 39 : Free energy of association between heparin and protamine ($\Delta G_{\text{binding}}$) in kcal.mol⁻¹. For the full set of computed complexes, see the appendix (Appendix III-V: Supporting information regarding the MD simulations of protamines with heparins)

The contribution of direct contacts to the overall association was investigated. The average binding energies between partners' residues and the percentage of simulation time during which direct contacts were present were plotted in 2D maps (Figure 40).

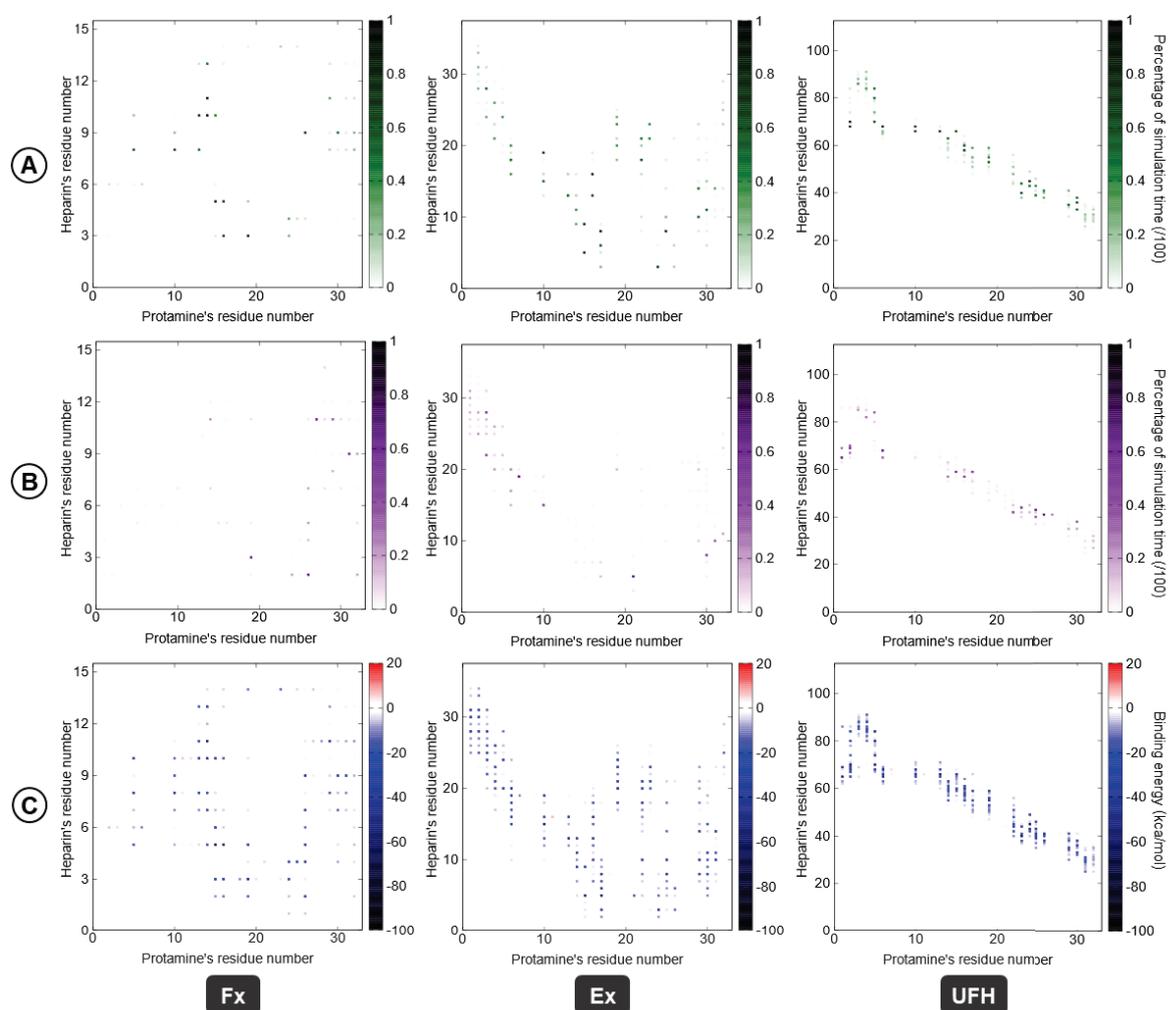


Figure 40: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine **P₁** and **UFH** (1), **Ex** (2), or **Fx** (3). For the full set of computed complexes, see the appendix (Appendix III-V: Supporting information regarding the MD simulations of protamines with heparins).

The contacts involved all the residues that are engaged in a H-bond. The contacts were decomposed into A) the charge/charge interactions, i.e. the residues that are engaged in salt bridges and B) the other H-bonds which were not salt bridges. The charge/charge interactions (A) were predominant and contributed mainly to the overall electrostatic interactions (C) between the three heparins and protamine. The electrostatic interaction maps (C) corresponded to the sum of the contacts maps (A) and (B), which was logical. However, the sums did not account for all the spots of electrostatic interactions (C), especially for **UFH**. Thus, long-range electrostatic non-contact interactions significantly stabilised the complexes of heparin with protamine and contributed to the selectivity of

protamine for different heparin size. That was coherent with the previous analysis of the number of non-contact interactions $N_{\text{non-contact}}$. We are currently developing a decomposition of the electrostatic interaction maps (C) into two contributions to the overall electrostatic interactions: 1) the contribution of direct contacts and 2) the contribution of the long-range non-contact electrostatic interactions. These maps will provide the exact location of the sites that contributed to the long-range stabilisation of the complexes

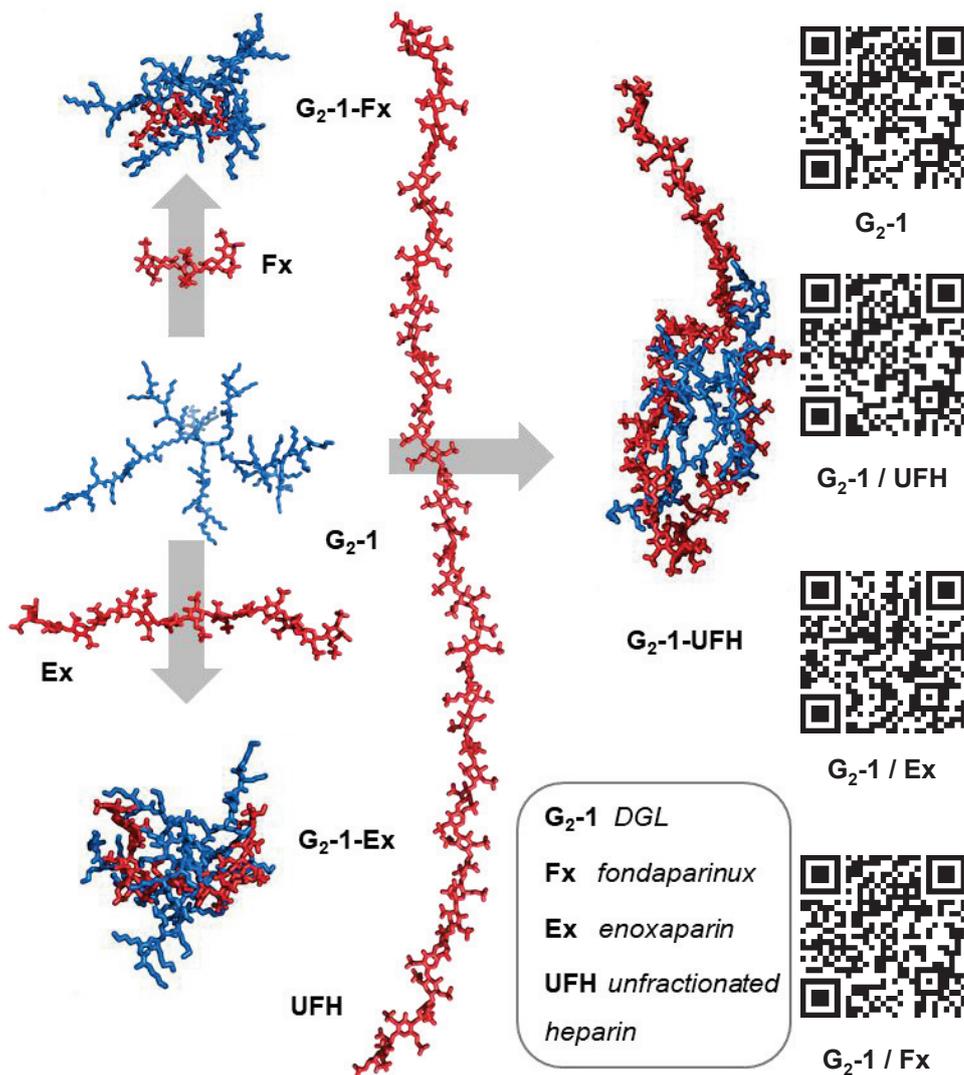
In a nutshell, the simulations highlighted the importance of non-contact interactions in the selectivity of protamine for different sizes of heparin. The number of electrostatic interactions explained the higher affinity of protamine for **Ex** and **UFH** compared to **Fx**, while the higher affinity for **UFH** compared to **Ex** came from the non-contact interactions.

b. Study of the binding of DGLs with heparin and comparison with protamine

As introduced, the evolution of complexes between 4 protamines $P_{1/4}$ and 4 heparins **Fx**, **Ex**, and **UFH** were simulated. Representative snapshots at the end of the simulations are presented below (Scheme 251).

The evolution of surface contacts and the 2D RMS maps are not yet available for these simulations, compared to the protamine study. The trajectories revealed that the oligosaccharides **Fx** and **Ex** were completely embraced by the DGL but had a great mobility in the complexes. **UFH** was less mobile and formed a more stiffened complex compared to the complex of **Ex**. RMS values would quantitatively support this analysis. As an alternative, average gyration radii of heparins were calculated to illustrate the DGL's ability to distort heparins (Figure 41). The radii indicated that shorter **Fx** was not deformed upon binding, similarly than in its complex with protamine. In addition, **Ex** was slightly less deformed upon binding with **G₂** compared to protamine. Longer **UFH** was highly deformed in its complex, more than with protamine. That was coherent with the snapshots from the simulations (Scheme 251) and should be associated with an increasing entropic penalty. This penalty should be counterbalanced by the more numerous electrostatic interactions between the partners as the heparin length increased. The higher deformation of **UFH** with **G₂** could be explained by the long stick shape of protamine compared to the spherical shape of DGL that forced the heparin to twist.

The formation of the complexes resulted from the electrostatic interactions between the **G₂**'s positive charges and the heparin's anionic charges (in which at least one pair of heavy atoms is within a hydrogen bonding distance (i.e., a salt bridge), or between a hydrogen atom from either protamine or heparin and any heavy atom from its partner (i.e., a hydrogen bond)). The average number per frame of salt bridges (N_E) and hydrogen bonds (N_H) engaged by either heparin (N_{hep}) or protamine (N_{pro}) in the complexes (from the simulations) revealed that salt bridges are not the main contributors to the binding events for heparin of all sizes, accounting for 76 % to 89 % of the number of contacts between partners (Figure 42). This tendency was farther pronounced than in the case of protamine.



Scheme 251 : Representative snapshots taken at the end of the MD trajectories for protamine P₁, and its stoichiometric complexes with fondaparinux Ff, enoxaparin Ex, and unfractionated heparin UFH. For the full set of computed complexes, see the appendix (Appendix III-VI: Supporting information regarding the MD simulations of DGLs with heparins)

The total number of contacts ($N_{\text{tot,hep/prot}} = N_{\text{E,hep/prot}} + N_{\text{H,hep/prot}}$) decreased progressively with the heparin length. Based on these values, **G₂** should form complexes less effectively from **UFH** to **Ff**, neutralising their anticoagulant activity less and less efficiently. Experimentally, **UFH** was completely reversed by **G₂**, **Ex** was neutralised at 90 % and **Ff** was neutralised at 60 % of its activity in human plasma. Therefore, the total number of contacts could account directly for the ability of **G₂** to inhibit heparin's anticoagulant property.

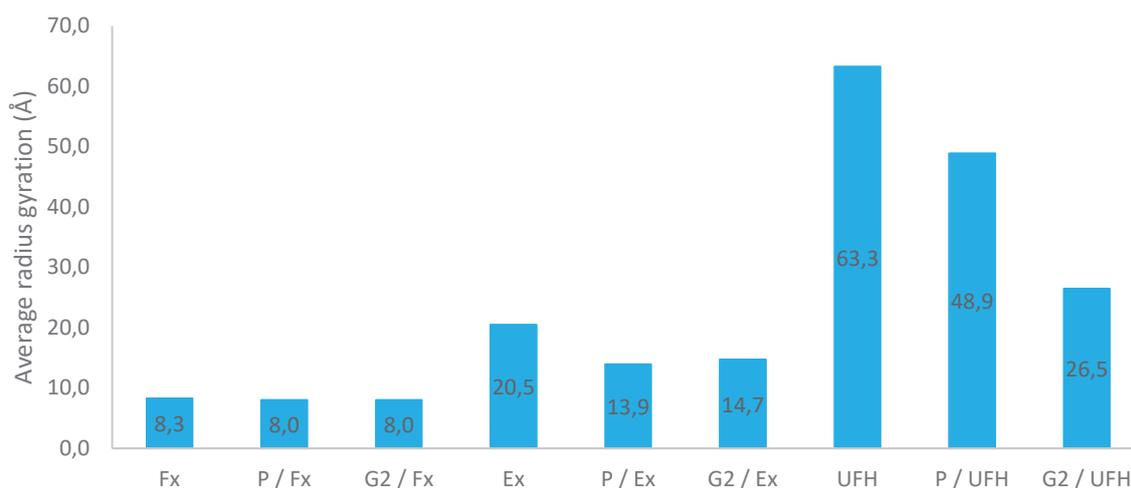


Figure 41: Average gyration radii of heparin alone and within its complexes with G_2 .

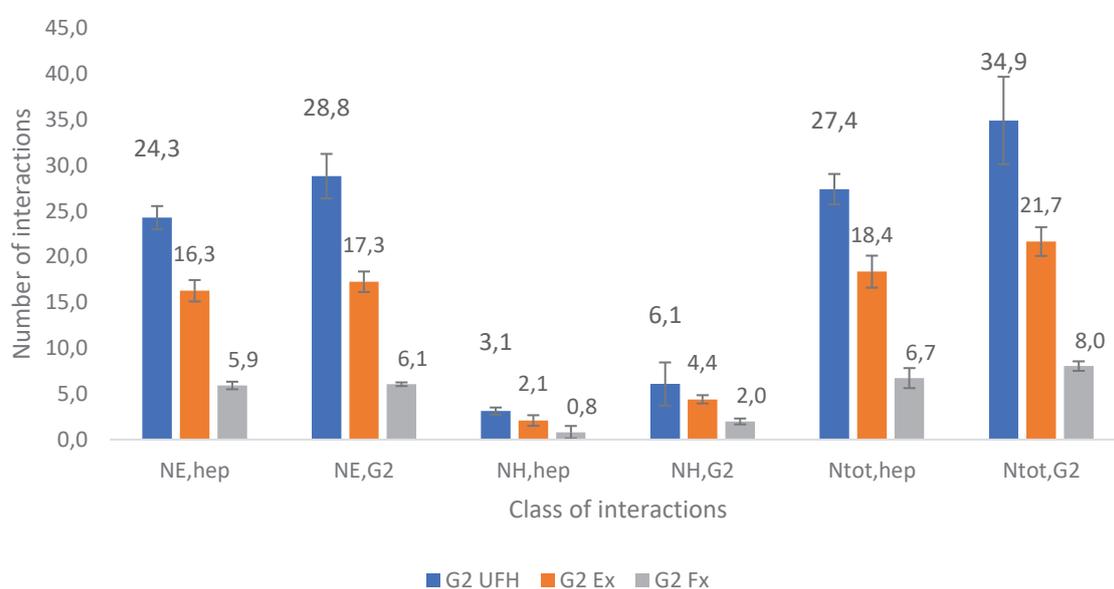


Figure 42: Average number of salt bridges per frame engaged by heparin ($N_{E,hep}$) or G_2 ($N_{E,G2}$), average number of hydrogen bonds (other than salt bridges) per frame engaged by heparin ($N_{H,hep}$) or G_2 ($N_{H,G2}$), total number of contacts engaged per frame by heparin ($N_{tot,hep}$) or G_2 ($N_{tot,G2}$).

MMPBSA method^{93a,221} provided the free energies of association ΔG_{bind} between heparin and protamine (Figure 43). ΔG_{bind} increased with the heparin length in agreement with the experimental ability of G_2 to neutralise the anticoagulant property of **Fx**, **Ex** and **UFH**. The ΔG_{bind} fitted nicely with the total number of contacts, even though we previously learned that non-contact interactions are also an important contributor to the selectivity. To date, we have not yet collected the average number of favourable interactions per frame engaged by heparin ($N_{LIE,hep}$) or G_2 ($N_{LIE,G2}$) to study the influence of non-contact interactions in the selectivity.

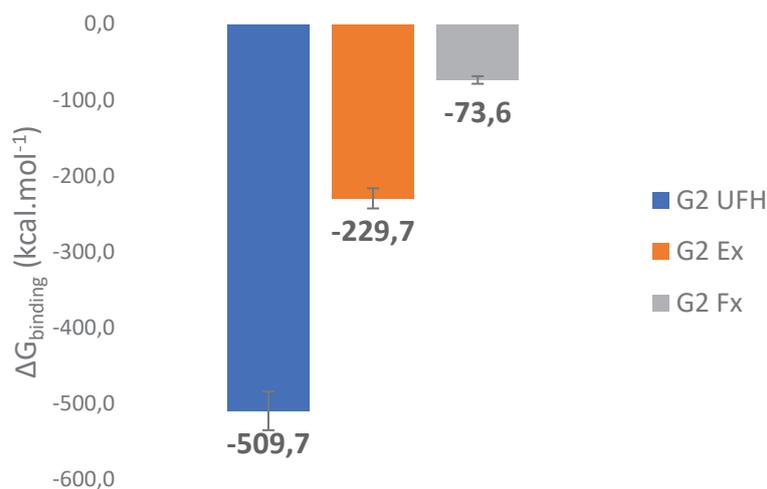


Figure 43 : Free energy of association between heparin and G_2 ($\Delta G_{\text{binding}}$) in kcal.mol⁻¹.

The contribution of direct contacts to the overall association was investigated. The average binding energies between partners' residues and the percentage of simulation time during which direct contacts were present were plotted in 2D maps (Figure 44). The contacts involved all the residues that are engaged in H-bonds. The contacts were decomposed into A) the charge/charge interactions, i.e. the residues that are engaged in salt bridges and B) the other H-bonds which were not salt bridges. The charge/charge interactions (A) and H-bonds (B) contributed similarly to the overall electrostatic interactions (C) between the three heparins and G_2 . The proportion of H-bonds involved in the association were higher than in the case of protamine. Interestingly, all the DGL's residues were involved in the associations, which illustrated the mobility of the heparin **Fx** and **Ex** in their respective complexes. That also highlighted the overall conformational flexibility of the dendrimer. Regarding **UFH**, the involvement of all its residues was a consequence of the wrapping of **UFH** around the dendrimer, in contrast with protamine. That was a direct result of the flexibility of the dendrimer that enabled an optimisation of the contacts on its cationic surface. This behaviour of the dendrimers was described as "adaptive multivalency"; the flexibility means individual surface ligands are not so well optimised locally to bind heparin chains, but the nanostructure can adapt more easily and maximise its binding contacts⁴³³. Long-range electrostatic non-contact interactions could significantly stabilise the complexes as previously discussed. We are currently developing a decomposition of the electrostatic interaction maps (C) into two contributions to the overall electrostatic interactions: 1) the contribution of direct contacts and 2) the contribution of the long-range non-contact electrostatic interactions. These maps will provide the exact location of the sites that could contribute to the long-range stabilisation of the complexes.

From this analysis, it is clear that the mechanism of association between protamine and DGL with heparins of three sizes were various. The comparison of the two antidotes's efficiency was possible thanks to the normalisation of their free energies of association ΔG_{bind} by their number of positive charges engaged in the complex. As testified by their relative ΔG_{eff} values (Figure 45), G_2 interacted with

each anticoagulant in a more effective way in comparison with protamine, being able to marshal its charges for binding as a result of its superior flexibility.

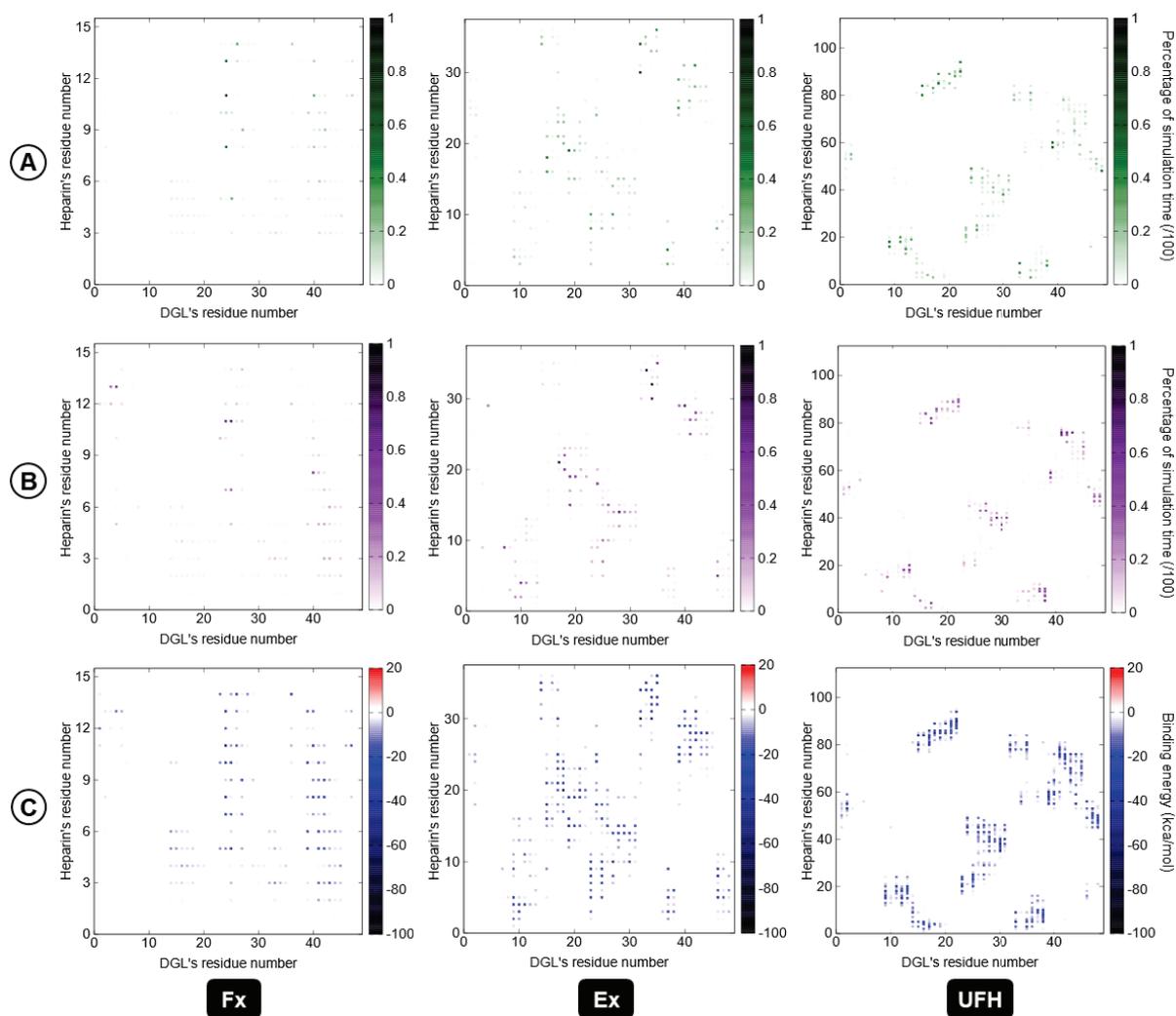


Figure 44: Two-dimensional maps of percentage of hydrogen bonds (A), percentages electrostatic interactions (B), average binding energies (C) for each of the residues involved in the complexes between protamine **G₂₋₁** and **UFH** (1), **Ex** (2), or **Fx** (3). For the full set of computed complexes, see the appendix (Appendix III-VI: Supporting information regarding the MD simulations of DGLs with heparins)

In a nutshell, **G₂** formed stable complexes with **Fx** and **Ex** in which the heparins had a great mobility. The **UFH** was wrapped around the dendrimer. The adaptive multivalency of **G₂** was visible from the contributions of all its residues in the association, which was driven similarly by salt bridges and H-bonds.

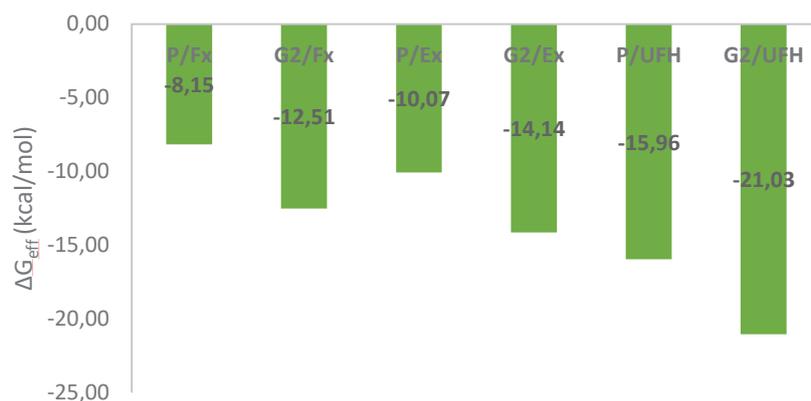


Figure 45 : Free energy of association between heparin and protamine/ G_2 (ΔG_{eff}) in kcal.mol⁻¹.

c. Conclusion

By carrying out long-time scale molecular dynamics simulations, we provided the first atomic level accounting of the binding events between heparin-based anticoagulants of various lengths and their antidotes (approved) protamine and (in study) DGL. Regarding protamine, this study showed that the relative size of the partners had a substantial influence on their corresponding binding affinities, where the number of direct contacts (i.e., electrostatic or hydrogen bonds) was not the only factor guiding the observed selectivities. Regarding G_2 , its adaptive multivalency was responsible for a high affinity for heparin that was similarly driven by H-bonds and salt bridges. The experimental difference between two antidotes's efficiency translated into their relative ΔG_{eff} values. They illustrated that the dendrimer marshalled more efficiently its charge in the complexation. This study was conducted on the hypothesis of a 1:1 stoichiometry, but the data suggested that a several antidotes could bind to **UFH**. The decomposition of electrostatic interactions into non-contact and contact contributions will enlighten the mechanism of association and is currently in progress.

6. Conclusion

The dendrigrafts of lysine are commercially available molecules that display unique features. They have a three-like structure and are polycationic at physiological pH. The study highlighted the inimitable positive charge density and the conformational plasticity of the DGL.

As a matter of fact, G_2 neutralised heparin's anticoagulant properties more efficiently than protamine did in human plasma. The dendrimer was especially active against fondaparinux. Molecular dynamic simulations revealed an unexpected decisive contribution of long range electrostatic interaction to the selectivity of protamine towards long heparin. Besides, the study highlighted the conformational plasticity of the DGL in its complexes with heparin of various sizes. This plasticity translated into an important contribution of H-bonds interactions to the overall association compared to salt bridges as well as more mobile heparin. The experimental difference between the two antidotes's efficiency was illustrated by their relative ΔG_{eff} values. The dendrimer marshalled more efficiently its charge in the complexation. The decomposition of electrostatic interactions into non-contact and contact

contributions will enlighten the mechanism of association for both protamine and DGL and is currently in progress. **G₂** should now be evaluated regarding its toxicity and in animal models, in the context of a concrete need for a surrogate of protamine. Indeed, they are not clinical translation regarding the molecules dedicated to the neutralisation of heparin that are proposed by the chemists.

Besides, DGLs interacted with different objects that presented anionic charge in water. DGL quenched the fluorescence of anionic homopolymers of aspartic acid and carboxyfluorescein upon aggregation. This fluorescence was turned on by the disruption of the complex following the introduction of spermine and Gram-negative bacteria with a limit of detection. These limits are, however, too high to enable concrete applications. The association of the exceptional optical properties of citrate-capped nanoparticles and DGL could improve our sensing platform. Finally, **G₂** mimicked the enzymatic activity of esterases and could be explored for catalysis.

Experimental section

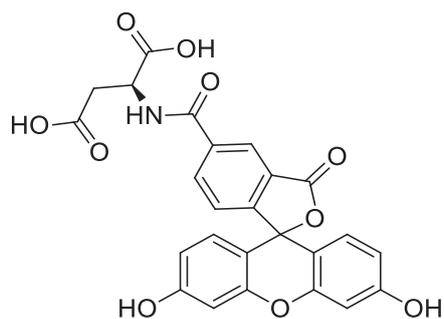
a. Synthetic procedures and characterisations

Only new compounds are fully characterised. Otherwise, only ¹H and ¹³C NMR were recorded. “Match with literature” means a reference for the compound spectra. The protocol references are given in the main text. Adaptations and optimisations were done without specific mention. The new compounds were fully characterised by HRMS, 1D, 2D NMR, and mp. The spectral data are described after their synthetic procedure. The corresponding spectra are not given in appendices for conciseness. They were compiled on previous reports and are electronically available. If a molecule was not fully characterised, for instance with macrocycles, spectral data could be delivered in appendices to support the discussion.

v. Materials and methods

All commercial reagents were used as received. Solvents were dried using conventional methods. All reactions were carried out under nitrogen or argon. Column chromatography was performed using silica gel (0.040–0.063 nm). Reactions were monitored via TLC on a silica gel plate and visualised under UV light. ¹H NMR and ¹³C NMR spectra were recorded on a Bruker DRX at 400 MHz or on a Bruker ALS at 300 MHz. J values are given in Hz. Mass spectra were acquired on an LCQ Advantage ion trap instrument, detecting positive or negative ions in the ESI mode. Melting points were determined by a Büchi Melting point B-540.

vi. Preparative procedures and characterisations

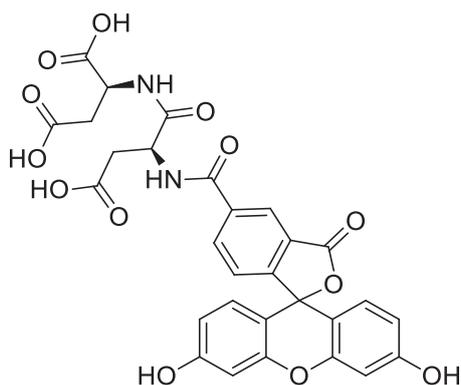


III-3₁ (S)-2-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-ylcarboxamido)succinic acid

Assembly of the linear protected peptide resin was carried out manually with a traditional Fmoc / *t*Bu strategy in a glass reactor with a glass frit. 500 mg of 2-chlorotrytil resin (0.44 mmol / g loading⁴¹) were made to swell in 10 mL of DCM. Fmoc-Asp(O*t*Bu)-OH (148 mg, 0.36 mmol) was added, followed by DIPEA (0.2 mL, 1.2 mmol). The suspension obtained was stirred for 90 mn at room temperature. The supernatant was removed by suction, and the resin was washed with DMF (5 * 6 mL). Then the resin was capped with DCM / MeOH / DIEA (17 : 2 : 1, 10 mL, 1 h), then DCM / MeOH / DIPEA (17 : 2 : 1, 10 mL, 10 mn). The supernatant was removed by suction, and the resin was washed with DMF (5*6 mL). Fmoc deprotection reactions were performed with DMF / piperidine (8:2, 3*6 mL, 5 mn each) and the resin was washed with DMF (5*6 mL, 1 mn each). The final introduction of the 5(6)-carboxyfluorescein label was performed as follows, reducing lighting exposure as much as possible: 50 mg of H₂N-Asp(O*t*Bu)-resin was made to swell in 7 mL of DMF. CF (56 mg, 0.15 mmol), EDCI (0.025 mL, 0.15 mmol) and HOBt (23 mg, 0.21 mmol) were added and the suspension was stirred for 1 h. The supernatant was removed by suction, and the resin was washed with DMF (5*4 mL). The peptide coupling was repeated once. The fully deprotected and labelled peptide was recovered directly upon acid cleavage with TFA / TIPS / H₂O (95 : 2.5 : 2.5, 5.6 mL, 1 h), followed by TFA / TIPS / H₂O (95 : 2.5 : 2.5, 2*4 mL, 1 min each). The combined filtrates were concentrated under reduced pressure. The residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give an orange powder (10 mg, 66 %).

Retention time (Method B) $t_R = 9.74$ mn; HRMS (ESI) $[M + H]^+$ found 492.0915 calculated 492.0925 for $[C_{25}H_{18}NO_{10}]^+$;

⁴¹ Loading was determined by UV titration of Fmoc deprotection layers at 299 nm. The resin was sold with a theoretical loading of 1 – 1.8 mmol/g.

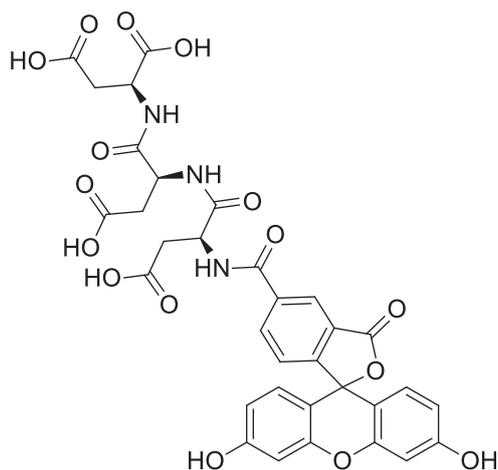


III-3₂ (S)-2-((S)-3-carboxy-2-((S)-3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-ylcarboxamido)propanamido)succinic acid

Assembly of the linear protected peptide resin was carried out manually with a traditional Fmoc / *t*Bu strategy in a glass reactor with a glass frit. 1000 mg of H₂N-Asp(*Ot*Bu)-resin (2.01 mmol / g loading⁴²) and a suspension of Fmoc-Asp(*Ot*Bu)-OH (822 mg, 3.0 mmol 1.5 eq), PyBop (1040 mg, 3.0 mmol, 1.5 eq) and DIPEA (0.35 mL, 6 mmol, 2 eq) in 15 mL of DMF was stirred for 45 min. The supernatant was removed by suction, and the resin was washed with DMF (5*5 mL). The peptide coupling was repeated once. Fmoc deprotection reactions were performed with DMF / piperidine (8:2, 3*10 mL, 5 mn each) and the resin was washed with DMF (5*6 mL, 1 mn each). The final introduction of the 5(6)-carboxyfluorescein label was performed as follows, reducing lighting exposure as much as possible: 320 mg of H₂N-Asp(*Ot*Bu)-Asp(*Ot*Bu)-resin was made to swell in 7 mL of DMF. CF (360 mg, 0.96 mmol), EDCI (0.16 mL, 0.96 mmol) and HOBt (129 mg, 0.96 mmol) were added and the suspension was stirred for 1 h. The supernatant was removed by suction, and the resin was washed with DMF (5*5 mL). The peptide coupling was repeated once. The fully deprotected and labelled peptide was recovered directly upon acid cleavage with TFA / TIPS / H₂O (95 : 2.5 : 2.5, 10 mL, 1 h), followed by TFA / TIPS / H₂O (95 : 2.5 : 2.5, 2*5 mL, 1 min each). The residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give an orange powder (400 mg, 90 %).

Retention time (Method B) $t_R = 9.21$ mn; HRMS (ESI) $[M + H]^+$ found 607.1188 calculated 607.1195 for $[C_{29}H_{23}N_2O_{13}]^+$;

⁴² Loading was determined by UV titration of Fmoc deprotection layers at 299 nm. The resin was sold with a theoretical loading of 1 – 1.8 mmol/g.

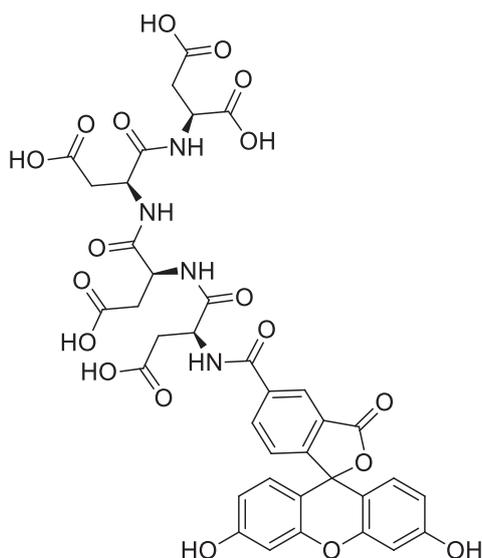


III-3₃ (S)-2-((S)-3-carboxy-2-((S)-3-carboxy-2-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-ylcarboxamido)propanamido)propanamido)succinic acid

Assembly of the linear protected peptide resin was carried out manually with a traditional Fmoc / tBu strategy in a glass reactor with a glass frit. 464 mg of H₂N-Asp(OtBu)-Asp(OtBu)-resin (0.44 mmol / g loading⁴³) and a suspension of Fmoc-Asp(OtBu)-OH (114 mg, 0.28 mmol 1.5 eq), PyBop (141 mg, 0.28 mmol, 1.5 eq) and DIEA (0.10 mL, 0.64 mmol, 2 eq) in 5 mL of DMF was stirred for 45 min. The supernatant was removed by suction, and the resin was washed with DMF (5*4 mL). The peptide coupling was repeated once. Fmoc deprotection reactions were performed with DMF / piperidine (8:2, 3*6 mL, 5 mn each) and the resin was washed with DMF (5*6 mL, 1 mn each). The final introduction of the 5(6)-carboxyfluorescein label was performed as follows, reducing lighting exposure as much as possible: 50 mg of H₂N-Asp(OtBu)-Asp(OtBu)-Asp(OtBu)-resin was made to swell in 7 mL of DMF. CF (56 mg, 0.15 mmol), EDCI (0.025 mL, 0.15 mmol) and HOBt (23 mg, 0.21 mmol) were added and the suspension was stirred for 1 h. The supernatant was removed by suction, and the resin was washed with DMF (5*4 mL). The peptide coupling was repeated once. The fully deprotected and labelled peptide was recovered directly upon acid cleavage with TFA / TIPS / H₂O (95 : 2.5 : 2.5, 5.6 mL, 1 h), followed by TFA / TIPS / H₂O (95 : 2.5 : 2.5, 2*4 mL, 1 min each). The residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give an orange powder (19 mg, 86 %).

Retention time (Method B) t_R = 9.10 mn; HRMS (ESI) [M + H]⁺ found 722.1456 calculated 722.1464 for [C₃₃H₂₈N₃O₁₆]⁺;

⁴³ Loading was determined by UV titration of Fmoc deprotection layers at 299 nm. The resin was sold with a theoretical loading of 1 – 1.8 mmol/g.

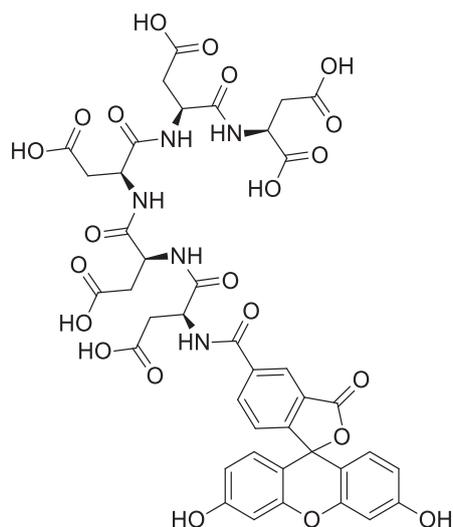


III-3₄ (S)-2-((S)-3-carboxy-2-((S)-3-carboxy-2-((S)-3-carboxy-2-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-ylcarboxamido)propanamido)propanamido)propanamido)succinic acid

Assembly of the linear protected peptide resin was carried out manually with a traditional Fmoc / *t*Bu strategy in a glass reactor with a glass frit. 250 mg of H₂N-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-resin (0.44 mmol / g loading⁴⁴) and a suspension of Fmoc-Asp(O*t*Bu)-OH (134 mg, 0.33 mmol 1.5 eq), PyBop (165 mg, 0.33 mmol, 1.5 eq) and DIEA (0.12 mL, 0.7 mmol, 2 eq) in 6 mL of DMF was stirred for 45 min. The supernatant was removed by suction, and the resin was washed with DMF (5*4 mL). The peptide coupling was repeated once. Fmoc deprotection reactions were performed with DMF / piperidine (8:2, 3*6 mL, 5 mn each) and the resin was washed with DMF (5*6 mL, 1 mn each). The final introduction of the 5(6)-carboxyfluorescein label was performed as follows, reducing lighting exposure as much as possible: 50 mg of H₂N-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-resin was made to swell in 7 mL of DMF. CF (56 mg, 0.15 mmol), EDCI (0.025 mL, 0.15 mmol) and HOBT (23 mg, 0.21 mmol) were added and the suspension was stirred for 1 h. The supernatant was removed by suction, and the resin was washed with DMF (5*4 mL). The peptide coupling was repeated once. The fully deprotected and labelled peptide was recovered directly upon acid cleavage with TFA / TIPS / H₂O (95 : 2.5 : 2.5, 5.6 mL, 1 h), followed by TFA / TIPS / H₂O (95 : 2.5 : 2.5, 2*4 mL, 1 min each). The combined filtrates were concentrated under reduced pressure. The residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give an orange powder (16 mg, 72 %).

Retention time (Method B) $t_R = 9.05$ mn; HRMS (ESI) $[M + H]^+$ found 837.1712 calculated 817.1734 for $[C_{37}H_{33}N_4O_{19}]^+$;

⁴⁴ Loading was determined by UV titration of Fmoc deprotection layers at 299 nm. The resin was sold with a theoretical loading of 1 – 1.8 mmol/g.

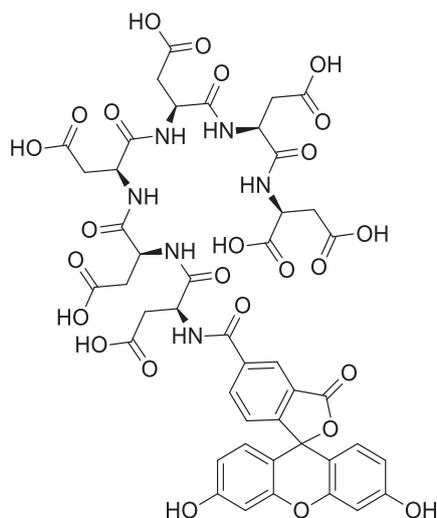


III-3₅ (2S,5S,8S,11S,14S)-5,8,11-tris(carboxymethyl)-14-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-ylcarboxamido)-4,7,10,13-tetraoxo-3,6,9,12-tetraazapentadecane-1,2,15-tricarboxylic acid

Assembly of the linear protected peptide resin was carried out manually with a traditional Fmoc / tBu strategy in a glass reactor with a glass frit. 1000 mg of H₂N-Asp(OtBu)-Asp(OtBu)-resin (2.01 mmol / g loading⁴⁵) and a suspension of Fmoc-Asp(OtBu)-OH (822 mg, 3.0 mmol 1.5 eq), PyBop (1040 mg, 3.0 mmol, 1.5 eq) and DIEA (0.35 mL, 6 mmol, 2 eq) in 15 mL of DMF was stirred for 45 min. The supernatant was removed by suction, and the resin was washed with DMF (5*5 mL). The peptide coupling was repeated once. Fmoc deprotection reactions were performed with DMF / piperidine (8:2, 3*10 mL, 5 mn each) and the resin was washed with DMF (5*6 mL, 1 mn each). The sequence was done three times, with the same amount of matter. The final introduction of the 5(6)-carboxyfluorescein label was performed as follows, reducing lighting exposure as much as possible: 360 mg of H₂N-Asp(OtBu)-Asp(OtBu)-Asp(OtBu)-Asp(OtBu)-resin was made to swell in 15 mL of DMF. CF (360 mg, 0.96 mmol), EDCI (0.16 mL, 0.96 mmol) and HOBT (129 mg, 0.96 mmol) were added and the suspension was stirred for 1 h. The supernatant was removed by suction, and the resin was washed with DMF (5*5 mL). The peptide coupling was repeated once. The fully deprotected and labelled peptide was recovered directly upon acid cleavage with TFA / TIPS / H₂O (95 : 2.5 : 2.5, 10 mL, 1 h), followed by TFA / TIPS / H₂O (95 : 2.5 : 2.5, 2*5 mL, 1 min each). The combined filtrates were concentrated under reduced pressure. The residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated, and the product was dried under vacuum to give an orange powder (38 mg, 56 %).

Retention time (Method B) $t_R = 9.03$ mn; HRMS (ESI) $[M + Na]^{2+}$ found 487.5950 calculated 487.5948 for $[C_{41}H_{38}N_5NaO_{22}]^{2+}$;

⁴⁵ Loading was determined by UV titration of Fmoc deprotection layers at 299 nm. The resin was sold with a theoretical loading of 1 – 1.8 mmol/g.

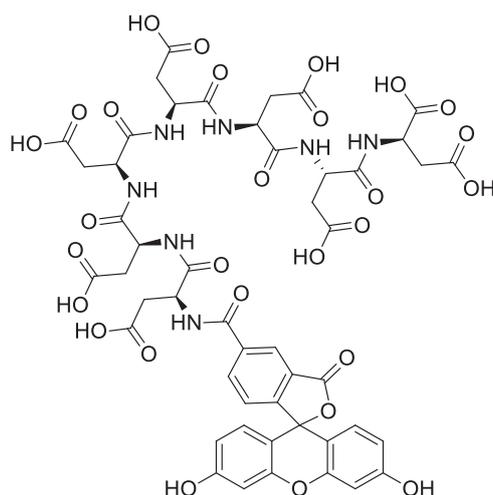


III-3₆ (2*S*,5*S*,8*S*,11*S*,14*S*,17*S*)-5,8,11,14-tetrakis(carboxymethyl)-17-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-ylcarboxamido)-4,7,10,13,16-pentaoxo-3,6,9,12,15-pentaazaocane-1,2,18-tricarboxylic acid

Assembly of the linear protected peptide resin was carried out manually with a traditional Fmoc / *t*Bu strategy in a glass reactor with a glass frit. 1000 mg of H₂N-Asp(*Ot*Bu)-Asp(*Ot*Bu)-Asp(*Ot*Bu)-Asp(*Ot*Bu)-Asp(*Ot*Bu)-resin (2.01 mmol / g loading⁴⁶) and a suspension of Fmoc-Asp(*Ot*Bu)-OH (822 mg, 3.0 mmol 1.5 eq), PyBop (1040 mg, 3.0 mmol, 1.5 eq) and DIEA (0.35 mL, 6 mmol, 2 eq) in 15 mL of DMF was stirred for 45 min. The supernatant was removed by suction, and the resin was washed with DMF (5*5 mL). The peptide coupling was repeated once. Fmoc deprotection reactions were performed with DMF / piperidine (8:2, 3*10 mL, 5 mn each) and the resin was washed with DMF (5*6 mL, 1 mn each). The sequence was done three times, with the same amount of matter. The final introduction of the 5(6)-carboxyfluorescein label was performed as follows, reducing lighting exposure as much as possible: 500 mg of H₂N-Asp(*Ot*Bu)-Asp(*Ot*Bu)-Asp(*Ot*Bu)-Asp(*Ot*Bu)-Asp(*Ot*Bu)-Asp(*Ot*Bu)-resin was made to swell in 15 mL of DMF. CF (504 mg, 1.34 mmol), EDCI (0.22 mL, 1.34 mmol) and HOBT (180 mg, 1.34 mmol) were added and the suspension was stirred for 1 h. The supernatant was removed by suction, and the resin was washed with DMF (5*5 mL). The peptide coupling was repeated once. The fully deprotected and labelled peptide was recovered directly upon acid cleavage with TFA / TIPS / H₂O (95 : 2.5 : 2.5, 10 mL, 1 h), followed by TFA / TIPS / H₂O (95 : 2.5 : 2.5, 2*5 mL, 1 min each). The combined filtrates were concentrated under reduced pressure. The residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated and the product was dried under vacuum to give an orange powder (52 mg, 78 %).

Retention time (Method B) $t_R = 9.12$ mn; HRMS (ESI) $[M + H]^+$ found 1067.2282 calculated 1067.2272 for $[C_{45}H_{43}N_6O_{25}]^+$;

⁴⁶ Loading was determined by UV titration of Fmoc deprotection layers at 299 nm. The resin was sold with a theoretical loading of 1 – 1.8 mmol/g.



III-37 (2R,5S,8S,11S,14S,17S,20S)-5,8,11,14,17-pentakis(carboxymethyl)-20-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-ylcarboxamido)-4,7,10,13,16,19-hexaoxo-3,6,9,12,15,18-hexaazahenicosane-1,2,21-tricarboxylic acid

Assembly of the linear protected peptide resin was carried out manually with a traditional Fmoc / *t*Bu strategy in a glass reactor with a glass frit. 500 mg of H₂N-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-resin (2.01 mmol / g loading⁴⁷) and a suspension of Fmoc-Asp(O*t*Bu)-OH (411 mg, 1.5 mmol, 1.5 eq), PyBop (520 mg, 1.5 mmol, 1.5 eq) and DIEA (0.17 mL, 3 mmol, 2 eq) in 10 mL of DMF was stirred for 45 min. The supernatant was removed by suction, and the resin was washed with DMF (5*5 mL). The peptide coupling was repeated once. Fmoc deprotection reactions were performed with DMF / piperidine (8:2, 3*10 mL, 5 mn each) and the resin was washed with DMF (5*6 mL, 1 mn each). The sequence was done three times, with the same amount of matter. The final introduction of the 5(6)-carboxyfluorescein label was performed as follows, reducing lighting exposure as much as possible: 500 mg of H₂N-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-Asp(O*t*Bu)-resin was made to swell in 15 mL of DMF. CF (504 mg, 1.34 mmol), EDCI (0.22 mL, 1.34 mmol) and HOBt (180 mg, 1.34 mmol) were added and the suspension was stirred for 1 h. The supernatant was removed by suction, and the resin was washed with DMF (5*5 mL). The peptide coupling was repeated once. The fully deprotected and labelled peptide was recovered directly upon acid cleavage with TFA / TIPS / H₂O (95 : 2.5 : 2.5, 10 mL, 1 h), followed by TFA / TIPS / H₂O (95 : 2.5 : 2.5, 2*5 mL, 1 min each). The combined filtrates were concentrated under reduced pressure. The residue was purified by column chromatography (C₁₈, reverse phase, water TFA 0.05 % / MeCN TFA 0.05 % = 7:3). The solvent was evaporated, and the product was dried under vacuum to give an orange powder (248 mg, 90 %).

Retention time (Method B) $t_R = 9.20$ mn; HRMS (ESI) [M + 2 H]²⁺ found 591.6293 calculated 591.6307 for [C₄₉H₄₉N₇NaO₂₈]²⁺;

b. HPLC analysis

⁴⁷ Loading was determined by UV titration of Fmoc deprotection layers at 299 nm. The resin was sold with a theoretical loading of 1 – 1.8 mmol/g.

i. HPLC analysis

Analysis are performed using a 1260 infinity Agilent Technologies system at 25 °C.

Method A: column Agilent Eclipse Plus C₈, 3.5 μm, 4.6 x 150 mm, V_{inj} = 3 μL, flow 1 mL/mn. Gradient water + 0.05 % TFA / MeCN + 0.05 % TFA in 25 mn method (t₀ 80/20, t₅ 60/40, t₈ 0/100).

Method B: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 μL, 2.7 μm, 2.1 x 50 mm flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 22 mn method (t₀ 100/0, t₁₅ 70/30, t₂₂ 0/100).

Method C: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 μL, 2.7 μm, 2.1 x 50 mm flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 28 mn method (t₀ 100/0, t₂ 100/00, t₂₀ 80/20, t₂₅ 0/100, t₂₈ 0/100)

Method D: Agilent Poroshell 120 EC C₈, 2.7 V_{inj} = 3 μL, 2.7 μm, 2.1 x 50 mm flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 25 mn method (t₀ 100/0, t_{1.5} 100/00, t₁₅ 85/15, t₂₄ 0/100, t₂₅ 0/100)

Method E: Agilent Nucleodur HILIC EC 250 4.6 V_{inj} = 5 μL, 5 μm, 4.6 x 250 mm flow 0.8 mL/mn. Gradient AcONH₄ 50 mM / MeCN in 25 mn method (t₀ 10/90, t₂₅ 60/40)

Method F: Restek Raptor Biphenyl, 2.7 μm, 3.00 x 50 mm, V_{inj} = 3 μL, flow 0.6 mL/mn. Gradient water + 0.1 % formic acid / MeOH + 0.1 % formic acid in 17 mn method (t₀ 99/1, t_{0.5} 99/1, t₁₅ 50/50, t₁₆ 1/99, t₁₇ 1/99)

Method G: Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 10 μL, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t₀ 90/10, t₁₀ 50/50, t₁₅ 40/60, t₁₇ 0/100, t₂₀ 0/100)

Method H: Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 10 μL, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t₀ 100/0, t₁ 100/0, t₁₅ 40/60, t₁₇ 0/100, t₁₉ 0/100)

Method I: Restek Raptor Biphenyl, 2.7 μm, 3.00 x 50 mm, V_{inj} = 5 μL, flow 0.8 mL/mn. Gradient water + 0.1 % formic acid / MeOH + 0.1 % formic acid in 17 mn method (t₀ 70/30, t₁₅ 0/100, t₁₇ 0/100)

Method PREP₁: Zorbax Eclipse XDB C₈ 5.0 μm 250 x 9.4 mm V_{inj} = 100 μL, flow 5.0 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 19 mn method (t₀ 92/8, t₂ 92/8, t₁₄ 50/50)

ii. UPLC analysis

Analysis are performed using a DIONEX ultimate 3000 Thermo Scientific system.

Method J: 30 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 4 μL, flow 0.5 mL/mn. Gradient water + 0.1 % formic acid / MeCN + 0.1 % formic acid in 20 mn method (t₀ 90/10, t₂₀ 50/50)

Method K: 45 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 4 μL, flow 0.4 mL/mn. Gradient water + 0.1 % TFA / MeCN + 0.1 % TFA in 8 mn method (t₀ 85/15, t₈ 58/42)

Method L: 45 °C, Phenomenex Kinetex C₁₈ 100 A 2.6 μm 100 x 2.1 mm V_{inj} = 4 μL, flow 0.4 mL/mn. Gradient water + 0.1 % TFA / MeCN + 0.1 % TFA in 8 mn method (t₀ 85/15, t₈ 40/60)

xi. UPLC/HRMS analysis

Analysis are performed using a U3000 Thermo Fisher Scientific / QTOF Impact II Bruker system. The column used was an Agilent Poroshell 120 EC C₈, 2.7 μm, 5 x 2.1 mn, V_{inj} = 3 μL, flow 0.8 mL/mn, Gradient water + 0.1 % formic acid / MeCN:MeOH (50/50) (or MeCN) + 0.1 % formic acid for typical run, other methods have been used. Mass spectrometry parameters depend on the analysis, UV detection performed at 250 and 350 nm.

c. DGL / Azur A / heparin UV / absorbance titrations

UV / absorbance measurements were performed on a UV-2401PC Shimadzu system at 25°C. UV measurements involving G₁ were performed on a VICTOR Nivo Multimode Microplate Reader at 25°C. A typical experiment with G₁ is described: PBS buffer was placed in the well, followed by Azur A, G₁, and finally heparin. The protocol of titration is given below. Heparin sodium salt from porcine intestinal mucosa (207 USP/mg, ref: H3399-50KU) and Azur A were purchased from Sigma-Aldrich. Solutions were prepared in PBS 1 M pH 7.4 buffer.

User interface program version	2.0.2.
Instrument serial number	1018013
PROTOCOL NAME	G1 AZA
PLATE TYPE	General 96
PLATE FORMAT	96 wells (8X12)
OPERATION	Shaker-Shake
SHAKE TYPE	Orbital
SHAKE SPEED (rpm)	600
SHAKE TIME	0 hours 1 minutes 0 seconds
OPERATION	Timing-Delay
DELAY	0 hours 0 minutes 30 seconds
OPERATION	Filter ABS-EndPoint
MEASUREMENT TYPE	Single label
EXCITATION FILTER	590/20 nm
MEASUREMENT TIME (ms)	500
FLASH ENERGY (mJ)	100

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

A	V Heparin [1.22 mM] (μL)	10	10	10	10	10	10	10	10	10	10
---	--------------------------	----	----	----	----	----	----	----	----	----	----

	V Azur A [1.01 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V PBS 1 M pH 7,4 (μL)	130	127,5	125	122,5	120	117,5	115	112,5	110	107,5
	V G_1 [38.7 μM] (μL)	0	2,5	5	7,5	10	12,5	15	17,5	20	22,5

B	V Heparin [1.22 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V Azur A [1.01 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V PBS 1 M pH 7,4 (μL)	105	102,5	100	97,5	95	92,5	90	87,5	85	82,5
	V G_1 [38.7 μM] (μL)	25	27,5	30	32,5	35	37,5	40	42,5	45	47,5

C	V Heparin [1.22 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V Azur A [1.01 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V PBS 1 M pH 7,4 (μL)	80	77,5	75	72,5	70	67,5	65	62,5	60	57,5
	V G_1 [38.7 μM] (μL)	50	52,5	55	57,5	60	62,5	65	67,5	70	72,5

D	V Heparin [1.22 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V Azur A [1.01 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V PBS 1 M pH 7,4 (μL)	55	52,5	50	47,5	45	42,5	40	37,5	35	32,5
	V G_1 [38.7 μM] (μL)	75	77,5	80	82,5	85	87,5	90	92,5	95	97,5

E	V Heparin [1.22 mM] (μL)	10	10	10	10	10	10	10	10	10	
	V Azur A [1.01 mM] (μL)	10	10	10	10	10	10	10	10	10	
	V PBS 1 M pH 7,4 (μL)	30	27,5	25	22,5	20	17,5	15	12,5	10	

	V G ₁ [38.7 μM] (μL)	100	102,5	105	107,5	110	112,5	115	117,5	120	
--	---------------------------------	-----	-------	-----	-------	-----	-------	-----	-------	-----	--

F	V Heparin [1.22 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V Azur A [1.01 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V PBS 1 M pH 7,4 (μL)	127,5	125	122,5	120	117,5	115	112,5	110	107,5	100,5
	V G ₁ [387 μM] (μL)	2,5	5	7,5	10	12,5	15	17,5	20	22,5	25

G	V Heparin [1.22 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V Azur A [1.01 mM] (μL)	10	10	10	10	10	10	10	10	10	10
	V PBS 1 M pH 7,4 (μL)	102,5	100	97,5	95	92,5	90	87,5	85	82,5	80
	V G ₁ [387 μM] (μL)	27,5	30	32,5	35	37,5	40	42,5	45	47,5	50

H	V Heparin [1.22 mM] (μL)	10	10	10							
	V Azur A [1.01 mM] (μL)	10	10	10							
	V PBS 1 M pH 7,4 (μL)	77,5	75	72,5							
	V G ₁ [387 μM] (μL)	52,5	55	57,5							

d. Fluorescence

Fluorescence measurements were performed on a VICTOR Nivo Multimode Microplate Reader (λ_{ex} = 480 nm and λ_{em} = 530 nm for carboxyfluorescein derivatives), at 25°C.

i. DGL / III-3_n titrations

A typical experiment with III-3₆ is described: III-3₆ was placed in the well, followed by HEPES buffer and finally G₄. The protocol of titration is given below. The antibodies variants were dialysed overnight at 4 °C and titrated by absorption at 280 nm⁴⁹³ after reception from UCB Pharma (absorbance at 280 nm,

1 cm, $A_{Fab} = 1.42$; $A_{Ig1} = 1.42$). Antibodies have the following characterisations ($MW_{Ig1} 144\ 248$ g/mol, $MW_{Fab} 46\ 906$ g/mol, major deamidation site and level: $Ig_{G1} - HC N_{389/390}$ 1% vs 0.5% in control, $F_{ab} - LC N_{53}$ 1.5% vs 1% control).

User interface program version	2.0.2.
Instrument serial number	1018013
PLATE TYPE	PerkinElmer OptiPlate
PLATE FORMAT	96 wells (8X12)
OPERATION	Shaker-Shake
SHAKE TYPE	Orbital
SHAKE SPEED (rpm)	600
SHAKE TIME	0 hours 2 minutes 0 seconds
OPERATION	Timing-Delay
DELAY	0 hours 1 minutes 0 seconds
OPERATION	FI-EndPoint
MEASUREMENT TYPE	Single label
EXCITATION FILTER	480/30nm
EMISSION FILTER	530/30nm
DICHROIC MIRROR	General (50/50)
MEASUREMENT DIRECTION	Top measurement
MEASUREMENT TIME (ms)	500
Z-FOCUS (mm)	8,5
EXCITATION SPOT SIZE (mm)	0,5
EMISSION SPOT SIZE (mm)	1
FLASH ENERGY (mJ)	10
PMT HV (V)	593
MEASUREMENT ORDER	Bi-directional

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

A	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V HEPES 10 mM pH 7,4 (μL)	139	138	137	136	135	134	133	132	131	130
	V D ₆ CF [9667 nM] (μL)	1	2	3	4	5	6	7	8	9	10

B	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
---	--------------------------------	----	----	----	----	----	----	----	----	----	----

	V HEPES 10 mM pH 7,4 (μL)	129	128	125	123	120	118	115	113	110	108
	V III-3₆ [9667 nM] (μL)	11	12	15	17	20	22	25	27	30	32

C	V G₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V HEPES 10 mM pH 7,4 (μL)	105	103	100	98	95	93	90	88	85	80
	V III-3₆ [9667 nM] (μL)	35	37	40	42	45	47	50	52	55	60

D	V G₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V HEPES 10 mM pH 7,4 (μL)	75	70	65	60	55	50	45	40	30	25
	V III-3₆ [9667 nM] (μL)	65	70	75	80	85	90	95	100	110	115

ii. DGL / **III-3_n** titrations **G₄** / **III-3_n** / BL21(DE3) titrations

A typical experiment with **III-3₁** is described: **III-3₁** was placed in the well, followed by HEPES buffer and finally **G₄**. The protocol of titration is given below.

User interface program version	2.0.2.
Instrument serial number	1018013
PLATE TYPE	PerkinElmer OptiPlate
PLATE FORMAT	96 wells (8X12)
OPERATION	Shaker-Shake
SHAKE TYPE	Orbital
SHAKE SPEED (rpm)	600
SHAKE TIME	0 hours 2 minutes 0 seconds
OPERATION	Timing-Delay
DELAY	0 hours 1 minutes 0 seconds
OPERATION	FI-EndPoint
MEASUREMENT TYPE	Single label
EXCITATION FILTER	480/30nm
EMISSION FILTER	530/30nm
DICHROIC MIRROR	General (50/50)
MEASUREMENT DIRECTION	Top measurement
MEASUREMENT TIME (ms)	500

Z-FOCUS (mm)	8,5
EXCITATION SPOT SIZE (mm)	0,5
EMISSION SPOT SIZE (mm)	1
FLASH ENERGY (mJ)	10
PMT HV (V)	593
MEASUREMENT ORDER	Bi-directional

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

A	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V III-3 ₁ [25.8 μM] (μL)	25	25	25	25	25	25	25	25	25	25
	V HEPES 10 mM pH 7,4 (μL)	110	105	100	95	90	85	80	75	70	65
	V BL21(DE3) [4 pM] (μL)	5	10	15	20	25	30	35	40	45	50

B	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V III-3 ₁ [25.8 μM] (μL)	25	25	25	25	25	25	25	25	25	25
	V HEPES 10 mM pH 7,4 (μL)	110	105	100	95	90	85	80	75	70	65
	V BL21(DE3) [0.4 pM] (μL)	5	10	15	20	25	30	35	40	45	50

C	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V III-3 ₁ [25.8 μM] (μL)	25	25	25	25	25	25	25	25	25	25
	V HEPES 10 mM pH 7,4 (μL)	110	105	100	95	90	85	80	75	70	65
	V BL21(DE3) [40 fM] (μL)	5	10	15	20	25	30	35	40	45	50

D	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V III-3 ₁ [25.8 μM] (μL)	25	25	25	25	25	25	25	25	25	25

	V HEPES 10 mM pH 7,4 (μL)	110	105	100	95	90	85	80	75	70	65
	V BL21(DE3) [4 fM] (μL)	5	10	15	20	25	30	35	40	45	50

E	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V III-3 ₁ [25.8 μM] (μL)	25	25	25	25	25	25	25	25	25	25
	V HEPES 10 mM pH 7,4 (μL)	110	105	100	95	90	85	80	75	70	65
	V BL21(DE3) [0.4 fM] (μL)	5	10	15	20	25	30	35	40	45	50

F	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V III-3 ₁ [25.8 μM] (μL)	25	25	25	25	25	25	25	25	25	25
	V HEPES 10 mM pH 7,4 (μL)	110	105	100	95	90	85	80	75	70	65
	V BL21(DE3) [40 aM] (μL)	5	10	15	20	25	30	35	40	45	50

G	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V III-3 ₁ [25.8 μM] (μL)	25	25	25	25	25	25	25	25	25	25
	V HEPES 10 mM pH 7,4 (μL)	110	105	100	95	90	85	80	75	70	65
	V BL21(DE3) [4 aM] (μL)	5	10	15	20	25	30	35	40	45	50

H	V G ₄ [817 nM] (μL)	10	10	10	10	10	10	10	10	10	10
	V III-3 ₁ [25.8 μM] (μL)	25	25	25	25	25	25	25	25	25	25
	V HEPES 10 mM pH 7,4 (μL)	110	105	100	95	90	85	80	75	70	65

V BL21(DE3) [4 pM] (μ L)	5	10	15	20	25	30	35	40	45	50
----------------------------------	---	----	----	----	----	----	----	----	----	----

Bacteria were grown in a minimum media M9 (glucose) overnight at 37.0°C. Then, measurement of turbidity at 600 nm on a Agilent Technologies Cary60 UV-VIS gave the approximative population of the colony ($A^{600} = 0.53$, corresponding to $4 \cdot 10^8$ bacteria / mL)⁴⁹⁴.

iii. **G₂CF** / BL21(DE3) measurements

A typical experiment is described: Stationary phase cultures of BL21(DE3) (5 mL, $4 \cdot 10^8$ cells / mL) were incubated with **G₂CF** (8,6 mg / mL) overnight at room temperature. Bacteria were then washed 5 times with HEPES buffer 10 mM pH 7,4 (1 mL) by centrifugation (5 mn, 5000 tr / mn). Then, bacteria were lysed by ultrasound (BANDELIN sonopuls GM 70, 5*30 s, 20 mHZ) and in presence of DNAase. The supernatant was analysed by fluorescence. The protocol of fluorescence is given below.

User interface program version	2.0.2.
Instrument serial number	1018013
PLATE TYPE	PerkinElmer OptiPlate
PLATE FORMAT	96 wells (8X12)
OPERATION	FI-EndPoint
MEASUREMENT TYPE	Single label
EXCITATION FILTER	480/30nm
EMISSION FILTER	530/30nm
DICHROIC MIRROR	General (50/50)
MEASUREMENT DIRECTION	Top measurement
MEASUREMENT TIME (ms)	500
Z-FOCUS (mm)	8,5
EXCITATION SPOT SIZE (mm)	0,5
EMISSION SPOT SIZE (mm)	1
FLASH ENERGY (mJ)	10
PMT HV (V)	593
MEASUREMENT ORDER	Bi-directional

e. pH titration

Experiments were performed on a Metrohm 848 Titrino plus at 25°C. The machine was calibrated prior use with adequate buffer solution (7.00; 10.01). NaOH solution was calibrated prior use with analytical grade potassium hydrogen phthalate. 11.0 mg.

f. ITC

Standard procedure^{80,94}: Typical experiments were performed at 298 K using a Malvern ITC200 instrument. The host solution in Tris buffer pH 7.4 (200mM) was placed into the calorimeter cell (200 μ l) and 20 successive aliquots (2 μ l) of guest solution (10 times more concentrated) were added via a computer-automated injector at 2 mn intervals (Reference power 5, Initial delay 60, Stirring speed 1000, tinj: 4.0 s, filter period 5 s, feedback mode/gain high, edit mode unique). Heat changes were recorded after each addition. Heats of dilution were subtracted from the titration data prior to curve fitting. The first injection was discarded from each dataset to remove the effect of guest diffusion across the syringe tip during the equilibration process. Titrations curves were fitted with the one binding site model using Origin v.5.0 software supplied by MicroCal. ITC apparatus was provided by the IBCP, Lyon 1.

g. Molecular dynamic

Molecular modelling and simulations of DGLs, protamines, heparins and their corresponding complexes have been performed using the AMBER biomolecular package^{95b}. Eight DGL samples structures were prepared as detailed in the corresponding paper^{464b}. Their primary structures are given in the main text. Four protamine samples have been built. Their primary structures⁴¹⁶ are presented in the main text. All these proteins have been modelled with the AMBER ff14SB protein force field⁴⁹⁵ whereas the GLYCAM06 carbohydrate force field⁴⁹⁶ was used for UFH, enoxaparin, and fondaparinux sugar molecules. After an initial 10 μ s molecular dynamics in implicit water solvent of each individual component, DGL/heparin and protamine/heparin complexes have been built in explicit TIP3P water²²⁰. Systems were solvated in truncated octahedrons with a minimum distance between any solute atom and the edge of the periodic box of 20 \AA , yielding systems ranging from around 62,000 to about 764,000 atoms depending on the size and the initial shape of the complexed moieties. Each system was neutralised by adding the right amount of counter sodium or chloride ions. First, all molecular systems have been energy minimized for 10,000 steps using consecutively 1,000 steps of steepest descent followed by 9,000 steps of conjugate gradients. Then they have been heated from 10 to 300 K during 250 ps at constant pressure (1 bar) with a time step of 1 fs. A 20 kcal/mol/ \AA^2 restrained harmonic potential was applied on all non-hydrogen atoms during heating to prevent the loss of the initial three-dimension complexed shape. Productions runs were performed for 2 microseconds, except for the protamine/UFH complexes whose sizes are comprised between 380,000 and 764,000 atoms and for which production molecular dynamics runs were limited to 1 microsecond. All molecular dynamics simulations were carried out using the GPU implementation⁴⁹⁷ of the pmemd program in AMBER18. For all equilibration and production runs, the Particle Mesh Ewald summation technique was used with the default 8 \AA cutoff. Bonds involving hydrogen atoms were constrained by the SHAKE algorithm⁴⁹⁸. Temperature was controlled with a Langevin thermostat at 300 K with a collision frequency $\gamma = 10.0 \text{ ps}^{-1}$. Constant pressure dynamics at 1 bar was applied with isotropic position scaling and controlled by a Monte Carlo barostat with volume change attempts every 100 steps. For production runs, a time step of 4 fs was used with hydrogen mass repartitioning. All trajectory analyses were carried out using the cpptraj module of AMBER^{95a}. The free energies of binding between any DGL or protamine on one side and any heparin on the other side were computed using the equation:

$$\Delta G_{\text{binding}} = \langle \text{GPH} \rangle - \langle \text{GP} \rangle - \langle \text{GH} \rangle$$

where PH, P and H stand for protein-heparin complex, protein (i.e. either DGL or protamine), and heparin (i.e. UFH, enoxaparin, or fondaparinux), respectively. The average free energy of each system is estimated as a sum of three terms:

$$G = E_{\text{MM}} + G_{\text{solv}} - TS_{\text{MM}}$$

where E_{MM} is the molecular mechanics energy of each system, including internal, non-bonded electrostatics, and van der Waals energies. G_{solv} is the solvation energy which consists of a polar and a non-polar part. The polar solvation free energy is calculated by a Generalized Born (GB) approach. The nonpolar solvation free energy is computed by a relation to the solvent-accessible surface area (SASA). The last term TS_{MM} is the product of absolute temperature and the entropy. In this study, the first two terms were calculated using the MMPBSA.py module of AMBER18 with all water and counter-ions stripped off. To evaluate the polar solvation free energy, the GB^{obc} solvation model was used⁴⁹⁹. The hydrophobic contribution has been approximated by the Linear Combinations of Pairwise Overlaps (LCPO) method⁵⁰⁰. The entropy term was not included in our calculations. Preliminary studies on a few protein/heparin complexes (data not shown) using normal-mode analysis showed us that the entropy contribution was small as compared to the other components of the free energy of binding whereas the computer times were increased by a large factor⁵⁰¹.

The primary structures of the modelled DGLs and protamine were the following (Table 22).

Molecule	Sequence
G₂-1	ACDDDDDDCCCh 8CCCCCCh 7CCCCCCh 6CCCh 5CCh 4CCCCCh 3CCh
G₂-2	ADDDDCDDCCCCCh 8CCCh 7CCCCCh 5CCCCCh 4CCCh 3CCCh 2CCh
G₂-3	BCDCDDDDCCCh 8CCCh 7CCCCCCh 6CCCh 5CCCh 3CCCh 1CCCh
G₂-4	BDCDCDDCCCh 8CCCCCh 7CCCCCCh 6CCCCCh 4CCCh 2CCCh 1CCCCCh
G₂-5	BDCDDDDCCCh 7CCCCCCh 6CCCh 5CCCh 4CCCh 2CCCCCh 1CCCh
G₂-6	BDDCDDCCCh 7CCCh 6CCCCCh 5CCh 3CCCCCCh 2CCCh 1CCh
G₂-7	BDDCCDDCCCCCh 8CCCh 7CCCCCh 4CCCh 3CCCCCh 2CCCh 1CCCh
G₂-8	BDDCDDCCCCCCh 7CCCh 6CCh 4CCh 3CCCCCCh 2CCCCCh 1CCCCCh
P1	PRRRRSSSR PIRRRRPRA SRRRRRGRR RR
P2	PRRRSSRRP VRRRRRPRVS RRRRRRGRR RR
P3	PRRRSSSRP VRRRRRPRVS RRRRRRGRR RR
P4	PRRRASRRI RRRRRPRVSR RRRRGRRRR

Table 22 : **G₂** and protamine primary sequences retained for MD simulations. Protamine letter refers to the one letter code for amino acid. DGL letter are given from C-terminal to N-terminal: A initiation/elongation, B initiation/branching/elongation, C elongation, D elongation/branching, e branching/termination, h termination, numbers addressed the locations of the residues where branched chains were growing starting from the C-terminal amino acid^{464b}

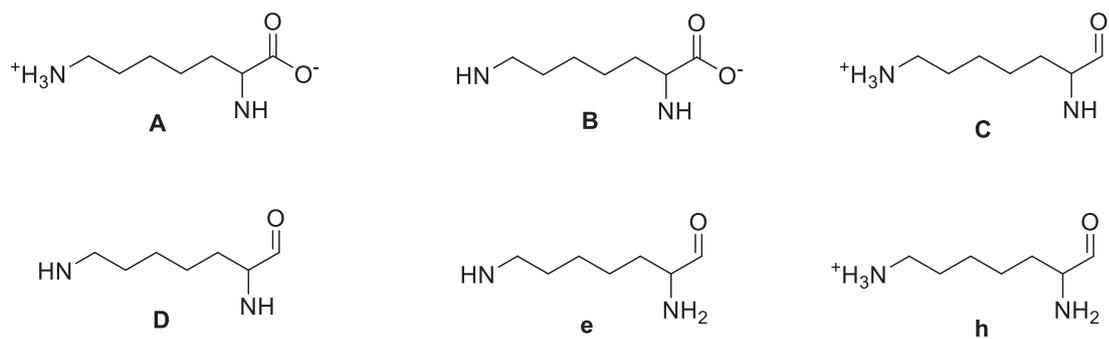
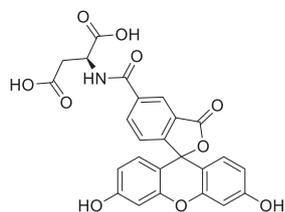


Figure 46: Building units of DGL encompassing all the possible arrangements of the lysine residues

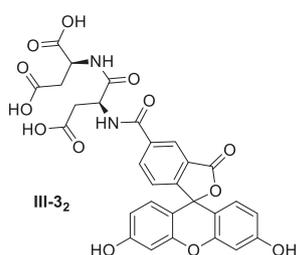
Appendices

a. Appendix III-I: List of molecules

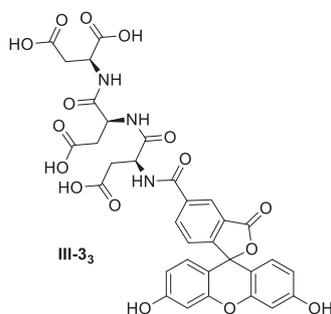
Synthesised molecules of chapter 3, for isolated and ESI-MS analysis + ^1H NMR identification on pure product. New and fully characterised molecules in a box.



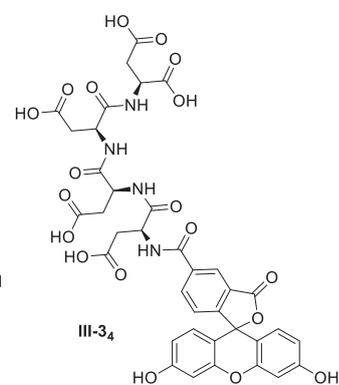
III-3₁



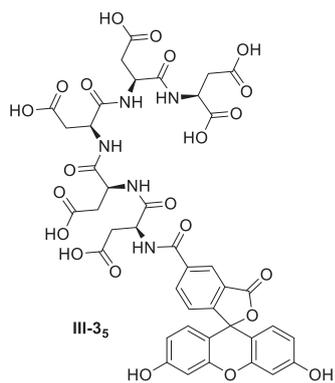
III-3₂



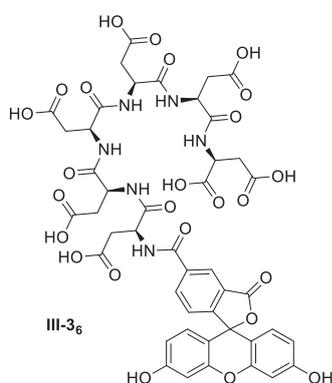
III-3₃



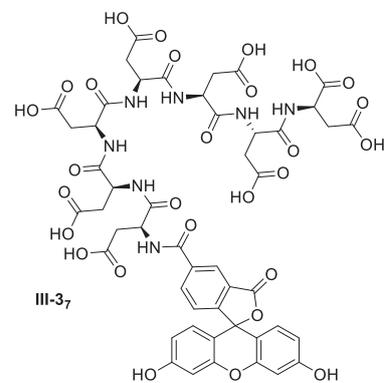
III-3₄



III-3₅



III-3₆



III-3₇

b. Appendix III-II: Fluorescence

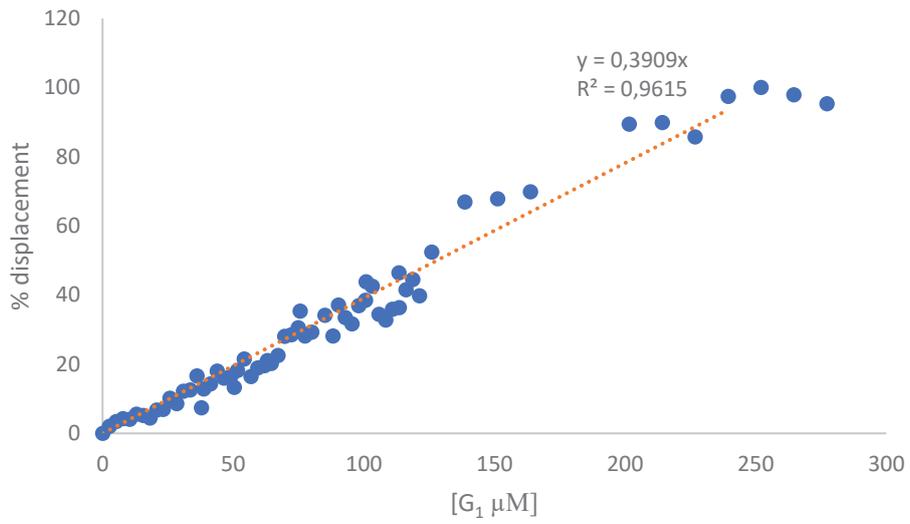


Figure 47: Restauration of UV absorbance at 590 nm of Azur A by addition of G₁ on a heparin ([81 μM]) / Azur A ([67.5 μM]) complex

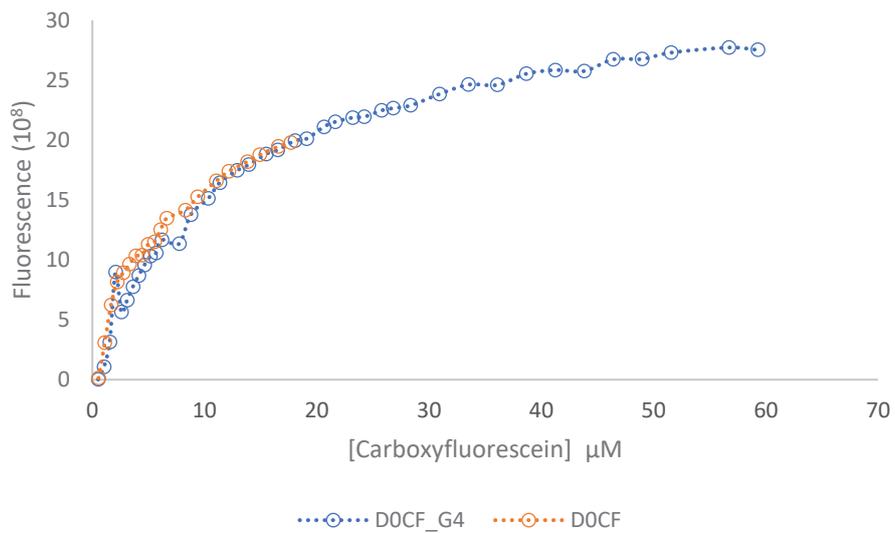


Figure 48: Fluorescence titration of Carboxyfluorescein = D₀CF with G₄ (54.0 nM)

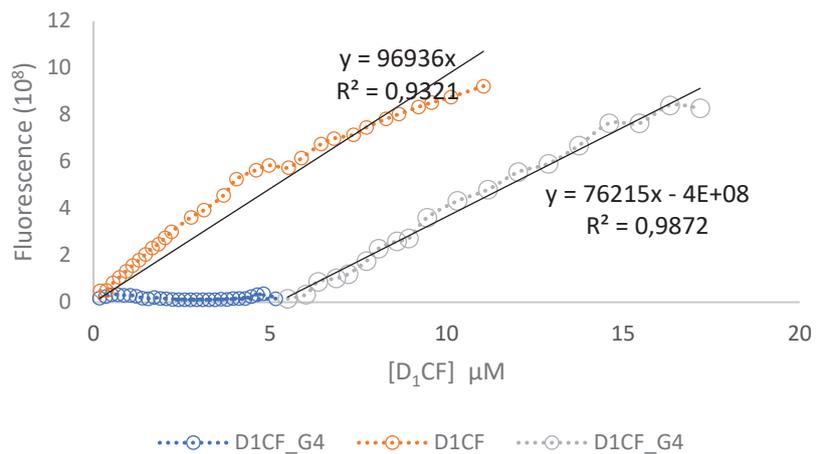


Figure 49: Fluorescence titration of III-3₁ = D₁CF with G₄ (54.0 nM)

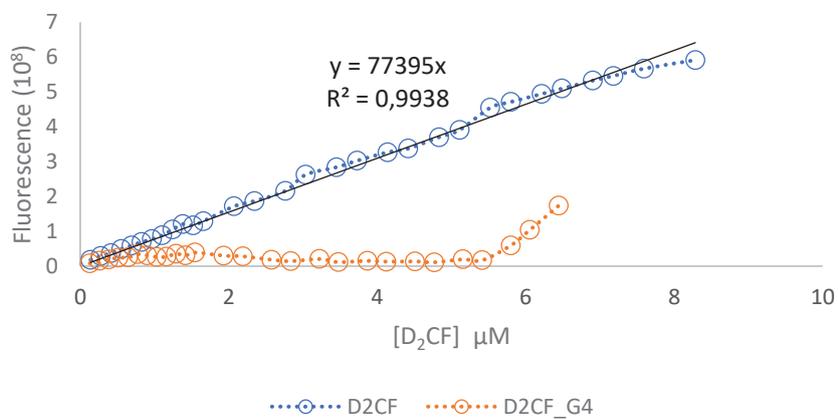


Figure 50: Fluorescence titration of III-3₂ = D₂CF with G₄ (54.0 nM)

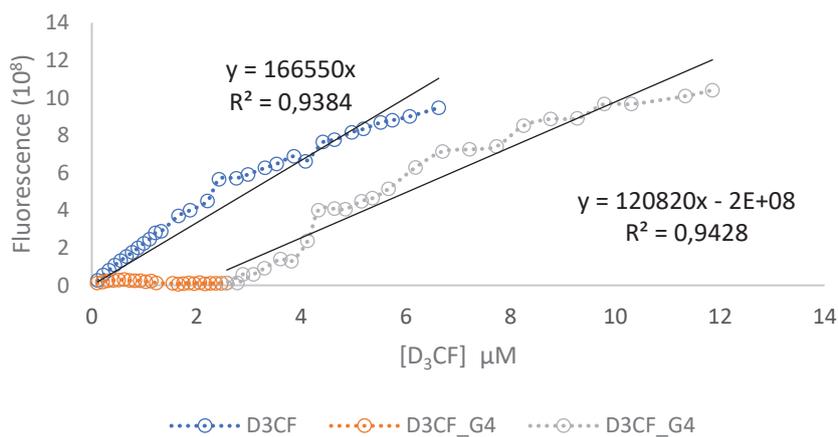


Figure 51: Fluorescence titration of III-3₃ = D₃CF with G₄ (54.0 nM)

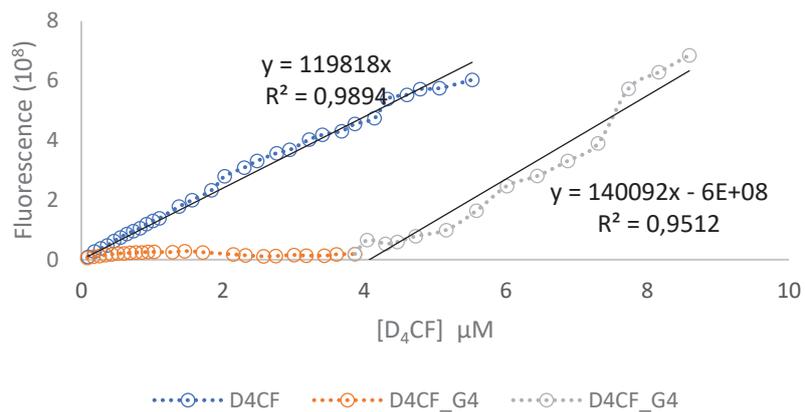


Figure 52: Fluorescence titration of III-34 = D4CF with G4 (54.0 nM)

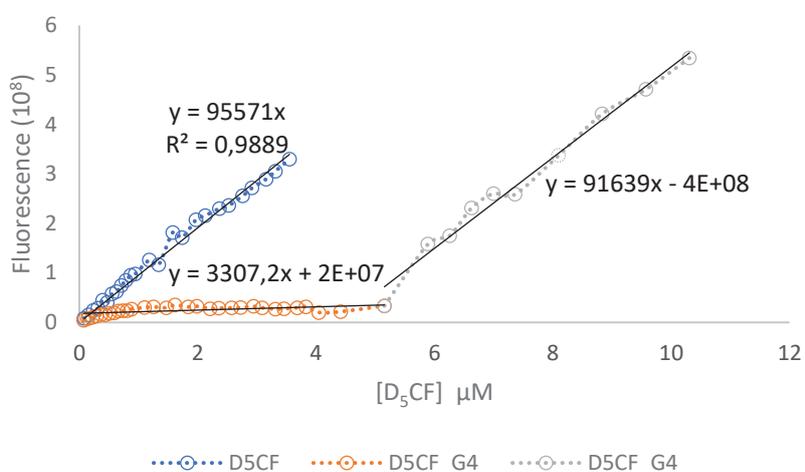


Figure 53: Fluorescence titration of III-35 = D5CF with G4 (54.0 nM)

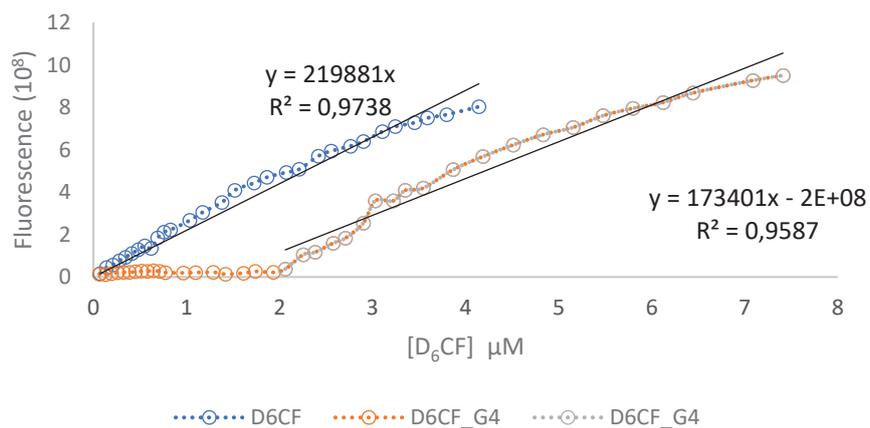


Figure 54: Fluorescence titration of III-36 = D6CF with G4 (54.0 nM)

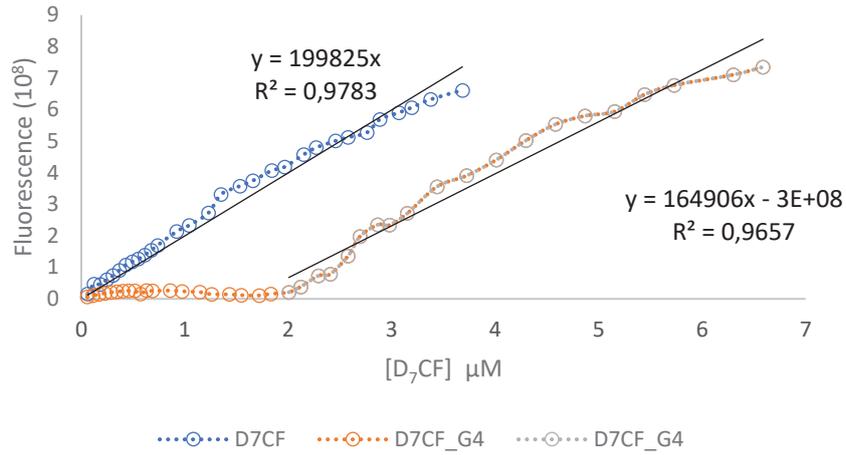


Figure 55: Fluorescence titration of III-37 = D7CF with G4 (54.0 nM)

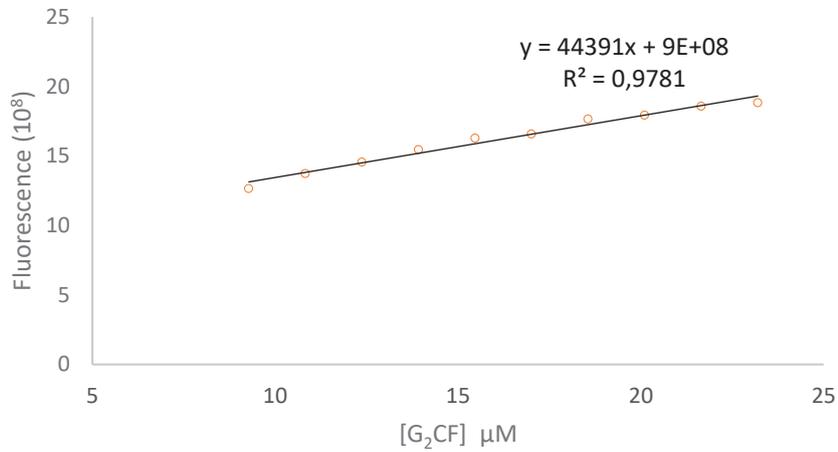


Figure 56: Fluorescence of G2-CF in HEPES buffer (10 mM, pH 7.4), range 2

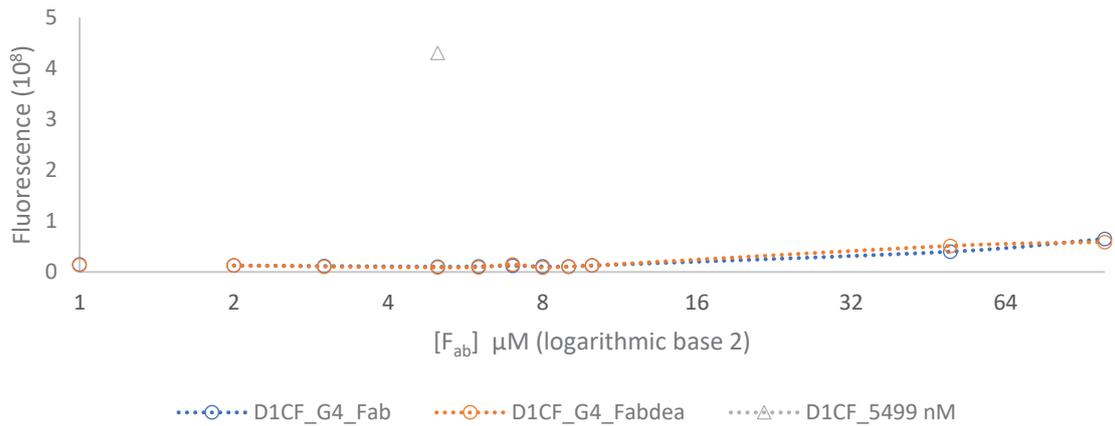


Figure 57: Fluorescence of III-31 = D1CF (5.50 μM) / G4 (54 nM) with Fab and Fab_dea

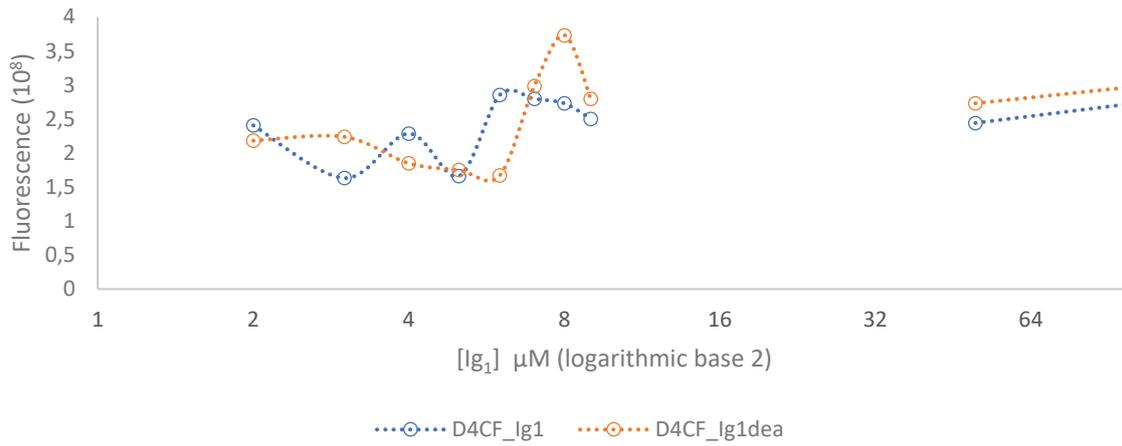


Figure 58: Fluorescence of III-3₄ = D₄CF (3.86 μM) with Ig₁ and Ig_{1_dea}

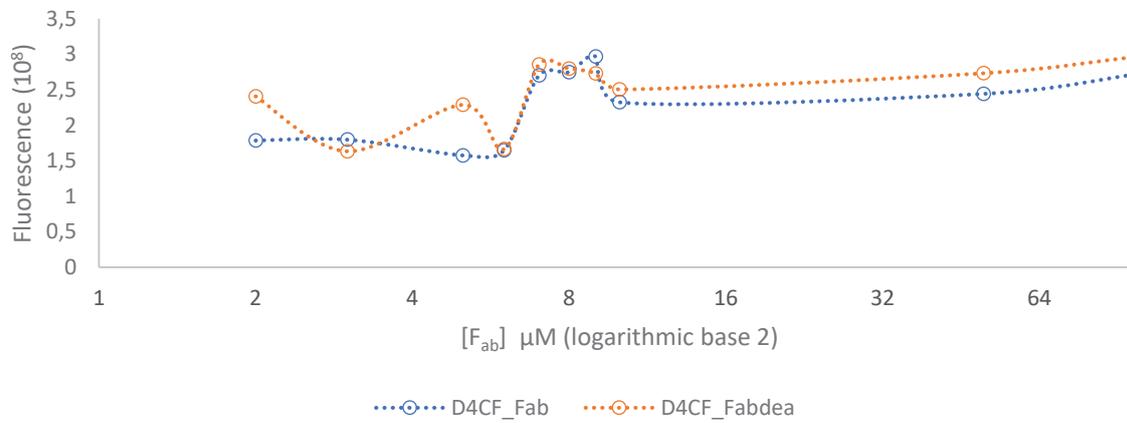


Figure 59: Fluorescence of III-3₄ = D₄CF (3.86 μM) with F_{ab} and F_{ab_dea}

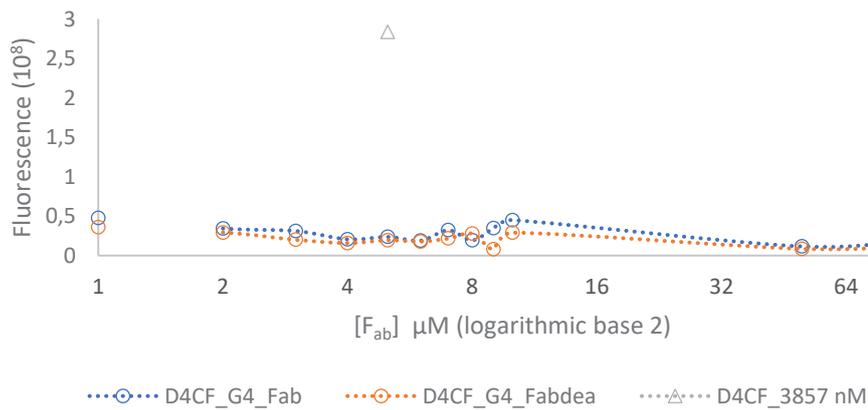


Figure 60: Fluorescence of III-3₄ = D₄CF (3.86 μM) / G₄ (54.0 nM) with F_{ab} and F_{ab_dea}

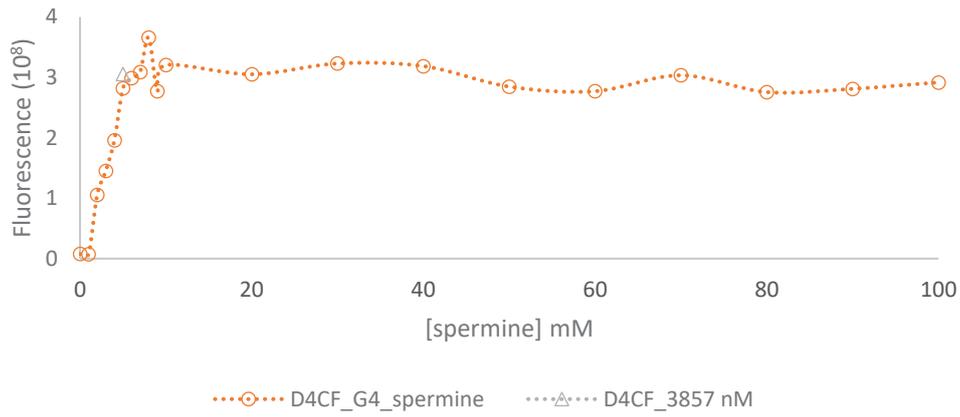


Figure 61: Fluorescence of III-34 = D₄CF (3.86 μM) / G₄ (54.0 nM) with spermine

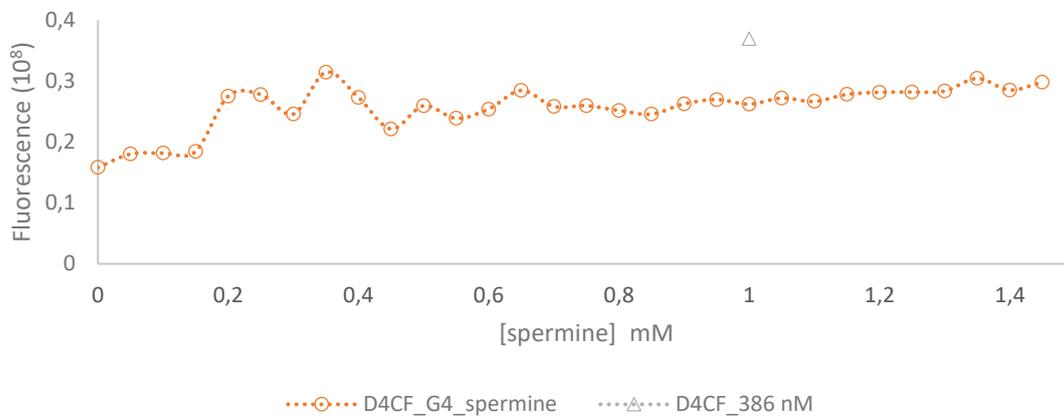


Figure 62: Fluorescence of III-34 = D₄CF (386 nM) / G₄ (5.40 nM) with spermine

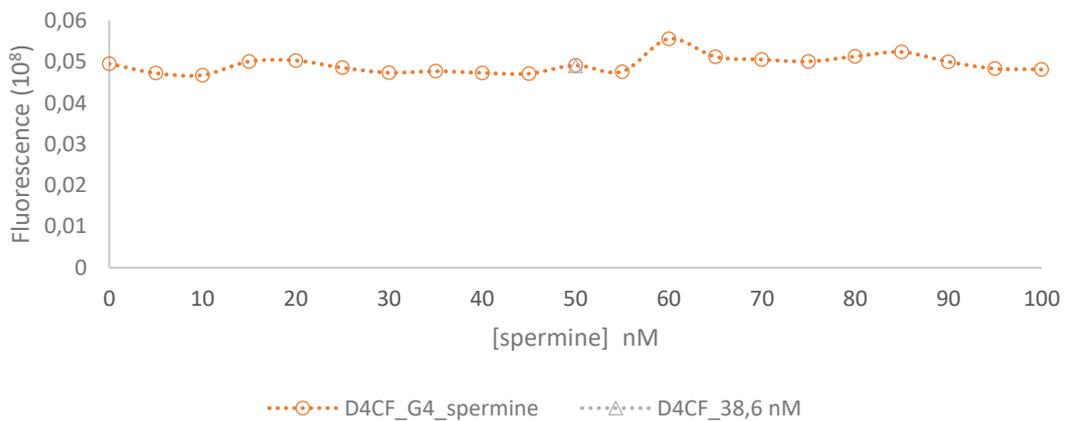


Figure 63: Fluorescence of III-34 = D₄CF (38.6 nM) / G₄ (0.540 nM) with spermine

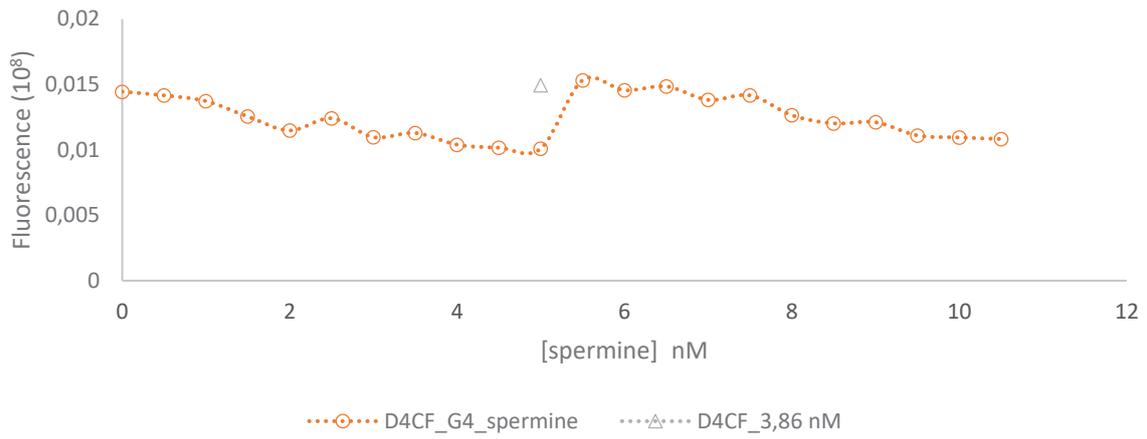


Figure 64: Fluorescence of III-3₄ = D₄CF (3,86 nM) / G₄ (0,0540 nM) with spermine

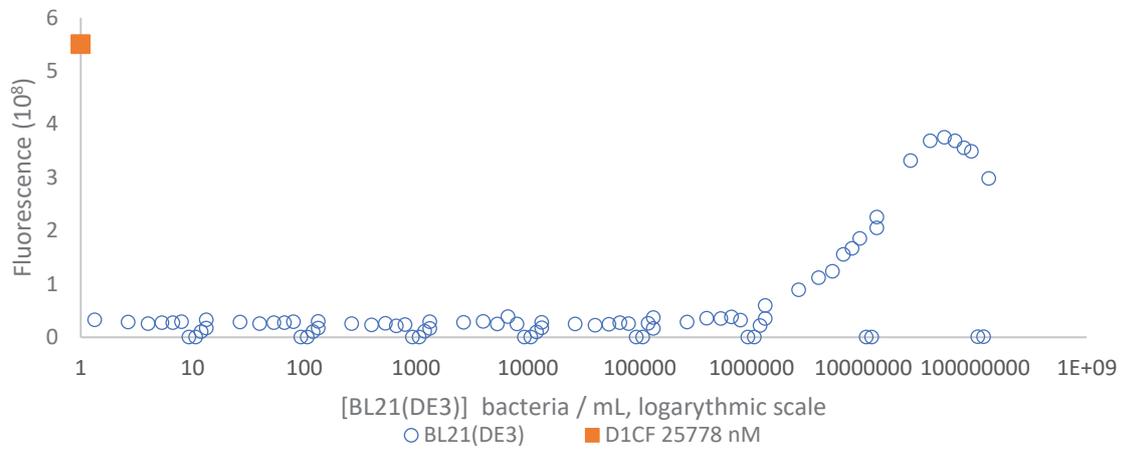


Figure 65: Fluorescence of III-3₁ = D₁CF (4,22 μM) / G₄ (54.0 nM) with BL21(DE3)

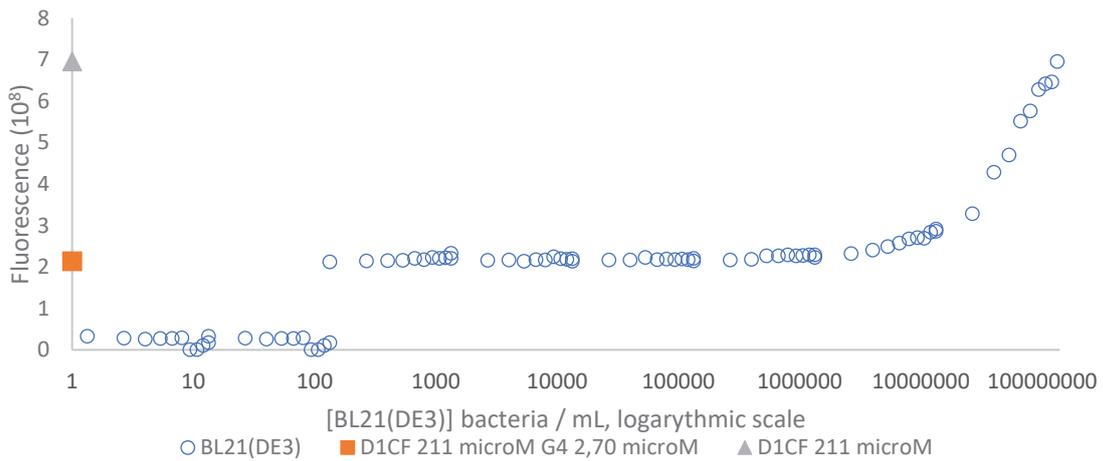


Figure 66: Fluorescence of III-3₁ = D₁CF (211 μM) / G₄ (2.70 μM) with BL21(DE3)

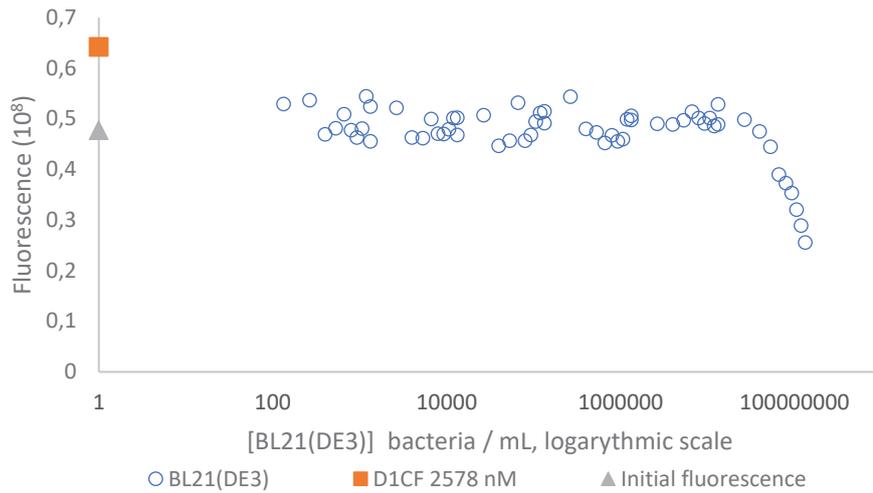


Figure 67: Fluorescence of **III-3₁** = D₁CF (422 nM) / G₄ (5.40 nM) with BL21(DE3)

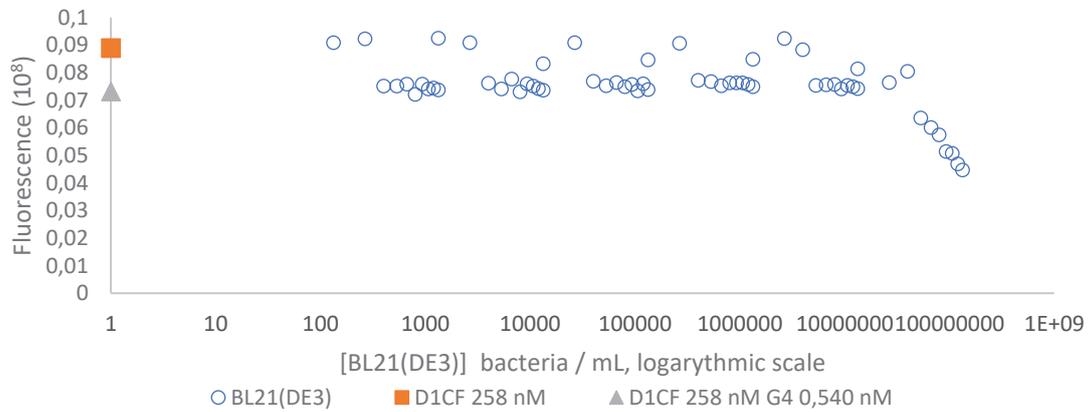


Figure 68: Fluorescence of **III-3₁** = D₁CF (42.2 nM) / G₄ (0.540 nM) with BL21(DE3)

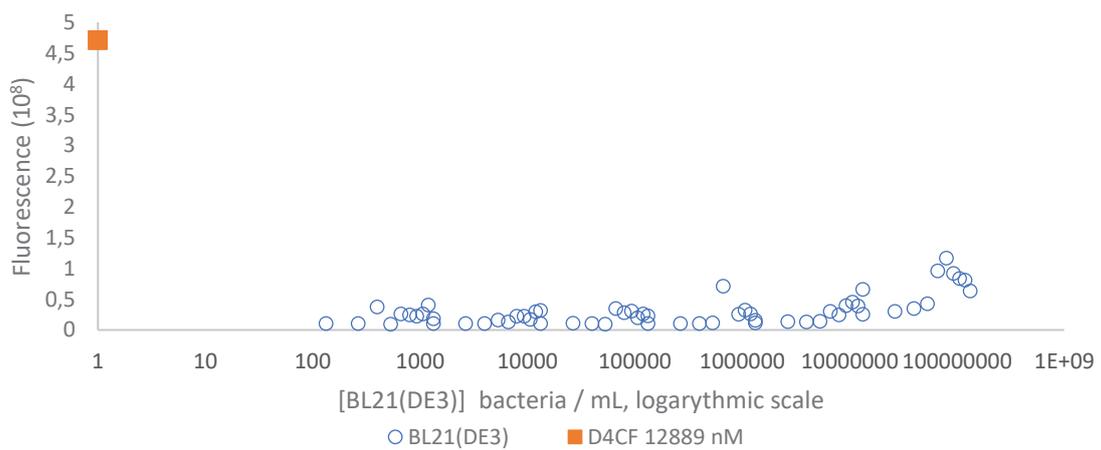


Figure 69: Fluorescence of **III-3₄** = D₄CF (12,9 μ M) / G₄ (54.0 nM) with BL21(DE3)

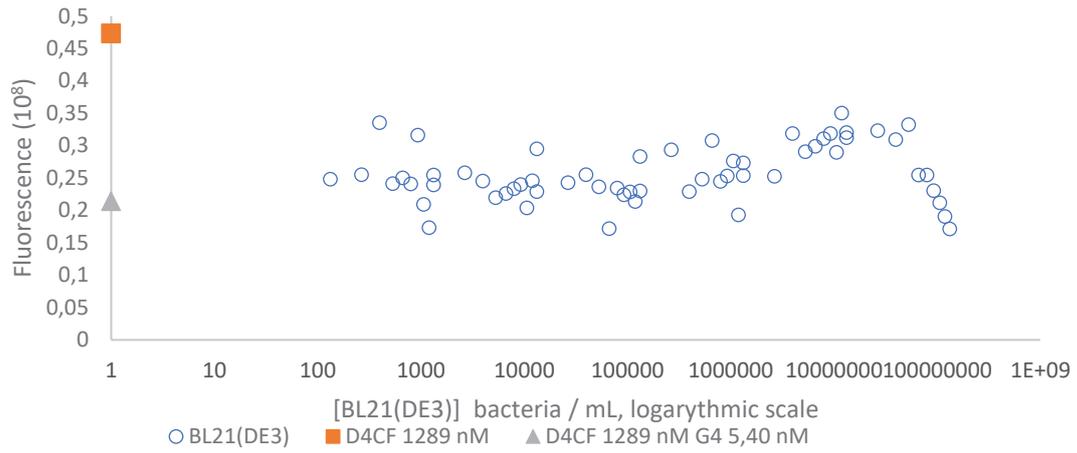


Figure 70: Fluorescence of **III-3₄** = D₄CF (1,29 μ M) / G₄ (5.40 nM) with BL21(DE3)

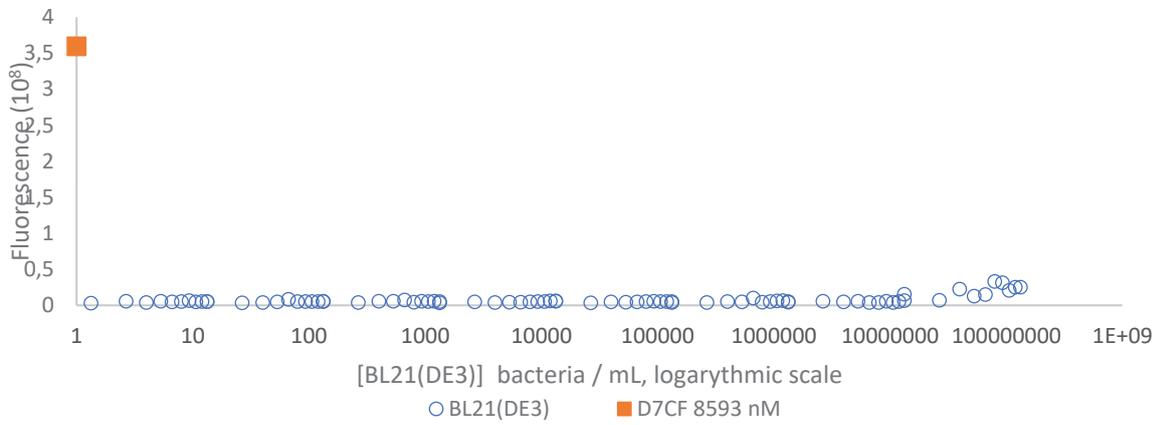


Figure 71: Fluorescence of **III-3₇** = D₇CF (8,59 μ M) / G₄ (54.0 nM) with BL21(DE3)

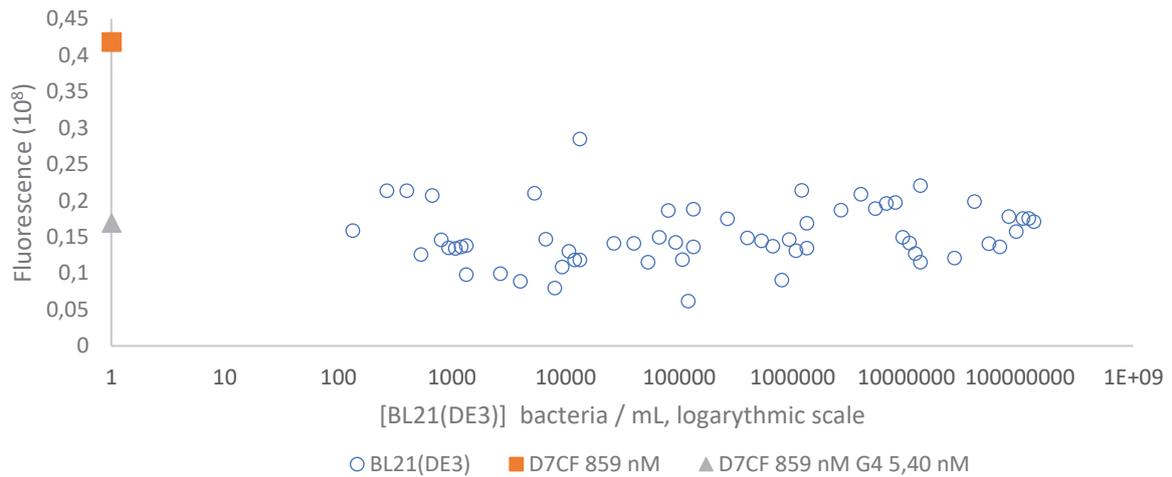


Figure 72 Fluorescence of **III-3₇** = D₇CF (859 nM) / G₄ (5.40 nM) with BL21(DE3)

c. Appendix III-III: pH titration

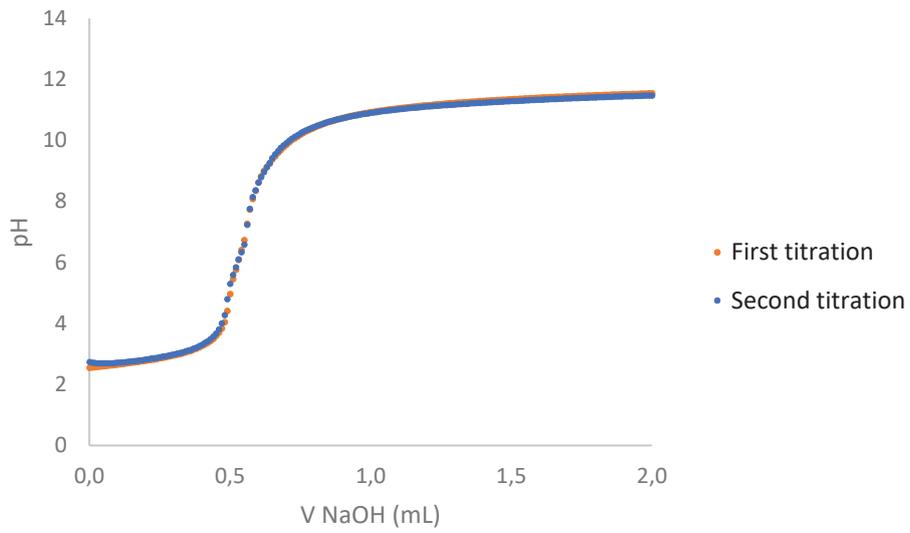


Figure 73: Titrations of G_1 25 mL [0.140 M] + 0.5 mL HCl [0.107 M] with NaOH [0.097 M]

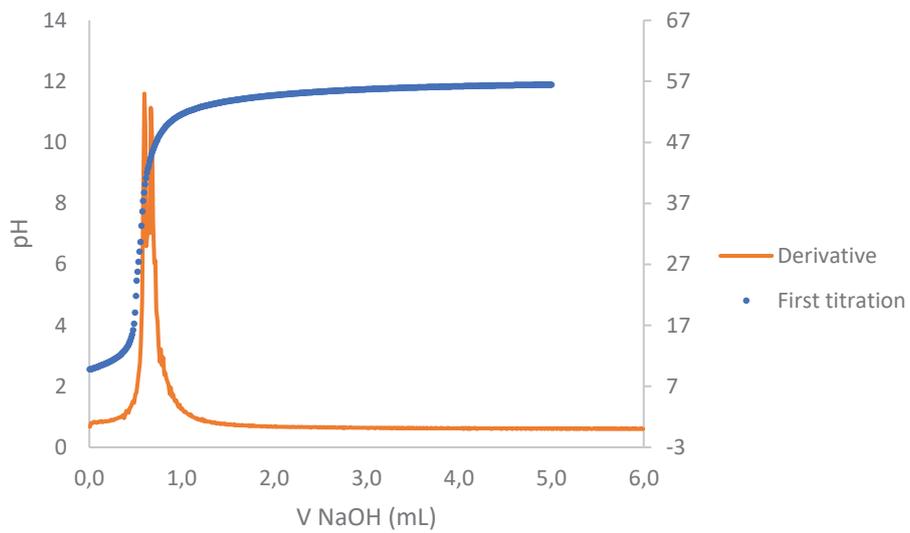


Figure 74: First titration of G_1 25 mL [0.140 M] + 0.5 mL HCl [0.107 M] with NaOH [0.097 M]. $V_{eq1} = 0.49$ mL, $V_{eq2} = 0.56$ mL

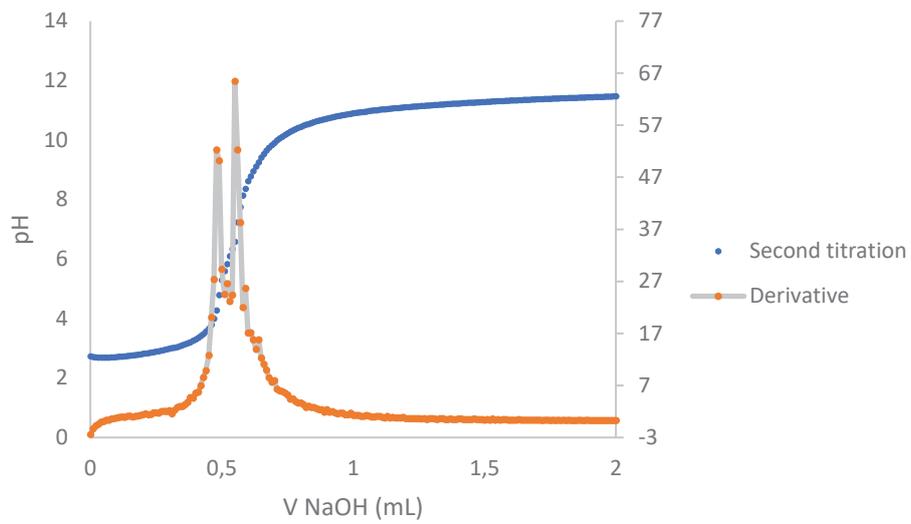


Figure 75: Second titration of G_1 25 mL [0.140 M] + 0.5 mL HCl [0.107 M] with NaOH [0.097 M]. $V_{eq1} = 0.49$ mL, $V_{eq2} = 0.56$ mL

d. Appendix III-IV: ITC

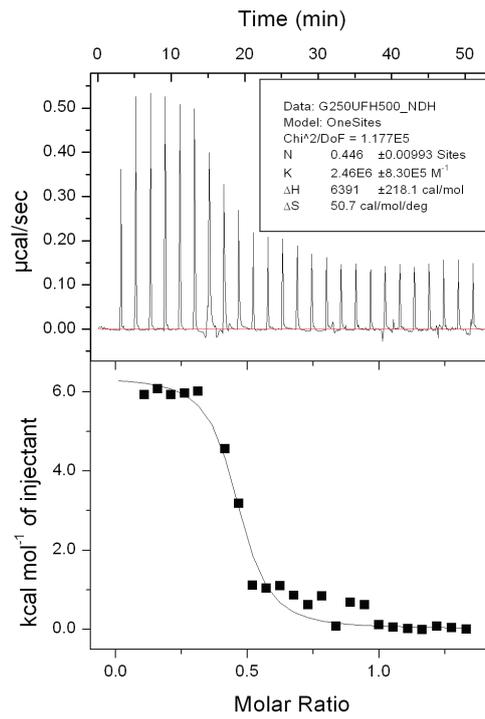


Figure 76 : Thermogram of G₂ (58 µM) and UFH (250 µM) in Tris 200 mM pH 7.4 (1)

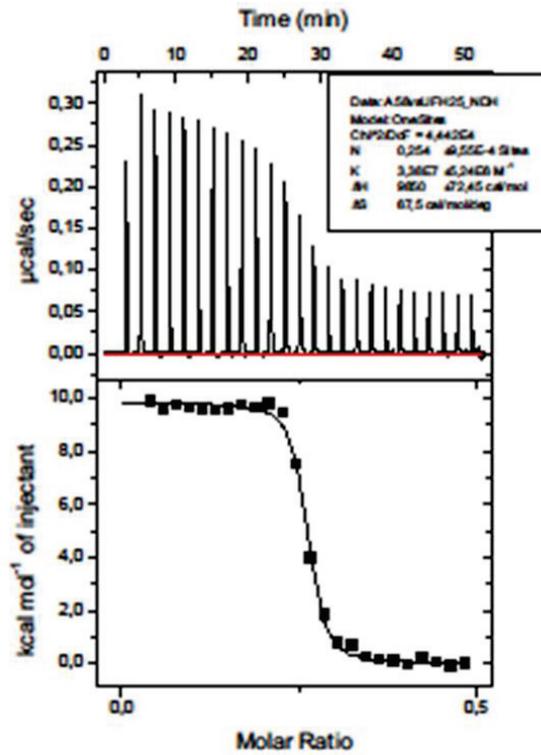
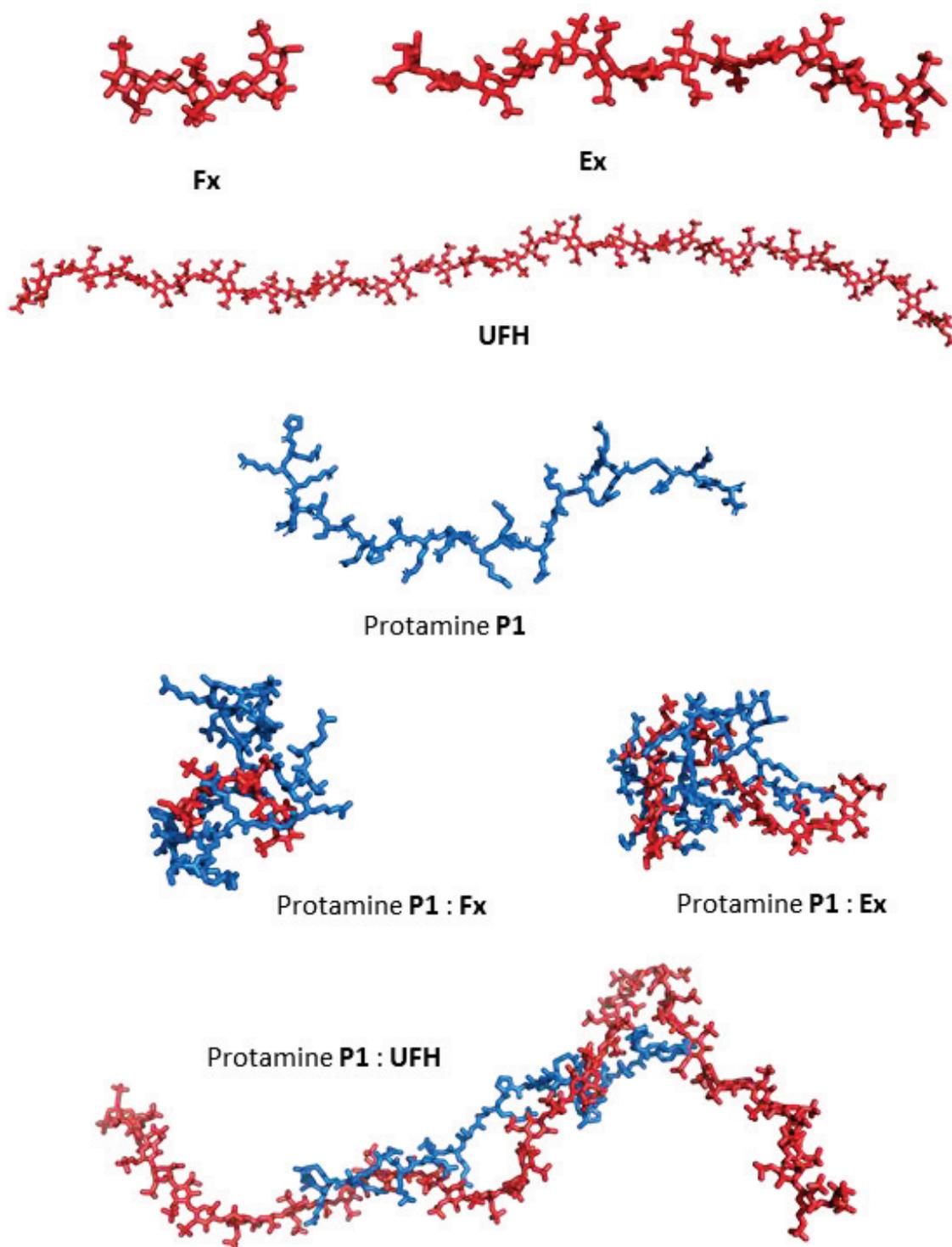
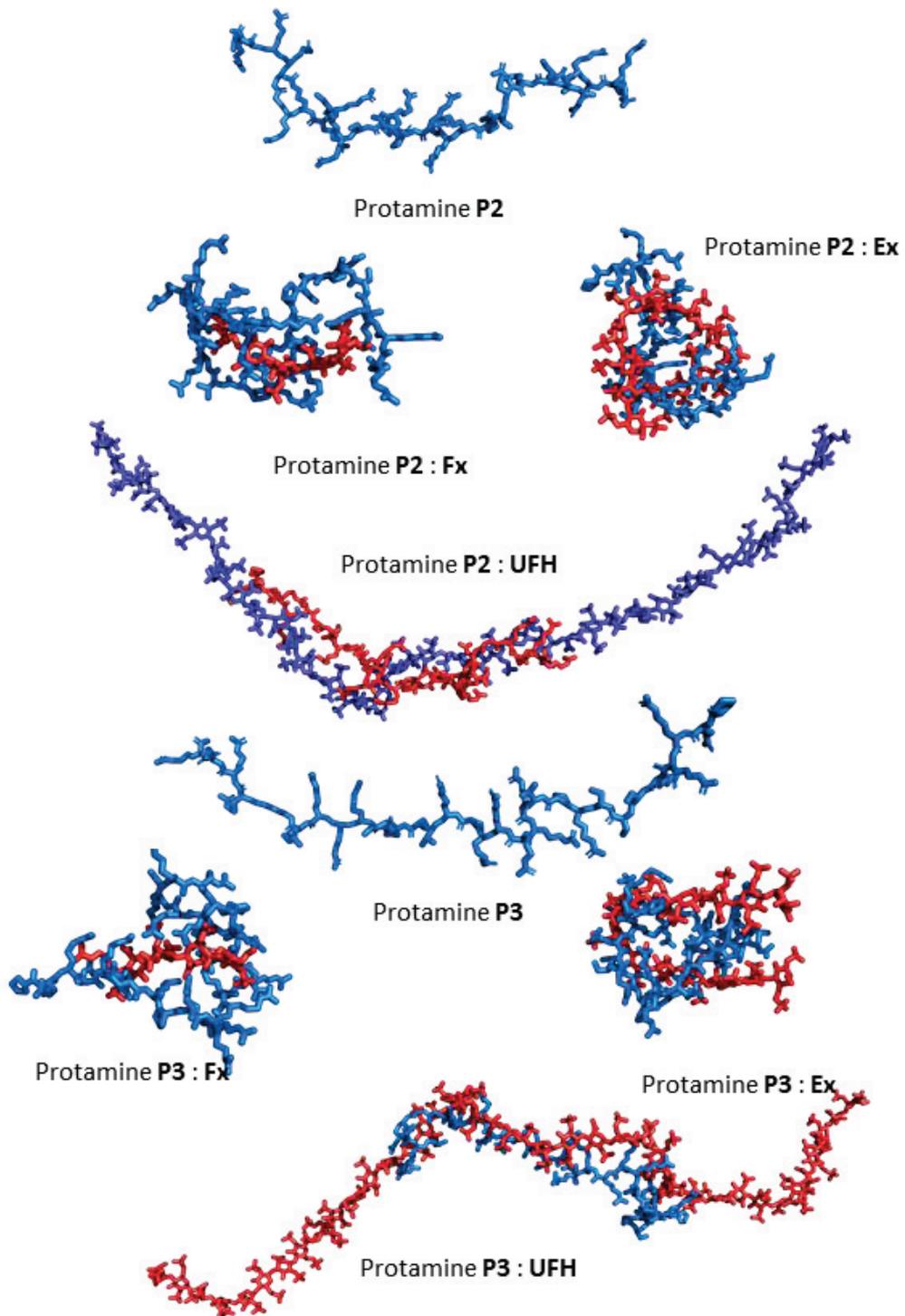


Figure 77: Thermogram of G₂ (58 µM) and UFH (250 µM) in Tris 200 mM pH 7.4 (2)

e. Appendix III-V: Supporting information regarding the MD simulations of protamines with heparins





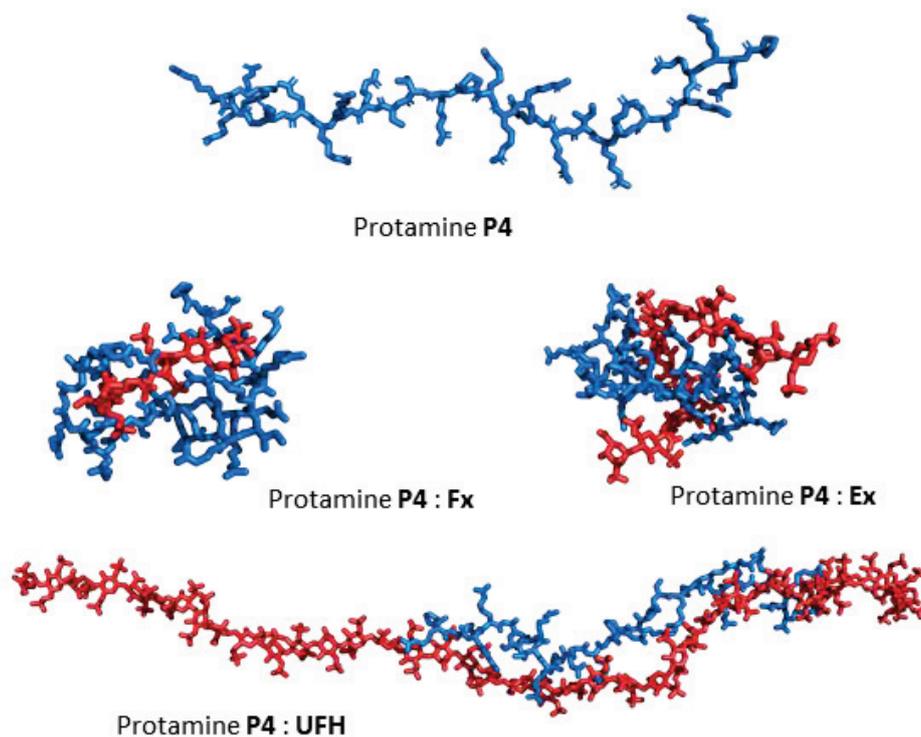


Figure 78 : Snapshots taken at the end of the MD trajectories for the stoichiometric protamine:fondaparinux, protamine:enoxaparin, and protamine:UFH. Heparins are represented as balls, and antidotes as sticks.

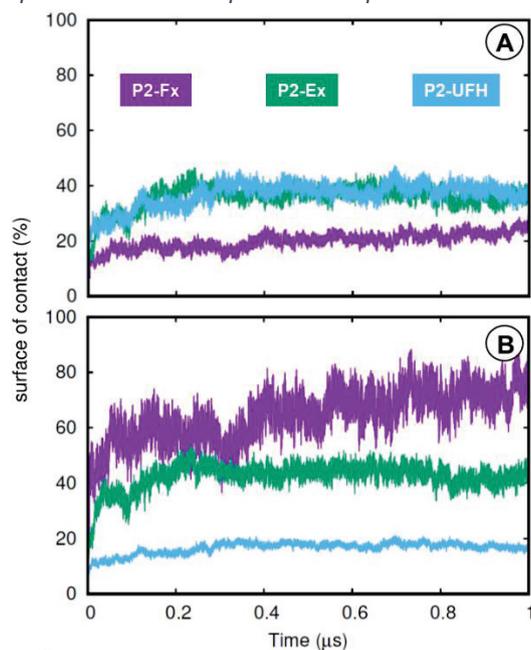


Figure 79 : Percentage of protamine **P2**'s surface (A) or heparin's surface (B) in contact with its partner within their complexes

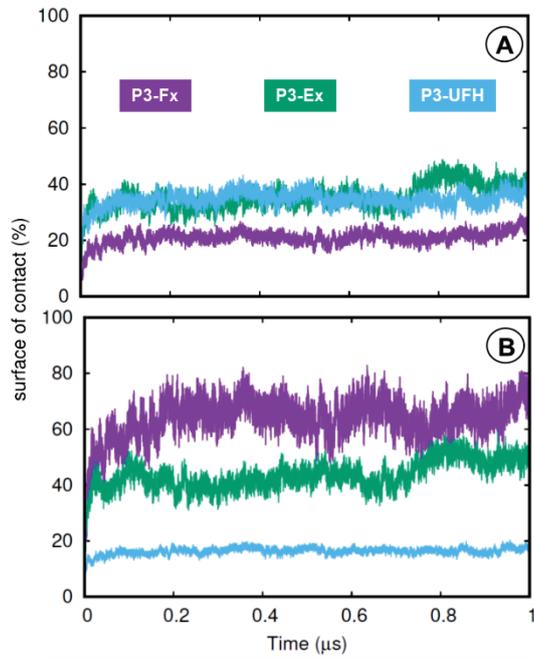


Figure 80: Percentage of protamine **P3**'s surface (A) or heparin's surface (B) in contact with its partner within their complexes

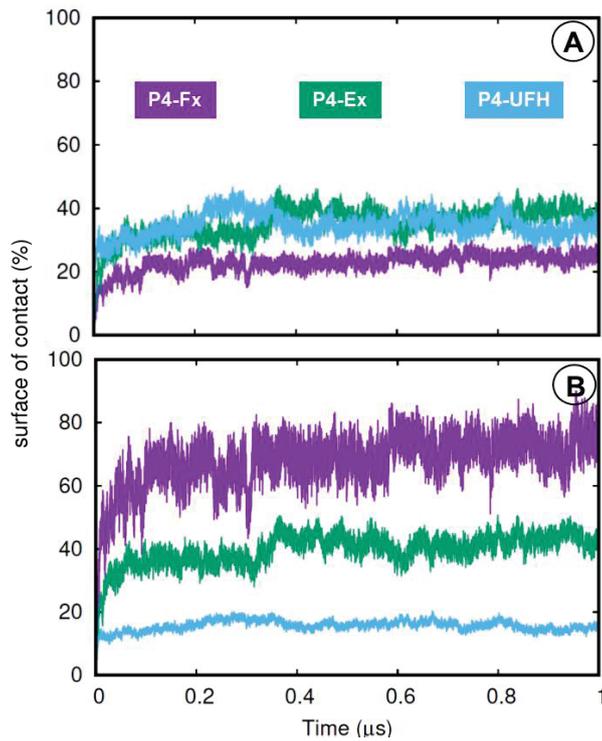


Figure 81 : Percentage of protamine **P4**'s surface (A) or heparin's surface (B) in contact with its partner within their complexes

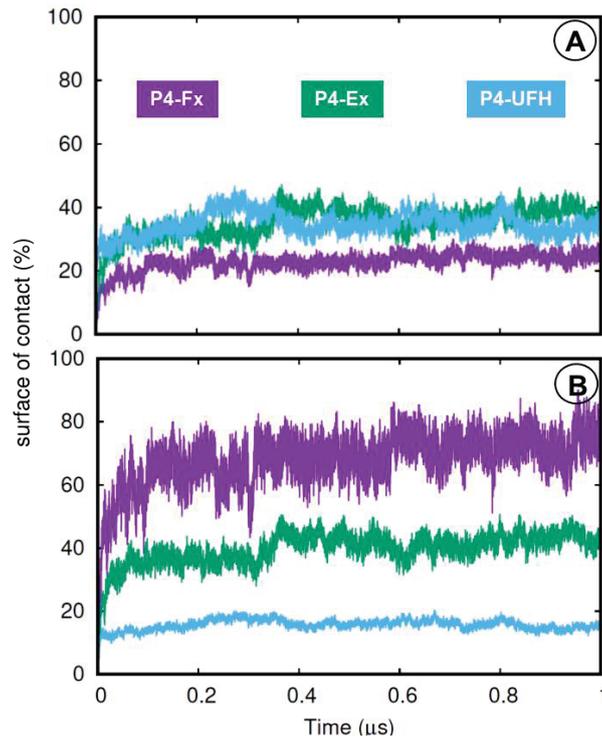


Figure 82 : Percentage of protamine $P_{2/4}$'s surface (A) or heparin's surface (B) in contact with its partner within their complexes

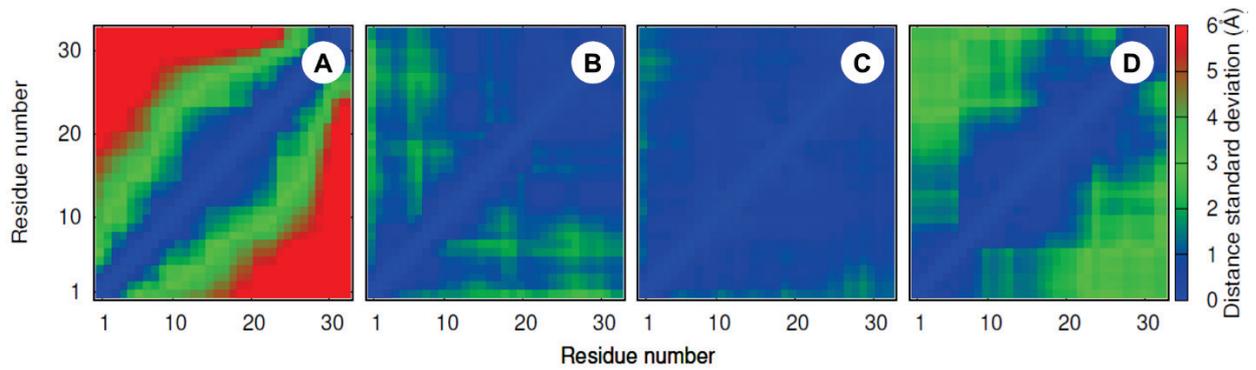


Figure 83: 2D RMS deviation maps of protamine P_2 alone (A), and within its complexes with Fx (B), Ex (C), and UFH (D).

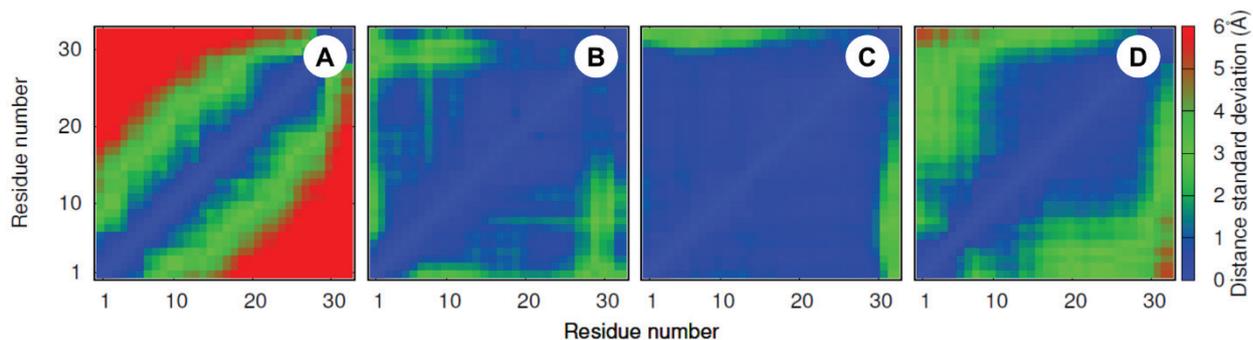


Figure 84: 2D RMS deviation maps of protamine P_3 alone (A), and within its complexes with Fx (B), Ex (C), and UFH (D).

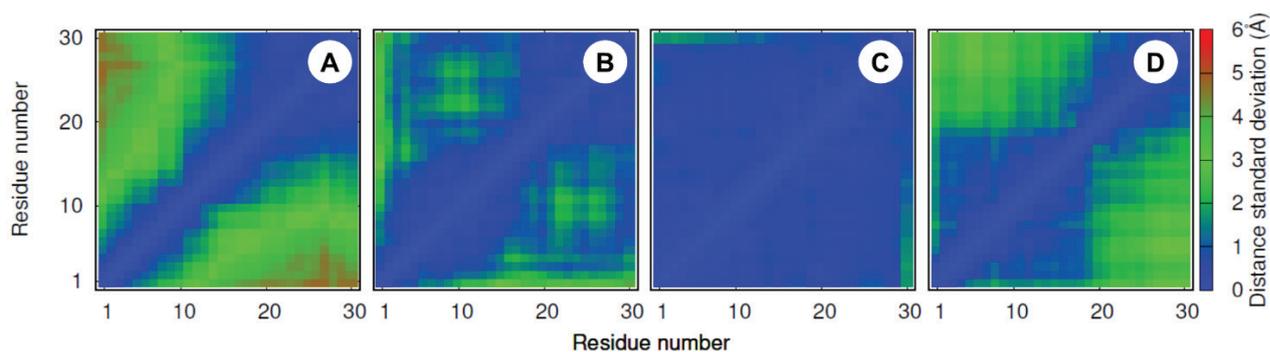


Figure 85: 2D RMS deviation maps of protamine P_4 alone (A), and within its complexes with Fx (B), Ex (C), and UFH (D).

Complex	$N_{E,hep}$	$N_{E,prot}$	$N_{H,hep}$	$N_{H,prot}$	$\Delta G_{binding}$	$N_{LIE,hep}$	$N_{LIE,prot}$
Protamine UFH	15.5 +/- 1.2	18.7 +/- 1.4	2.4 +/- 0.4	0.8 +/- 0.2	-247 +/- 29	34.3 +/- 0.4	20.0 +/- 1.1
Protamine Ex	15.4 +/- 0.4	17.4 +/- 1.0	1.5 +/- 0.3	0.5 +/- 0.1	-155 +/- 11	24.6 +/- 1.4	19.3 +/- 0.9
Protamine Fx	9.9 +/- 0.2	8.3 +/- 0.5	0.5 +/- 0.1	0.3 +/- 0.1	-80.8 +/- 4.1	11.0 +/- 0.4	14.0 +/- 0.8

Table 23: Average number of electrostatic interactions per frame engaged by heparin ($N_{E,hep}$) or protamine ($N_{E,prot}$), average number of hydrogen bonds per frame engaged by heparin ($N_{H,hep}$) or protamine ($N_{H,prot}$), average number of favorable interactions per frame engaged by heparin ($N_{LIE,hep}$) or protamine ($N_{LIE,prot}$), free energy of binding events between heparin and protamine ($\Delta G_{binding}$), energies are in $kcal.mol^{-1}$.

Hep.	Prot.	$N_{E,hep}$	$N_{E,prot}$	$N_{H,hep}$	$N_{H,prot}$	$\Delta G_{binding}$	$N_{LIE,hep}$	$N_{LIE,prot}$
UFH	P1	17.1	20.8	2.9	0.7	-268	35.0	21.1
	P2	15.4	18.0	2.1	0.5	-276	34.0	19.9
	P3	14.9	18.1	2.5	0.9	-230	34.3	20.3
	P4	14.6	17.7	2.2	1.0	-215	34.1	18.5
		15.5 ± 1.2	18.7 ± 1.4	2.4 ± 0.4	0.8 ± 0.2	-247 ± 29	34.3 ± 0.4	20.0 ± 1.1
Ex	P1	15.6	16.7	1.1	0.5	-141	22.6	19.8
	P2	15.9	18.8	1.4	0.6	-152	25.7	20.4
	P3	15.2	17.4	1.7	0.5	-167	25.5	18.6
	P4	15.0	16.7	1.7	0.3	-160	24.9	18.6
		15.4 ± 0.4	17.4 ± 1.0	1.5 ± 0.3	0.5 ± 0.1	-155 ± 11	24.6 ± 1.4	19.3 ± 0.9
Fx	P1	10.1	7.6	0.5	0.2	-82.3	10.4	13.0
	P2	9.9	8.5	0.3	0.2	-83.3	11.2	14.4
	P3	9.6	8.4	0.6	0.3	-74.6	11.3	14.0
	P4	10.0	8.7	0.5	0.3	-82.8	10.9	14.8
		9.9 ± 0.2	8.3 ± 0.5	0.5 ± 0.1	0.3 ± 0.1	$-80,8 \pm 4.1$	11.0 ± 0.4	14.0 ± 0.8

Table 24: Average number of electrostatic interactions per frame engaged by heparin ($N_{E,hep}$) or protamine ($N_{E,prot}$), average number of hydrogen bonds per frame engaged by heparin ($N_{H,hep}$) or protamine ($N_{H,prot}$), average number of favorable interactions per frame engaged by heparin ($N_{LIE,hep}$) or protamine ($N_{LIE,prot}$), free energy of binding events between heparin and protamine ($\Delta G_{binding}$), means values in bold . energies are in $kcal.mol^{-1}$

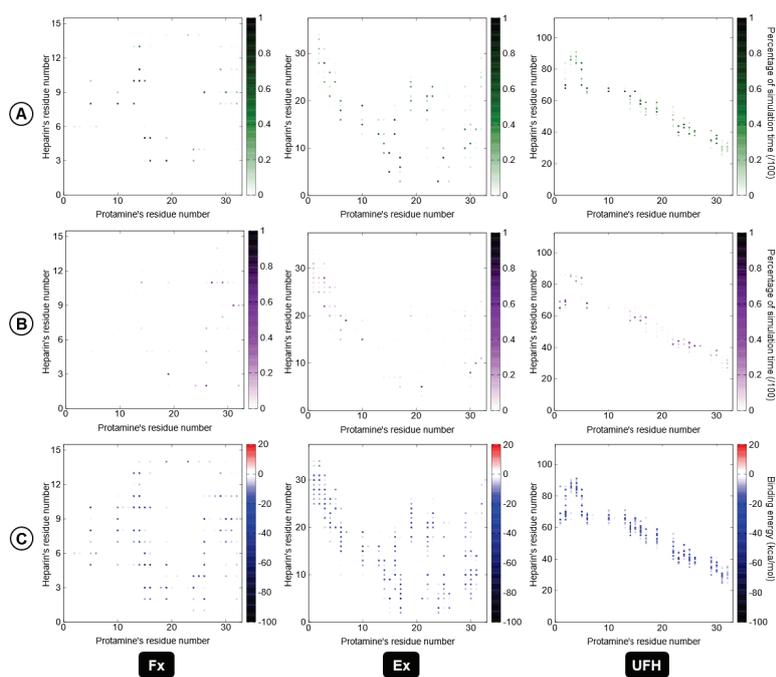


Figure 86: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine **P1** and **UFH** (1), **Ex** (2), or **Fx** (3).

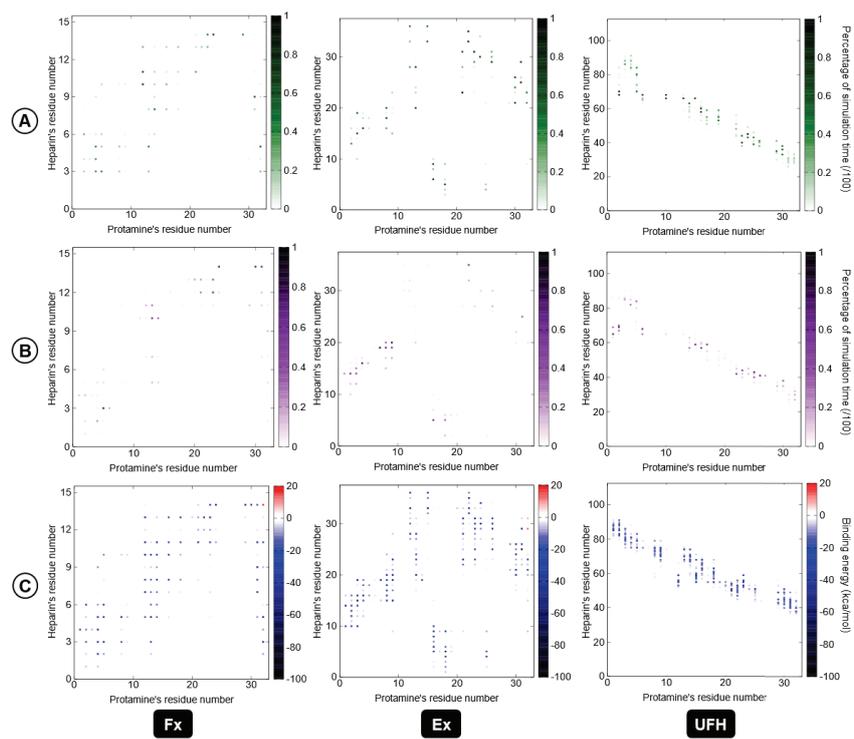


Figure 87: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine P_2 and UFH (1), Ex (2), or Fx (3).

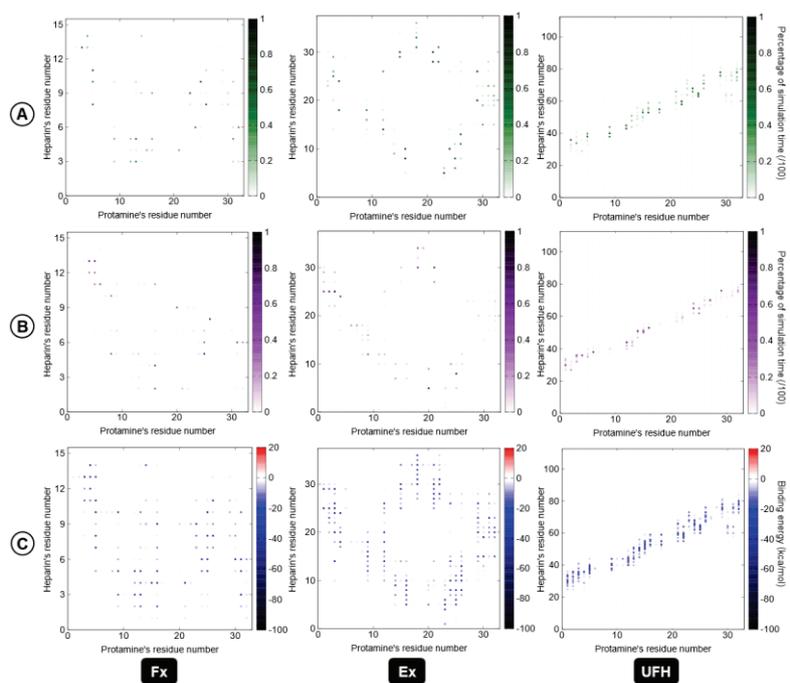


Figure 88: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine P_3 and UFH (1), Ex (2), or Fx (3).

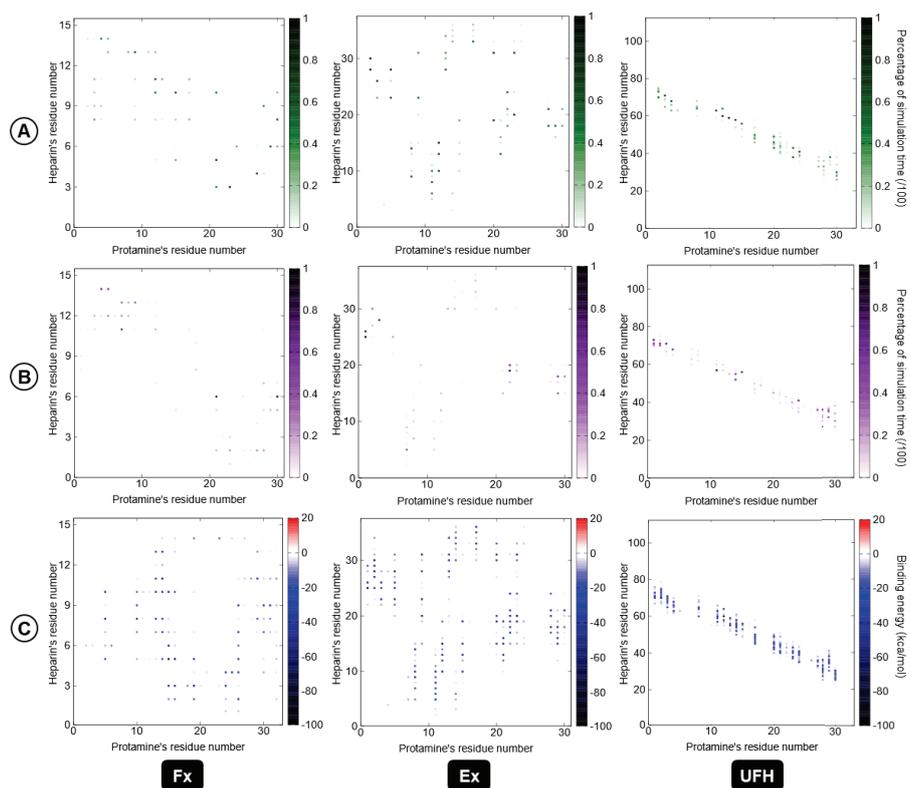
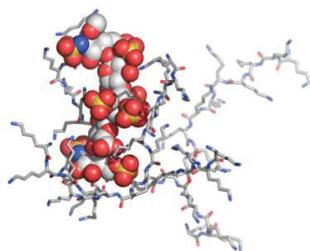
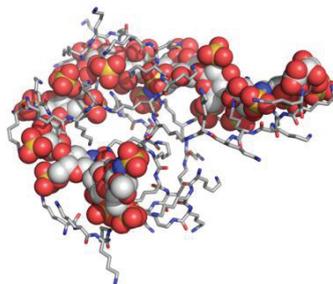


Figure 89: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine P_4 and UFH (1), Ex (2), or Fx (3).

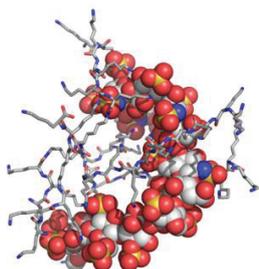
f. Appendix III-VI: Supporting information regarding the MD simulations of DGLs with heparins



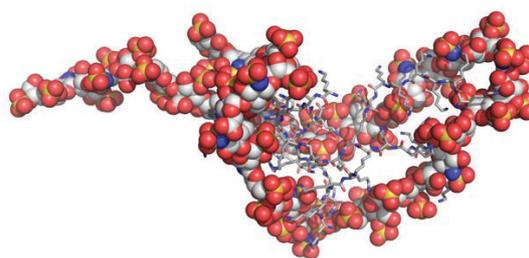
DGL G2-1:fondaparinux



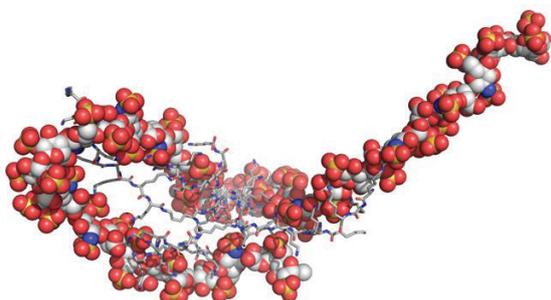
DGL G2-2:enoxaparin



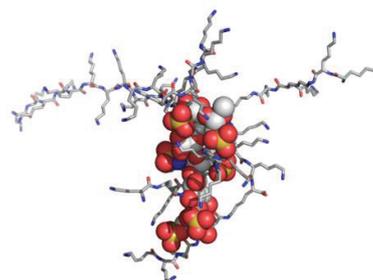
DGL G2-1:enoxaparin



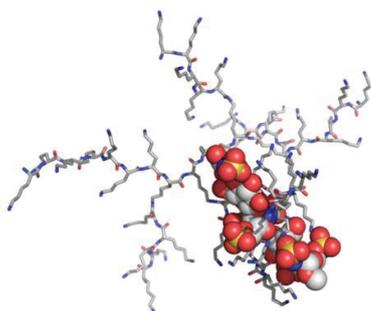
DGL G2-2:UFH



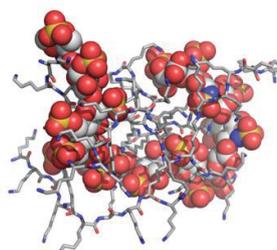
DGL G2-1:UFH



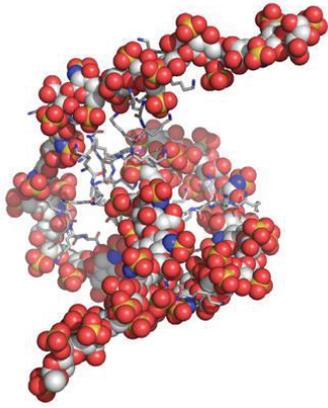
DGL G2-3:fondaparinux



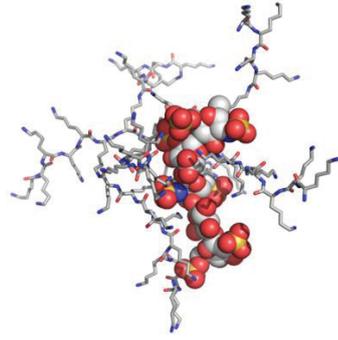
DGL G2-2:fondaparinux



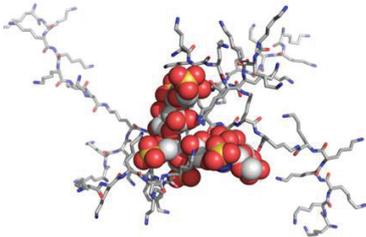
DGL G2-3:enoxaparin



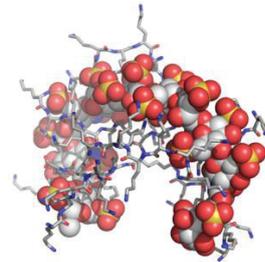
DGL G2-3:UFH



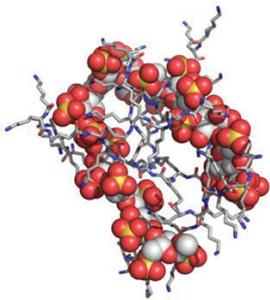
DGL G2-5:fondaparinux



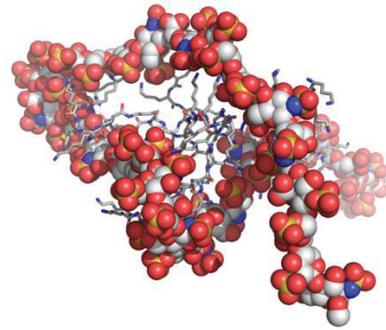
DGL G2-4:fondaparinux



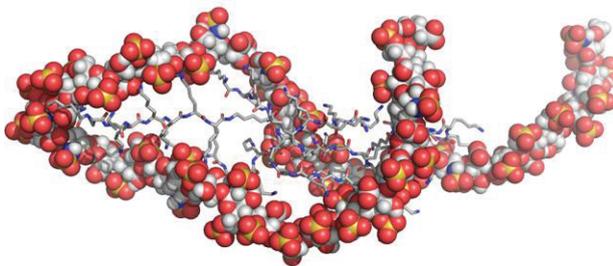
DGL G2-5:enoxaparin



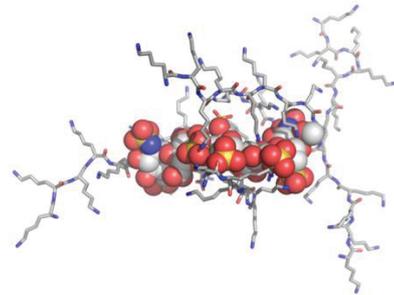
DGL G2-4:enoxaparin



DGL G2-5:UFH



DGL G2-4:UFH



DGL G2-6:fondaparinux

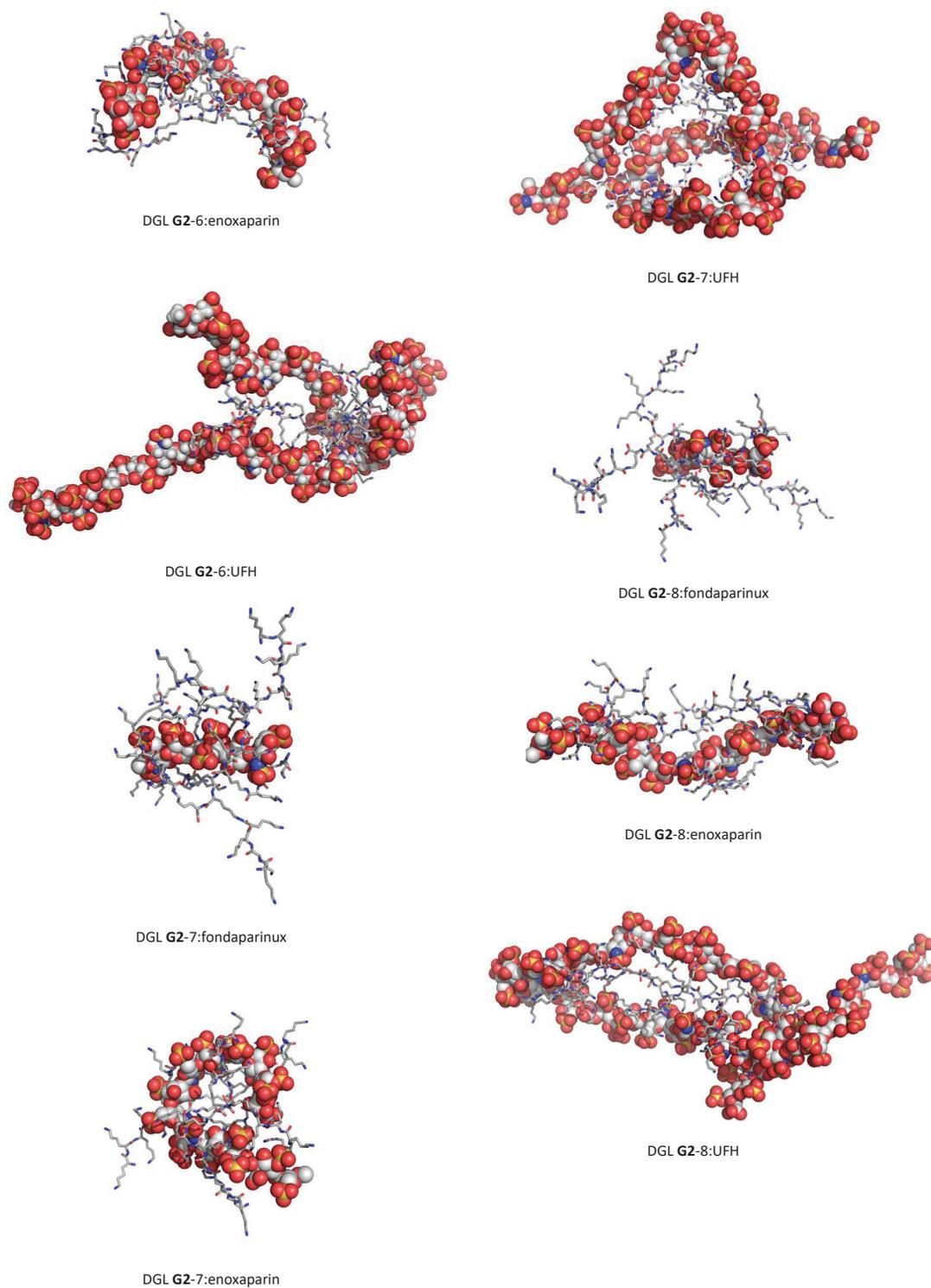


Figure 90 : Snapshots taken at the end of the MD trajectories for the stoichiometric complexes DGL G2:fondaparinux, DGL G2:enoxaparin, DGL G2:UFH, Heparins are represented as balls, and antidotes as sticks.

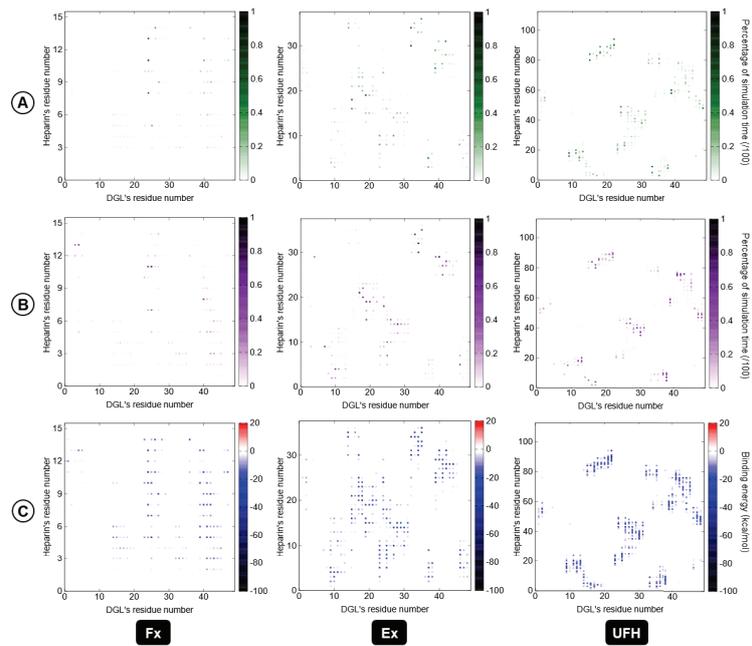


Figure 91: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine **G₂-1** and **UFH** (1), **Ex** (2), or **Fx** (3).

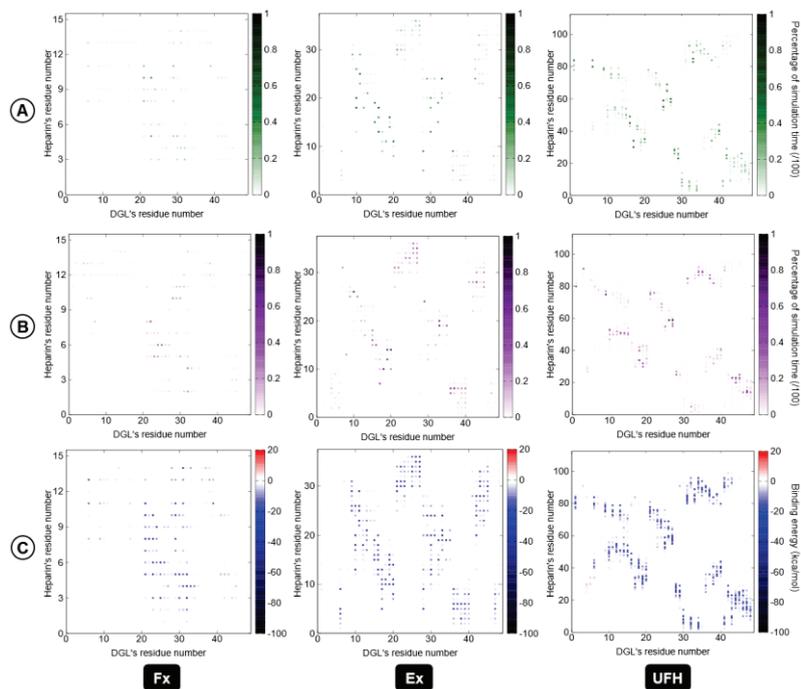


Figure 92: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine **G₂-2** and **UFH** (1), **Ex** (2), or **Fx** (3)

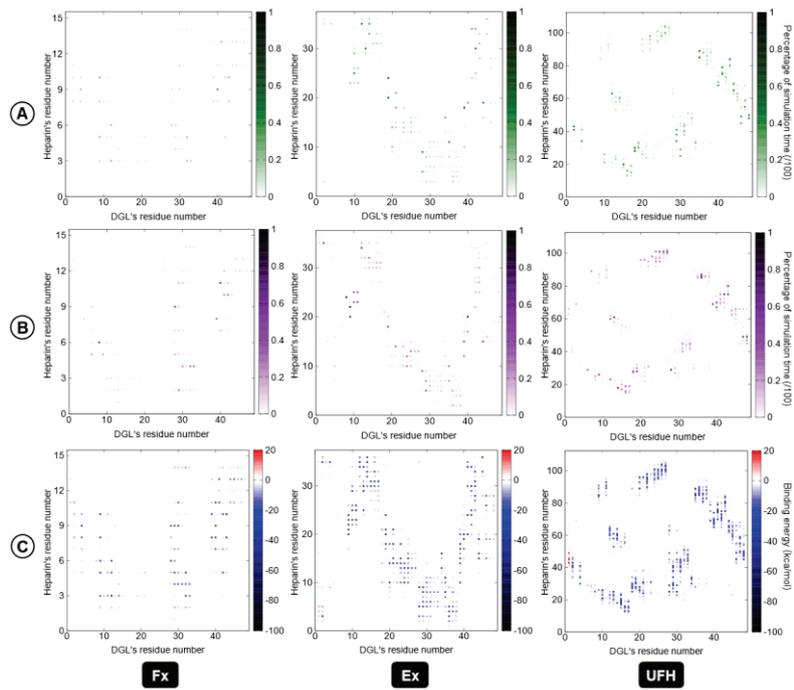


Figure 93: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine **G₂₋₃** and **UFH (1), Ex (2), or Fx (3)**

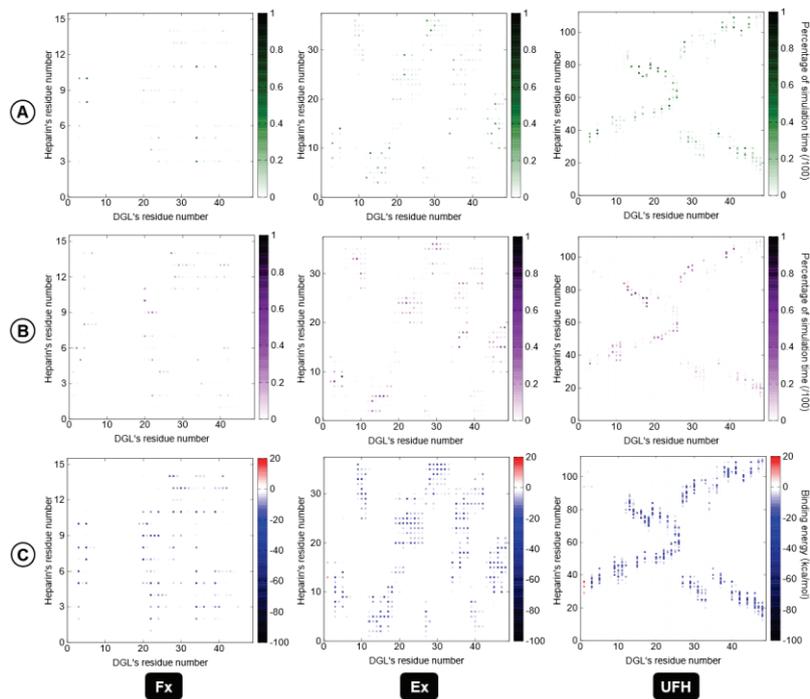


Figure 94: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine **G₂₋₄** and **UFH (1), Ex (2), or Fx (3)**

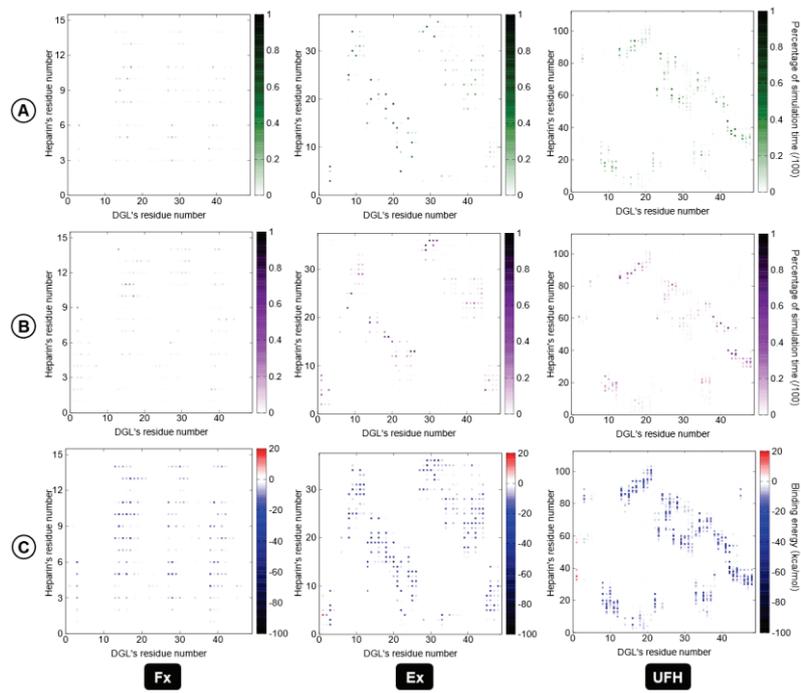


Figure 95: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine G_{2-5} and UFH (1), Ex (2), or Fx (3)

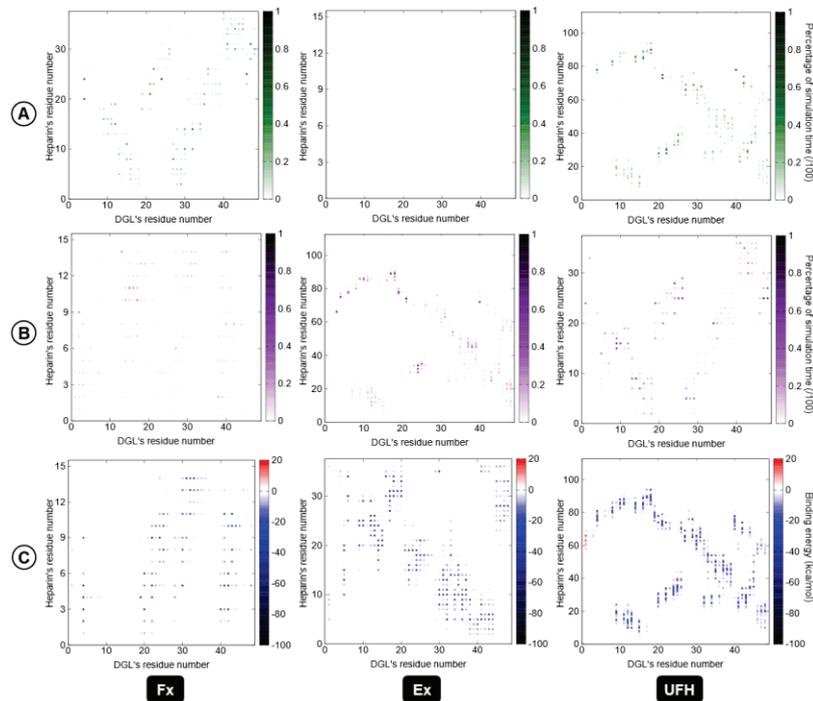


Figure 96: Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine G_{2-6} and UFH (1), Ex (2), or Fx (3)

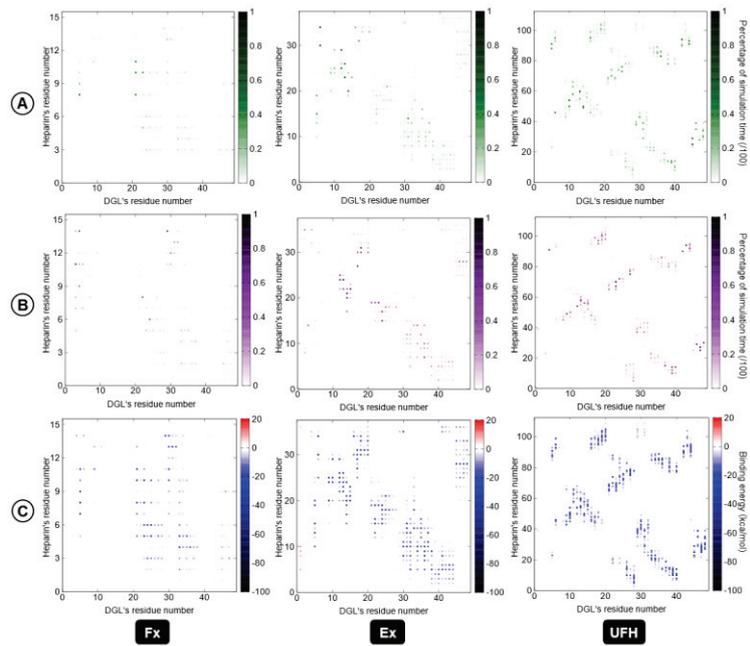


Figure 97 : Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine G_2-7 and UFH (1), Ex (2), or Fx (3)

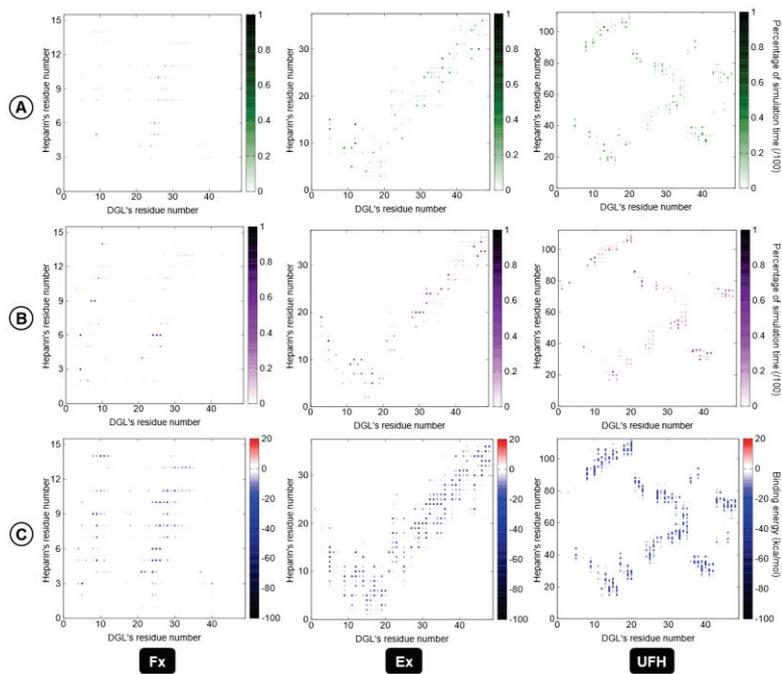


Figure 98 : Two-dimensional maps of percentage electrostatic interactions (A), percentage of hydrogen bonds (B), average binding energies (C) for each of the residues involved in the complexes between protamine G_2-8 and UFH (1), Ex (2), or Fx (3)

Appendix: Activities during the PhD

- List of publications and patents

Ourri, B., and Vial, L. Lost in (Clinical) Translation: Recent Advances in Heparin Neutralization and Monitoring, *ACS Chem. Biol.* **2019**, *14*, 2512-2526.

Ourri, B.; Francoia, J. P.; Monard, G.; Gris, J. C.; Leclaire, J.; Vial, L., Dendrigraft of Poly-L-lysine as a Promising Candidate To Reverse Heparin-based Anticoagulants in Clinical Settings. *ACS Med Chem Lett* **2019**, *10* (6), 917-922

Jeamet, E.; Septavaux, J.; Héloin, A.; Donnier-Maréchal, M.; Dumartin, M.; Ourri, B.; Mandal, P.; Huc, I.; Bignon, E.; Dumont, E.; Morell, C.; Francoia, J.-P.; Perret, F.; Vial, L.; Leclaire, J. *Chem. Sci.* **2019**, *10*, 277

Dauvergne, J.; Jawhari, A.; Lessault, M.; Ourri, B. Steroidic calixarene compounds, their synthesis and use for extraction, solubilization and/or stabilization of native and functional membrane proteins. 2017 WO2017134087

Ren, H.; Ourri, B.; Jeanneau, E.; Tu, T.; Bonnamour, I.; Darbost, U., Synthesis and crystal structures of palladium(II) complexes of bis-(N-heterocyclic carbenes) on a calix[4]arene platform. *Transition Metal Chemistry* **2016**, *41*, 827

Ourri, B.; Tillement, O.; Tu, T.; Jeanneau, E.; Darbost, U.; Bonnamour, I., Copper complexes bearing NHC-calixarene unit: synthesis and application in click chemistry. *New Journal of Chemistry* **2016**, *40* (11), 9477 - 9485

De Proft, F.; Forquet, V.; Ourri, B.; Chermette, H.; Geerlings, P.; Morell, C., Investigation of electron density changes at the onset of a chemical reaction using the state-specific dual descriptor from conceptual density functional theory. *Physical Chemistry Chemical Physics* **2015**, *17* (14), 9359-9368

Dailier, D.; Danoun, G.; Ourri, B.; Baudoin, O., Divergent Synthesis of Aeruginosins Based on a C(sp³)-H Activation Strategy. *Chemistry – A European Journal* **2015**, *21* (26), 9370-9379

Invited: Actualité chimique, un point sur: Ourri, B.; Vial, L.

Manuscript in preparation : Ourri, B.; Monnard, G; Vial, L. About the Binding of Protamine with Heparin of Various Lengths from Molecular Dynamics Simulations

- List of posters and communications

SupraLyon, december, 2018, poster

French chemical society, conference of Rhône Alpes region, october, 2018, poster

Olympiad of chemistry, invited speaker, september, 2018

ICBMS day, event of the institute, july, 2018, poster

Chemical doctoral school of Lyon conference, june, 2018, poster

International Symposium on Macrocyclic and Supramolecular Chemistry (ISMSC) in conjunction with ISACS: Challenges in Organic Materials & Supramolecular Chemistry, Cambridge, July, 2017, poster

Oxford Synthesis Summer Conference, August 2016, poster

- Teaching activities

Part-time lecture (vacataire), CPE Lyon, practical courses of synthetic chemistry 136 hours, 2018-2019

Part-time lecture (vacataire), UCBL, DAEU-B general chemistry 40 hours, 2017-2018

Part-time lecture (vacataire), Lycée La Martinière Monplaisir, oral examination 80 hours, 2017-2018

Invited speaker, Olympiades de la Chimie, September, 2018

- Courses

Ecole de l'innovation thérapeutique, Rochefort en Yvelines, July, 2018

Recherche et industrie : les enjeux de la Propriété Intellectuelle, April, 2017

Le brevet, une mine d'information. Comment rechercher, lire vite et bien un brevet ? Quels outils gratuits utiliser pour quels résultats ?, April, 2017

- Miscellaneous

Expert étudiant/Student assessor, HCERES, 2017-2019. Participation à l'évaluation du CNAM (2017), de l'Université de Lille (2018), du réseau Figure (2019) et de l'Université de Skoltech (Russie, 2019)

References

- (1) Voet, D.; Voet, J. G.; Pratt, C. W. *Fundamentals of biochemistry life at the molecular level*, 2016.
- (2) Ma, J. C.; Dougherty, D. A. *Chem. Rev.* **1997**, *97*, 1303.
- (3) (a) Izatt, R. M.; Pawlak, K.; Bradshaw, J. S.; Bruening, R. L. *Chem. Rev.* **1995**, *95*, 2529(b) Mordasini Denti, T. Z.; van Gunsteren, W. F.; Diederich, F. *J. Am. Chem. Soc.* **1996**, *118*, 6044(c) Piatnitski, E. L.; Flowers II, R. A.; Deshayes, K. *Chemistry – A European Journal* **2000**, *6*, 999(d) Hillyer, M. B.; Gibb, B. C. *Annu. Rev. Phys. Chem.* **2016**, *67*, 307.
- (4) (a) Liu, Y.; An, S.; Li, J.; Kuang, Y.; He, X.; Guo, Y.; Ma, H.; Zhang, Y.; Ji, B.; Jiang, C. *Biomaterials* **2016**, *80*, 33(b) Romestand, B.; Rolland, J. L.; Commeyras, A.; Coussot, G.; Desvignes, I.; Pascal, R.; Vandenabeele-Trambouze, O. *Biomacromolecules* **2010**, *11*, 1169(c) Sun, T.; Jiang, X.; Wang, Q.; Chen, Q.; Lu, Y.; Liu, L.; Zhang, Y.; He, X.; Ruan, C.; Zhang, Y.; Guo, Q.; Liu, Y.; Jiang, C. *ACS Appl. Mater. Interfaces* **2017**, *9*, 34603.
- (5) Wolle, P.; Müller, M. P.; Rauh, D. *ACS Chem. Biol.* **2019**.
- (6) (a) Werner, T. *Journal für Praktische Chemie/Chemiker-Zeitung* **1994**, 336, 474(b) Diederich, F. *Cyclophanes*; Royal Society of Chemistry, 1991.
- (7) (a) Biedermann, F.; Uzunova, V. D.; Scherman, O. A.; Nau, W. M.; De Simone, A. *J. Am. Chem. Soc.* **2012**, *134*, 15318(b) Ramsey, S.; Nguyen, C.; Salomon-Ferrer, R.; Walker, R. C.; Gilson, M. K.; Kurtzman, T. *J. Comput. Chem.* **2016**, *37*, 2029(c) Nguyen, C. N.; Kurtzman, T.; Gilson, M. K. *J. Chem. Theory Comput.* **2016**, *12*, 414.
- (8) Cram, D. J. *Science* **1983**, *219*, 1177.
- (9) Cheng, L.; Zhang, H.; Dong, Y.; Zhao, Y.; Yu, Y.; Cao, L. *Chem. Commun.* **2019**, 55, 2372.
- (10) (a) Chin-Yang, Y.; L., T. M. *Angew. Chem. Int. Ed.* **2006**, *45*, 7797(b) Watson, M. D.; Jäckel, F.; Severin, N.; Rabe, J. P.; Müllen, K. *J. Am. Chem. Soc.* **2004**, *126*, 1402.
- (11) Michel, P. *Recueil des Travaux Chimiques des Pays-Bas et de la Belgique* **1899**, *18*, 457.
- (12) (a) Brown, C. J.; Farthing, A. C. *Nature* **1949**, *164*, 915(b) Cram, D. J.; Steinberg, H. *J. Am. Chem. Soc.* **1951**, *73*, 5691.
- (13) Mullins, S. *J. Chem. Technol. Biotechnol.* **1991**, *50*, 293.
- (14) (a) Ogoshi, T.; Yamagishi, T.-a.; Nakamoto, Y. *Chem. Rev.* **2016**, *116*, 7937(b) Yang, K.; Pei, Y.; Wen, J.; Pei, Z. *Chem. Commun.* **2016**, 52, 9316.
- (15) Cram, D. J.; Cram, J. M. *Acc. Chem. Res.* **1971**, *4*, 204.
- (16) Vial, L.; Ludlow, R. F.; Leclaire, J.; Pérez-Fernández, R.; Otto, S. *J. Am. Chem. Soc.* **2006**, *128*, 10253.
- (17) (a) Dobrowolski, M. A.; Cyrański, M. K.; Merner, B. L.; Bodwell, G. J.; Wu, J. I.; Schleyer, P. v. R. *The Journal of Organic Chemistry* **2008**, *73*, 8001(b) Bodwell, G. J.; Bridson, J. N.; Cyrański, M. K.; Kennedy, J. W. J.; Krygowski, T. M.; Mannion, M. R.; Miller, D. O. *The Journal of Organic Chemistry* **2003**, *68*, 2089.
- (18) (a) Mireles, N.; Salcedo, R.; Sansores, L.; Martinez, A. *Int. J. Quantum Chem.* **2000**, *80*, 258(b) Nozawa, R.; Kim, J.; Oh, J.; Lamping, A.; Wang, Y.; Shimizu, S.; Hisaki, I.; Kowalczyk, T.; Fliegl, H.; Kim, D.; Shinokubo, H. *Nature Communications* **2019**, *10*, 3576.
- (19) (a) In *Modern Cyclophane Chemistry*(b) Atwood, J. L.; Lehn, J. M. *Comprehensive supramolecular chemistry*; Pergamon: New York, 1996.
- (20) Diederich, F.; Griebel, D. *J. Am. Chem. Soc.* **1984**, *106*, 8037.
- (21) Murakami, Y.; Hayashida, O.; Ito, T.; Hisaeda, Y. In *Pure Appl. Chem.*, 1993; Vol. 65.
- (22) Dougherty, D. A. *Acc. Chem. Res.* **2013**, *46*, 885.
- (23) (a) Meyer, E. A.; Castellano, R. K.; Diederich, F. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 1210(b) Zacharias, N.; Dougherty, D. A. *Trends Pharmacol. Sci.* **2002**, *23*, 281.

- (24) Rekharsky, M. and Inoue, Y. **2012**. Solvation Effects in Supramolecular Recognition. In *Supramolecular Chemistry* (eds P.A. Gale and J.W. Steed). doi:[10.1002/9780470661345.smc009](https://doi.org/10.1002/9780470661345.smc009)
- (25) Nugent, M. J.; Weigang, O. E. *J. Am. Chem. Soc.* **1969**, *91*, 4556.
- (26) Liu, X.; Ma, Y.; Duan, W.; He, F.; Zhao, L.; Song, C. *The Journal of Organic Chemistry* **2011**, *76*, 1953.
- (27) (a) Pye, P. J.; Rossen, K.; Reamer, R. A.; Tsou, N. N.; Volante, R. P.; Reider, P. J. *J. Am. Chem. Soc.* **1997**, *119*, 6207(b) Wack, H.; France, S.; Hafez, A. M.; Drury, W. J.; Weatherwax, A.; Lectka, T. *The Journal of Organic Chemistry* **2004**, *69*, 4531.
- (28) (a) Laganis, E. D.; Finke, R. G.; Boekelheide, V. *Tetrahedron Lett.* **1980**, *21*, 4405(b) Jens, G.; Gabriele, H.; Astrid, S.; Jörg, H.; Fritz, V.; Holger, S.; Karsten, G.; Benedikt, A.; Karl, C.; Kari, R. *Chemistry – A European Journal* **1996**, *2*, 1585.
- (29) (a) Zhang, P.; Yang, B.; Fang, X.; Cheng, Z.; Yang, M. *J. Braz. Chem. Soc.* **2012**, *23*, 1771(b) Buffet, K.; Nierengarten, I.; Galanos, N.; Gillon, E.; Holler, M.; Imberty, A.; Matthews, S. E.; Vidal, S.; Vincent, S. P.; Nierengarten, J.-F. *Chemistry – A European Journal* **2016**, *22*, 2955.
- (30) F., G. W. *Journal of Polymer Science Part A-1: Polymer Chemistry* **1966**, *4*, 3027.
- (31) Kotha, S.; Shirbhate, M. E.; Waghule, G. T. *Beilstein J. Org. Chem.* **2015**, *11*, 1274.
- (32) (a) Paudler, W. W.; Bezoari, M. D. In *Organic Chemistry: A Series of Monographs*; Keehn, P. M., Rosenfeld, S. M., Eds.; Academic Press, 1983; Vol. 45(b) Collins, M. S.; Carnes, M. E.; Nell, B. P.; Zakharov, L. N.; Johnson, D. W. *Nat Commun* **2016**, *7*.
- (33) In *Modern Cyclophane Chemistry*.
- (34) (a) Li, J.; Nowak, P.; Otto, S. *J. Am. Chem. Soc.* **2013**, *135*, 9222(b) Holloway, L. R.; Bogie, P. M.; Lyon, Y.; Julian, R. R.; Hooley, R. J. *Inorg. Chem.* **2017**, *56*, 11435(c) Ronson, T. K.; Pilgrim, B. S.; Nitschke, J. R. *J. Am. Chem. Soc.* **2016**, *138*, 10417(d) Roberts, D. A.; Pilgrim, B. S.; Cooper, J. D.; Ronson, T. K.; Zarra, S.; Nitschke, J. R. *J. Am. Chem. Soc.* **2015**, *137*, 10068(e) Roberts, D. A.; Castilla, A. M.; Ronson, T. K.; Nitschke, J. R. *J. Am. Chem. Soc.* **2014**, *136*, 8201(f) Pilgrim, B.; Roberts, D.; Lohr, T.; Ronson, T.; Nitschke, J. *Nat. Chem.* **2017**, *9*.
- (35) Moulin, E.; Cormos, G.; Giuseppone, N. *Chem. Soc. Rev.* **2012**, *41*, 1031.
- (36) Beeren, S.R. and Sanders, J.K.M. **2010**, History and Principles of Dynamic Combinatorial Chemistry. In *Dynamic Combinatorial Chemistry* (eds J.N.H. Reek and S. Otto). doi:[10.1002/9783527629701.ch1](https://doi.org/10.1002/9783527629701.ch1)
- (37) Corbett, P. T.; Leclaire, J.; Vial, L.; West, K. R.; Wietor, J.-L.; Sanders, J. K. M.; Otto, S. *Chem. Rev.* **2006**, *106*, 3652.
- (38) (a) Matache, M.; Bogdan, E.; Hădade, N. D. *Chemistry – A European Journal* **2014**, *20*, 2106(b) Jin, Y.; Yu, C.; Denman, R. J.; Zhang, W. *Chem. Soc. Rev.* **2013**, *42*, 6634.
- (39) (a) Leclaire, J.; Vial, L.; Otto, S.; Sanders, J. K. M. *Chem. Commun.* **2005**, 1959(b) West, K. R.; Bake, K. D.; Otto, S. *Org. Lett.* **2005**, *7*, 2615(c) Sarma, R. J.; Otto, S.; Nitschke, J. R. *Chemistry – A European Journal* **2007**, *13*, 9542(d) Black, S. P.; Sanders, J. K. M.; Stefankiewicz, A. R. *Chem. Soc. Rev.* **2014**, *43*, 1861(e) Rasmussen, B.; Ulfkjær, A.; Gotfredsen, H.; Pittelkow, M. *Dynamic combinatorial chemistry with diselenides and disulfides in water*, 2014(f) Bracchi, M. E.; Fulton, D. A. *Chem. Commun.* **2015**, *51*, 11052(g) Diemer, S.; Kristensen, M.; Rasmussen, B.; Beeren, S.; Pittelkow, M. *International Journal of Molecular Sciences* **2015**, *16*, 21858(h) Klepel, F.; Ravoo, B. J. *Org. Biomol. Chem.* **2017**, *15*, 3840(i) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Science* **2002**, *297*, 590(j) Brisig, B.; Sanders, J. K. M.; Otto, S. *Angew. Chem. Int. Ed.* **2003**, *42*, 1270(k) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Angew. Chem. Int. Ed.* **2007**, *46*, 8858.
- (40) (a) Skowron, P.-T.; Dumartin, M.; Jeamet, E.; Perret, F.; Gourlaouen, C.; Baudouin, A.; Fenet, B.; Naubron, J.-V.; Fotiadu, F.; Vial, L.; Leclaire, J. *The Journal of Organic Chemistry* **2016**, *81*,

- 654(b) Donnier-Maréchal, M.; Septavaux, J.; Jeamet, E.; Héloin, A.; Perret, F.; Dumont, E.; Rossi, J.-C.; Ziarelli, F.; Leclaire, J.; Vial, L. *Org. Lett.* **2018**.
- (41) Vial, L.; Dumartin, M.; Donnier-Marechal, M.; Perret, F.; Francoia, J.-P.; Leclaire, J. *Chem. Commun.* **2016**, 52, 14219.
- (42) Nguyen, C. N.; Young, T. K.; Gilson, M. K. *The Journal of Chemical Physics* **2012**, 137, 044101.
- (43) Biedermann, F.; Vendruscolo, M.; Scherman, O. A.; De Simone, A.; Nau, W. M. *J. Am. Chem. Soc.* **2013**, 135, 14879.
- (44) Yao, H.; Ke, H.; Zhang, X.; Pan, S. J.; Li, M. S.; Yang, L. P.; Schreckenbach, G. **2018**, 140, 13466.
- (45) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R.; Sanders, J. K.; Stoddart, J. F. *Angew. Chem. Int. Ed. Engl.* **2002**, 41, 898.
- (46) Huc, I.; Lehn, J.-M. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, 94, 2106.
- (47) (a) Corbett, P. T.; Otto, S.; Sanders, J. K. M. *Org. Lett.* **2004**, 6, 1825(b) Lehn, J.-M. *Chemistry – A European Journal* **1999**, 5, 2455.
- (48) Rosario, M. S.-M.; Stifun, M.; Mark, B. *Curr. Top. Med. Chem.* **2004**, 4, 653.
- (49) Lindell, S. D.; Pattenden, L. C.; Shannon, J. *Biorg. Med. Chem.* **2009**, 17, 4035.
- (50) Ji, Q.; Lirag, R. C.; Miljanic, O. S. *Chem. Soc. Rev.* **2014**, 43, 1873.
- (51) Otto, S.; Furlan, R. L. E.; Sanders, J. K. M. *Curr. Opin. Chem. Biol.* **2002**, 6, 321.
- (52) Frei, P.; Hevey, R.; Ernst, B. *Chemistry – A European Journal* **2019**, 25, 60.
- (53) (a) Saur, I.; Severin, K. *Chem. Commun.* **2005**, 1471(b) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *J. Am. Chem. Soc.* **2005**, 127, 9390.
- (54) (a) Miljanić, O. Š. *Chem* **2017**, 2, 502(b) Ladame, S. *Org. Biomol. Chem.* **2008**, 6, 219(c) Hunt, R. A. R.; Otto, S. *Chem. Commun.* **2011**, 47, 847(d) Rozenman, M. M.; McNaughton, B. R.; Liu, D. R. *Curr. Opin. Chem. Biol.* **2007**, 11, 259.
- (55) (a) Ulrich, S. *Acc. Chem. Res.* **2019**, 52, 510(b) Ramström, O.; Lehn, J.-M. *Nat. Rev. Drug Discov.* **2002**, 1, 26.
- (56) (a) Zhang, H.; Cai, C.; Liu, W.; Li, D.; Zhang, J.; Zhao, N.; Xu, J. *Sci. Rep.* **2017**, 7, 11833(b) Lehn, J.-M. *Prog. Polym. Sci.* **2005**, 30, 814(c) Nguyen, R.; Allouche, L.; Buhler, E.; Giuseppone, N. *Angew. Chem. Int. Ed.* **2009**, 48, 1093.
- (57) (a) Clipson, A. J.; Bhat, V. T.; McNae, I.; Caniard, A. M.; Campopiano, D. J.; Greaney, M. F. *Chemistry – A European Journal* **2012**, 18, 10562(b) Bunyapaiboonsri, T.; Ramström, O.; Lohmann, S.; Lehn, J.-M.; Peng, L.; Goeldner, M. *ChemBioChem* **2001**, 2, 438(c) Hochgürtel, M.; Kroth, H.; Piecha, D.; Hofmann, M. W.; Nicolau, C.; Krause, S.; Schaaf, O.; Sonnenmoser, G.; Eliseev, A. V. *Proceedings of the National Academy of Sciences* **2002**, 99, 3382(d) Lins, R. J.; Flitsch, S. L.; Turner, N. J.; Irving, E.; Brown, S. A. *Angew. Chem. Int. Ed.* **2002**, 41, 3405(e) Larsson, R.; Pei, Z.; Ramström, O. *Angew. Chem. Int. Ed.* **2004**, 43, 3716(f) Lins, R. J.; Flitsch, S. L.; Turner, N. J.; Irving, E.; Brown, S. A. *Tetrahedron* **2004**, 60, 771(g) Lam, R. T. S.; Belenguer, A.; Roberts, S. L.; Naumann, C.; Jarrosson, T.; Otto, S.; Sanders, J. K. M. *Science* **2005**(h) Larsson, R.; Ramström, O. *European J. Org. Chem.* **2006**, 2006, 285(i) Poulsen, S.-A.; Bornaghi, L. F. *Biorg. Med. Chem.* **2006**, 14, 3275(j) Shi, B.; Stevenson, R.; Campopiano, D. J.; Greaney, M. F. *J. Am. Chem. Soc.* **2006**, 128, 8459(k) Voshell, S. M.; Lee, S. J.; Gagné, M. R. *J. Am. Chem. Soc.* **2006**, 128, 12422(l) Das, A. K.; Collins, R.; Ulijn, R. V. *Small* **2008**, 4, 279(m) Bhat, V. T.; Caniard, A. M.; Luksch, T.; Brenk, R.; Campopiano, D. J.; Greaney, M. F. *Nat. Chem.* **2010**, 2, 490.
- (58) Mondal, M.; Radeva, N.; Fanlo-Virgós, H.; Otto, S.; Klebe, G.; Hirsch, A. K. H. *Angewandte Chemie (International ed. in English)* **2016**, 55, 9422.
- (59) (a) Ludlow, R. F.; Otto, S. *J. Am. Chem. Soc.* **2010**, 132, 5984(b) Ludlow, R. F.; Otto, S. *J. Am. Chem. Soc.* **2008**, 130, 12218.

- (60) (a) G. Hamilton, D.; Feeder, N.; J. Teat, S.; K. M. Sanders, J. *New J. Chem.* **1998**, *22*, 1019(b) Storm, O.; Lüning, U. *Chemistry – A European Journal* **2002**, *8*, 793(c) Ziach, K.; Ceborska, M.; Jurczak, J. *Tetrahedron Lett.* **2011**, *52*, 4452(d) Atcher, J.; Moure, A.; Bujons, J.; Alfonso, I. *Chemistry – A European Journal* **2015**, *21*, 6869.
- (61) (a) Otsuka, H.; Nagano, S.; Kobashi, Y.; Maeda, T.; Takahara, A. *Chem. Commun.* **2010**, *46*, 1150(b) Fritze, U. F.; von Delius, M. *Chem. Commun.* **2016**, *52*, 6363(c) Behrendt, F. N.; Schlaad, H. *Macromol. Rapid Commun.* **2018**, *39*, 1700735(d) Caraballo, R.; Rahm, M.; Vongvilai, P.; Brinck, T.; Ramstrom, O. *Chem. Commun. (Camb.)* **2008**, 6603.
- (62) Chauvin, J.-P. R.; Pratt, D. A. *Angew. Chem. Int. Ed.* **2017**, *56*, 6255.
- (63) Holler, M.; Allenbach, N.; Sonet, J.; Nierengarten, J.-F. *Chem. Commun.* **2012**, *48*, 2576.
- (64) (a) Rasmussen, B.; Sorensen, A.; Gotfredsen, H.; Pittelkow, M. *Chem. Commun.* **2014**, *50*, 3716(b) Sola, J.; Lafuente, M.; Atcher, J.; Alfonso, I. *Chem. Commun.* **2014**, *50*, 4564(c) Besenius, P.; Cormack, P. A. G.; Liu, J.; Otto, S.; Sanders, J. K. M.; Sherrington, D. C. *Chemistry – A European Journal* **2008**, *14*, 9006(d) Belenguer, A. M.; Friscic, T.; Day, G. M.; Sanders, J. K. M. *Chem. Sci.* **2011**, *2*, 696(e) Cougnon, F. B. L.; Sanders, J. K. M. *Acc. Chem. Res.* **2012**, *45*, 2211(f) Stefankiewicz, A. R.; Sambrook, M. R.; Sanders, J. K. M. *Chem. Sci.* **2012**, *3*, 2326(g) Stefankiewicz, A. R.; Sanders, J. K. M. *Chem. Commun.* **2013**, *49*, 5820.
- (65) Combes, J.-M. **1977**, *The Born-Oppenheimer Approximation*, pp 139-159, Springer Vienna, Vienna.
- (66) Lieb, E. H.; Simon, B. *Advances in Mathematics* **1977**, *23*, 22.
- (67) Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864.
- (68) Engel, E.; Dreizler, R. M. In *Density Functional Theory: An Advanced Course*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011.
- (69) Parr, R. G. **1980**, *Density Functional Theory of Atoms and Molecules*, pp 5-15, Springer Netherlands, Dordrecht.
- (70) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. *Chem. Rev.* **2012**, *112*, 289.
- (71) Engel, E.; Dreizler, R. M. In *Density Functional Theory: An Advanced Course*; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011.
- (72) Lee, C.; Yang, W.; Parr, R. G. *Physical Review B* **1988**, *37*, 785.
- (73) (a) Burke, K. *The Journal of Chemical Physics* **2012**, *136*, 150901(b) Jones, R. O. *Reviews of Modern Physics* **2015**, *87*, 897.
- (74) Monticelli, L.; Tieleman, D. P. *Methods Mol. Biol.* **2013**, *924*, 197.
- (75) Satoh, A. In *Introduction to Practice of Molecular Simulation*; Satoh, A., Ed.; Elsevier: London, 2011.
- (76) Satoh, A. In *Introduction to Practice of Molecular Simulation*; Satoh, A., Ed.; Elsevier: London, 2011.
- (77) (a) Perricone, U.; Gulotta, M. R.; Lombino, J.; Parrino, B.; Cascioferro, S.; Diana, P.; Cirrincione, G.; Padova, A. *MedChemComm* **2018**, *9*, 920(b) Hospital, A.; Goñi, J. R.; Orozco, M.; Gelpí, J. L. *Advances and applications in bioinformatics and chemistry : AABC* **2015**, *8*, 37.
- (78) (a) Ramstrom, O.; Lehn, J. M. *Nat. Rev. Drug Discov.* **2002**, *1*, 26(b) Otto, S. *Current opinion in drug discovery & development* **2003**, *6*, 509.
- (79) (a) Otto, S. *Acc. Chem. Res.* **2012**, *45*, 2200(b) West, K. R.; Ludlow, R. F.; Corbett, P. T.; Besenius, P.; Mansfeld, F. M.; Cormack, P. A. G.; Sherrington, D. C.; Goodman, J. M.; Stuart, M. C. A.; Otto, S. *J. Am. Chem. Soc.* **2008**, *130*, 10834(c) Corbett, P. T.; Sanders, J. K. M.; Otto, S. *Chemistry – A European Journal* **2008**, *14*, 2153(d) Gober, I. N.; Waters, M. L. *Org. Biomol. Chem.* **2017**(e) Pinkin, N. K.; N. Power, A.; Waters, M. L. *Org. Biomol. Chem.* **2015**, *13*, 10939.

- (80) E. Beaver, J.; C. Peacor, B.; V. Bain, J.; James, L. I.; L. Waters, M. *Org. Biomol. Chem.* **2015**, *13*, 3220.
- (81) Lloyd-Jones, G. C.; Moseley, J. D.; Renny, J. S. *Synthesis* **2008**, *2008*, 661.
- (82) Thèses soutenues: (a) Dumartin, M., Université Claude Bernard Lyon 1, 2017(b) Jeamet, E., 2018(c) Pascal, Y., Université Claude Bernard Lyon 1, 2018(d) Heloin, A., University of Lyon, 2019.
- (83) Panijpan, B. *J. Chem. Educ.* **1977**, *54*, 670.
- (84) (a) Lunazzi; Placucci, G.; Cerioni, G.; Foresti, E.; Plumitallo, A. *The Journal of Organic Chemistry* **1997**, *62*, 4924(b) Cerioni, G.; Cremonini, M. A.; Lunazzi, L.; Placucci, G.; Plumitallo, A. *The Journal of Organic Chemistry* **1998**, *63*, 3933.
- (85) Koval, I. V. *Russian Chemical Reviews* **1994**, *63*, 735.
- (86) McMurry, J. *Organic chemistry*, **2016**, Boston, MA, USA : Cengage Learning
- (87) (a) Gao, L.-J.; De Jonghe, S.; Daelemans, D.; Herdewijn, P. *Bioorg. Med. Chem. Lett.* **2016**, *26*, 2142(b) Li, M.; Stojković, M. R.; Ehlers, M.; Zellermann, E.; Piantanida, I.; Schmuck, C. *Angew. Chem. Int. Ed.* **2016**, *55*, 13015(c) Cheruku, P.; Keffer, J. L.; Dogo-Isonagie, C.; Bewley, C. A. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4108.
- (88) Miller-Fleming, L.; Olin-Sandoval, V.; Campbell, K.; Ralser, M. *J. Mol. Biol.* **2015**, *427*, 3389.
- (89) Barbier, A. J.; Guenaneche, F.; Lefebvre, R. A. *J. Auton. Pharmacol.* **1989**, *9*, 279.
- (90) Pegg, A. E. *IUBMB Life* **2009**, *61*, 880.
- (91) Mandal, S.; Mandal, A.; Johansson, H. E.; Orjalo, A. V.; Park, M. H. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 2169.
- (92) Jeamet, E.; Septavaux, J.; Héloin, A.; Donnier-Maréchal, M.; Dumartin, M.; Ourri, B.; Mandal, P.; Huc, I.; Bignon, E.; Dumont, E.; Morell, C.; Francoia, J.-P.; Perret, F.; Vial, L.; Leclair, J. *Chem. Sci.* **2018**.
- (93) (a) Genheden, S.; Ryde, U. *Expert opinion on drug discovery* **2015**, *10*, 449(b) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z. H.; Hou, T. *Chem. Rev.* **2019**.
- (94) (a) Pegg, A. E. *IUBMB Life* **2014**, *66*, 8(b) Igarashi, K.; Kashiwagi, K. *The International Journal of Biochemistry & Cell Biology* **2010**, *42*, 39(c) Nowotarski, S. L.; Woster, P. M.; Casero, R. A. *Expert Rev. Mol. Med.* **2013**, *15*, e3.
- (95) (a) Roe, D. R.; Cheatham, T. E. *J. Chem. Theory Comput.* **2013**, *9*, 3084(b) D.A. Case, I. Y. B.-S., S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R Harris, N. Homeyer, S. Izadi, A. Kovalenko, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, R. Salomon-Ferrer, J. Swails, R.C. Walker, J. Wang, H. Wei, R.M. Wolf, X. Wu, L. Xiao, D.M. York and P.A. Kollman; University of California, San Francisco, 2018.
- (96) Homeyer, N.; Gohlke, H. *J. Comput. Chem.* **2013**, *34*, 965.
- (97) Darby, J. F.; Hopkins, A. P.; Shimizu, S.; Roberts, S. M.; Brannigan, J. A.; Turkenburg, J. P.; Thomas, G. H.; Hubbard, R. E.; Fischer, M. *J. Am. Chem. Soc.* **2019**.
- (98) (a) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. *J. Am. Chem. Soc.* **2008**, *130*, 2817(b) Abel, R.; Salam, N. K.; Shelley, J.; Farid, R.; Friesner, R. A.; Sherman, W. *ChemMedChem* **2011**, *6*, 1049(c) Sindhikara, D. J.; Hirata, F. *The Journal of Physical Chemistry B* **2013**, *117*, 6718(d) Haider, K.; Cruz, A.; Ramsey, S.; Gilson, M. K.; Kurtzman, T. *J. Chem. Theory Comput.* **2018**, *14*, 418.
- (99) Heyden, M. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2019**, *9*, e1390.
- (100) Cui, G.; Swails, J. M.; Manas, E. S. *J. Chem. Theory Comput.* **2013**, *9*, 5539.

- (101) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. *Nature* **2001**, *414*, 188.
- (102) (a) Lazaridis, T. *The Journal of Physical Chemistry B* **1998**, *102*, 3531(b) Zielkiewicz, J. *The Journal of Chemical Physics* **2008**, *128*, 196101.
- (103) (a) Edwards, P. J.; Allart, B.; Andrews, M. J.; Clase, J. A.; Menet, C. *Curr Opin Drug Discov Devel* **2006**, *9*, 425(b) Grebner, C.; Iegre, J.; Ulander, J.; Edman, K.; Hogner, A.; Tyrchan, C. *J. Chem. Inf. Model.* **2016**, *56*, 774(c) Yugandhar, K.; Gromiha, M. M. In *Prediction of Protein Secondary Structure*; Zhou, Y., Kloczkowski, A., Faraggi, E., Yang, Y., Eds.; Springer New York: New York, NY, 2017(d) Vilseck, J. Z.; Armacost, K. A.; Hayes, R. L.; Goh, G. B.; Brooks, C. L. *The Journal of Physical Chemistry Letters* **2018**, *9*, 3328.
- (104) (a) Ji, S.; Cao, W.; Yu, Y.; Xu, H. *Angew. Chem. Int. Ed.* **2014**, *53*, 6781(b) Pleasants, J. C.; Guo, W.; Rabenstein, D. L. *J. Am. Chem. Soc.* **1989**, *111*, 6553.
- (105) (a) Koketsu, M.; Fukuta, Y.; Ishihara, H. *The Journal of Organic Chemistry* **2002**, *67*, 1008(b) Koketsu, M.; Nada, F.; Hiramatsu, S.; Ishihara, H. *J. Chem. Soc., Perkin Trans. 1* **2002**, 737(c) Ishihara, H.; Koketsu, M.; Fukuta, Y.; Nada, F. *J. Am. Chem. Soc.* **2001**, *123*, 8408.
- (106) Schmidt, J.; Silks, L. *Seleno- and Tellurocarbonic Acids and Derivatives*, 2006.
- (107) (a) Sørensen, A.; Rasmussen, B.; Agarwal, S.; Schau-Magnussen, M.; Sjølling, T. I.; Pittelkow, M. *Angew. Chem. Int. Ed.* **2013**, *52*, 12346(b) Eriksen, K.; Ulfkjær, A.; Ivan Sjølling, T.; Pittelkow, M. *Benzylic Thio and Seleno Newman-Kwart Rearrangements*, 2018.
- (108) (a) Fujihara, H.; Chiu, J.-J.; Furukawa, N. *Tetrahedron Lett.* **1989**, *30*, 7441(b) Houk, J.; Whitesides, G. M. *Tetrahedron* **1989**, *45*, 91.
- (109) Danehy, J. P.; Doherty, B. T.; Egan, C. P. *The Journal of Organic Chemistry* **1971**, *36*, 2525.
- (110) Kirihara, M.; Asai, Y.; Ogawa, S.; Noguchi, T.; Hatano, A.; Hirai, Y. *Synthesis-stuttgart* **2007**, 2007, 3286.
- (111) (a) Danehy, J. P.; Oester, M. Y. *The Journal of Organic Chemistry* **1967**, *32*, 1491(b) Danehy, J. P.; Egan, C. P.; Switalski, J. *The Journal of Organic Chemistry* **1971**, *36*, 2530.
- (112) Ulatowski, F.; Sadowska-Kuzioła, A.; Jurczak, J. *The Journal of Organic Chemistry* **2014**, *79*, 9762.
- (113) Cholewiak, A.; Dobrzycki, L.; Jurczak, J.; Ulatowski, F. *Org. Biomol. Chem.* **2018**, *16*, 2411.
- (114) (a) Collins, M. S.; Phan, N.-M.; Zakharov, L. N.; Johnson, D. W. *Inorg. Chem.* **2018**, *57*, 3486(b) Johnson, D. W.; Collins, M.; Google Patents, 2016; Vol. US20160137598 A1(c) Collins, M. S.; Carnes, M. E.; Nell, B. P.; Zakharov, L. N.; Johnson, D. W. *Nature Communications* **2016**, *7*, 11052(d) Collins, M. S.; Carnes, M. E.; Sather, A. C.; Berryman, O. B.; Zakharov, L. N.; Teat, S. J.; Johnson, D. W. *Chem. Commun.* **2013**, *49*, 6599(e) Shear, T. A.; Lin, F.; Zakharov, L. N.; Johnson, D. W. *Angew. Chem. Int. Ed.*, *n/a*.
- (115) Watt, M. M.; Collins, M. S.; Johnson, D. W. *Acc. Chem. Res.* **2013**, *46*, 955.
- (116) Phan, N.-M.; Zakharov, L. N.; Johnson, D. W. *Chem. Commun.* **2018**, *54*, 13419.
- (117) Groom, C. R.; Bruno, I. J.; Lightfoot, M. P.; Ward, S. C. *Acta Crystallographica Section B* **2016**, *72*, 171.
- (118) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605.
- (119) Rahman, R.; Safe, S.; Taylor, A. *Quarterly Reviews, Chemical Society* **1970**, *24*, 208.
- (120) Hynes, J. B. *J. Med. Chem.* **1970**, *13*, 1235.
- (121) Muranaka, A.; Shibahara, M.; Watanabe, M.; Matsumoto, T.; Shinmyozu, T.; Kobayashi, N. *The Journal of Organic Chemistry* **2008**, *73*, 9125.
- (122) Djerassi, C. *Chem. Rev.* **1948**, *43*, 271.
- (123) Su, X.-C.; Man, B.; Beeren, S.; Liang, H.; Simonsen, S.; Schmitz, C.; Huber, T.; Messerle, B. A.; Otting, G. *J. Am. Chem. Soc.* **2008**, *130*, 10486.

- (124) Végh, D.; Morel, J.; Decroix, B.; Zálupský, P. *Synth. Commun.* **1992**, *22*, 2057.
- (125) (a) Karimi, M.; Ignasiak, M. T.; Chan, B.; Croft, A. K.; Radom, L.; Schiesser, C. H.; Pattison, D. I.; Davies, M. J. *Sci. Rep.* **2016**, *6*, 38572(b) Anglada, J. M.; Crehuet, R.; Adhikari, S.; Francisco, J. S.; Xia, Y. *Phys. Chem. Chem. Phys.* **2018**, *20*, 4793.
- (126) (a) (b) Swinehart, D. F. *J. Chem. Educ.* **1962**, *39*, 333.
- (127) (a) Eghbali, N.; Harpp, D. N. *The Journal of Organic Chemistry* **2007**, *72*, 3906(b) Harpp, D. N.; Gleason, J. G. *J. Am. Chem. Soc.* **1971**, *93*, 2437(c) Williams, C. R.; Harpp, D. N. *Sulfur Reports* **1990**, *10*, 103.
- (128) (a) Nagy, P. *Kinetics and Mechanisms of Thiol–Disulfide Exchange Covering Direct Substitution and Thiol Oxidation-Mediated Pathways*, 2012(b) Fernandes, P. A.; Ramos, M. J. *Chemistry – A European Journal* **2004**, *10*, 257.
- (129) Matysiak, B. M.; Nowak, P.; Cvrtila, I.; Pappas, C. G.; Liu, B.; Komáromy, D.; Otto, S. *J. Am. Chem. Soc.* **2017**, *139*, 6744.
- (130) Joshi, G.; Anslyn, E. V. *Org. Lett.* **2012**, *14*, 4714.
- (131) Mao, H.; You, B.-X.; Zhou, L.-J.; Xie, T.-T.; Wen, Y.-H.; Lv, X.; Wang, X.-X. *Org. Biomol. Chem.* **2017**, *15*, 6157.
- (132) (a) Macrae, C. F.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Shields, G. P.; Taylor, R.; Towler, M.; van de Streek, J. *J. Appl. Crystallogr.* **2006**, *39*, 453(b) Macrae, C. F.; Bruno, I. J.; Chisholm, J. A.; Edgington, P. R.; McCabe, P.; Pidcock, E.; Rodriguez-Monge, L.; Taylor, R.; van de Streek, J.; Wood, P. A. *J. Appl. Crystallogr.* **2008**, *41*, 466.
- (133) Tabushi, I.; Yamada, H.; Kuroda, Y. *The Journal of Organic Chemistry* **1975**, *40*, 1946.
- (134) Tsuji, T. **2005**, Highly Strained Cyclophanes. In *Modern Cyclophane Chemistry* (eds R. Gleiter and H. Hopf). doi:[10.1002/3527603964.ch3](https://doi.org/10.1002/3527603964.ch3)
- (135) Bruno, I. J.; Cole, J. C.; Kessler, M.; Luo, J.; Motherwell, W. D. S.; Purkis, L. H.; Smith, B. R.; Taylor, R.; Cooper, R. I.; Harris, S. E.; Orpen, A. G. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 2133.
- (136) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. *Anal. Biochem.* **1989**, *179*, 131.
- (137) Konigsberg, W. In *Methods Enzymol.*; Academic Press, 1972; Vol. 25.
- (138) Wagner, J. P.; Schreiner, P. R. *Angew. Chem. Int. Ed.* **2015**, *54*, 12274.
- (139) Pitzer, K. S.; Catalano, E. *J. Am. Chem. Soc.* **1956**, *78*, 4844.
- (140) Hanlon, S. *Biochem. Biophys. Res. Commun.* **1966**, *23*, 861.
- (141) London, F. *Transactions of the Faraday Society* **1937**, *33*, 8b.
- (142) Bursch, M.; Caldeweyher, E.; Hansen, A.; Neugebauer, H.; Ehlert, S.; Grimme, S. *Acc. Chem. Res.* **2019**, *52*, 258.
- (143) (a) Riplinger, C.; Sandhoefer, B.; Hansen, A.; Neese, F. *J. Chem. Phys.* **2013**, *139*, 134101(b) Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 034106.
- (144) (a) Weigend, F.; Häser, M. *Theor. Chem. Acc.* **1997**, *97*, 331(b) Weigend, F.; Häser, M.; Patzelt, H.; Ahlrichs, R. *Chem. Phys. Lett.* **1998**, *294*, 143(c) Ishikawa, T.; Kuwata, K. *The Journal of Physical Chemistry Letters* **2012**, *3*, 375.
- (145) (a) Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. *J. Phys. Org. Chem.* **2009**, *22*, 1127(b) Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175(c) Pérez-Jordá, J.; Becke, A. D. *Chem. Phys. Lett.* **1995**, *233*, 134(d) Sturniolo, S.; Liborio, L. **2019**, *150*, 154301.
- (146) Klimeš, J.; Michaelides, A. *The Journal of Chemical Physics* **2012**, *137*, 120901.
- (147) (a) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215(b) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157.
- (148) (a) Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.* **2006**, *125*, 194101(b) Schreiner, P. R. *Angew. Chem. Int. Ed.* **2007**, *46*, 4217(c) Hohenstein, E. G.; Chill, S. T.; Sherrill, C. D. *J. Chem. Theory Comput.* **2008**, *4*, 1996.

- (149) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.
- (150) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456.
- (151) Liu, Y.; Goddard, W.; A, W. *MATERIALS TRANSACTIONS* **2009**, *50*.
- (152) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463.
- (153) (a) Goerigk, L. In *Non-Covalent Interactions in Quantum Chemistry and Physics*; Otero de la Roza, A., DiLabio, G. A., Eds.; Elsevier, 2017(b) Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. *The Journal of Physical Chemistry Letters* **2016**, *7*, 2197.
- (154) Grimme, S. *Angew. Chem. Int. Ed.* **2008**, *47*, 3430.
- (155) Caldeweyher, E.; Bannwarth, C.; Grimme, S. **2017**, *147*, 034112.
- (156) (a) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401(b) Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2005**, *95*, 109902(c) Vydrov, O. A.; Van Voorhis, T. *Phys. Rev. Lett.* **2009**, *103*, 063004(d) Lee, K.; Murray, É. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Physical Review B* **2010**, *82*, 081101(e) Vydrov, O. A.; Voorhis, T. V. *The Journal of Chemical Physics* **2010**, *133*, 244103.
- (157) (a) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865(b) Murray, É. D.; Lee, K.; Langreth, D. C. *J. Chem. Theory Comput.* **2009**, *5*, 2754.
- (158) Hujo, W.; Grimme, S. *J. Chem. Theory Comput.* **2013**, *9*, 308.
- (159) Grimme, S. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011**, *1*, 211.
- (160) Najibi, A.; Goerigk, L. *J. Chem. Theory Comput.* **2018**, *14*, 5725.
- (161) Mehta, N.; Casanova-Páez, M.; Goerigk, L. *Phys. Chem. Chem. Phys.* **2018**, *20*, 23175.
- (162) Iron, M. A.; Janes, T. *The Journal of Physical Chemistry A* **2019**, *123*, 3761.
- (163) Ehrlich, S.; Bettinger, H. F.; Grimme, S. *Angew. Chem. Int. Ed.* **2013**, *52*, 10892.
- (164) Schmidt, M.W., Ivanic, J. and Ruedenberg, K. **2014**, The Physical Origin of Covalent Bonding. In *The Chemical Bond* (eds G. Frenking and S. Shaik). doi:[10.1002/9783527664696.ch1](https://doi.org/10.1002/9783527664696.ch1)
- (165) Wang, C.; Mo, Y.; Wagner, J. P.; Schreiner, P. R.; Jemmis, E. D.; Danovich, D.; Shaik, S. *J. Chem. Theory Comput.* **2015**, *11*, 1621.
- (166) (a) Oki, M. *Acc. Chem. Res.* **1984**, *17*, 154(b) Mo, Y.; Gao, J. *Acc. Chem. Res.* **2007**, *40*, 113(c) Schreiner, P. R. *Angew. Chem. Int. Ed.* **2002**, *41*, 3579(d) Hoffmann, R. W. *Chem. Rev.* **1989**, *89*, 1841.
- (167) Lu, Q.; Neese, F.; Bistoni, G. *Angew. Chem. Int. Ed.* **2018**, *57*, 4760.
- (168) Rösel, S.; Becker, J.; Allen, W. D.; Schreiner, P. R. *J. Am. Chem. Soc.* **2018**, *140*, 14421.
- (169) Song, L.; Schoening, J.; Wölper, C.; Schulz, S.; Schreiner, P. R. *Organometallics* **2019**, *38*, 1640.
- (170) Goerigk, L.; Grimme, S. *Phys. Chem. Chem. Phys.* **2011**, *13*, 6670.
- (171) Krüger, J.; Wölper, C.; John, L.; Song, L.; Schreiner, P. R.; Schulz, S. *Eur. J. Inorg. Chem.* **2019**, *2019*, 1669.
- (172) Wagner, J. P.; Schreiner, P. R. *J. Chem. Theory Comput.* **2016**, *12*, 231.
- (173) Yang, L.; Adam, C.; Nichol, G. S.; Cockroft, S. L. *Nat. Chem.* **2013**, *5*, 1006.
- (174) Geraskina, M. R.; Dutton, A. S.; Juetten, M. J.; Wood, S. A.; Winter, A. H. *Angew. Chem. Int. Ed.* **2017**, *56*, 9435.
- (175) Fumino, K.; Fossog, V.; Stange, P.; Paschek, D.; Hempelmann, R.; Ludwig, R. *Angew. Chem. Int. Ed.* **2015**, *54*, 2792.
- (176) (a) Raj, M.; Singh, V. K. *Chem. Commun.* **2009**, 6687(b) Lindström, U. M.; Andersson, F. *Angew. Chem. Int. Ed.* **2006**, *45*, 548.
- (177) Dalrymple, D. L.; Reinheimer, J. D.; Barnes, D.; Baker, R. *The Journal of Organic Chemistry* **1964**, *29*, 2647.

- (178) (a) Lu, G.; Liu, R. Y.; Yang, Y.; Fang, C.; Lambrecht, D. S.; Buchwald, S. L.; Liu, P. *J. Am. Chem. Soc.* **2017**, *139*, 16548(b) Meyer, T. H.; Liu, W.; Feldt, M.; Wuttke, A.; Mata, R. A.; Ackermann, L. *Chem. Eur. J.* **2017**, *23*, 5443.
- (179) Yamakawa, M.; Yamada, I.; Noyori, R. *Angew. Chem. Int. Ed.* **2001**, *40*, 2818.
- (180) Krenske, E. H.; Houk, K. N.; Harmata, M. *Org. Lett.* **2010**, *12*, 444.
- (181) Karton, A.; Goerigk, L. *J. Comput. Chem.* **2015**, *36*, 622.
- (182) (a) Müller, C. E.; Wanka, L.; Jewell, K.; Schreiner, P. R. *Angew. Chem. Int. Ed.* **2008**, *47*, 6180(b) Seebach, D.; Grošelj, U.; Schweizer, W. B.; Grimme, S.; Mück-Lichtenfeld, C. *Helv. Chim. Acta* **2010**, *93*, 1.
- (183) Procházková, E.; Kolmer, A.; Ilgen, J.; Schwab, M.; Kaltschnee, L.; Fredersdorf, M.; Schmidts, V.; Wende, R. C.; Schreiner, P. R.; Thiele, C. M. *Angew. Chem. Int. Ed.* **2016**, *55*, 15754.
- (184) Strauss, M. A.; Wegner, H. A. *European J. Org. Chem.* **2019**, *2019*, 295.
- (185) Xie, M.; Lu, W. *RSC adv.* **2018**, *8*, 2240.
- (186) Moussa, M. E.; Evariste, S.; Krämer, B.; Réau, R.; Scheer, M.; Lescop, C. *Angew. Chem. Int. Ed.* **2018**, *57*, 795.
- (187) Zhang, Y., Ecole Centrale de Marseille, 2013.
- (188) (a) Constantino, M. G.; Beatriz, A.; da Silva, G. V. J. *Tetrahedron Lett.* **2000**, *41*, 7001(b) Dickson, R. S.; Dobney, B. J.; Eastwood, F. W. *J. Chem. Educ.* **1987**, *64*, 898.
- (189) Klärner, F.-G.; Panitzky, J.; Bläser, D.; Boese, R. *Tetrahedron* **2001**, *57*, 3673.
- (190) Staiger, C. L.; Loy, D. A.; Jamison, G. M.; Schneider, D. A.; Cornelius, C. J. *J. Am. Chem. Soc.* **2003**, *125*, 9920.
- (191) (a) Mitsudo, T.-a.; Naruse, H.; Kondo, T.; Ozaki, Y.; Watanabe, Y. *Angewandte Chemie International Edition in English* **1994**, *33*, 580(b) Mitsudo, T.; Kokuryo, K.; Shinsugi, T.; Nakagawa, Y.; Watanabe, Y.; Takegami, Y. *The Journal of Organic Chemistry* **1979**, *44*, 4492.
- (192) Lide, D. R. *CRC handbook of chemistry and physics, 1992-1993 : a ready-reference book of chemical and physical data*; CRC Press: Boca Raton (Florida), 1992.
- (193) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams, Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. Wallingford, CT, 2016.
- (194) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456.
- (195) Pollice, R.; Chen, P. *Angew. Chem. Int. Ed.* **2019**, *58*, 9758.
- (196) Lu, Q.; Neese, F.; Bistoni, G. *Phys. Chem. Chem. Phys.* **2019**, *21*, 11569.
- (197) (a) Contreras-García, J.; Johnson, E. R.; Keinan, S.; Chaudret, R.; Piquemal, J.-P.; Beratan, D. N.; Yang, W. *J. Chem. Theory Comput.* **2011**, *7*, 625(b) Alonso, M.; Woller, T.; Martín-Martínez, F. J.; Contreras-García, J.; Geerlings, P.; De Proft, F. *Chemistry – A European Journal* **2014**, *20*, 4931.
- (198) Delano, W. L., **2002**, The PyMOL Molecular Graphics System.
- (199) Heyl, D.; Fessner, W.-D. *Synthesis* **2014**, *46*, 1463.

- (200) Kunishima, M.; Kawachi, C.; Monta, J.; Terao, K.; Iwasaki, F.; Tani, S. *Tetrahedron* **1999**, *55*, 13159.
- (201) (a) Rauschenberg, M.; Fritz, E.-C.; Schulz, C.; Kaufmann, T.; Ravoo, B. J. *Beilstein J. Org. Chem.* **2014**, *10*, 1354(b) Eisenführ, A.; Arora, P. S.; Sengle, G.; Takaoka, L. R.; Nowick, J. S.; Famulok, M. *Bioorg. Med. Chem.* **2003**, *11*, 235.
- (202) Gilmartin, B. P.; McLaughlin, R. L.; Williams, M. E. *Chem. Mater.* **2005**, *17*, 5446.
- (203) Pícha, J.; Fabre, B.; Buděšínský, M.; Hajduch, J.; Abdellaoui, M.; Jiráček, J. *European J. Org. Chem.* **2018**, *2018*, 5180.
- (204) Boger, D. L.; Goldberg, J.; Jiang, W.; Chai, W.; Ducray, P.; Lee, J. K.; Ozer, R. S.; Andersson, C.-M. *Bioorg. Med. Chem.* **1998**, *6*, 1347.
- (205) Das, S.; Das, U.; Varela-Ramírez, A.; Lema, C.; Aguilera, R. J.; Balzarini, J.; De Clercq, E.; Dimmock, S. G.; Gorecki, D. K. J.; Dimmock, J. R. *ChemMedChem* **2011**, *6*, 1892.
- (206) Sauer, J.; Heldmann, D. K.; Hetzenegger, J.; Krauthan, J.; Sichert, H.; Schuster, J. *European J. Org. Chem.* **1998**, *1998*, 2885.
- (207) Zhu, J.; Hiltz, J.; Lennox, R. B.; Schirrmacher, R. *Chem. Commun.* **2013**, *49*, 10275.
- (208) (a) Lei, S.; Jin, B.; Peng, R.; Zhang, Q.; Chu, S. *J. Coord. Chem.* **2017**, *70*, 2384(b) Stetsiuk, O.; El-Ghayoury, A.; Lloret, F.; Julve, M.; Avarvari, N. *Eur. J. Inorg. Chem.* **2018**, *2018*, 449.
- (209) (a) Nhu, D.; Duffy, S.; Avery, V. M.; Hughes, A.; Baell, J. B. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 4496(b) Mao, W.; Shi, W.; Li, J.; Su, D.; Wang, X.; Zhang, L.; Pan, L.; Wu, X.; Wu, H. *Angew. Chem. Int. Ed.* **2019**, *58*, 1106.
- (210) Lukesh Iii, J. C.; Andersen, K. A.; Wallin, K. K.; Raines, R. T. *Org. Biomol. Chem.* **2014**, *12*, 8598.
- (211) Clark, R. C.; Reid, J. S. *Acta Crystallographica Section A* **1995**, *51*, 887.
- (212) Altomare, A.; Burla, M. C.; Camalli, M.; Cascarano, G. L.; Giacovazzo, C.; Guagliardi, A.; Moliterni, A. G. G.; Polidori, G.; Spagna, R. *J. Appl. Crystallogr.* **1999**, *32*, 115.
- (213) Prencipe, M. *J. Appl. Crystallogr.* **2003**, *36*, 1486.
- (214) (a) Dolomanov, O.; Bourhis, L.; Gildea, R.; A. K. Howard, J.; Puschmann, H. *OLEX2: A complete structure solution, refinement and analysis program*, 2009(b) Sheldrick, G. M. *Acta Crystallogr A Found Adv* **2015**, *71*, 3.
- (215) Sheldrick, G. M. *Acta crystallographica. Section C, Structural chemistry* **2015**, *71*, 3.
- (216) (a) Tellinghuisen, J. *Biochimica et Biophysica Acta (BBA) - General Subjects* **2016**, *1860*, 861(b) Schmidtchen, F. P. In *Supramol. Chem.*; John Wiley & Sons, Ltd, 2012(c) Martinez, J. C.; Murciano-Calles, J.; Cobos, E. S.; Iglesias-Bexiga, M.; Luque, I.; Ruiz-Sanz, J. In *Applications of Calorimetry in a Wide Context - Differential Scanning Calorimetry, Isothermal Titration Calorimetry and Microcalorimetry*; Elkordy, A. A., Ed.; InTech: Rijeka, 2013(d) Lewis, E. A.; Murphy, K. P. In *Protein-Ligand Interactions: Methods and Applications*; Ulrich Nienhaus, G., Ed.; Humana Press: Totowa, NJ, 2005(e) Indyk, L.; Fisher, H. F. In *Methods Enzymol.*; Academic Press, 1998; Vol. 295(f) Freyer, M. W.; Lewis, E. A. In *Methods Cell Biol.*; Academic Press, 2008; Vol. 84(g) Doyle, M. L. *Curr. Opin. Biotechnol.* **1997**, *8*, 31(h) Callies, O.; Hernandez Daranas, A. *Nat. Prod. Rep.* **2016**, *33*, 881(i) Bundle, D. R.; Sigurskjold, B. W. In *Methods Enzymol.*; Academic Press, 1994; Vol. 247(j) In *Methods in Molecular Biophysics: Structure, Dynamics, Function*; Serdyuk, I. N., Zaccai, J., Zaccai, N. R., Eds.; Cambridge University Press: Cambridge, 2007.
- (217) Austin, A.; Petersson, G. A.; Frisch, M. J.; Dobek, F. J.; Scalmani, G.; Throssell, K. *J. Chem. Theory Comput.* **2012**, *8*, 4989.
- (218) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
- (219) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157.

- (220) Mark, P.; Nilsson, L. *The Journal of Physical Chemistry A* **2001**, *105*, 9954.
- (221) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. *J. Chem. Theory Comput.* **2012**, *8*, 3314.
- (222) Mitchell, R. H.; Lai, Y.-H.; Williams, R. V. *The Journal of Organic Chemistry* **1979**, *44*, 4733.
- (223) Ding, P.-P.; Gao, M.; Mao, B.-B.; Cao, S.-L.; Liu, C.-H.; Yang, C.-R.; Li, Z.-F.; Liao, J.; Zhao, H.; Li, Z.; Li, J.; Wang, H.; Xu, X. *Eur. J. Med. Chem.* **2016**, *108*, 364.
- (224) Jiang, X.; Shen, M.; Tang, Y.; Li, C. *Tetrahedron Lett.* **2005**, *46*, 487.
- (225) Beck-Sickinger, A. G.; Mörl, K. *Angew. Chem. Int. Ed.* **2006**, *45*, 1020.
- (226) Huang, K. Y.; Su, M. G.; Kao, H. J.; Hsieh, Y. C.; Jhong, J. H.; Cheng, K. H.; Huang, H. D.; Lee, T. Y. *Nucleic Acids Res.* **2016**, *44*, D435.
- (227) Khoury, G. A.; Baliban, R. C.; Floudas, C. A. *Sci. Rep.* **2011**, *1*, 90.
- (228) (a) Duan, G.; Walther, D. *PLoS Comput. Biol.* **2015**, *11*, e1004049(b) Sabari, B. R.; Zhang, D.; Allis, C. D.; Zhao, Y. *Nature Reviews Molecular Cell Biology* **2016**, *18*, 90.
- (229) Wang, Y.-C.; Peterson, S. E.; Loring, J. F. *Cell Res.* **2013**, *24*, 143.
- (230) (a) Minguez, P.; Parca, L.; Diella, F.; Mende, D. R.; Kumar, R.; Helmer-Citterich, M.; Gavin, A. C.; van Noort, V.; Bork, P. *Mol. Syst. Biol.* **2012**, *8*, 599(b) Minguez, P.; Letunic, I.; Parca, L.; Bork, P. *Nucleic Acids Res.* **2013**, *41*, D306.
- (231) Cournoyer, J. J.; O'Connor, P. B. In *Comprehensive Analytical Chemistry*; Whitelegge, J. P., Ed.; Elsevier, 2008; Vol. 52.
- (232) Blanquet, S.; Marol-Bonnin, S.; Beyssac, E.; Pompon, D.; Renaud, M.; Alric, M. *Trends Biotechnol.* **2001**, *19*, 393.
- (233) Dimitrov, D. S. In *Therapeutic Proteins: Methods and Protocols*; Voynov, V., Caravella, J. A., Eds.; Humana Press: Totowa, NJ, 2012.
- (234) Adler, M. J.; Dimitrov, D. S. *Hematol. Oncol. Clin. North Am.* **2012**, *26*, 447.
- (235) (a) Eichler, H.-G.; Pignatti, F.; Flamion, B.; Leufkens, H.; Breckenridge, A. *Nat. Rev. Drug Discov.* **2008**, *7*, 818(b) Walsh, G.; Jefferis, R. *Nat Biotech* **2006**, *24*, 1241(c) Doyle, H. A.; Mamula, M. J. *Trends Immunol.* **2001**, *22*, 443.
- (236) Černý, M.; Skalák, J.; Cerna, H.; Brzobohatý, B. *J. Proteomics* **2013**, *92*, 2.
- (237) (a) Laemmli, U. K. *Nature* **1970**, *227*, 680(b) Irvine, G. B. *Anal. Chim. Acta* **1997**, *352*, 387.
- (238) (a) Stoll, D. R.; Li, X.; Wang, X.; Carr, P. W.; Porter, S. E. G.; Rutan, S. C. *J. Chromatogr. A* **2007**, *1168*, 3(b) Klose, J. *Hum. Genet.* **1975**, *26*, 231(c) O'Farrell, P. H. *J. Biol. Chem.* **1975**, *250*, 4007.
- (239) Myung, J.-K.; Lubec, G. *J. Proteome Res.* **2006**, *5*, 1267.
- (240) (a) Priego Capote, F.; Sanchez, J.-C. *Mass Spectrom. Rev.* **2009**, *28*, 135(b) Balderrama, G. D.; Meneses, E. P.; Orihuela, L. H.; Hernández, O. V.; Franco, R. C.; Robles, V. P.; Batista, C. V. F. *Rapid Commun. Mass Spectrom.* **2011**, *25*, 1017(c) Kettenbach, A. N.; Gerber, S. A. *Anal. Chem.* **2011**, *83*, 7635.
- (241) (a) Yarmush, M. L.; Weiss, A. M.; Antonsen, K. P.; Odde, D. J.; Yarmush, D. M. *Biotechnol. Adv.* **1992**, *10*, 413(b) Qian, W.-J.; Kaleta, D. T.; Petritis, B. O.; Jiang, H.; Liu, T.; Zhang, X.; Mottaz, H. M.; Varnum, S. M.; Camp, D. G.; Huang, L.; Fang, X.; Zhang, W.-W.; Smith, R. D. *Molecular & Cellular Proteomics* **2008**, *7*, 1963.
- (242) Cleaves, H. J. In *Encyclopedia of Astrobiology*; Gargaud, M., Amils, R., Quintanilla, J. C., Cleaves, H. J., Irvine, W. M., Pinti, D. L., Viso, M., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 2011.
- (243) Marmé, D.; Fusenig, N. *Tumor angiogenesis: Basic mechanisms and cancer therapy*, 2008.
- (244) Miettinen, M.; Rikala, M.-S.; Rys, J.; Lasota, J.; Wang, Z.-F. *The American journal of surgical pathology* **2012**, *36*, 629.
- (245) Sahoo, N.; Choudhury, K.; Manchikanti, P. *Biodrugs* **2009**, *23*, 217.

- (246) Lanners, S.; Norrant, E.; Pasau, P. In *Green Processing and Synthesis*, 2012; Vol. 1.
- (247) (a) Gupta, R.; Srivastava, O. P. *J. Biol. Chem.* **2004**, *279*, 44258(b) Ahmed, Z. M.; Yousaf, R.; Lee, B. C.; Khan, S. N.; Lee, S.; Lee, K.; Husnain, T.; Rehman, A. U.; Bonneux, S.; Ansar, M.; Ahmad, W.; Leal, S. M.; Gladyshev, V. N.; Belyantseva, I. A.; Van Camp, G.; Riazuddin, S.; Friedman, T. B.; Riazuddin, S. *Am. J. Hum. Genet.* **2011**, *88*, 19(c) Ringman, J. M.; Fithian, A. T.; Gylys, K.; Cummings, J. L.; Coppola, G.; Elashoff, D.; Pratico, D.; Moskovitz, J.; Bitan, G. *Dement. Geriatr. Cogn. Disord.* **2012**, *33*, 219.
- (248) Bae, N.; Yang, J.-W.; Sitte, H.; Pollak, A.; Marquez, J.; Lubec, G. *Anal. Biochem.* **2012**, *428*, 1.
- (249) Reissner, K. J.; Aswad, D. W. *Cellular and Molecular Life Sciences CMLS* **2003**, *60*, 1281.
- (250) (a) Yang, H.; Zubarev, R. A. *Electrophoresis* **2010**, *31*, 1764(b) Hao, P.; Adav, S. S.; Gallart-Palau, X.; Sze, S. K. *Mass Spectrom. Rev.* **2017**, *36*, 677(c) Zhang, Y. T.; Hu, J.; Pace, A. L.; Wong, R.; Wang, Y. J.; Kao, Y.-H. *J. Chromatogr. B* **2014**, *965*, 65(d) Wang, W.; Meeler, A. R.; Bergerud, L. T.; Hesselberg, M.; Byrne, M.; Wu, Z. *Int. J. Mass spectrom.* **2012**, *312*, 107(e) Chelius, D.; Rehder, D. S.; Bondarenko, P. V. *Anal. Chem.* **2005**, *77*, 6004(f) Zhen, J.; Kim, J.; Zhou, Y.; Gaidamauskas, E.; Subramanian, S.; Feng, P. *mAbs* **2018**, *10*, 951.
- (251) Jia, L.; Sun, Y. *PLoS One* **2017**, *12*, e0181347.
- (252) (a) Gervais, D. *J. Chem. Technol. Biotechnol.* **2016**, *91*, 569(b) Zhao, J.; Li, J.; Xu, S.; Feng, P. *J. Virol.* **2016**, *90*, 4262.
- (253) Robinson, N. E.; Robinson, A. *Molecular Clocks Deamidation of Asparaginyl and Glutaminyl Residues in Peptides and Proteins*, 2004.
- (254) Catak, S.; Monard, G.; Aviyente, V.; Ruiz-López, M. F. *The Journal of Physical Chemistry A* **2006**, *110*, 8354.
- (255) Catak, S.; Monard, G.; Aviyente, V.; Ruiz-López, M. F. *The Journal of Physical Chemistry A* **2008**, *112*, 8752.
- (256) Catak, S.; Monard, G.; Aviyente, V.; Ruiz-López, M. F. *The Journal of Physical Chemistry A* **2009**, *113*, 1111.
- (257) Pace, A. L.; Wong, R. L.; Zhang, Y. T.; Kao, Y.-H.; Wang, Y. J. *J. Pharm. Sci.* **2013**, *102*, 1712.
- (258) (a) Ugur, I.; Aviyente, V.; Monard, G. *The Journal of Physical Chemistry B* **2012**, *116*, 6288(b) Ugur, I.; Marion, A.; Aviyente, V.; Monard, G. *Biochemistry* **2015**, *54*, 1429.
- (259) Robinson, N. E.; Robinson, A. B. *Proceedings of the National Academy of Sciences* **2001**, *98*, 12409.
- (260) Schöneich, C.; Zhao, F.; Yang, J.; Miller, B. L. In *Therapeutic Protein and Peptide Formulation and Delivery*; American Chemical Society, 1997; Vol. 675.
- (261) Betteridge, D. J. *Metabolism* **2000**, *49*, 3.
- (262) (a) Pan, B.; Abel, J.; Ricci, M. S.; Brems, D. N.; Wang, D. I. C.; Trout, B. L. *Biochemistry* **2006**, *45*, 15430(b) Liu, F.; Lu, W.; Yin, X.; Liu, J. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 59.
- (263) Stadtman, E. R.; Moskovitz, J.; Levine, R. L. *Antioxidants & Redox Signaling* **2003**, *5*, 577.
- (264) Ryan, B. J.; Nissim, A.; Winyard, P. G. *Redox Biology* **2014**, *2*, 715.
- (265) (a) Levine, R. L.; Moskovitz, J.; Stadtman, E. R. *IUBMB Life* **2000**, *50*, 301(b) Kim, G.; Weiss, S. J.; Levine, R. L. *Biochim. Biophys. Acta* **2014**, *1840*, 901.
- (266) Mo, J.; Yan, Q.; So, C. K.; Soden, T.; Lewis, M. J.; Hu, P. *Anal. Chem.* **2016**, *88*, 9495.
- (267) Veredas, F. J.; Cantón, F. R.; Aledo, J. C. *Sci. Rep.* **2017**, *7*, 40403.
- (268) Stephan, J. R.; Yu, F.; Costello, R. M.; Bleier, B. S.; Nolan, E. M. *J. Am. Chem. Soc.* **2018**, *140*, 17444.
- (269) Berti, P. J.; Ekiel, I.; Lindahl, P.; Abrahamson, M.; Storer, A. C. *Protein Expression Purif.* **1997**, *11*, 111.
- (270) Regl, C.; Wohlschlager, T.; Holzmann, J.; Huber, C. G. *Anal. Chem.* **2017**, *89*, 8391.

- (271) Balakrishnan, G.; Barnett, G. V.; Kar, S. R.; Das, T. K. *Anal. Chem.* **2018**, *90*, 6959.
- (272) (a) Oien, D. B.; Canello, T.; Gabizon, R.; Gasset, M.; Lundquist, B. L.; Burns, J. M.; Moskovitz, J. *Arch. Biochem. Biophys.* **2009**, *485*, 35(b) Liang, X.; Kaya, A.; Zhang, Y.; Le, D. T.; Hua, D.; Gladyshev, V. N. *BMC Biochem.* **2012**, *13*, 21.
- (273) (a) Gianazza, E.; Crawford, J.; Miller, I. *Amino Acids* **2007**, *33*, 51(b) Hedayati, M. H.; Norouzian, D.; Aminian, M.; Teimourian, S.; Ahangari Cohan, R.; Khorramizadeh, M. R. *Prep. Biochem. Biotechnol.* **2017**, *47*, 990(c) Pei, J.; Hsu, C.-C.; Yu, K.; Wang, Y.; Huang, G. *Anal. Chim. Acta* **2018**, *1011*, 59.
- (274) Tarrago, L.; Péterfi, Z.; Lee, B. C.; Michel, T.; Gladyshev, V. N. *Nat. Chem. Biol.* **2015**, *11*, 332.
- (275) Aledo, J. C.; Cantón, F. R.; Veredas, F. J. *BMC Bioinformatics* **2017**, *18*, 430.
- (276) (a) Martin, C.; Zhang, Y. *Nature Reviews Molecular Cell Biology* **2005**, *6*, 838(b) Greer, E. L.; Shi, Y. *Nature reviews. Genetics* **2012**, *13*, 343(c) Han, D.; Huang, M.; Wang, T.; Li, Z.; Chen, Y.; Liu, C.; Lei, Z.; Chu, X. *Cell Death Dis.* **2019**, *10*, 290.
- (277) (a) Baril, S. A.; Koenig, A. L.; Krone, M. W.; Albanese, K. I.; He, C. Q.; Lee, G. Y.; Houk, K. N.; Waters, M. L.; Brustad, E. M. *J. Am. Chem. Soc.* **2017**, *139*, 17253(b) Daze, K. D.; Hof, F. *Acc. Chem. Res.* **2013**, *46*, 937.
- (278) Ngola, S. M.; Kearney, P. C.; Mecozzi, S.; Russell, K.; Dougherty, D. A. *J. Am. Chem. Soc.* **1999**, *121*, 1192.
- (279) (a) Fraser Hof, G. A. E. G.; B01J 20/281 (20060101); G01N 30/06 (20060101); C08F 257/02 (20060101); C08B 37/02 (20060101); C08B 15/00 (20060101) ed.; Inc., U. I. P., Ed., 2016(b) Allen, Hillary F.; Daze, Kevin D.; Shimbo, T.; Lai, A.; Musselman, Catherine A.; Sims, Jennifer K.; Wade, Paul A.; Hof, F.; Kutateladze, Tatiana G. *Biochem. J.* **2014**, *459*, 505(c) Ali, M.; Daze, K. D.; Strongin, D. E.; Rothbart, S. B.; Rincon-Arano, H.; Allen, H. F.; Li, J.; Strahl, B. D.; Hof, F.; Kutateladze, T. G. *J. Biol. Chem.* **2015**, *290*, 22919(d) Hof, F. *Chem. Commun.* **2016**.
- (280) Garnett, G. A. E.; Starke, M. J.; Shaurya, A.; Li, J.; Hof, F. *Anal. Chem.* **2016**, *88*, 3697.
- (281) (a) Milosevich, N.; Hof, F. *Biochemistry* **2016**, *55*, 1570(b) Milosevich, N.; Warmerdam, Z.; Hof, F. *Future Med. Chem.* **2016**, *8*, 1681(c) James, L. I.; Beaver, J. E.; Rice, N. W.; Waters, M. L. *J. Am. Chem. Soc.* **2013**, *135*, 6450(d) Mullins, A. G.; Pinkin, N. K.; Hardin, J. A.; Waters, M. L. *Angew. Chem.* **2019**, *131*, 5336.
- (282) (a) Pinkin, N. K.; Liu, I.; Abron, J. D.; Waters, M. L. *Chemistry – A European Journal* **2015**, *21*, 17981(b) Pinkin, N. K.; Waters, M. L. *Org. Biomol. Chem.* **2014**, *12*, 7059(c) Ingerman, L. A.; Cuellar, M. E.; Waters, M. L. *Chem. Commun.* **2010**, *46*, 1839.
- (283) (a) Mathijs Hartman, A.; Matthias Gierse, R.; Hirsch, A. *Protein-Templated Dynamic Combinatorial Chemistry: Brief Overview and Experimental Protocol*, 2019(b) Mondal, M.; Radeva, N.; Koster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 3259(c) van Dun, S.; Ottmann, C.; Milroy, L.-G.; Brunsveld, L. *J. Am. Chem. Soc.* **2017**, *139*, 13960(d) Mondal, M.; Radeva, N.; Köster, H.; Park, A.; Potamitis, C.; Zervou, M.; Klebe, G.; Hirsch, A. K. *Angew. Chem. Int. Ed.* **2014**, *53*, 3259.
- (284) (a) L. Sessler, J.; Gale, P.; Cho, W. S. *Anion Receptor Chemistry*, 2006(b) Kubik, S. *Chem. Soc. Rev.* **2010**, *39*, 3648(c) Gale, Philip A.; Howe, Ethan N. W.; Wu, X. *Chem* **2016**, *1*, 351.
- (285) Morita, T.; Westh, P.; Nishikawa, K.; Koga, Y. *The Journal of Physical Chemistry B* **2014**, *118*, 8744.
- (286) Giese, M.; Albrecht, M.; Rissanen, K. *Chem. Commun.* **2016**, *52*, 1778.
- (287) Langton, M. J.; Serpell, C. J.; Beer, P. D. *Angew. Chem. Int. Ed.* **2016**, *55*, 1974.
- (288) Bencini, A.; Coluccini, C.; Garau, A.; Giorgi, C.; Lippolis, V.; Messori, L.; Pasini, D.; Puccioni, S. *Chem. Commun.* **2012**, *48*, 10428.

- (289) (a) Sansone, F.; Baldini, L.; Casnati, A.; Lazzarotto, M.; Ugozzoli, F.; Ungaro, R. *Proceedings of the National Academy of Sciences* **2002**, *99*, 4842(b) Ludwig, R. *Microchim. Acta* **2005**, *152*, 1(c) Bozkurt, S.; Türkmen, M. B. *Tetrahedron: Asymmetry* **2016**, *27*, 443(d) Neri, P.; L. Sessler, J.; Wang, M.-X. *Calixarenes and Beyond*, 2016(e) Prata, J. V.; Barata, P. D. *RSC adv.* **2016**, *6*, 1659(f) Hayashida, O.; Uchiyama, M. *Tetrahedron Lett.* **2006**, *47*, 4091.
- (290) Mostovaya, O. A.; Agafonova, M. N.; Galukhin, A. V.; Khayrutdinov, B. I.; Islamov, D.; Kataeva, O. N.; Antipin, I. S.; Konovalov, A. I.; Stoikov, I. I. *J. Phys. Org. Chem.* **2014**, *27*, 57.
- (291) Gordo, S.; Martos, V.; Santos, E.; Menéndez, M.; Bo, C.; Giralt, E.; de Mendoza, J. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16426.
- (292) Kovalenko, E.; Vilaseca, M.; Díaz-Lobo, M.; Masliy, A. N.; Vicent, C.; Fedin, V. P. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 265.
- (293) (a) Fox, J. M.; Kang, K.; Sherman, W.; Héroux, A.; Sastry, G. M.; Baghbanzadeh, M.; Lockett, M. R.; Whitesides, G. M. *J. Am. Chem. Soc.* **2015**, *137*, 3859(b) Sokkalingam, P.; Shraberg, J.; Rick, S. W.; Gibb, B. C. *J. Am. Chem. Soc.* **2016**, *138*, 48.
- (294) Hu, Y.; Fan, C.-P.; Fu, G.; Zhu, D.; Jin, Q.; Wang, D.-C. *J. Mol. Biol.* **2008**, *382*, 99.
- (295) Özcelik, D.; Barandun, J.; Schmitz, N.; Sutter, M.; Guth, E.; Damberger, F. F.; Allain, F. H. T.; Ban, N.; Weber-Ban, E. *Nature Communications* **2012**, *3*, 1014.
- (296) (a) *Crystal Structure of the Deamidase from Legionella pneumophila*; <http://www.rcsb.org/structure/5YM9>(b) Yao, Q.; Cui, J.; Wang, J.; Li, T.; Wan, X.; Luo, T.; Gong, Y. N.; Xu, Y.; Huang, N.; Shao, F. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20395.
- (297) (a) Lewis, A. K.; Dunleavy, K. M.; Senkow, T. L.; Her, C.; Horn, B. T.; Jersett, M. A.; Mahling, R.; McCarthy, M. R.; Perell, G. T.; Valley, C. C.; Karim, C. B.; Gao, J.; Pomerantz, W. C. K.; Thomas, D. D.; Cembran, A.; Hinderliter, A.; Sachs, J. N. *Nat. Chem. Biol.* **2016**, *12*, 860(b) Gruez, A.; Libiad, M.; Boschi-Muller, S.; Branlant, G. *J. Biol. Chem.* **2010**, *285*, 25033(c) Ranaivoson, F. M.; Neiers, F.; Kauffmann, B.; Boschi-Muller, S.; Branlant, G.; Favier, F. *J. Mol. Biol.* **2009**, *394*, 83(d) Ranaivoson, F. M.; Antoine, M.; Kauffmann, B.; Boschi-Muller, S.; Aubry, A.; Branlant, G.; Favier, F. *J. Mol. Biol.* **2008**, *377*, 268(e) Taylor, A. B.; Benglis, J.; David M.; Dhandayuthapani, S.; Hart, P. J. *J. Bacteriol.* **2003**, *185*, 4119(f) Lee, E. H.; Kwak, G.-H.; Kim, M.-J.; Kim, H.-Y.; Hwang, K. Y. *Arch. Biochem. Biophys.* **2014**, *545*, 1.
- (298) Taylor, A. B.; Benglis, D. M., Jr.; Dhandayuthapani, S.; Hart, P. J. *J. Bacteriol.* **2003**, *185*, 4119.
- (299) Waters, M. L. *Nat. Chem. Biol.* **2016**, *12*, 768.
- (300) (a) Ranaivoson, F. M.; Neiers, F.; Kauffmann, B.; Boschi-Muller, S.; Branlant, G.; Favier, F. *J. Mol. Biol.* **2009**, *394*, 83(b) Gruez, A.; Libiad, M.; Boschi-Muller, S.; Branlant, G. *J. Biol. Chem.* **2010**, *285*, 25033(c) Ma, X.-X.; Guo, P.-C.; Shi, W.-W.; Luo, M.; Tan, X.-F.; Chen, Y.; Zhou, C.-Z. *J. Biol. Chem.* **2011**, *286*, 13430.
- (301) Long, A.; Fantozzi, N.; Pinet, S.; Genin, E.; Pétuya, R.; Bégué, D.; Robert, V.; Dutasta, J.-P.; Gosse, I.; Martinez, A. *Org. Biomol. Chem.* **2019**, *17*, 5253.
- (302) Long, A.; Lefevre, S.; Guy, L.; Robert, V.; Dutasta, J.-P.; Chevallier, M. L.; Della-Negra, O.; Saaidi, P.-L.; Martinez, A. *New J. Chem.* **2019**, *43*, 10222.
- (303) (a) Tromans, R. A.; Carter, T. S.; Chabanne, L.; Crump, M. P.; Li, H.; Matlock, J. V.; Orchard, M. G.; Davis, A. P. *Nat. Chem.* **2019**, *11*, 52(b) Rios, P.; Mooibroek, T. J.; Carter, T. S.; Williams, C.; Wilson, M. R.; Crump, M. P.; Davis, A. P. *Chem. Sci.* **2017**, *8*, 4056(c) Rios, P.; Carter, T. S.; Mooibroek, T. J.; Crump, M. P.; Lisbjerg, M.; Pittelkow, M.; Supekar, N. T.; Boons, G.-J.; Davis, A. P. *Angew. Chem. Int. Ed.* **2016**, *55*, 3387(d) Mooibroek, T. J.; Crump, M. P.; Davis, A. P. *Org. Biomol. Chem.* **2016**, *14*, 1930(e) Mooibroek, T. J.; Casas-Solvas, J. M.; Harniman, R. L.; Renney, C. M.; Carter, T. S.; Crump, M. P.; Davis, A. P. *Nat. Chem.* **2016**, *8*, 69(f) Carter, T. S.; Mooibroek, T. J.; Stewart, P. F. N.; Crump, M. P.; Galan, M. C.; Davis, A. P. *Angew. Chem. Int. Ed.* **2016**,

- n/a(g) Destecroix, H.; Renney, C. M.; Mooibroek, T. J.; Carter, T. S.; Stewart, P. F. N.; Crump, M. P.; Davis, A. P. *Angew. Chem. Int. Ed.* **2015**, *54*, 2057(h) Howgego, J. D.; Butts, C. P.; Crump, M. P.; Davis, A. P. *Chem. Commun.* **2013**, *49*, 3110(i) Sookcharoenpinyo, B.; Klein, E.; Ferrand, Y.; Walker, D. B.; Brotherhood, P. R.; Ke, C.; Crump, M. P.; Davis, A. P. *Angew. Chem. Int. Ed.* **2012**, *51*, 4586(j) Ke, C.; Destecroix, H.; Crump, M. P.; Davis, A. P. *Nat. Chem.* **2012**, *4*, 718(k) Barwell, N. P.; Davis, A. P. *The Journal of Organic Chemistry* **2011**, *76*, 6548(l) Walker, D. B.; Joshi, G.; Davis, A. P. *Cell. Mol. Life Sci.* **2009**, *66*, 3177.
- (304) Ulrich, S.; Dumy, P. *Chem. Commun.* **2014**, *50*, 5810.
- (305) (a) Liu, J.; West, K. R.; Bondy, C. R.; Sanders, J. K. M. *Org. Biomol. Chem.* **2007**, *5*, 778(b) Zhang, D.; Martinez, A.; Dutasta, J.-P. *Chem. Rev.* **2017**, *117*, 4900.
- (306) Khan, S.; Sur, S.; Dankers, P. Y. W.; da Silva, R. M. P.; Boekhoven, J.; Poor, T. A.; Stupp, S. I. *Bioconjug. Chem.* **2014**, *25*, 707.
- (307) Kakuta, T.; Yamagishi, T.; Ogoshi, T. *Chem. Commun.* **2017**, *53*, 5250.
- (308) WATERS MARCEY L; GOBER ISIAH N; **2017**, SYNTHETIC RECEPTOR COMPOUNDS FOR DETECTION OF MODIFIED AMINO ACIDS IN PEPTIDES AND PROTEINS (UNIV NORTH CAROLINA CHAPEL HILL), WO2016US60499
- (309) (a) Clipson, A. J.; Bhat, V. T.; McNae, I.; Caniard, A. M.; Campopiano, D. J.; Greaney, M. F. *Chem. Eur. J.* **2012**, *18*, 10562(b) bhat, v.; Caniard, A.; Luksch, T.; Brenk, R.; Campopiano, D.; F Greaney, M. *Nucleophilic catalysis of acylhydrazone equilibration for protein-directed dynamic covalent chemistry*, 2010(c) Belowich, M. E.; Stoddart, J. F. *Chem. Soc. Rev.* **2012**, *41*, 2003(d) Schaufelberger, F.; Hu, L.; Ramström, O. *Chemistry – A European Journal* **2015**, *21*, 9776(e) Hochgürtel, M.; Biesinger, R.; Kroth, H.; Piecha, D.; Hofmann, M. W.; Krause, S.; Schaaf, O.; Nicolau, C.; Eliseev, A. V. *J. Med. Chem.* **2003**, *46*, 356(f) Nguyen; Ivan Huc, R. *Chem. Commun.* **2003**, 942.
- (310) Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem. Int. Ed.* **2002**, *41*, 898.
- (311) (a) Beatty, M. A.; Pye, A. T.; Shaurya, A.; Kim, B.; Selinger, A. J.; Hof, F. *Org. Biomol. Chem.* **2019**, *17*, 2081(b) Beatty, M. A.; Busmann, J. A.; Fagen, N. G.; Garnett, G. A. E.; Hof, F. *Supramol. Chem.* **2019**, *31*, 101(c) Liu, Y.; Lehn, J.-M.; Hirsch, A. K. H. *Acc. Chem. Res.* **2017**, *50*, 376(d) Yu, Y.; Lin, J.; Lei, S. *RSC adv.* **2017**, *7*, 11496(e) von Delius, M.; Geertsema, E. M.; Leigh, D. A.; Slawin, A. M. Z. *Org. Biomol. Chem.* **2010**, *8*, 4617(f) Liu, Y.; Sengupta, A.; Raghavachari, K.; Flood, A. H. *Chem* **2017**, *3*, 411.
- (312) (a) Orrillo, A. G.; Escalante, A. M.; Furlan, R. L. E. *Chemistry – A European Journal* **2016**, *22*, 6746(b) Berkovich-Berger, D.; Gabriel Lemcoff, N. *Chem. Commun.* **2008**, 1686.
- (313) (a) Krenske, E. H.; Petter, R. C.; Houk, K. N. *The Journal of Organic Chemistry* **2016**, *81*, 11726(b) Woll, M. G.; Gellman, S. H. *J. Am. Chem. Soc.* **2004**, *126*, 11172(c) Boul, P. J.; Reutenauer, P.; Lehn, J.-M. *Org. Lett.* **2005**, *7*, 15(d) Dohno, C.; Okamoto, A.; Saito, I. *J. Am. Chem. Soc.* **2005**, *127*, 16681(e) Zhang, B.; Chakma, P.; Shulman, M. P.; Ke, J.; Digby, Z. A.; Konkolewicz, D. *Org. Biomol. Chem.* **2018**, *16*, 2725(f) Furlan, R. L. E.; Cousins, G. R. L.; Sanders, J. K. M. *Chem. Commun.* **2000**, 1761(g) Yang, G.; Kochovski, Z.; Ji, Z.; Lu, Y.; Chen, G.; Jiang, M. *Chem. Commun.* **2016**(h) Di Stefano, S.; Mazzonna, M.; Bodo, E.; Mandolini, L.; Lanzalunga, O. *Org. Lett.* **2011**, *13*, 142(i) Di Stefano, S. *J. Phys. Org. Chem.* **2010**, *23*, 797.
- (314) Wilson, A.; Gasparini, G.; Matile, S. *Chem. Soc. Rev.* **2014**, *43*, 1948.
- (315) Drożdż, W.; Bouillon, C.; Kotras, C.; Richeter, S.; Barboiu, M.; Clément, S.; Stefankiewicz, A. R.; Ulrich, S. *Chemistry – A European Journal* **2017**, *23*, 18010.
- (316) Lascano, S.; Zhang, K.-D.; Wehlauch, R.; Gademann, K.; Sakai, N.; Matile, S. *Chem. Sci.* **2016**, *7*, 4720.

- (317) Seifert, H. M.; Ramirez Trejo, K.; Anslyn, E. V. *J. Am. Chem. Soc.* **2016**, *138*, 10916.
- (318) Sathiyajith, C.; Shaikh, R. R.; Han, Q.; Zhang, Y.; Meguellati, K.; Yang, Y.-W. *Chem. Commun.* **2017**, *53*, 677.
- (319) Wang, M.-X. *Science China Chemistry* **2018**, *61*, 993.
- (320) (a) Guo, Q. H.; Fu, Z. D.; Zhao, L.; Wang, M. X. *Angew. Chem. Int. Ed.* **2014**, *53*, 13548(b) Qing-Hui, G.; Liang, Z.; Mei-Xiang, W. *Angew. Chem. Int. Ed.* **2015**, *54*, 8386(c) Guo, Q.-H.; Zhao, L.; Wang, M.-X. *Chemistry – A European Journal* **2016**, *22*, 6947(d) Lu, Y.; Fu, Z.-D.; Guo, Q.-H.; Wang, M.-X. *Org. Lett.* **2017**, *19*, 1590.
- (321) (a) Zhao, M.-Y.; Wang, D.-X.; Wang, M.-X. *The Journal of Organic Chemistry* **2018**, *83*, 1502(b) Wu, Z. C.; Guo, Q. H.; Wang, M. X. *Angew. Chem. Int. Ed.* **2017**, *56*, 7151(c) Fu, Z.-D.; Guo, Q.-H.; Zhao, L.; Wang, D.-X.; Wang, M.-X. *Org. Lett.* **2016**, *18*, 2668.
- (322) (a) Molander, G. A.; Cadoret, F. *Tetrahedron Lett.* **2011**, *52*, 2199(b) DEBARGE, S. I.; ERDMAN, D. T. U.; O'NEILL, P. M. I.; KUMAR, R. U.; KARMILOWICZ, M. J. U.; PHARMACEUTICALS, P. I., Ed.; LANE, Graham; Pfizer Limited Ramsgate Road Sandwich Kent CT13 9NJ, GB; Vol. WO/2014/155291(c) Breslow, R.; Guo, T. *J. Am. Chem. Soc.* **1988**, *110*, 5613(d) Blackwell, H. E.; O'Leary, D. J.; Chatterjee, A. K.; Washenfelder, R. A.; Busmann, D. A.; Grubbs, R. H. *J. Am. Chem. Soc.* **2000**, *122*, 58(e) Breed, P. G.; Ramsden, J. A.; Brown, J. M. *Can. J. Chem.* **2001**, *79*, 1049(f) Wan, Z.-K.; Woo, G. H. C.; Snyder, J. K. *Tetrahedron* **2001**, *57*, 5497(g) Pozharskii, A. F.; Ozeryanskii, V. A.; Vistorobskii, N. V. *Russ. Chem. Bull.* **2003**, *52*, 218(h) Ma, Z.; Ni, F.; Woo, G. H. C.; Lo, S.-M.; Roveto, P. M.; Schaus, S. E.; Snyder, J. K. *Beilstein J. Org. Chem.* **2012**, *8*, 829(i) Nagarkar, A. A.; Crochet, A.; Fromm, K. M.; Kilbinger, A. F. M. *Macromolecules* **2012**, *45*, 4447(j) Ye, Q.; Neo, W. T.; Cho, C. M.; Yang, S. W.; Lin, T.; Zhou, H.; Yan, H.; Lu, X.; Chi, C.; Xu, J. *Org. Lett.* **2014**, *16*, 6386(k) Chen, C.-P.; Huang, C.-Y.; Chuang, S.-C. *Adv. Funct. Mater.* **2015**, *25*, 207(l) ULRIKE[CH], L. N. C. R.; ZÜRICH, U., Ed., 2015(m) Eising, S.; Lelivelt, F.; Bongers, K. M. *Angew. Chem. Int. Ed.* **2016**, *55*, 12243(n) Jouha, J.; Buttard, F.; Lorion, M.; Berthonneau, C.; Khouili, M.; Hiebel, M.-A.; Guillaumet, G.; Brière, J.-F.; Suzenet, F. *Org. Lett.* **2017**, *19*, 4770.
- (323) (a) Oliveira, B. L.; Guo, Z.; Bernardes, G. J. L. *Chem. Soc. Rev.* **2017**, *46*, 4895(b) Png, Z. M.; Zeng, H.; Ye, Q.; Xu, J. *Chemistry – An Asian Journal* **2017**, *12*, 2142.
- (324) Lu, Y.; Liang, D. D.; Fu, Z. D.; Guo, Q. H.; Wang, M. X. *Chin. J. Chem.* **2018**, *36*, 630.
- (325) Zhao, M.-Y.; Guo, Q.-H.; Wang, M.-X. *Organic Chemistry Frontiers* **2018**, *5*, 760.
- (326) Daniela, A.-A.; G., S. J.; William, T.; Álvaro, C.; E., A. M. *J. Phys. Org. Chem.* **2014**, *27*, 670.
- (327) Müller, H.-M.; Stan, H.-J. *J. High Resolut. Chromatogr.* **1990**, *13*, 759.
- (328) Kessler, S. N.; Wegner, H. A. *Org. Lett.* **2010**, *12*, 4062.
- (329) (a) Neumann, C. N.; Hooker, J. M.; Ritter, T. *Nature* **2016**, *534*, 369(b) Neumann, C. N.; Ritter, T. *Acc. Chem. Res.* **2017**, *50*, 2822(c) Kwan, E. E.; Zeng, Y.; Besser, H. A.; Jacobsen, E. N. *Nat. Chem.* **2018**, *10*, 917.
- (330) Singh, A.; Goel, N. *New J. Chem.* **2015**, *39*, 4351.
- (331) A. J. J. Lennox, *Angew. Chem. Int. Ed.* **2018**, *57*, 14686.
- (332) Svatoněk, D.; Denk, C.; Mikula, H. *Monatshefte für Chemie - Chemical Monthly* **2018**, *149*, 833.
- (333) (a) Abdel-Halim, H.; Cowan, D. O.; Robinson, D. W.; Wiygul, F. M.; Kimura, M. *The Journal of Physical Chemistry* **1986**, *90*, 5654(b) Petek, H.; Akdemir, N.; Özil, M.; Ağar, E.; Şenel, İ. *Acta Crystallographica Section E* **2004**, *60*, o621(c) Vangala, V. R.; Mondal, R.; Broder, C. K.; Howard, J. A. K.; Desiraju, G. R. *Crystal Growth & Design* **2005**, *5*, 99.
- (334) Zimmer, M. *Chem. Rev.* **2002**, *102*, 759.
- (335) (a) Kašička, V. *Electrophoresis* **2009**, *30*, S40(b) Turgeon, R. T.; Bowser, M. T. *Anal. Bioanal. Chem.* **2009**, *394*, 187.
- (336) Protection for the Thiol Group, In *Greene's Protective Groups in Organic Synthesis*, pp 837-894

- (337) Goethals, F.; Frank, D.; Du Prez, F. *Prog. Polym. Sci.* **2017**, *64*, 76.
- (338) Satyanarayana, M.; Vitali, F.; Frost, J. R.; Fasan, R. *Chem. Commun.* **2012**, *48*, 1461.
- (339) Mansfeld, F. M.; Feng, G.; Otto, S. *Org. Biomol. Chem.* **2009**, *7*, 4289.
- (340) Hamieh, S., University of Groningen, 2015.
- (341) In *Greene's Protective Groups in Organic Synthesis*.
- (342) Brun, M. A.; Tan, K.-T.; Griss, R.; Kielkowska, A.; Reymond, L.; Johnsson, K. *J. Am. Chem. Soc.* **2012**, *134*, 7676.
- (343) (a) Park, K.-H.; Berrier, C.; Lebaupain, F.; Pucci, B.; Popot, J.-L.; Ghazi, A.; Zito, F. *The Biochemical Journal* **2007**, *403*, 183(b) Boussambe, G. N. M.; Guillet, P.; Mahler, F.; Marconnet, A.; Vargas, C.; Cornut, D.; Soulie, M.; Ebel, C.; Le Roy, A.; Jawhari, A.; Bonnete, F.; Keller, S.; Durand, G. *Methods* **2018**, *147*, 84.
- (344) Protection for the Thiol Group, In *Protective Groups in Organic Synthesis*, pp 454-493.
- (345) Belle, C.; Bougault, C.; Averbuch, M.-T.; Durif, A.; Pierre, J.-L.; Latour, J.-M.; Le Pape, L. *J. Am. Chem. Soc.* **2001**, *123*, 8053.
- (346) Hori, Y.; Hirayama, S.; Sato, M.; Kikuchi, K. *Angew. Chem. Int. Ed.* **2015**, *54*, 14368.
- (347) Mehta, A.; Jaouhari, R.; Benson, T. J.; Douglas, K. T. *Tetrahedron Lett.* **1992**, *33*, 5441.
- (348) Kimura, T.; Yamamoto, A.; Namauo, T. *Phosphorus, Sulfur, and Silicon and the Related Elements* **2010**, *185*, 1008.
- (349) Ornelas, C.; Weck, M. *Chem. Commun.* **2009**, 5710.
- (350) (a) Romagnoli, B.; Ashton, P. R.; Harwood, L. M.; Philp, D.; Price, D. W.; Smith, M. H.; Hayes, W. *Tetrahedron* **2003**, *59*, 3975(b) Adejare, A.; Shen, J.; Ogunbadeniya, A. M. *J. Fluor. Chem.* **2000**, *105*, 107(c) Jiang, X.; García-Fortanet, J.; De Brabander, J. K. *J. Am. Chem. Soc.* **2005**, *127*, 11254(d) Coman, A. G.; Paraschivescu, C. C.; Hadade, N. D.; Juncu, A.; Vlaicu, O.; Popescu, C.-I.; Matache, M. *Synthesis* **2016**, *48*, 3917(e) Takemoto, T.; Yasuda, K.; Ley, S. V. *Synlett* **2001**, *2001*, 1555.
- (351) Passaniti, P.; Browne, W. R.; Lynch, F. C.; Hughes, D.; Nieuwenhuyzen, M.; James, P.; Maestri, M.; Vos, J. G. *J. Chem. Soc., Dalton Trans.* **2002**, 1740.
- (352) (a) Albrecht, B. K.; Gehling, V. S.; Harmange, J. C.; Lai, T.; Liang, J.; Dragovich, P.; ORTWINE, D.; Labadie, S.; Zhang, B.; KIEFER, J.; INC[US];, G. I. U. C. P., Ed., 2015(b) Kelly, C. B.; Mercadante, M. A.; Wiles, R. J.; Leadbeater, N. E. *Org. Lett.* **2013**, *15*, 2222.
- (353) (a) Chenot, E. D.; Bernardi, D.; Comel, A.; Kirsch, G. *Synth. Commun.* **2007**, *37*, 483(b) Khan, A. A.; Kamena, F.; Timmer, M. S. M.; Stocker, B. L. *Org. Biomol. Chem.* **2013**, *11*, 881(c) Hidaka, K.; Gohda, K.; Teno, N.; Wanaka, K.; Tsuda, Y. *Biorg. Med. Chem.* **2016**, *24*, 545(d) Guan, X.; Luo, P.; He, Q.; Hu, Y.; Ying, H. *Molecules* **2017**, *22*, 32.
- (354) Price, K. E.; Armstrong, C. M.; Imlay, L. S.; Hodge, D. M.; Pidathala, C.; Roberts, N. J.; Park, J.; Mikati, M.; Sharma, R.; Lawrenson, A. S.; Tolia, N. H.; Berry, N. G.; O'Neill, P. M.; John, A. R. *O. Sci. Rep.* **2016**, *6*, 36777.
- (355) Bravo, Y.; Teriete, P.; Dhanya, R.-P.; Dahl, R.; Lee, P. S.; Kiffer-Moreira, T.; Ganji, S. R.; Sergienko, E.; Smith, L. H.; Farquharson, C.; Millán, J. L.; Cosford, N. D. P. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 4308.
- (356) (a) Wang, Z.; Kuninobu, Y.; Kanai, M. *The Journal of Organic Chemistry* **2013**, *78*, 7337(b) Yu, T.-Q.; Hou, Y.-S.; Jiang, Y.; Xu, W.-X.; Shi, T.; Wu, X.; Zhang, J.-C.; He, D.; Wang, Z. *Tetrahedron Lett.* **2017**, *58*, 2084.
- (357) Guo, J.-R.; Gong, J.-F.; Song, M.-P. *Org. Biomol. Chem.* **2019**, *17*, 5029.
- (358) Xiong, J.; Zhong, G.; Liu, Y. *Adv. Synth. Catal.* **2019**, *361*, 550.
- (359) (a) Schilter, D. *Nature Reviews Chemistry* **2017**, *1*, 0013(b) Rayner, C. M. *Contemporary Organic Synthesis* **1994**, *1*, 191.

- (360) (a) Lexa, D.; Savéant, J. M. In *Redox Chemistry and Interfacial Behavior of Biological Molecules*; Dryhurst, G., Niki, K., Eds.; Springer US: Boston, MA, 1988(b) Schmittl, M.; He, B.; Kalsani, V.; Bats, J. W. *Org. Biomol. Chem.* **2007**, *5*, 2395.
- (361) Azzaroni, O.; Yameen, B.; Knoll, W. *Phys. Chem. Chem. Phys.*, **2008**, *10*, 7031-7038
- (362) Pattabiraman, V. R.; Bode, J. W. *Nature* **2011**, *480*, 471.
- (363) Radzicka, A.; Wolfenden, R. *J. Am. Chem. Soc.* **1996**, *118*, 6105.
- (364) Smith, R. M.; Hansen, D. E. *J. Am. Chem. Soc.* **1998**, *120*, 8910.
- (365) Mahesh, S.; Tang, K.-C.; Raj, M. *Molecules (Basel, Switzerland)* **2018**, *23*, 2615.
- (366) (a) In *Biopolymers Online*(b) Mcmanus, S. P.; Karaman, R. M.; Sedaghat-Herati, M. R.; Shannon, T. G.; Hovatter, T. W.; Harris, J. M. *J. Polym. Sci., Part A: Polym. Chem.* **1990**, *28*, 3337(c) Giroto, J. A.; Teixeira, A. C. S. C.; Nascimento, C. A. O.; Guardani, R. *Ind. Eng. Chem. Res.* **2010**, *49*, 3200(d) Glastrup, J. *Polym. Degradation Stab.* **1996**, *52*, 217(e) Han, S.; Kim, C.; Kwon, D. *Polymer* **1997**, *38*, 317.
- (367) (a) Li, G.; Szostak, M. *Nature Communications* **2018**, *9*, 4165(b) Loomis, W. D.; Stumpf, P. K. In *Der Stickstoffumsatz / Nitrogen Metabolism*; Allen, E. K., Allen, O. N., Böttger, I., Caspersson, T., Dillemann, G., Engel, H., Fischer, H., Guggenheim, M., Haas, P., Haurowitz, F., Loomis, W. D., Manshard, E., McKee, H. S., McQuellen, K., Mevius, W., Moritz, O., Mothes, K., Nielsen, N., Rautanen, N., Romeike, A., Scheffer, F., Schmalfuss, K., Schramm, G., Schröter, H.-B., Sheat, D. E. G., Spencer, D., Street, H. E., Stumpf, P. K., Täufel, K., Thomas, M., Waldschmidt-Leitz, E., Waley, S. G., Wilson, P. W., Yemm, E. W., Eds.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1958.
- (368) Mohammadi Ziarani, G.; Hassanzadeh, Z.; Gholamzadeh, P.; Asadi, S.; Badiei, A. *RSC adv.* **2016**, *6*, 21979.
- (369) (a) Staros, J. V.; Bayley, H.; Standring, D. N.; Knowles, J. R. *Biochem. Biophys. Res. Commun.* **1978**, *80*, 568(b) Zhang, D.; Yang, Z.; Li, H.; Pei, Z.; Sun, S.; Xu, Y. *Chem. Commun.* **2016**, *52*, 749(c) Saegusa, T.; Ito, Y.; Shimizu, T. *The Journal of Organic Chemistry* **1970**, *35*, 2979(d) Sasmal, P. K.; Carregal-Romero, S.; Han, A. A.; Streu, C. N.; Lin, Z.; Namikawa, K.; Elliott, S. L.; Köster, R. W.; Parak, W. J.; Meggers, E. *ChemBioChem* **2012**, *13*, 1116.
- (370) (a) Hoyle, C. E.; Lowe, A. B.; Bowman, C. N. *Chem. Soc. Rev.* **2010**, *39*, 1355(b) Lowe, A. B. *Polymer Chemistry* **2014**, *5*, 4820.
- (371) (a) Nagy, P. *Antioxidants & redox signaling* **2013**, *18*, 1623(b) Rabenstein, D. L.; Yeo, P. L. *The Journal of Organic Chemistry* **1994**, *59*, 4223.
- (372) (a) Mardirossian, N.; Head-Gordon, M. *Mol. Phys.* **2017**, *115*, 2315(b) Izgorodina, E. I.; Brittain, D. R. B.; Hodgson, J. L.; Krenske, E. H.; Lin, C. Y.; Namazian, M.; Coote, M. L. *The Journal of Physical Chemistry A* **2007**, *111*, 10754.
- (373) (a) Priya, A. M.; Lakshmi pathi, S. *J. Phys. Org. Chem.* **2017**, *30*, e3713(b) Kyne, S. H.; Schiesser, C. H.; Matsubara, H. *Org. Biomol. Chem.* **2011**, *9*, 3217(c) Bernardi, F.; Bottoni, A.; Garavelli, M. *Quant. Struct.-Act. Relat.* **2002**, *21*, 128.
- (374) López, R. V.; Faza, O. N.; López, C. S. *The Journal of Organic Chemistry* **2015**, *80*, 11206.
- (375) (a) Cardey, B.; Foley, S.; Enescu, M. *The Journal of Physical Chemistry A* **2007**, *111*, 13046(b) Tarbell, D. S. In *Organic Sulfur Compounds*; Pergamon, 1961(c) Oswald, A. A.; Wallace, T. J. In *The Chemistry of Organic Sulfur Compounds*; Kharasch, N., Meyers, C. Y., Eds.; Pergamon, 1966.
- (376) (a) Borden, W. T.; Hoffmann, R.; Stuyver, T.; Chen, B. *J. Am. Chem. Soc.* **2017**, *139*, 9010(b) Song, X.; Fanelli, M. G.; Cook, J. M.; Bai, F.; Parish, C. A. *The Journal of Physical Chemistry A* **2012**, *116*, 4934.
- (377) Marenich, A. V.; Cramer, C. J.; Truhlar, D. G. *The Journal of Physical Chemistry B* **2009**, *113*, 6378.

- (378) Baker, J.; Scheiner, A.; Andzelm, J. *Chem. Phys. Lett.* **1993**, *216*, 380.
- (379) Levine, B. G.; Ko, C.; Quenneville, J.; Martínez, T. J. *Mol. Phys.* **2006**, *104*, 1039.
- (380) (a) Aziz, S. G.; Elroby, S. K.; Alyoubi, A.; Hilal, R. *Procedia Computer Science* **2014**, *29*, 1384(b) Mei, Y.; Yang, W. *The Journal of Physical Chemistry Letters* **2019**, *10*, 2538.
- (381) Marcus, R. A.; Sutin, N. *Biochimica et Biophysica Acta (BBA) - Reviews on Bioenergetics* **1985**, *811*, 265.
- (382) (a) Salmeen, A.; Andersen, J. N.; Myers, M. P.; Meng, T.-C.; Hinks, J. A.; Tonks, N. K.; Barford, D. *Nature* **2003**, *423*, 769(b) van Montfort, R. L. M.; Congreve, M.; Tisi, D.; Carr, R.; Jhoti, H. *Nature* **2003**, *423*, 773.
- (383) Sarma, B. K.; Muges, G. *J. Am. Chem. Soc.* **2007**, *129*, 8872.
- (384) Sarma, B. K. *J. Mol. Struct.* **2013**, *1048*, 410.
- (385) Gupta, V.; Carroll, K. S. *Chem. Sci.* **2016**, *7*, 400.
- (386) Hunter, C. A. *Angew. Chem. Int. Ed.* **2004**, *43*, 5310.
- (387) Reichardt, C. and Welton, T. **2010**,. Solute-Solvent Interactions. In *Solvents and Solvent Effects in Organic Chemistry* (eds C. Reichardt and T. Welton). doi:[10.1002/9783527632220.ch2](https://doi.org/10.1002/9783527632220.ch2)
- (388) (a) Sonnenschein, H. *Angew. Chem.* **1997**, *109*, 1854(b) Yudin, A. K. *Chem. Sci.* **2015**, *6*, 30.
- (389) Sandford, G. *Chemistry – A European Journal* **2003**, *9*, 1464.
- (390) De Leger, W.; Adriaensen, K.; Robeyns, K.; Van Meervelt, L.; Thomas, J.; Meijers, B.; Smet, M.; Dehaen, W. *Beilstein J. Org. Chem.* **2018**, *14*, 2190.
- (391) Ashram, M.; Miller, D. O.; Bridson, J. N.; Georghiou, P. E. *The Journal of Organic Chemistry* **1997**, *62*, 6476.
- (392) Rajakumar, P.; Rasheed, A. M. A. *Tetrahedron* **2005**, *61*, 5351.
- (393) Xiong, Z.; Li, B.; Li, L.; Wan, L.; Peng, X.; Yin, Y. *Molecules* **2017**, *22*, 888.
- (394) (a) Kempin, U.; Hennig, L.; Knoll, D.; Welzel, P.; Müller, D.; Markus, A.; van Heijenoort, J. *Tetrahedron* **1997**, *53*, 17669(b) Tedaldi, L. M.; Smith, M. E. B.; Nathani, R. I.; Baker, J. R. *Chem. Commun.* **2009**, 6583(c) Nicponski, D. R.; Marchi, J. M. *Synthesis* **2014**, *46*, 1725.
- (395) O'Brien, J.; Lee, S.-H.; Onogi, S.; Shea, K. J. *J. Am. Chem. Soc.* **2016**, *138*, 16604.
- (396) Sakamoto, K. K., Satoshi; Okuyama, Kumi; C07C251/86; C07D295/30; C08F22/22; G02B1/04 ed.; Corporation, Z., Ed., 2015.
- (397) Dangerfield, E. M.; Plunkett, C. H.; Win-Mason, A. L.; Stocker, B. L.; Timmer, M. S. M. *The Journal of Organic Chemistry* **2010**, *75*, 5470.
- (398) (a) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70(b) Aitken, A.; Learmonth, M. In *The Protein Protocols Handbook*; Walker, J. M., Ed.; Humana Press: Totowa, NJ, 2009.
- (399) He, Z.; Ye, G.; Jiang, W. *Chemistry – A European Journal* **2015**, *21*, 3005.
- (400) Lott, R. S.; Chauhan, V. S.; Stammer, C. H. *J. Chem. Soc., Chem. Commun.* **1979**, 495.
- (401) Gjoka, B.; Romano, F.; Zonta, C.; Licini, G. *European J. Org. Chem.* **2011**, *2011*, 5636.
- (402) Favre, A.; Grugier, J.; Brans, A.; Joris, B.; Marchand-Brynaert, J. *Tetrahedron* **2012**, *68*, 10818.
- (403) (a) Hanessian, S.; Marcotte, S.; Machaalani, R.; Huang, G. *Org. Lett.* **2003**, *5*, 4277(b) Guo, Z.; Reddy, M. V.; Goh, B. M.; San, A. K. P.; Bao, Q.; Loh, K. P. *RSC adv.* **2013**, *3*, 19051.
- (404) Kuhnert, N.; Rossignolo, G. M.; Lopez-Periago, A. *Org. Biomol. Chem.* **2003**, *1*, 1157.
- (405) Okada, Y.; Sugai, M.; Chiba, K. *The Journal of Organic Chemistry* **2016**, *81*, 10922.
- (406) Hattori, T.; Kagawa, K.; Kakimoto, M.; Imai, Y. *Macromolecules* **1993**, *26*, 4089.
- (407) Silva, F.; Campello, M. P. C.; Gano, L.; Fernandes, C.; Santos, I. C.; Santos, I.; Ascenso, J. R.; Joao Ferreira, M.; Paulo, A. *Dalton Transactions* **2015**, *44*, 3342.
- (408) (a) Buhleier, E.; Wehner, W.; VÖGtle, F. *Synthesis* **1978**, *1978*, 155(b) Newkome, G. R.; Yao, Z.; Baker, G. R.; Gupta, V. K. *The Journal of Organic Chemistry* **1985**, *50*, 2003(c) Tomalia, D. A.;

- Baker, H.; Dewald, J.; Hall, M.; Kallos, G.; Martin, S.; Roeck, J.; Ryder, J.; Smith, P. *Polym. J.* **1985**, *17*, 117.
- (409) (a) Li, T. R.; Yu, M. H.; Huang, X. B.; Yang, Z. J.; Lu, G. M.; Li, Y. J. *Chin. Med. J. (Engl.)* **2017**, *130*, 2591(b) Rossi, J. C.; Maret, B.; Vidot, K.; Francoia, J. P.; Cangiotti, M.; Lucchi, S.; Coppola, C.; Ottaviani, M. F. *Macromol. Biosci.* **2015**, *15*, 275(c) Lorion, C.; Faye, C.; Maret, B.; Trimaille, T.; Regnier, T.; Sommer, P.; Debret, R. *J. Biomater. Sci. Polym. Ed.* **2014**, *25*, 136(d) Kojima, C.; Fusaoka-Nishioka, E.; Imai, T.; Nakahira, A.; Onodera, H. *Journal of Biomedical Materials Research Part A* **2016**, *104*, 2744.
- (410) Wolberg, A. S.; Rosendaal, F. R.; Weitz, J. I.; Jaffer, I. H.; Agnelli, G.; Baglin, T.; Mackman, N. *Nature Reviews Disease Primers* **2015**, *1*, 15006.
- (411) Alquwaizani, M.; Buckley, L.; Adams, C.; Fanikos, J. *Current emergency and hospital medicine reports* **2013**, *1*, 83.
- (412) Mackman, N. *Nature* **2008**, *451*, 914.
- (413) Oduah, E. I.; Linhardt, R. J.; Sharfstein, S. T. *Pharmaceuticals (Basel)* **2016**, *9*, 38.
- (414) Capila, I.; Linhardt, R. J. *Angew. Chem. Int. Ed.* **2002**, *41*, 390.
- (415) Lim, G. B. *Nature Reviews Cardiology* **2017**.
- (416) Sokolowska, E.; Kalaska, B.; Miklosz, J.; Mogielnicki, A. *Expert Opin. Drug Metab. Toxicol.* **2016**, *12*, 897.
- (417) (a) Kalathottukaren, M. T.; Abraham, L.; Kapopara, P. R.; Lai, B. F. L.; Shenoi, R. A.; Rosell, F. I.; Conway, E. M.; Prydzial, E. L. G.; Morrissey, J. H.; Haynes, C. A.; Kizhakkedathu, J. N. *Blood* **2017**, *129*, 1368(b) Kalathottukaren, M. T.; Creagh, A. L.; Abbina, S.; Lu, G.; Karbarz, M. J.; Pandey, A.; Conley, P. B.; Kizhakkedathu, J. N.; Haynes, C. *Blood advances* **2018**, *2*, 2104.
- (418) Nimesh, S. In *Gene Ther.*; Nimesh, S., Ed.; Woodhead Publishing, 2013.
- (419) (a) Noriega-Luna, B.; God; #xed; nez, L. A.; Rodr; #xed; guez, F. J.; Rodr; #xed; guez, A.; Zald; #xed; var-Lelo de Larrea, G.; Sosa-Ferreyra, C. F.; Mercado-Curiel, R. F.; Manr; #xed; quez, J.; Bustos, E. *Journal of Nanomaterials* **2014**, *2014*, 19(b) Cloninger, M. J. *Curr. Opin. Chem. Biol.* **2002**, *6*, 742.
- (420) (a) Lee, C. C.; MacKay, J. A.; Fréchet, J. M. J.; Szoka, F. C. *Nat. Biotechnol.* **2005**, *23*, 1517(b) Abbasi, E.; Aval, S. F.; Akbarzadeh, A.; Milani, M.; Nasrabadi, H. T.; Joo, S. W.; Hanifehpour, Y.; Nejati-Koshki, K.; Pashaei-Asl, R. *Nanoscale research letters* **2014**, *9*, 247.
- (421) Esko, J. D.; Kimata, K.; Lindahl, U. In *Essentials of Glycobiology* Cold Spring Harbor (NY), 2009.
- (422) Bianchini, E. P.; Sebestyen, A.; Abache, T.; Bourti, Y.; Fontayne, A.; Richard, V.; Tamion, F.; Plantier, J.-L.; Doguet, F.; Borgel, D. *Br. J. Haematol.* **2018**, *180*, 715.
- (423) Bourti, Y.; Fazavana, J.; Armand, M.; Saller, F.; Lasne, D.; Borgel, D.; Bianchini, E. P. *Thromb. Haemost.* **2016**, *116*, 452.
- (424) (a) Connolly, S. J.; Crowther, M.; Eikelboom, J. W.; Gibson, C. M.; Curnutte, J. T.; Lawrence, J. H.; Yue, P.; Bronson, M. D.; Lu, G.; Conley, P. B.; Verhamme, P.; Schmidt, J.; Middeldorp, S.; Cohen, A. T.; Beyer-Westendorf, J.; Albaladejo, P.; Lopez-Sendon, J.; Demchuk, A. M.; Pallin, D. J.; Concha, M.; Goodman, S.; Leeds, J.; Souza, S.; Siegal, D. M.; Zotova, E.; Meeks, B.; Ahmad, S.; Nakamya, J.; Milling, T. J. *N. Engl. J. Med.*, *0*, null(b) Lu, G.; Lin, J.; Curnutte, J. T.; Conley, P. B. *Blood* **2015**, *126*, 2329(c) Siegal, D. M.; Curnutte, J. T.; Connolly, S. J.; Lu, G.; Conley, P. B.; Wiens, B. L.; Mathur, V. S.; Castillo, J.; Bronson, M. D.; Leeds, J. M.; Mar, F. A.; Gold, A.; Crowther, M. A. *N. Engl. J. Med.* **2015**, *373*, 2413(d) Connolly, S. J.; Milling, T. J.; Eikelboom, J. W.; Gibson, C. M.; Curnutte, J. T.; Gold, A.; Bronson, M. D.; Lu, G.; Conley, P. B.; Verhamme, P.; Schmidt, J.; Middeldorp, S.; Cohen, A. T.; Beyer-Westendorf, J.; Albaladejo, P.; Lopez-Sendon, J.; Goodman, S.; Leeds, J.; Wiens, B. L.; Siegal, D. M.; Zotova, E.; Meeks, B.; Nakamya, J.; Lim, W. T.; Crowther, M. *N. Engl. J. Med.* **2016**, *375*, 1131.

- (425) Longstaff, C.; Hogwood, J.; Gray, E.; Komorowicz, E.; Varjú, I.; Varga, Z.; Kolev, K. *Thromb. Haemost.* **2016**, *115*, 591.
- (426) Cheong, H. Y.; Groner, M.; Hong, K.; Lynch, B.; Hollingsworth, W. R.; Polonskaya, Z.; Rhee, J.-K.; Baksh, M. M.; Finn, M. G.; Gale, A. J.; Udit, A. K. *Biomacromolecules* **2017**, *18*, 4113.
- (427) Corredor, M.; Bonet, R.; Moure, A.; Domingo, C.; Bujons, J.; Alfonso, I.; Pérez, Y.; Messeguer, À. *Biophys. J.* **2016**, *110*, 1291.
- (428) Li, T.; Meng, Z.; Zhu, X.; Gan, H.; Gu, R.; Wu, Z.; Li, J.; Zheng, Y.; Liu, T.; Han, P.; Han, S.; Dou, G. *Biochem. Biophys. Res. Commun.* **2015**, *467*, 497.
- (429) Ansell, J. E.; Laulicht, B. E.; Bakhru, S. H.; Hoffman, M.; Steiner, S. S.; Costin, J. C. *Thromb. Res.* **2016**, *146*, 113.
- (430) Sullivan, D. W., Jr.; Gad, S. C.; Laulicht, B.; Bakhru, S.; Steiner, S. *Int. J. Toxicol.* **2015**, *34*, 308.
- (431) Montalvo, G. L.; Zhang, Y.; Young, T. M.; Costanzo, M. J.; Freeman, K. B.; Wang, J.; Clements, D. J.; Magavern, E.; Kavash, R. W.; Scott, R. W.; Liu, D.; DeGrado, W. F. *ACS Chem. Biol.* **2014**, *9*, 967.
- (432) Weiss, R. J.; Gordts, P. L. S. M.; Le, D.; Xu, D.; Esko, J. D.; Tor, Y. *Chem. Sci.* **2015**, *6*, 5984.
- (433) Smith, D. K. *Chem. Commun.* **2018**, *54*, 4743.
- (434) Vieira, V. M. P.; Liljeström, V.; Posocco, P.; Laurini, E.; Pricl, S.; Kostianen, M. A.; Smith, D. K. *J. Mater. Chem. B* **2017**, *5*, 341.
- (435) Bromfield, S. M.; Posocco, P.; Chan, C. W.; Calderon, M.; Guimond, S. E.; Turnbull, J. E.; Pricl, S.; Smith, D. K. *Chem. Sci.* **2014**, *5*, 1484.
- (436) Albanyan, B.; Laurini, E.; Posocco, P.; Pricl, S.; Smith, D. K. *Chem. Eur. J.* **2017**, *23*, 6391.
- (437) (a) Chan, C. W.; Laurini, E.; Posocco, P.; Pricl, S.; Smith, D. K. *Chem. Commun.* **2016**, *52*, 10540 (b) Thornalley, K. A.; Laurini, E.; Pricl, S.; Smith, D. K. *Angew. Chem. Int. Ed.* **2018**, *57*, 8530.
- (438) (a) Rodrigo, A. C.; Bromfield, S. M.; Laurini, E.; Posocco, P.; Pricl, S.; Smith, D. K. *Chem. Commun.* **2017**, *53*, 6335 (b) Fechner, L. E.; Albanyan, B.; Vieira, V. M. P.; Laurini, E.; Posocco, P.; Pricl, S.; Smith, D. K. *Chem. Sci.* **2016**, *7*, 4653.
- (439) Abbina, S.; Vappala, S.; Kumar, P.; Siren, E. M. J.; La, C. C.; Abbasi, U.; Brooks, D. E.; Kizhakkedathu, J. N. *J. Mater. Chem. B* **2017**, *5*, 9249.
- (440) Kalaska, B.; Kaminski, K.; Sokolowska, E.; Czaplicki, D.; Kujdowicz, M.; Stalinska, K.; Bereta, J.; Szczubialka, K.; Pawlak, D.; Nowakowska, M.; Mogielnicki, A. *PLoS One* **2015**, *10*, e0119486.
- (441) Lorkowska-Zawicka, B.; Kamiński, K.; Ciejka, J.; Szczubialka, K.; Białas, M.; Okoń, K.; Adamek, D.; Nowakowska, M.; Jawień, J.; Olszanecki, R.; Korbut, R. *Mar. Drugs* **2014**, *12*, 3953.
- (442) Kamiński, K.; Kałaska, B.; Koczurkiewicz, P.; Michalik, M.; Szczubialka, K.; Mogielnicki, A.; Buczko, W.; Nowakowska, M. *MedChemComm* **2014**, *5*, 489.
- (443) Robison, A. D.; Sun, S.; Poyton, M. F.; Johnson, G. A.; Pellois, J. P.; Jungwirth, P.; Vazdar, M.; Cremer, P. S. *J. Phys. Chem. B* **2016**, *120*, 9287.
- (444) Mitchell, D. J.; Kim, D. T.; Steinman, L.; Fathman, C. G.; Rothbard, J. B. *J. Pept. Res.* **2000**, *56*, 318.
- (445) Pasut, G.; Veronese, F. M. *J. Control. Release* **2012**, *161*, 461.
- (446) Välimäki, S.; Khakalo, A.; Ora, A.; Johansson, L.-S.; Rojas, O. J.; Kostianen, M. A. *Biomacromolecules* **2016**, *17*, 2891.
- (447) Kalaska, B.; Kaminski, K.; Miklosz, J.; Yusa, S.-i.; Sokolowska, E.; Blazejczyk, A.; Wietrzyk, J.; Kasacka, I.; Szczubialka, K.; Pawlak, D.; Nowakowska, M.; Mogielnicki, A. *Trans. Res.* **2016**, *177*, 98.
- (448) Bromfield, S. M.; Posocco, P.; Fermeglia, M.; Tolosa, J.; Herreros-López, A.; Pricl, S.; Rodríguez-López, J.; Smith, D. K. *Chem. Eur. J.* **2014**, *20*, 9666.

- (449) Jones, C. F.; Campbell, R. A.; Franks, Z.; Gibson, C. C.; Thiagarajan, G.; Vieira-de-Abreu, A.; Sukavaneshvar, S.; Mohammad, S. F.; Li, D. Y.; Ghandehari, H.; Weyrich, A. S.; Brooks, B. D.; Grainger, D. W. *Mol. Pharm.* **2012**, *9*, 1599.
- (450) Shenoi, R. A.; Kalathottukaren, M. T.; Travers, R. J.; Lai, B. F. L.; Creagh, A. L.; Lange, D.; Yu, K.; Weinhart, M.; Chew, B. H.; Du, C.; Brooks, D. E.; Carter, C. J.; Morrissey, J. H.; Haynes, C. A.; Kizhakkedathu, J. N. *Sci. Transl. Med.* **2014**, *6*, 260ra150.
- (451) Kalathottukaren, M. T.; Abbina, S.; Yu, K.; Shenoi, R. A.; Creagh, A. L.; Haynes, C.; Kizhakkedathu, J. N. *Biomacromolecules* **2017**, *18*, 3343.
- (452) Mlynarczyk, D., Kocki, T., Goslinski, T., **2017**, Dendrimer Structure Diversity and Tailorability as a Way to Fight Infectious Diseases, 10.5772/67660
- (453) Bosman, A. W.; Janssen, H. M.; Meijer, E. W. *Chem. Rev.* **1999**, *99*, 1665.
- (454) Tomalia, D.; Christensen, J.; Boas, U. *Dendrimers, dendrons, and dendritic polymers: Discovery, applications, and the future*, 2009.
- (455) Tomalia, D. A. In *Physical Properties of Polymers Handbook*; Mark, J. E., Ed.; Springer New York: New York, NY, 2007.
- (456) Gorman, C. In *Encyclopedia of Materials: Science and Technology*; Buschow, K. H. J., Cahn, R. W., Flemings, M. C., Ilschner, B., Kramer, E. J., Mahajan, S., Veyssi re, P., Eds.; Elsevier: Oxford, 2001.
- (457) (a) Tomalia, D.; R. Swanson, D., 2002(b) Maiti, P. K.;  ađın, T.; Wang, G.; Goddard, W. A. *Macromolecules* **2004**, *37*, 6236.
- (458) (a) Hawker, C.; Fr chet, J. M. J. *J. Chem. Soc., Chem. Commun.* **1990**, 1010(b) In *Dendrimers and Other Dendritic Polymers*.
- (459) Hawker, C. J.; Fr chet, J. M. J. *J. Am. Chem. Soc.* **1990**, *112*, 7638.
- (460) Wooley, K. L.; Hawker, C. J.; Fr chet, J. M. J. *J. Am. Chem. Soc.* **1991**, *113*, 4252.
- (461) Collet, H.; Souaid, E.; Cottet, H.; Deratani, A.; Boiteau, L.; Dessalces, G.; Rossi, J. C.; Commeyras, A.; Pascal, R. *Chemistry (Easton)* **2010**, *16*, 2309.
- (462) Denkewalter, R. G. (Westfield, NJ), Kolc, J. (Randolph Twp., Morris County, NJ), Lukasavage, W. J. (Harrison, NJ), **1981** Macromolecular highly branched homogeneous compound based on lysine units (Morris Township, Morris County, NJ) USA patent 4289872
- (463) Klok, H.-A.; Rodr guez-Hern ndez, J. *Macromolecules* **2002**, *35*, 8718.
- (464) (a) Francoia, J.-P., 2016(b) Francoia, J.-P.; Rossi, J.-C.; Monard, G.; Vial, L. *J. Chem. Inf. Model.* **2017**, *57*, 2173.
- (465) Haag, R. *Beilstein J. Org. Chem.* **2015**, *11*, 848.
- (466) Fasting, C.; Schalley, C. A.; Weber, M.; Seitz, O.; Hecht, S.; Kokscho, B.; Dervedde, J.; Graf, C.; Knapp, E.-W.; Haag, R. *Angew. Chem.* **2012**, *124*, 10622.
- (467) Patel, P. *International Journal Of Pharma and Bioscience* **2013**, *4*, 454.
- (468) (a) Ara jo, R. V. d.; Santos, S. d. S.; Igne Ferreira, E.; Giarolla, J. *Molecules (Basel, Switzerland)* **2018**, *23*, 2849(b) Palmerston Mendes, L.; Pan, J.; Torchilin, V. P. *Molecules (Basel, Switzerland)* **2017**, *22*, 1401.
- (469) Sadler, K.; Tam, J. P. *Rev. Mol. Biotechnol.* **2002**, *90*, 195.
- (470) Kodama, Y.; Nakamura, T.; Kurosaki, T.; Egashira, K.; Mine, T.; Nakagawa, H.; Muro, T.; Kitahara, T.; Higuchi, N.; Sasaki, H. *Eur. J. Pharm. Biopharm.* **2014**, *87*, 472.
- (471) (a) Chen, H.; Wang, Y.; Yao, Y.; Qiao, S.; Wang, H.; Tan, N. *Theranostics* **2017**, *7*, 3781(b) Zhang, C.; Ling, C. L.; Pang, L.; Wang, Q.; Liu, J. X.; Wang, B. S.; Liang, J. M.; Guo, Y. Z.; Qin, J.; Wang, J. X. *Theranostics* **2017**, *7*, 3260.

- (472) (a) Francoia, J.-P.; Vial, L. *Chem. Commun.* **2015**, *51*, 17544(b) Couturaud, B.; Molero Bondia, A.; Faye, C.; Garrelly, L.; Mas, A.; Robin, J. J. *J. Colloid Interface Sci.* **2013**, *408*, 242(c) Francoia, J.-P.; Pascal, R.; Vial, L. *Chem. Commun.* **2015**, *51*, 1953.
- (473) Svenson, S. *Chem. Soc. Rev.* **2015**, *44*, 4131.
- (474) Thiruvankatarajan, V.; Pruet, A.; Adhikary, S. D. *Indian J. Anaesth.* **2014**, *58*, 565.
- (475) van den Besselaar, A. M.; Sturk, A.; Reijnierse, G. L. *Thromb. Res.* **2002**, *107*, 235.
- (476) Taylor, C. T.; Petros, W. P.; Ortel, T. L. *Pharmacotherapy* **1999**, *19*, 383.
- (477) Eikelboom, J. W.; Hirsh, J. *Thromb. Haemost.* **2006**, *96*, 547.
- (478) Marson, D.; Laurini, E.; Fermeglia, M.; Smith, D. K.; Pricl, S. *Fluid Phase Equilib.* **2018**, *470*, 259.
- (479) Warttinger, U.; Giese, C.; Krämer, R. In *arXiv e-prints*, 2017.
- (480) Lima, M.; Nader, H. *BMC Proc.* **2014**, *8*, O10.
- (481) Ourri, B.; Francoia, J. P.; Monard, G.; Gris, J. C.; Leclaire, J.; Vial, L. *ACS Med. Chem. Lett.* **2019**, *10*, 917.
- (482) Gehrie, E.; Laposata, M. *Am. J. Hematol.* **2012**, *87*, 194.
- (483) Chan, C. W.; Smith, D. K. *Supramol. Chem.* **2017**, *29*, 688.
- (484) Magnes, C.; Fauland, A.; Gander, E.; Narath, S.; Ratzer, M.; Eisenberg, T.; Madeo, F.; Pieber, T.; Sinner, F. *Journal of chromatography. A* **2014**, *1331*, 44.
- (485) Sisavath, N.; Got, P.; Charrière, G. M.; Destoumieux-Garzon, D.; Cottet, H. *Anal. Chem.* **2015**, *87*, 6761.
- (486) Hodgson, E. In *Pesticide Biotransformation and Disposition*; Hodgson, E., Ed.; Academic Press: Boston, 2012.
- (487) Busto, E.; Gotor-Fernández, V.; Gotor, V. *Chem. Soc. Rev.* **2010**, *39*, 4504.
- (488) Wolfbeis, O. S.; Koller, E. *Anal. Biochem.* **1983**, *129*, 365.
- (489) Kumar, A.; Dhar, K.; Kanwar, S. S.; Arora, P. K. *Biol. Proced. Online* **2016**, *18*, 2.
- (490) Jones, J. B. In *Pure Appl. Chem.*, 1990; Vol. 62.
- (491) (a) Khan, S.; Gor, J.; Mulloy, B.; Perkins, S. J. *J. Mol. Biol.* **2010**, *395*, 504(b) Bromfield, S. M.; Posocco, P.; Fermeglia, M.; Pricl, S.; Rodríguez-López, J.; Smith, D. K. *Chem. Commun.* **2013**, *49*, 4830.
- (492) Bromfield, S. M.; Wilde, E.; Smith, D. K. *Chem. Soc. Rev.* **2013**, *42*, 9184.
- (493) Layne, E. *Spectrophotometric and Turbidimetric Methods for Measuring Proteins*, 1957.
- (494) (a) Koch, A. L. *Biochim. Biophys. Acta* **1961**, *51*, 429(b) Koch, A. L. *Anal. Biochem.* **1970**, *38*, 252(c) Stevenson, K.; McVey, A. F.; Clark, I. B. N.; Swain, P. S.; Pilizota, T. *Sci. Rep.* **2016**, *6*, 38828.
- (495) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *J. Chem. Theory Comput.* **2015**, *11*, 3696.
- (496) Kirschner, K. N.; Yongye, A. B.; Tschampel, S. M.; González-Outeiriño, J.; Daniels, C. R.; Foley, B. L.; Woods, R. J. *J. Comput. Chem.* **2008**, *29*, 622.
- (497) Gotz, A. W.; Williamson, M. J.; Xu, D.; Poole, D.; Le Grand, S.; Walker, R. C. *J. Chem. Theory Comput.* **2012**, *8*, 1542.
- (498) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. *Numerical-Integration of Cartesian Equations of Motion of a System with Constraints – Molecular-Dynamics of N-Alkanes*, 1977.
- (499) Onufriev, A.; Bashford, D.; Case, D. A. *Proteins: Structure, Function, and Bioinformatics* **2004**, *55*, 383.
- (500) Weiser, J.; Shenkin, P. S.; Still, W. C. *J. Comput. Chem.* **1999**, *20*, 217.
- (501) Genheden, S.; Kuhn, O.; Mikulskis, P.; Hoffmann, D.; Ryde, U. *J. Chem. Inf. Model.* **2012**, *52*, 2079.

General conclusion

The aim of this thesis was to explore fundamental forces that govern the self-assembling of multimolecular edifices in water thanks to two different class of receptors, cylindric and ramified.

Firstly, a large family of various 1,4-bisthiophenols were synthesised. They self-assembled in cylindric polycarboxylate disulfide bridge cyclophanes through DCC. Among them, the polyanionic macrocycle **I-14** showed tremendous affinity for polycationic amines in solutions with K_a up to 10^{11} M^{-1} . The selectivity of the receptor for 1) a series of α,ω -alkyldiammoniums, 2) spermine and spermidine was studied by a combination of experimental and theoretical studies. The driving forces of the selectivity were analysed and the influence of solvation effects on the associations were highlighted.

The formation of heterocyclo-oligomers that are mutants of **I-14**, regarding its chemical function, spartially met synthetic issues. Nevertheless, a successful expansion of the diameter of the cavity was obtained by the self-assembly through DCC of a benzylic analogue of **I-1**. Discrete macrocycles were observed, and a strained dimer was isolated, crystallised and post-functionalised. The distribution of macrocycles was influenced by templating. A further modification of the link between the aromatic units, turning disulfide bridges into thioether was observed but could not be managed.

An expansion of the height of the macrocyclic cavity was obtained by using bulk 1,4-bisthiophenols with distant thiols and carboxylates. Two diastereoisomers showed opposite self-assemblies in aqueous solutions. Especially, a single cyclic tetramer over all the oligomers and over the 24 accessible macrocyclic configurations was obtained by a spectacular selective self-assembly of **I-105**. The structural elucidation resulted from a combination of experimental and computational study and will be confirmed by a crystallographic structure. The driving force of the self-assembly was first supposed to be the London forces, which was invalidated by DFT. The assembly was probably governed by repulsive steric effects.

In addition, the synthesis of 1,4-bisthiophenols bearing side chains on their 2,5 positions with functional groups such as alcohols and amines was successfully realised from **I-1** thanks to trityl and cyanopropyl protecting groups. This building block were dedicated to the formation of disulfide bridge cyclophane by DCC, that would be effective in the recognition of therapeutic proteins that are affected by PTM. The design of the receptors were inspired by the analysis of specific enzymes and reader proteins of the targeted PTMs, deamidation and methionine oxidation. However, an unexpected oxidation of thiols into cyclic sulfenyl amides prevented from an efficient oligomerisation of the monomers. The family of corona[n]arenes and the simultaneous exchange of disulfide and acylhydrazone bonds was also explored in this context but met synthetic issues. Water-soluble cyclophanes will continue to play a major role in molecular recognition and fundamental studies, as highlighted by the investigation of solvation and steric effects. Self-assembly through DCC provided a variety of macrocycles, which are promising platform for different applications.

Finally, ramified dendrigrafts of lysine (DGL) were assessed for a variety of applications in water, focussing on the detection of bio(macro)molecules and the neutralisation of heparin in replacement of protamine, the only FDA-approved drug in case of heparin overdose. A second generation of DGL

inhibited the anticoagulant activity of unfractionated heparin (UFH), low molecular-weight heparin (LMWH) and fondaparinux more efficiently than protamine did in human plasma. A computational study by molecular dynamic simulations revealed that this superior effect of DGL was due to its conformational flexibility. In fact, the branched architectures enabled the dendrimer to marshal more efficiently its charges in the association with heparin compared to protamine. The simulations revealed that mid size and long heparin formed a more stiffened complex with DGL, which explains its better neutralisation of UFH and LMWH compared to protamine. Regarding the binding of heparin with protamine, the study showed that the relative size of the partners has a substantial influence on their corresponding binding affinities, where the number of direct contacts (i.e., electrostatic or hydrogen bonds) is not the only factor guiding the observed selectivity. The positive surface of DGLs was also used to quench the fluorescence of homopolymers of acid aspartic and fluorescein upon aggregation. The disruption of the complex by an external partner restored the fluorescence and was used as a strategy of detection, resulting in various limits of detection for spermine and bacteria. The DGLs' ability to mimic esterases was also validated.

The two family of receptors, cylindric disulfide cyclophanes and ramified DGLs, illustrated the complexity of the supramolecular assemblies in water. In fact, solvation effects could affect the selectivity of a binding event, as well as long range interactions and steric effects. The combination of theoretical and computational study will continue to help the chemist to improve the design of new molecules and to better understand the natural systems.

Complex molecular architectures for the recognition of therapeutic bio(macro)molecules

The recognition of biomolecules in complex biological media is a challenge associated with various therapeutic applications. The chemists can address this issue following two approaches: either they design and synthesise molecules or they select commercially available or natural molecules. Following the last strategy, dendrigraft of lysine (DGL) efficiently neutralised all classes of the anticoagulant heparin, with a superior effect compared to protamine, the only FDA-approved drug in case of heparin overdose. A study by molecular dynamic revealed the mechanism of binding between heparins and DGL and protamine respectively. At the opposite of this approach, we used dynamic combinatorial chemistry in order to obtain disulfide bridged cyclophanes from the self-assembly of various 1,4-bisthiophenols by oxidation of thiols into disulfide bonds. By a combination of theoretical (DFT and molecular dynamic) and experimental studies, we investigated the driving forces and the influences of fundamental concepts such as solvation and steric effects for the self-assembly of these polythiols and the binding of the corresponding cavitands with therapeutic biomolecules.

Keywords: Dynamic combinatorial chemistry, 1,4-bisthiophenol-2,5-functionalised building blocks, post-translational modifications, molecular recognition, heparin neutralisation, dendrigraft of lysine, computational chemistry

Architectures moléculaires complexes pour la reconnaissance de bio(macro)molécules d'intérêt thérapeutique

La reconnaissance de biomolécules dans des milieux biologiques complexes est un réel défi pour les chimistes et les biologistes, associé à des enjeux médicaux majeurs. Face à cette problématique, le chimiste peut choisir d'utiliser des molécules désignées par ses soins, ou encore de sélectionner et d'utiliser directement des structures commerciales ou naturelles. Suivant cette dernière approche, les dendrigrafts de lysines (DGL) ont montré une neutralisation des héparines de différentes tailles supérieures à l'action de la protamine, le seul médicament autorisé en cas de surdosage de l'anticoagulant. Une étude par dynamique moléculaire a permis de mettre en avant le mécanisme d'interaction entre les héparines d'une part, et les DGLs et la protamine d'autre part. Par ailleurs, suivant la première approche de design et synthèse, nous avons utilisé la chimie combinatoire dynamique pour obtenir des nouveaux récepteurs synthétiques à partir de briques moléculaires diverses de type 1,4-dithiophénols. Des études à la fois théorique, en DFT et dynamique moléculaire, et expérimentale, ont été menés pour comprendre les phénomènes régissant l'auto-assemblage de ces briques en oligomères cycliques et la complexation de ces cavitands avec des biomolécules d'intérêt thérapeutiques.

Mots-clés: Chimie combinatoire dynamique, briques 1,4-dithiophénol-2,5-fonctionnalisées, modifications post-translationnelles, reconnaissance moléculaire, neutralisation des heparines, dendrigrafte de lysine, chimie computationnelle