



HAL
open science

Chimie combinatoire dynamique pour l'identification de mimes de protéines

Benjamin Zagiel

► **To cite this version:**

Benjamin Zagiel. Chimie combinatoire dynamique pour l'identification de mimes de protéines. Chimie organique. Sorbonne Université, 2018. Français. NNT : 2018SORUS520 . tel-03001244

HAL Id: tel-03001244

<https://theses.hal.science/tel-03001244v1>

Submitted on 12 Nov 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.

Sorbonne Université

École doctorale 406 – Chimie Moléculaire de Paris Centre

*Laboratoire des BioMolécules, UMR7203 – Équipe 1- Peptides, glycoconjugués et métaux en
biologie*

Chimie combinatoire dynamique pour l'identification de mimes de protéines

Par Benjamin ZAGIEL

Thèse de doctorat de Chimie Bioorganique

Dirigée par le Dr. Roba Mourné

Présentée et soutenue publiquement le 19 Octobre 2018

Devant un jury composé de :

Maria Duca	Chargée de recherche	Rapportrice
Laurent Micouin	Directeur de recherche	Rapporteur
Bernold Hasenknopf	Professeur	Président du jury
Roba Mourné	Maître de conférences	Directrice de thèse
Philippe Karoyan	Professeur	Invité

Remerciements

Je tiens tout d'abord à remercier le Docteur Sandrine Sagan pour son accueil au sein du laboratoire, il y a trois ans (et un peu plus) et pour avoir fait en sorte que je puisse y rester pour conduire ma thèse.

Je tiens à remercier le Docteur Maria Duca et le Docteur Laurent Micouin pour le temps qu'ils ont consacré à l'examen de ce manuscrit en tant que rapportrice et rapporteur. Je souhaite également remercier le Professeur Bernold Hansenknopf et le Professeur Philippe Karoyan pour avoir accepté de participer à mon jury de thèse et ainsi d'évaluer mon travail. Je remercie collectivement les membres du jury pour la conversation scientifique animée et passionnante qui s'est déroulée lors de ma soutenance et pour leurs remarques constructives.

Je ne l'ai pas encore citée car elle mérite grandement un paragraphe pour elle toute seule : ma directrice de thèse, ma « boss », le Docteur Roba Moumné. Je souhaite te remercier de tout mon cœur pour m'avoir proposé de commencer à travailler avec toi il y a bientôt quatre ans lors de mon stage de M2. Je ne me voyais pas ailleurs et sur un autre sujet que celui-ci. Je te remercie profondément de m'avoir fait confiance et de m'avoir accepté pour continuer en thèse. *Le deal* était que tu passes ton HDR avant la fin de ma thèse. Mission acceptée, mission accomplie ! Félicitations à toi, je sais que tu t'es beaucoup investie pour l'accomplir.

Ce travail, c'est aussi ton bébé en quelque sorte, ta thématique, et je suis fier d'avoir pu participer à la développer. Il y a eu des hauts et des bas, des obstacles que l'on n'avait pas prévu, mais au final ça a été un vrai plaisir de travailler avec toi, d'apprendre avec toi (depuis assez longtemps au final !) et de grandir avec toi en tant que scientifique. Merci pour tout ce que tu m'as appris, au labo bien sûr mais aussi en dehors. En particulier, de ne jamais baisser les bras, de toujours s'accrocher et d'aller le plus loin possible. Je te souhaite tout le meilleur et du succès pour continuer dans ce monde fantastique de la DCC et continuer à cultiver ces idées géniales dont tu as le secret. J'y ai été sensible et le suis encore... D'autres seront également emballés par ces projets innovants aussi !

Je crois me souvenir qu'à l'occasion d'un oral en M2, probablement pour l'UE 5C406, une personne du jury t'avais donné la parole en te qualifiant de chimiste bioorganicienne. C'était, il me semble, la première fois que j'entendais un chimiste se définir ainsi. Ca me plaisait bien : ce domaine intermédiaire entre la chimie organique de synthèse, les molécules biologiques et les outils chimiques pour jouer avec et tout un tas d'applications potentielles pour aller jouer avec les copains biologistes. Merci vraiment de m'avoir aidé et guidé dans ma

conception de la chimie et dans la construction du type de chimiste que je suis. Tu as été, es et seras encore longtemps un modèle pour moi pour cela. Je ne sais pas si je peux m'auto-qualifier comme ton premier thésard, ça n'est pas très important... Je suis simplement très fier d'avoir été ton étudiant, ton élève, merci.

Je remercie les différentes personnes qui ont participé à ce travail de thèse et qui m'ont transmis leur savoir-faire. Je commence par le Docteur Emmanuelle Sachon, Manu, qui s'est plongée dans les aspects d'analyses en MS et en digestion enzymatique, parfois à s'en arracher les cheveux. Merci pour tout le temps que tu as consacré à ce travail, que tu m'as consacré. C'était déjà un plaisir de t'avoir en tant qu'enseignante en L3 mais c'était un vrai plaisir de travailler avec toi malgré toutes les difficultés que l'on a rencontrées.

Je remercie également le Docteur Emeric Miclet pour son investissement sur les aspects RMN et MD de ce travail. Merci pour le temps passé au spectromètre à m'apprendre à programmer mes expériences, à décortiquer les spectres et à faire mes premiers pas en simulation. Le temps nous a un peu manqué et je sais que tu as sacrifié du temps cet été pour finaliser ce travail, merci beaucoup.

Je remercie du fond du cœur Rodrigue Marquant dit Rodrigou. Merci pour tout ce que tu m'as appris en synthèse peptidique et en HPLC, merci aussi pour le temps que tu as consacré à ces petites hélices. Tu étais toujours là quand un souci survenait et il y'en a eu ... Sans toi le labo ne serait pas ce qu'il est et merci de tout ton investissement et ton dévouement qui m'a permis (et a permis et permet à tous les doctorants du labo) de travailler correctement.

Je remercie également les personnels de la plateforme de MS : Lucrèce et Gilles pour leur temps et leurs conseils ainsi que Christophe de la plateforme de synthèse peptidique pour les synthèses.

Je remercie les collaborateurs de l'extérieur du labo en particulier le Docteur Nathalie Gagey-Eilstein avec qui j'ai passé deux jours de manip toutes nouvelles pour moi et que j'ai adoré. Je regrette que le projet ne soit pas allé assez loin pour travailler davantage avec toi.

Je remercie tous les stagiaires que j'ai pu encadrer durant ce travail de thèse à commencer par Taleen Peker. Ma petite Taleennette, merci pour tous ces moments de rire et de joie au labo durant la fin de ma thèse. Merci pour tout le travail que tu as accompli avec passion, bonne humeur et acharnement parfois. Sans toi ce travail n'aurait pas été ce qu'il est et je pense que ma fin de thèse aurait été bien plus difficile à vivre. Merci pour ce souffle de motivation en

fin de parcours, merci du fond du cœur ... Je te félicite encore pour ton M2 et pour la thèse que tu entames après la mienne. Plein de succès et de belles découvertes à toi avec Roba et Chahrazade ! Garde toujours ce bel état d'esprit et cette bonne humeur ! J'espère te revoir car tu vas me manquer.

Merci aux différents stagiaires de L3 et L3 Pro, les petites mains du labo, Gabrielle, Laetitia et Florent, qui ont été vraiment sympathiques et acharnés au travail. Enfin, merci à Renaud pour ce semestre passé à tes côtés. Ces bons moments passé à se casser la tête sur ces synthèses.

Je souhaite remercier le Professeur Philippe Karoyan, le professeur Solange Lavielle et le professeur Bernold Hasenknopf. Le premier pour les premiers cours de chimie organique que j'ai eu à l'université et qui malgré des débuts difficiles m'ont ouvert l'esprit à ce domaine que j'affectionne et puis plus tard à l'interface avec la biologie en M1.

Entre temps, j'ai connu Solange que je remercie également pour les cours de L3 et des toutes les années suivantes. Je me revois, petit L3 sur les bancs de l'amphi, à me dire que les biomolécules et leur labo, « ça me tentait bien ». Merci infiniment de m'avoir appris et montré une chimie qui me faisait me dire « c'est ça que je veux faire ». Merci également à Annie et Dominique.

Enfin, merci à Bernie, pour son merveilleux cours de M2 et sa passion de l'enseignement qu'il a su me transmettre je pense. Merci de m'avoir transmis l'envie de faire de la chimie d'une façon que je ne connaissais pas et qui continue aujourd'hui de me donner des idées et de m'ouvrir l'esprit. Sans cela, je n'aurais pas choisi ce sujet de thèse.

Je tiens à remercier mes anciens camarades doctorants : Mathou, Seb, Fred, Aude, Lucille et Emilie pour ces très bons moments passés ensemble ici où à boire des bières. Sans vous cette thèse n'aurait pas été aussi folle et agréable. Merci particulièrement à Mathou, partager ces années avec toi a été un pur plaisir et une amitié forte en est né. Des bisous ma belle ! Même si on ne se voit plus autant, tu es toujours un petit peu là quelque part dans ma tête et mon cœur.

Merci à mes camarades de promo et aux arrivants des années suivantes pour les bons moments partagés également : Damien toujours à l'heure pour déjeuner, Didier, mon camarade de la Fête de la Science, Simon pour cette belle amitié, Lucrezia, Roxanne et Laurent toujours partants pour aller boire un coup, Leïla pour les folles soirées trop peu nombreuses à mon goût, Pauline (Paulinou) pour ton entrain, ta bonne humeur et ta passion des bons restos, Antoine

pour ces trop courts et trop rares moments de rire, de discussions tout en détente et parfois de science. A bientôt mon petit chat. Eva pour ton sourire, ta bonne humeur à chaque fois que tu passais. Antonio adorable visiteur d'Italie, à bientôt je l'espère. Agathe P., Lucas, Amandine, Alex, Bingwei, Shadaï, Soha, Martha. Agathe B. et Paula, rencontres de dernière minutes mais attachantes. J'espère vous revoir. Prenez soin de ma petite Taleennette, les super thésardes. Merci pour les déjeuners pour les uns, les bières pour les autres, les têtes parfois, les retours à pied pour certains et à tous merci pour l'écoute quand j'avais besoin de râler un bon coup. Merci à mes camarades du haut de la montagne à l'ENS pour avoir maintenu le contact et pour les sorties quasi hebdomadaires pour boire un coup au moins et malgré parfois la difficulté d'être un peu isolé. Tâchons de tous garder contact.

Je réitère un merci à Mathou et j'y ajoute Ludo. Merci à toi Mathou pour la reprise du sport et merci à vous deux pour la motivation et les bonnes séances de dépense.

Je termine en remerciant tous les membres du labo pour tous les conseils et les discussions scientifiques qu'ils ont pu m'apporter durant ce temps au sein du LBM. Merci à Cillian pour tous ces conseils en synthèse et en HPLC et pour les cafés (et Solange et Christelle aussi). Merci à Claire pour sa bonne humeur dès le matin autour d'un bon café. Merci à Karine pour les cafés également mais aussi pour les échanges au labo. Merci à Astrid et Nicolas pour les cafés aussi (oui j'en ai piqué un peu partout) et merci de nouveau à toi Astrid pour les quelques échanges scientifiques et autour de l'enseignement. Merci à Jean-Jacques pour le boulot qu'on a fait autour de l'agrégation, merci pour les clichés de MET. Merci à toi et Lucie pour la rigolade. Merci à nouveau à Ludo, en plus du sport. Même si l'on n'a pas travaillé ensemble, toujours une oreille disponible pour discuter. Merci aux personnes du 4^e étage : Sophie, Fabienne, Françoise, Delphine, Thierry pour un bonjour, un rire partagé ou une discussion scientifique animée.

Pour finir, un merci aux amis de longue date ou pas des autres labos, en particulier Jorge. On se suit depuis un moment et merci d'avoir été là tout au long de ces années, merci pour les apéros bières, les soirées et les vacances. Nos chemins s'éloignent mais pas facile de se débarrasser de moi, tu vas voir !! Merci aussi à Lorien pour le SOS DMF et également à Flo pour les apéros et les cafés sur le parvis, une amitié faite de petites choses mais de bons moments que je tiens à garder. Une petite pensée pour Sha et Mariecka, merci pour ces moments de respiration et de détente autour d'un verre, d'un bon repas de quelques jours de vacances.

J'adresse un dernier mais immense merci à mes parents sans qui je ne serai pas là aujourd'hui à écrire ces mots. Merci à vous de m'avoir toujours poussé à aller jusqu'au bout de mes envies, sans restriction et toujours dans le soutien. Merci de m'avoir appris à apprendre et d'aimer ça, toujours. Merci de m'avoir permis de cultiver mon goût pour les sciences, de m'avoir transmis l'envie et la curiosité de comprendre le monde et de m'avoir encouragé, quel que soit le chemin que j'ai décidé d'emprunter (et il a changé pas mal de fois). Merci d'avoir toujours été là, de l'être encore, comme soutien, comme socle et comme force dans la vie.

Ce travail m'a fait grandir, m'a réjoui et m'a permis d'apprendre encore et toujours de nouvelles choses. Il a été plein de bons moments et de moments plus difficiles. De hauts et de bas. Je suis très fier de l'avoir réalisé.

A Mamie Jojo
A Mémé Hélène
A Pépé Sewek
A Papy Pierrot

Sommaire

REMERCIEMENTS	1
SOMMAIRE	7
ABREVIATIONS	9
INTRODUCTION	15
I. <i>Les interactions protéine-protéine et leurs inhibiteurs</i>	15
1) Rôle physiologique & implications pathologiques	15
2) Conception d'inhibiteurs d'interactions protéine-protéine (PPI).....	17
a) Des cibles difficiles « undruggable »	17
b) Conception d'inhibiteurs d'interactions protéine-protéine.....	18
i) Criblage aléatoire	19
ii) Conception rationnelle.....	21
II. <i>Développement de Mimes Protéiques par Chimie Combinatoire Dynamique</i>	26
1) La Chimie Combinatoire Dynamique	27
2) Conception de mimes protéiques par DCC : État de l'art.....	28
III. <i>Objectifs</i>	37
CHAPITRE 1 : CONCEPTION ET SYNTHÈSE DE BLOCS ET SCAFFOLDS PEPTIDIQUES POUR LA CHIMIE COMBINATOIRE DYNAMIQUE.....	41
I. <i>Choix de la réaction de DCC</i>	41
II. <i>Synthèse des blocs thioesters</i>	43
1) Considérations structurales.....	43
2) Schéma général de synthèse	44
3) Blocs à chaînes aliphatiques et aromatiques	44
4) Blocs à chaînes fonctionnalisées et potentiellement réactives.....	45
5) Perspectives	51
III. <i>Synthèses de scaffolds peptidiques</i>	52
1) Structures tridimensionnelles	52
a) Scaffold en hélice α	52
b) Scaffold en β -hairpin	53
2) Séquence des scaffolds peptidiques.....	55
a) Résidus impliqués dans la DCC.....	56
b) Sonde spectroscopique.....	57
c) Solubilité	63
3) Synthèse des scaffolds.....	64
a) Scaffolds en hélice- α	64
b) Scaffold β -hairpin	67
4) Bilan et perspectives	69
CHAPITRE 2 : CHIMIE COMBINATOIRE DYNAMIQUE PAR TRANSTHIOESTERIFICATION SUR SCAFFOLDS EN EPINGLE A CHEVEUX β.....	71
I. <i>Généralités</i>	71
II. <i>Fonctionnalisation d'un peptide modèle à une cystéine</i>	72
1) Mise au point des conditions expérimentales.....	72
2) Influence de la nature des blocs sur la constante d'équilibre	77
3) Librairie issue d'une plateforme à un site fonctionnalisable.....	80
4) Réversibilité du système.....	82
5) Combinatoire dynamique sur un système à deux cystéines	84
6) Combinatoire dynamique sur un systèmes à quatre cystéines : Hairp4-Lys	86
a) Exemple d'une librairie simple à un bloc	86

CHAPITRE 3 : APPLICATION A L'IDENTIFICATION DE NOUVELLES STRUCTURES CPP	91
I. Contexte et choix de la cible	91
II. Librairies à un bloc	92
III. Librairie mixte à deux blocs.....	94
IV. Discussion	96
V. Conclusion et perspectives.....	97
CHAPITRE 4 : METHODOLOGIES DE LINEARISATION DE PEPTIDES CYCLIQUES β-HAIRPIN POUR LEUR ANALYSE EN SPECTROMETRIE DE MASSE.....	101
I. Généralités et problématiques.....	101
II. Ouvertures sélectives du cycle peptidique	105
1) Ouverture par ISD.....	105
2) Ouverture chimique au bromure de cyanogène (CNBr).....	110
3) Ouverture enzymatique par hydrolyse trypsique	112
a) Mise au point des conditions	112
4) Évolution de la structure du <i>scaffold</i> et nouveaux défis	117
III. Conclusion.....	118
CHAPITRE 5 : ÉTUDES STRUCTURALES.....	125
I. Problématiques.....	125
II. Simulation de dynamique moléculaire sur <i>Hairp4-Lys</i>	126
III. Étude RMN sur la plateforme non fonctionnalisée <i>Hairp4-Lys</i>	135
IV. Étude RMN comparative sur <i>Hairp4-tetra(BBGly)</i> et <i>Hairp0</i>	138
V. Discussion et perspectives.....	139
CONCLUSION ET PERSPECTIVES	143
EXPERIMENTAL SECTION	147
I. Building Blocks Synthesis	147
II. Peptide Synthesis	173
1) Synthesis of unnatural amino acids.....	174
2) Synthesis of β -Hairpin Peptides	180
3) Synthesis of Stapled Helix Peptides.....	184
III. Dynamic Combinatorial Chemistry	187
IV. Structural Investigations.....	188
ANNEXES.....	191
RÉFÉRENCES BIBLIOGRAPHIQUES	203

Abréviations

Ac :acétyle

ACN : acétonitrile

ADN : acide désoxyribonucléique

Ala, A : alanine

Alloc : allyloxy-carbonyle

Aq. : aqueux (aqueous)

APS : Persulfate d'ammonium

Arg, R : arginine

ARN : acide ribonucléique

ARNm : acide ribonucléique messenger

Asp, D : Acide aspartique, aspartate

Asn, N : Asparagine

BAIB : Diacétate d'iodobenzène

BB : building block

Bn : benzyle

Boc : *tert*-butyloxy-carbonyle

Bt : benzotriazole

CID : fragmentation induite par collision

2-Cl-Z : 2-Chlorobenzoyl

CPP : peptide pénétrateur de cellule

C-ter : C-terminal

Cys, C : cystéine

1,5-DAN : 1,5-Diaminonaphthalène

Dap : acide diaminopropionique

4-DAPA : 4-diméthylaminophthalimide

DBU : 1,8-diazabicyclo[5.4.0]undéc-7-ène

DCC : chimie combinatoire dynamique

DCL : librairie combinatoire dynamique

DCM : dichlorométhane

2,5-DHB : acide 2,5-dihydroxybenzoïque

DIEA : diisopropyléthylamine

DMF : *N,N*-diméthylformamide

DSS : acide 4,4-diméthyl-4-silapentane-1-sulfonique

ECD : fragmentation induite par capture d'électron
EDT : 1,2-éthanedithiol
Equiv. ou éq. : équivalent
ESI : ionisation par électronébuliseur
Et : éthyle
ETD : fragmentation induite par transfert d'électron
Fmoc : 9-fluorénylméthylloxycarbonyle
GAG : glycoaminoglycane
Glu, E : acide glutamique, glutamate
Gln, Q : glutamine
Gly, G : glycine
HCCA : acide α -cyano-4-hydroxycinnamique
HATU : hexafluorophosphate de (diméthylamino)-N,N-diméthyl(3H-[1,2,3]triazolo[4,5-b]pyridin-3-yloxy)méthaniminium
HBTU : hexafluorophosphate de (1H-benzotriazol-1-yloxy)(diméthylamino)-N,N-diméthylméthaniminium
His, H : histidine
HOBT : 1-hydroxybenzotriazole
HF : acide fluorhydrique
HFIPA : hexafluoropropan-2-ol
*i*Bu : isobutyle
IM-MS : spectrométrie de mobilité ionique-spectrométrie de masse
ISD : fragmentation en source
Ile, I : isoleucine
K' : constante d'équilibre apparente
LC : chromatographie liquide
Leu, L : leucine
Lys, K : lysine
MALDI : désorption ionisation laser assistée par matrice
MBHA : 4-méthylbenzyldiamine
MD : dynamique moléculaire
Me : méthyle
Met, M : méthionine
Mob : méthoxybenzyle
MS : spectrométrie de masse

Mtt : 4-méthyltrityle
Naphth : 2-Naphthoyle
NCL : ligation chimique native
n.d. : non déterminé
2-Npa : 2-Nitrophénylalanine
2-Npg : 2-Nitrophénylglycine
N-ter : N-terminal
PAGE : électrophorèse sur gel de polyacrylamide
Pbf : 2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyle
Pd/C : palladium sur charbon
Ph : phényle
Phe, F : phénylalanine
2-PhiPr : 2-phénylisopropyle
*i*Pr : isopropyle
*n*Pr : n-propyle
pNA : para-nitroanilide
PNA : acide nucléique peptidique
PPI : interaction protéine-protéine
Pro, P : proline
PSD : fragmentation post-source
RMN : résonance magnétique nucléaire
RMSD : variation quadratique moyenne
RP-HPLC : chromatographie liquide haute performance en phase inverse
RX : rayons X
SDS : dodécylsulfate de sodium
Ser, S : sérine
SPPS : synthèse peptidique sur support solide
TBAF : fluorure de tétrabutylammonium
TBS : *tert*-butyldiméthylsilyl
*t*Bu : *tert*-butyle
TCEP : tris(2-carboxyethyl)phosphine
TEMPO : 1- λ^1 -oxidanyl-2,2,6,6-tétraméthylpipéridine
TFA : acide trifluoroacétique
THF : tétrahydrofurane
Thr, T : thréonine

TIS : triisopropylsilane
TLC : chromatographie sur couche mince
TNBS : acide 2,4,6-trinitrobenzenesulfonique
Trt : trityle
Trp, W : tryptophane
ToF : temps de vol
Tos : tosyle
Tyr, Y : tyrosine
TNF : tumor necrosis factor
UPLC : ultra performance liquid chromatography
UV : ultraviolet
Val, V : valine
VEGF(R) : vascular endothelial growth factor (receptor)
VIH : virus de l'immunodéficience humaine

Introduction

I. Les interactions protéine-protéine et leurs inhibiteurs

1) Rôle physiologique & implications pathologiques

Les protéines sont les macromolécules du Vivant ayant pour rôle d'assurer les différentes fonctions nécessaires au bon fonctionnement de l'organisme à l'échelle moléculaire. Qu'il s'agisse de transmission de signal, de catalyse pour le métabolisme, de régulation de différents phénomènes physiologiques, de reconnaissance moléculaire, toutes ces fonctions sont réalisées à l'aide de protéines et plus précisément grâce aux interactions que ces protéines établissent entre elles ou avec d'autres macromolécules biologiques comme certains oligosaccharides ou acides nucléiques¹ (Figure 1).

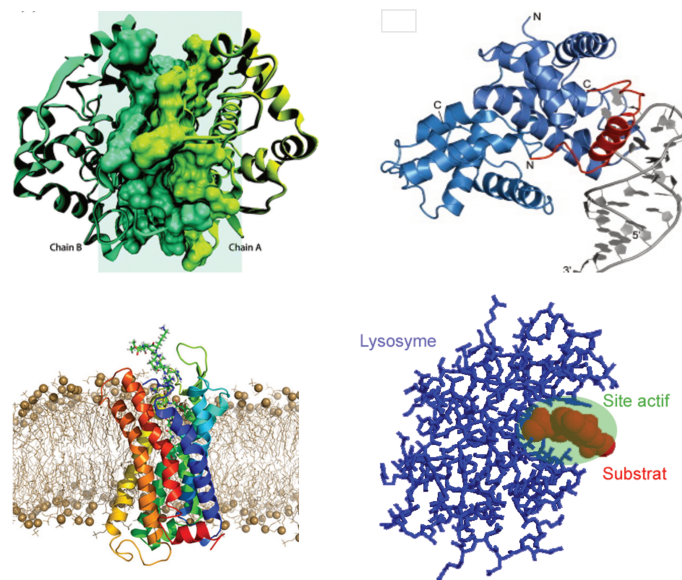


Fig. 1 : Illustrations d'interaction protéine-protéine (en haut à gauche), protéine-ADN (en haut à droite), protéine-lipides (en bas à gauche), enzyme-substrat (en bas à droite).

Ces interactions forment des réseaux complexes au sein de la cellule et leurs perturbations peuvent entraîner des situations pathologiques, qu'elles proviennent de causes génétiques ou bien environnementales². La protéine p53, facteur de transcription suppresseur de tumeur et son réseau d'interactions constituent un exemple très étudié. Cette protéine est au cœur d'un système de régulation lié notamment au système de réparation des dommages de l'ADN. Comme le montre la figure 2, de nombreux signaux ou voies sont capables d'activer cette protéine qui à son tour va stimuler l'expression de différents gènes en aval impliqués dans

de nombreuses fonctions comme la régulation du cycle cellulaire, la réparation de l'ADN, l'apoptose ou l'angiogénèse. Dans près de 50% des cancers, cette protéine est mutée³. Si cela a lieu au niveau de son domaine d'interaction avec l'ADN, elle ne peut plus stimuler l'expression des gènes impliqués dans les mécanismes anti-cancer cités plus haut.

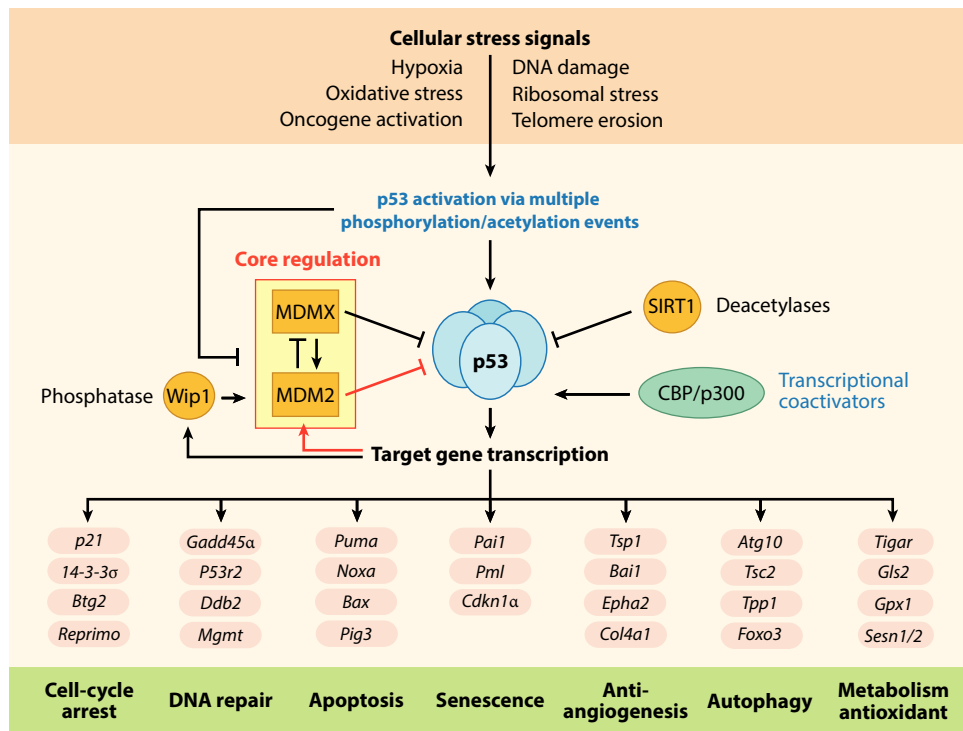


Fig. 2 : Schéma des voies de régulations autour de la protéine p53. Différentes modalités de stress cellulaire participe à l'activation de p53 qui sur le plan moléculaire se déclenche suite à de multiples modifications post-traductionnelles. Une fois activée, p53 va participer à l'activation de l'expression de gènes impliqués dans la régulation de ce stress. Des protéines annexes comme MDMX et MDM2 mais aussi Wip1 et SIRT1 constituent des systèmes de régulation de p53.

Cette situation est commune à un bon nombre de pathologies⁴:

- Soit les protéines mutées ou anormales entraînent des interactions dysfonctionnelles. C'est le cas du cancer avec p53 comme vu précédemment. On peut également citer la maladie d'Hungtinton, dans laquelle l'hungtintine mutée tend à former des agrégats dans les neurones conduisant à la dégénérescence de ces cellules et aux symptômes de la maladie^{5,6,7,8}. Il en va de même pour d'autres maladies neurodégénératives comme la maladie d'Alzheimer avec la protéine Tau par exemple^{9,10}. Dans ce cas, c'est le système de régulation de l'état de phosphorylation de la protéine Tau qui est défectueux entraînant des formes pathologiques hyper phosphorylées de Tau qui participent à la dégénérescence neuronale.

- Soit les fonctions sont correctement réalisées mais dans un contexte pathologique, ce qui conduit à favoriser la maladie et à en promouvoir les symptômes. Dans un contexte tumorale, le VEGF (*vascular endothelial growth factor*) est souvent surexprimé. Cette protéine stimule l'angiogenèse, c'est à dire la croissance de nouveaux vaisseaux sanguins, en interagissant avec son récepteur membranaire le VEGFR (*vascular endothelial growth factor receptor*). Ainsi, les cellules tumorales utilisent une interaction physiologique afin de pallier le manque de ressources en nutriments qu'entraîne leur multiplication anarchique. On retrouve l'implication de cette interaction dans d'autres pathologies comme des maladies cardiovasculaires, la pré éclampsie ou la dégénérescence maculaire liée à l'âge¹¹. Ce mécanisme est également la cause de toutes les maladies infectieuses pour lesquelles les interactions entre les pathogènes et les protéines de l'hôte permettent le détournement de la machinerie métabolique de l'hôte au profit de l'agent infectieux¹² : l'infection au VIH (*virus de l'immunodéficience humaine*)^{13,14}, la rougeole¹⁵ et l'hépatite C¹⁶ pour les infections virales ; la maladie de Lyme¹⁷ ou la méningite bactérienne¹⁸ pour les infections bactériennes par exemple.

Dans un cas comme dans l'autre, cibler ces interactions dysfonctionnelles ou indésirables peut être un moyen de lutter contre la maladie en restaurant une fonction défectueuse ou en empêchant la pathologie. Pour parvenir à cela, il est tout d'abord nécessaire d'identifier ou de concevoir des molécules capables d'interagir avec les protéines cibles.

2) Conception d'inhibiteurs d'interactions protéine-protéine (PPI)

a) *Des cibles difficiles « undruggable »*

Comme vu précédemment, les interactions protéine-protéine sont au cœur de nombreuses pathologies et représentent de ce fait des cibles thérapeutiques potentielles de premier intérêt. Elles sont malheureusement encore sous exploitées puisque moins de 0,01 % des interactions protéine-protéine humaines ont été ciblées par un inhibiteur. Elles sont d'ailleurs souvent qualifiées d'« undruggable », un terme que l'on pourrait traduire par « inmédicamentable »^{19,20}. Pour comprendre ce terme, il est plus simple d'essayer de définir ce qu'est une cible « druggable » ou « médicamentable ». Deux critères sont à considérer pour qualifier une cible de « druggable »^{21,22}. Le premier est l'existence d'un effet thérapeutique à la suite de l'interaction entre la cible et une molécule inhibitrice. Le second est sa « ligandabilité », autrement dit sa capacité à lier de petites molécules avec une forte affinité²³. Cette propriété de la cible est évaluée par des méthodes computationnelles ou de biophysique expérimentale. Elles

permettent d'identifier de potentiels sites de liaison en analysant la géométrie des reliefs (poches, cavités) à la surface de la cible ou en repérant des interactions favorables sur le plan énergétique.

La difficulté vient du fait que les surfaces d'interaction entre protéines et petites molécules sont plutôt petites (entre 300 et 1000 Å²)^{24,25} alors que les surfaces d'interaction entre protéines sont plus vastes (entre 1500 et 3000 Å²)^{26,27}. De plus, elles ne possèdent que peu de reliefs définis contrairement aux interfaces entre protéines et petites molécules. Les surfaces entre protéines sont plutôt plates alors que celles entre protéines et ligands présentent des creux et des poches (Figure 3). Enfin, quand il s'agit d'inhiber une interaction entre une protéine et une petite molécule, comme par exemple une interaction enzyme-substrat ou ligand-récepteur, la petite molécule est en générale utilisée comme point de départ pour la conception d'un nouveau ligand plus performant. Cette conception rationnelle est en général beaucoup plus complexe lorsque le ligand est lui-même une protéine.

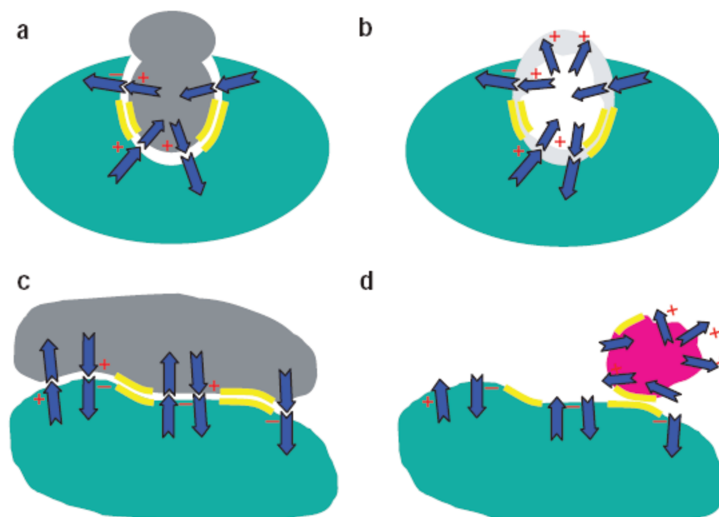


Fig. 3 : Illustrations des différences au niveau des surfaces d'interaction entre a) une enzyme et son substrat, b) une enzyme et une petite molécule inhibitrice, c) deux protéines et d) une protéine et une petite molécule.

b) Conception d'inhibiteurs d'interactions protéine-protéine

Deux types d'approches peuvent être envisagées pour développer des inhibiteurs d'interactions protéine-protéine : l'identification par criblage aléatoire de banques de molécules et la conception rationnelle basée sur la cible (Fig. 4).

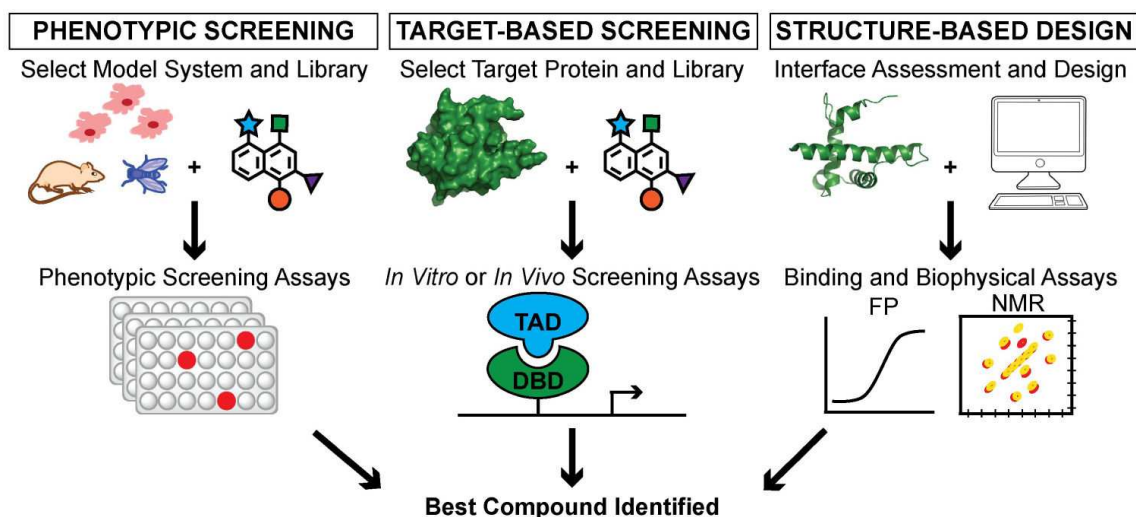


Fig. 4 : Les différentes stratégies d'identification ou conception d'inhibiteurs d'interactions protéine-protéine. A gauche : le criblage phénotypique. Au milieu : le criblage basé sur la cible. A droite : la conception basé sur la structure¹⁹.

i) Criblage aléatoire

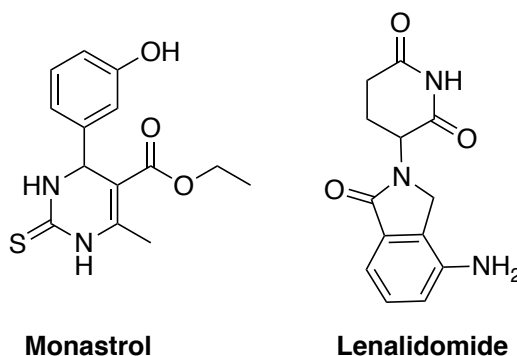
La première solution consiste à jouer sur le nombre de composés impliqués dans le criblage. En augmentant le nombre de candidats étudiés, la probabilité de trouver un partenaire affin augmente en conséquence. Ces méthodes sont en général effectuées à haut débit, ce qui implique donc l'accès à des chimiothèques de larges tailles. Les molécules de sources naturelles représentent souvent de bons candidats puisqu'elles possèdent une grande complexité structurale mais la difficulté d'accès à des analogues pour l'optimisation des structures limite leur utilisation et rend le développement de molécules synthétiques nécessaire. Des collections historiques, synthétisées au fil des années par les laboratoires et stockées au sein de chimiothèques sont criblées mais elles restent cependant limitées et leurs diversités structurales faibles. C'est pourquoi de nouveaux composés sont constamment synthétisés pour la découverte de tête de séries. Celles-ci peuvent être obtenues par des méthodes de synthèses chimiques, combinatoires ou parallèles, ou encore par voie biologique comme par exemple dans les méthodes « display » ou des molécules de séquences diverses sont synthétisés par les machineries cellulaires et présentées à la surface de biomolécules (bactériophage, ARNm).

En ce qui concerne la méthode de criblage, plusieurs approches peuvent être envisagées. Le criblage phénotypique consiste à identifier des composés conduisant à une réponse biologique, un phénotype souhaité : l'inhibition de la réplication cellulaire, la modulation de l'expression d'un gène ou encore l'inhibition d'une voie de signalisation par exemple²⁸. Elle consiste à cribler une bibliothèque de composés dans un système biologique modèle animal ou cellulaire et de sélectionner le ou les composés conduisant au phénotype souhaité. C'est ici

l'activité biologique des molécules qui est directement évaluée. La connaissance de la cible et le mécanisme d'action ne sont pas nécessairement connus au préalable lors de l'utilisation de cette approche.

A titre d'exemples de molécules identifiées par cette stratégie, l'on peut citer le monastrol, inhibiteur de la formation du fuseau mitotique²⁹ et dont la cible n'a été identifiée *qu'a posteriori* comme étant la kinésine Eg5³⁰.

Le lenalidomide dont l'activité inhibitrice de la production de TNF (*tumor necrosis factor*) et les faibles impacts négatifs sur le plan neuropathique ont été découverts sans aucune connaissance de son mécanisme d'action. Ce n'est qu'ensuite que la protéine Cereblon a été identifiée comme sa cible^{31,32}. Il est actuellement utilisé comme immuno-modulateur et dans le traitement des myélomes multiples notamment.

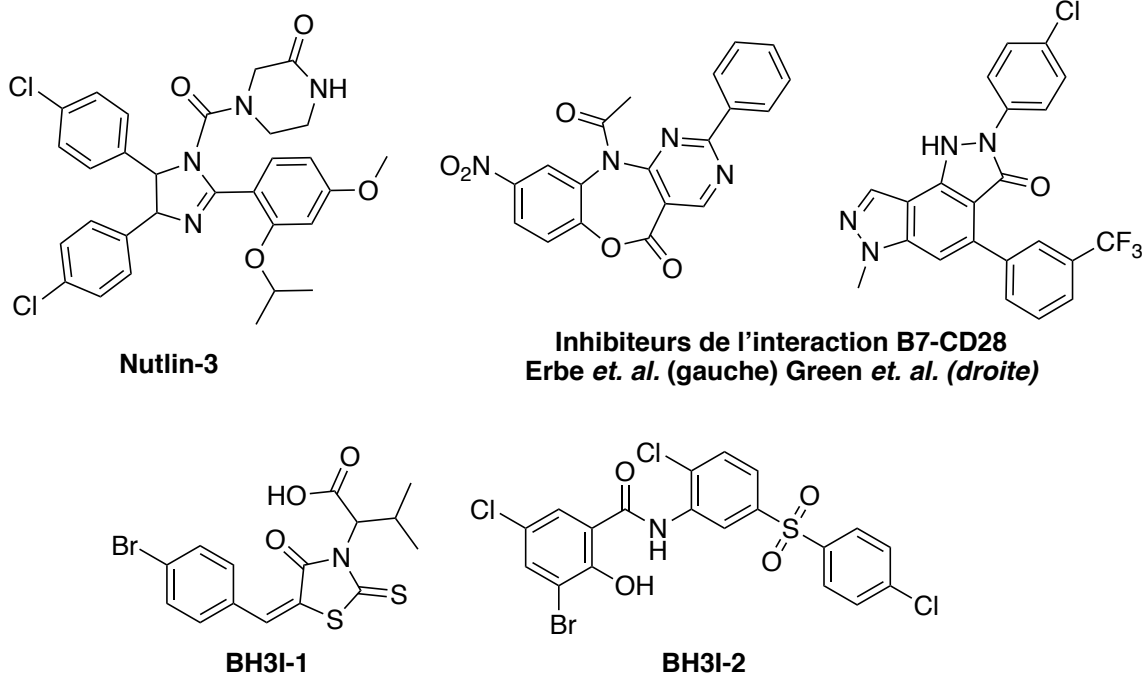


La seconde approche est quant à elle basée sur la cible. Pour cela, la cible doit être clairement identifiée et validée. Les aspects structuraux de la cible ne sont par contre pas nécessairement requis à l'emploi de cette méthodologie. Ici des tests d'affinités et d'activités peuvent à la fois être envisagés. En 2004, Vassilev *et. al.* ont identifié notamment le Nutlin-3 grâce à cette approche en ciblant l'interaction p53-MDM2 impliquée dans de nombreux cancers³³.

L'interaction entre les protéines de surface B7 des cellules présentatrices d'antigènes et les récepteurs CD28 des lymphocytes T entraîne une activation de ces cellules immunitaires alors que l'interaction avec la protéine CTLA4 portée elle aussi par des lymphocytes conduit à leur désactivation³⁴. Parvenir à cibler ce système d'interactions permettrait de moduler l'activité du système immunitaire, option intéressante dans le contexte des rejets de greffe ou des maladies auto-immunes. C'est le cas des molécules découvertes par Erbe *et. al.*³⁵. et par Green *et. al.*³⁶.

Enfin, la protéine Bcl-X_L (et la famille des protéines Bcl-2 plus généralement) possède une activité anti-apoptotique. Elle est régulée par sa liaison à des protéines anti- ou pro-

apoptotiques comme la protéine pro-apoptotique BAK³⁷. Dans un contexte tumoral, inhiber la liaison de Bcl-X_L à BAK en ciblant Bcl-X_L conduit à favoriser l'apoptose induite par BAK et donc limiter voire contrer la croissance tumorale comme l'ont montré Degterev *et. al.* avec les molécules BH3I-1 et BH3I-2³⁸.



ii) Conception rationnelle

La seconde approche pour identifier des molécules susceptibles de cibler les interactions protéine-protéine est d'employer une approche de conception rationnelle en se basant sur des modèles structuraux obtenus par RMN ou diffraction des rayons X sur cristal ou encore des approches biophysiques ou computationnelles quant à l'implication de résidus dans l'interaction. Cette conception rationnelle repose sur l'hypothèse dite des *hot spots* ou points chauds formulée par Clackson et Wells à la fin des années 1990³⁹. Ils mettent en avant le fait que bien que les surfaces d'interactions entre les protéines soient vastes, un nombre plus ou moins restreint de résidus sont véritablement prépondérants dans l'interaction entre les partenaires macromoléculaires sur le plan énergétique, le repliement tridimensionnel de la protéine entraînant leur rapprochement dans l'espace de manière à former des *clusters* ou amas, surfaces énergétiquement indispensables à l'interaction au global d'où l'appellation « points chauds » (Fig. 5).

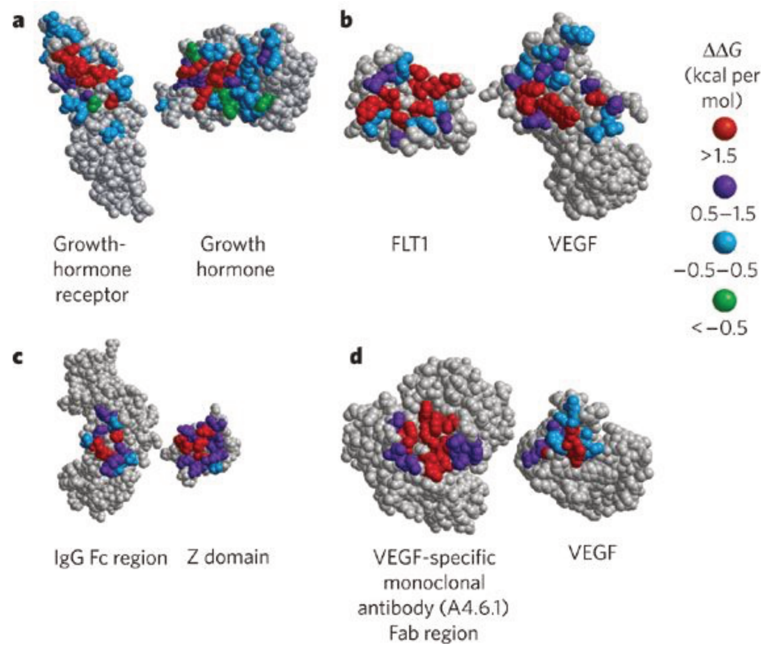


Fig. 5 : Illustrations de couples de protéines formant des complexes *in vivo*. Les hot spots/épitoes identifiés comme support de ces interactions sont symbolisés par les résidus colorés. La couleur dépend de l'implication énergétique de chaque résidu dans l'énergie d'interaction.

Deux grandes approches de conceptions rationnelles basées sur la structure de la cible sont décrites : l'approche par assemblage de fragments (*fragment-based design*) et l'imitation de domaine protéique ou d'épitope protéique (*protein domain mimics* ou *protein epitope mimics*). Dans une conception basée sur des fragments, l'idée centrale est de considérer une molécule modulatrice d'interaction protéine-protéine comme un assemblage de plusieurs briques ayant des propriétés de médicaments. Une fois plusieurs de ces briques identifiées, en général par criblage aléatoire, la molécule modulatrice de l'interaction peut être obtenue soit en reliant deux fragments ayant des sites d'interactions différents mais à proximité. Une autre façon de faire est de faire croître la structure du modulateur d'interaction à partir d'un fragment initial en cherchant à lui joindre de nouvelles parties de façon à augmenter son affinité. Là, il est nécessaire de procéder à l'optimisation à chaque étape de croissance de la molécule^{40, 41}. La figure 6 schématise ces deux processus.

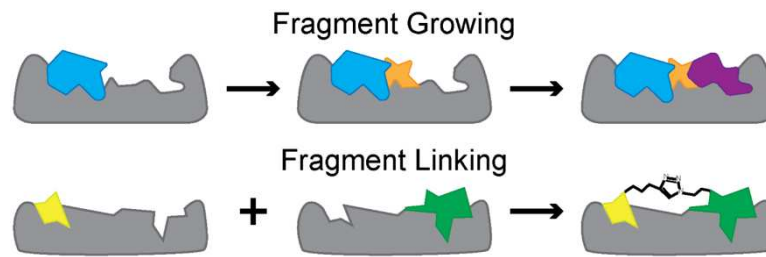


Fig. 6 : Conception basée sur des fragments. En haut, l’approche par croissance de fragment : après avoir optimisé le fragment bleu pour son affinité pour la cible, on lui adjoint un nouveau fragment pour étendre sa surface d’interaction avec la protéine et à chaque étape de croissance, il est nécessaire d’optimiser la nouvelle structure. En bas l’approche par connexion de fragments : après avoir optimisé chacun des fragments ici jaune et vert, on les connecte par un lien dont la nature est elle aussi sujette à optimisation¹⁹.

Dans une approche d’imitation de domaine protéique, l’idée est d’utiliser les structures secondaires des protéines comme base pour concevoir un modulateur de PPI. En effet, l’interaction d’une protéine avec son partenaire implique la plupart du temps des chaînes latérales d’acides aminés inclus dans des segments de structures secondaires bien définies, telles que des hélices ou encore des feuillettes, et qui de ce fait sont présentées dans les orientations particulières nécessaires à la reconnaissance (Fig. 7).

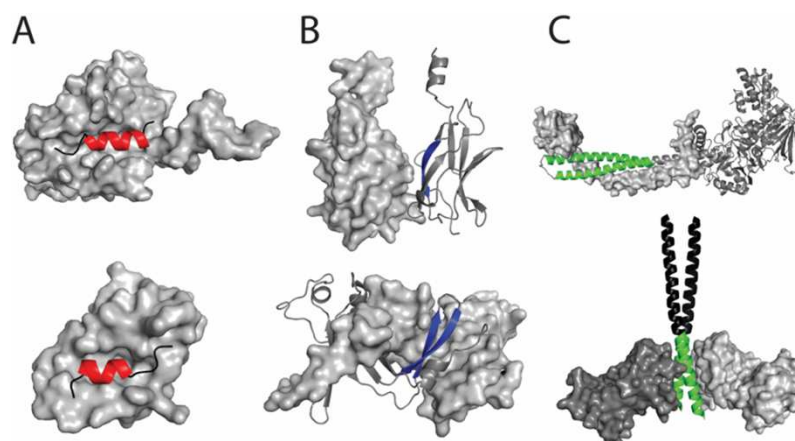


Fig. 7 : Exemples de protéines interagissant par le biais de structures secondaires A) Hélices α B) Feuillet β C) Dimères d’hélices⁴².

Le rôle crucial de ces repliements a conduit à l’hypothèse que des segments isolés reproduisant les mêmes structures secondaires ou tertiaires pouvaient représenter de bons mimés de la protéine entière, les PDM pour Protein Domain Mimics⁴². Différents exemples de ces molécules imitant des structures secondaires ont pu être utilisées et il s’agit souvent de

peptides^{43,44,45}. En effet les peptides représentent une classe intéressante de composés qui occupent un espace chimique intermédiaire entre celui des molécules biologiques, telles que les anticorps monoclonaux (>5kDa) et des petites molécules types médicaments (<500 Da) (Fig.8).

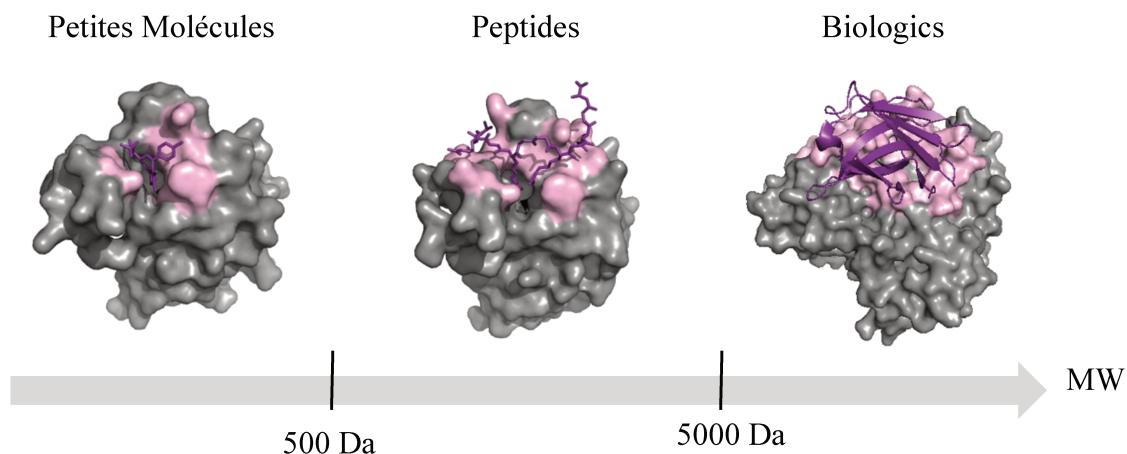


Fig. 8 : Évolution du poids moléculaire des petites molécules médicaments, typiquement moins de 500 Da, aux « biologics », typiquement plus de 5000 Da. Les peptides présente un poids moléculaire intermédiaire.

Ils présentent l'avantage d'être facilement synthétisables grâce aux techniques de synthèse sur support solide modernes et leurs séquences peuvent être directement dérivées de celle de la protéine à imiter. Mais dans la majorité des cas, une fois isolés de leur contexte protéique, les segments peptidiques perdent leur structuration ce qui conduit le plus souvent à une perte d'affinité pour la cible et plus généralement d'activité biologique. Pour accéder à des petites séquences peptidiques de conformation définie, deux stratégies peuvent être envisagées. La première consiste en l'introduction de contraintes conformationnelles⁴⁶ par macrocyclisation du squelette ou des chaînes latérales (stapled peptides), incorporation d'acides aminés contraints ou bien encore de motifs permettant la nucléation du repliement (Fig. 9).

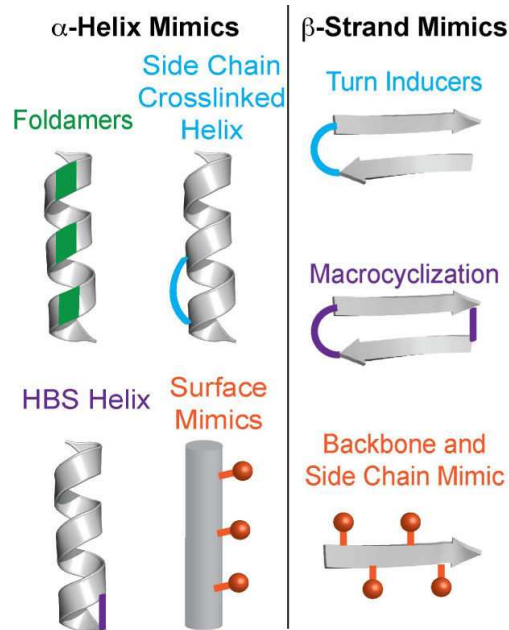


Fig. 9 : Approches pour obtenir des mimes de structures secondaires (hélice α et feuillet β)¹⁹.

La seconde stratégie qui est conceptuellement différente consiste à utiliser un châssis moléculaire ou *scaffold* possédant une conformation stable sur lequel des résidus impliqués dans la reconnaissance moléculaire sont greffés dans une orientation bien définie, stratégie aussi désignée sous le nom de *protein grafting*⁴⁷. De nombreuses protéines de petites tailles ont été utilisées à cette fin comme par exemple le polypeptide pancréatique aviaire (APP)⁴⁸, les toxines de scorpion⁴⁹, les domaines à glissière à leucines⁵⁰. Dans le but d'accéder à des structures plus petites et plus stables des *scaffolds* artificiels ont également été développés constitués de peptides conçu pour adopter une structure particulière, ou encore des polymères non peptidiques qui ont la capacité à adopter des structures régulières, communément appelés foldamères (Fig. 10).

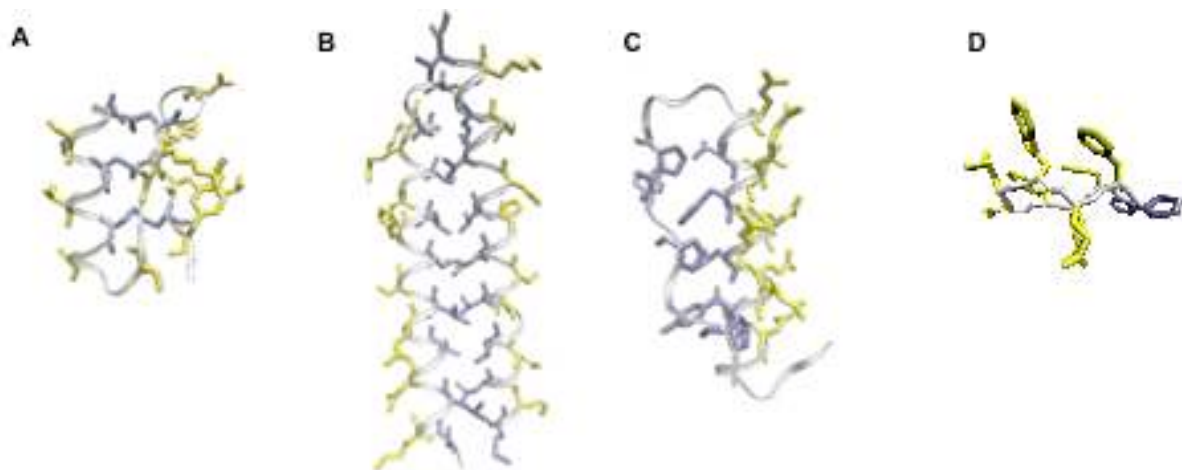


Fig. 10 : Exemples de *scaffolds* peptidiques : A) toxine de scorpion BmBKTx1 (PDB 1R1G)⁵¹, B) domaine en glissière à leucine (PDB 2ZTA)⁵², C) polypeptide pancréatique aviaire (PDB 1PPT)⁵³, D) peptide 10-mer cyclique avec inducteur de coude β D-Pro-L-Pro (PDB 2AXI)⁵⁴. Les résidus cruciaux pour la structures sont colorés en gris. Les résidus colorés en jaune peuvent être modifiés sans que la structure ne soit significativement affectée.

Qu'il s'agisse d'une stratégie reposant sur des fragments ou sur des mimes de domaines protéiques, plusieurs obstacles peuvent freiner la conception d'inhibiteurs d'interaction protéine-protéine. Tout d'abord, ces approches basées sur la cible impliquent d'avoir accès à des détails sur la structure de la cible et ses modalités d'interaction avec son partenaire. Ces connaissances peuvent être obtenues par différentes approches d'études structurales et biophysiques, mais à l'heure actuelle la majorité des interactions protéine-protéine manquent encore de caractérisations aussi poussées. Par ailleurs, ces approches impliquent des étapes de synthèse et évaluation en tant qu'inhibiteur qui seront répétées au travers de cycles d'optimisation. Ce travail constitue également un frein dans la conception d'inhibiteurs de PPI.

II. Développement de Mimes Protéiques par Chimie Combinatoire Dynamique

Parvenir à contourner ou outrepasser ces obstacles implique le développement de stratégies nouvelles. Une approche qui rassemblerait les forces de différentes stratégies mentionnées précédemment s'avèrerait tout à fait intéressante. En particulier, utiliser la puissance des techniques de criblage dites à haut débit et ne nécessitant pas de prérequis structuraux, à partir d'un nombre restreint de molécules de structures contrôlées qui auraient leur place dans des approches de conception basée sur la cible serait un choix judicieux. Il

s'agirait de trouver un juste milieu entre l'aléatoire et les grands nombres offerts par le haut débit et la conception rationnelle basée sur la cible. L'objectif est d'exploiter la chimie combinatoire dynamique (DCC) pour le développement d'une telle stratégie.

1) La Chimie Combinatoire Dynamique

La chimie combinatoire dynamique a vu le jour avec l'essor de la chimie supramoléculaire dans les années 1990 et les travaux de Jean-Marie Lehn et Jeremy Keith Morris Sanders notamment^{55,56,57}. Elle repose sur la génération de bibliothèques combinatoires dites dynamiques ou DCL grâce à des transformations réversibles entre des blocs initiaux. Ces transformations peuvent être covalentes ou non (Fig. 11).

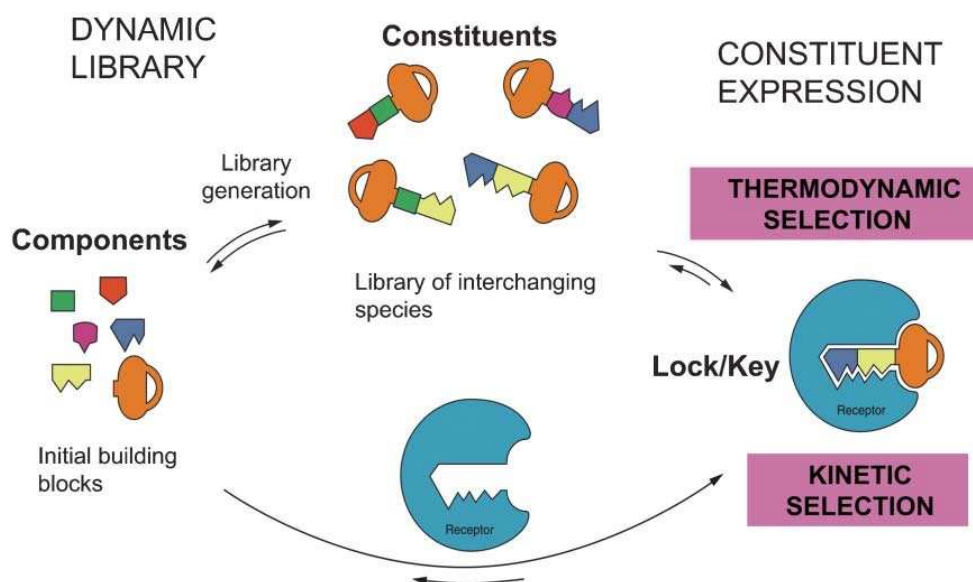


Fig. 11 : Principe de la chimie combinatoire dynamique⁵⁵

Une fois les blocs de départ choisis, le système sera autonome dans la manière de les associer afin de générer potentiellement toutes les combinaisons attendues de blocs. Toutes ces combinaisons se retrouvent en équilibre les unes avec les autres au sein du mélange complexe que constitue la librairie dynamique. A l'équilibre, la proportion des différentes combinaisons de blocs sera le reflet du paysage énergétique global du système en terme d'énergie libre.

Un tel système restera donc en équilibre dynamique tant que les conditions physico-chimiques imposées le lui permettent et l'introduction de stimuli ou de perturbations entraînera la réorganisation de la DCL afin de retrouver un équilibre répondant à ces nouvelles conditions. L'introduction d'une molécule cible par exemple, si elle interagit avec un composé généré par la librairie, entraînera son propre réarrangement suivant le principe de Le Chatelier et exprimera

le partenaire de la cible de manière prépondérante au dépend des autres membres de la DCL. Ce phénomène est parfois appelé amplification.

Ainsi, l'emploi de la DCC permet de combiner certaines forces des méthodologies plus traditionnelles et également de contourner certains obstacles qui leur sont associés. La taille et la complexité des DCL sont contrôlées par la nature et le nombre de blocs initiaux utilisés et en jouant sur ces paramètres, il est possible de se rapprocher de l'efficacité des techniques dites à haut débit. Par ailleurs, en choisissant des blocs qui possèdent une structure tridimensionnelle contrôlée, il est possible d'apporter un minimum de conception rationnelle au système. Enfin, la possibilité de réaliser les deux étapes de chimie combinatoire et de criblage en une permet de s'affranchir des étapes des cycles de synthèse, évaluation et optimisation nécessaires dans d'autres stratégies.

2) Conception de mimes protéiques par DCC : État de l'art

L'utilisation de la combinatoire dynamique pour la formation de mimes structuraux et fonctionnels de protéines peut être envisagée sous différents angles⁵⁸. La première approche qui pourrait s'apparenter à une approche par fragments consiste en l'assemblage de plusieurs fragments peptidiques de manière combinatoire dynamique de manière à permettre la formation d'une structure plus grande. De tels systèmes conduisent à des structures auto-assemblées proches de protéines naturelles et l'assemblage le plus stable sera celui majoritairement représenté (Fig. 12).

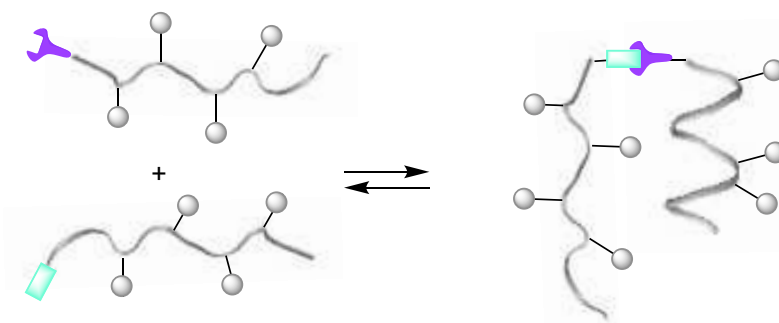


Fig. 12 : Assemblage de séquences peptidiques de manière combinatoire dynamique pour obtenir des structures de type protéines. Les fragments peptidiques réagissent via les groupements bleu et violet pour former une nouvelle fonction reliant les fragments précédents de manière réversible.

Roy *et. al.* ont ainsi identifié une structure supramoléculaire de type faisceau d'hélices α présentant la meilleure compaction grâce à un assemblage médié par la présence d'un ion métallique⁵⁹. Un travail similaire a été réalisé par Bilgiçer *et al.* mais dans ce cas l'assemblage de séquences hélicoïdales s'est fait de manière covalente grâce à l'échange réversible entre thiols et disulfures⁶⁰. Le groupe de Gellman a développé une méthode permettant d'évaluer la stabilité d'un assemblage protéique supramoléculaire par chimie combinatoire dynamique⁶¹. Cette méthode est basée sur l'étude d'échanges thiol/thioester entre deux fragments peptidiques, la constante d'équilibre de cet échange pouvant être directement reliée à la stabilité thermodynamique de la structure tridimensionnelle. Cette méthode a été appliquée à l'étude d'hélices⁶², de structures coiled-coil parallèle^{63,64} et antiparallèle^{65,66} et feuillettes β ⁶⁷ (Fig. 13).

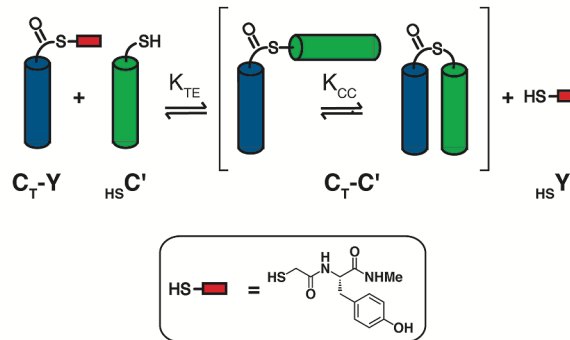


Fig. 13 : Exemple de système développé par le groupe de Gellman pour la formation de structures coiled-coil via une association covalente par échange thiol/thioester entre deux fragments peptidiques⁶².

Le groupe de Albericio a montré que le groupement thioester était iso-structural d'une fonction amide dans la structure coiled-coil par des études comparatives de RX et que cette stratégie pourrait donc être exploitée pour cribler différents segments protéiques susceptibles de former le coiled-coil et déterminer la séquence optimale⁶⁸. Une stratégie similaire a été récemment décrite par Bartus *et. al.*, utilisant cette fois des échanges thiol/disulfure, pour l'assemblage de deux segments β -peptidiques hélicoïdaux courts, mimant deux « hot spots » distincts, en une structure supramoléculaire pour cibler la calmoduline (Fig. 14)⁶⁹.

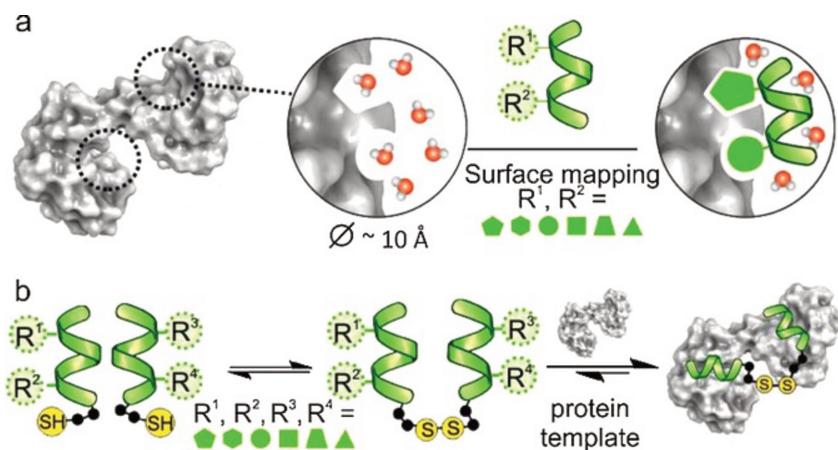


Fig. 14 : Illustration du système combinatoire dynamique développé par Bartus *et. al.* dans lequel deux fragments β -peptidiques hélicoïdaux ont été optimisés pour cibler deux épitopes de la calmoduline. Par la suite, ces fragments optimisés sont associés par chimie combinatoire dynamique. Cette approche ayant permis d'identifier un dimère afin ciblant les deux épitopes simultanément⁶⁹.

Le groupe de Balasubramanian s'est également appuyé sur cet échange thiols disulfures pour l'auto-assemblage de séquences peptidiques en structure mimant des épingle à cheveux β (Fig. 15)⁷⁰.

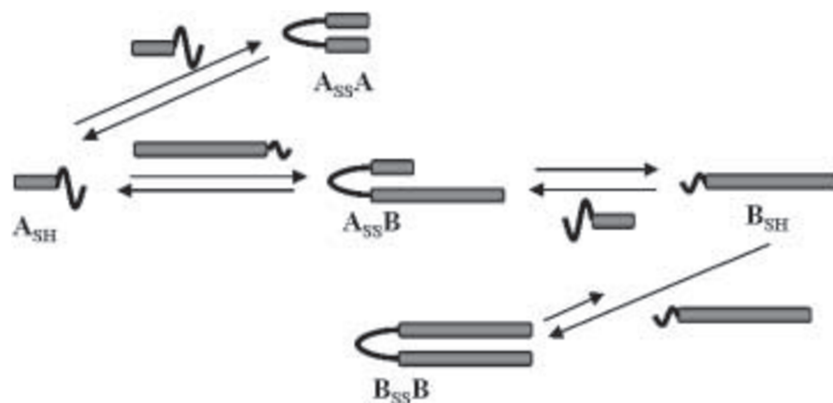


Fig. 15 : Illustration du système combinatoire dynamique développé par Balasubramanian reposant sur l'échange thiol/disulfure et permettant l'auto-assemblage de séquences peptidiques en mimes d'épingle à cheveux β ⁷⁰.

Enfin, le groupe de Otto quant à lui a pu générer des bibliothèques dynamiques de peptides portant des thiols et formant des cycles grâce à l'échange thiols disulfure. Ils ont mis en évidence une auto-amplification d'un des cycles formés par un auto-assemblage de ce dernier par empilement de feuilletts β ⁷¹.

Une seconde approche consiste à assembler de manière combinatoire dynamique des aminoacides ou leurs analogues pour former le squelette peptidique lui-même de manière dynamique (Fig. 16).

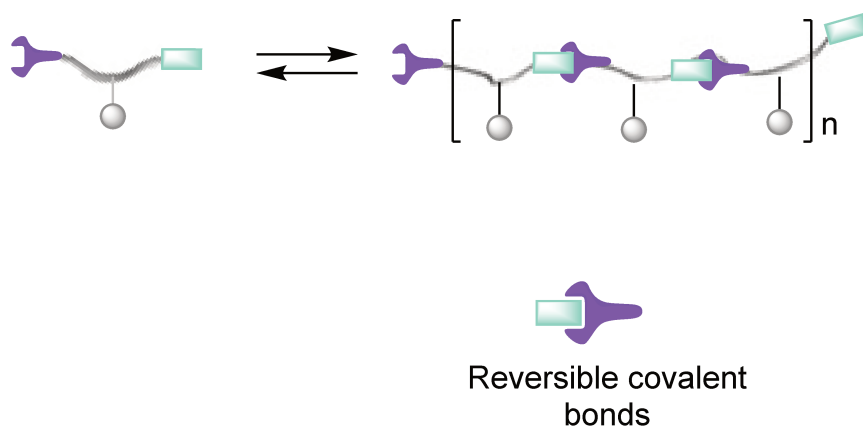


Fig. 16 : Assemblage du squelette peptidique de manière combinatoire dynamique. Les monomères imitant les acides aminés peuvent s'assembler de manière réversible par l'association des fonctions chimiques violette et bleue afin de former des liens covalents réversibles formant ainsi un squelette peptidique ou imitant cette structure.

La formation de liens amides étant très défavorisée en milieu aqueux des substitués d'amides sont le plus souvent utilisés à cette fin. Sanders *et. al.* ont ainsi généré des bibliothèques dynamiques reposant sur l'échange hydrazide acylhydrazone à partir de blocs hydrazides peptidiques (Fig. 17)⁷². Ainsi, ils ont pu identifier un composé de type [2]-catenane comme ligand de l'enzyme acétylcholinestérase.

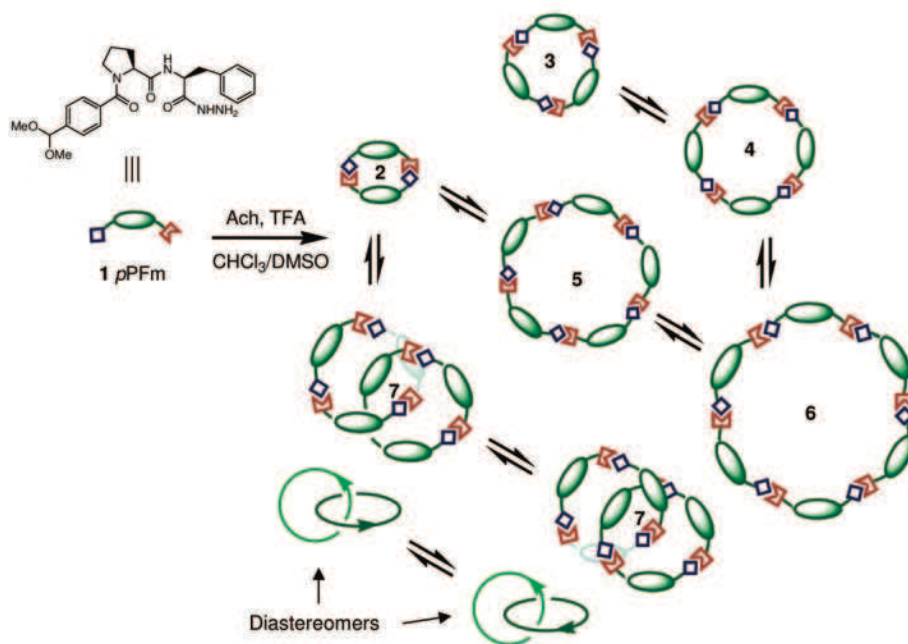


Fig. 17 : Librairie combinatoire dynamique générée par échange hydrazide acylhydrazone⁷².

L'échange thiols disulfure a été utilisé par Rauschenberg *et. al.* pour générer des bibliothèques de dimères cycliques à partir de tripeptides et ainsi identifier des structures minimales imitant des récepteurs d'oses de type lectine montrant une sélectivité importante (Fig. 18)⁷³.

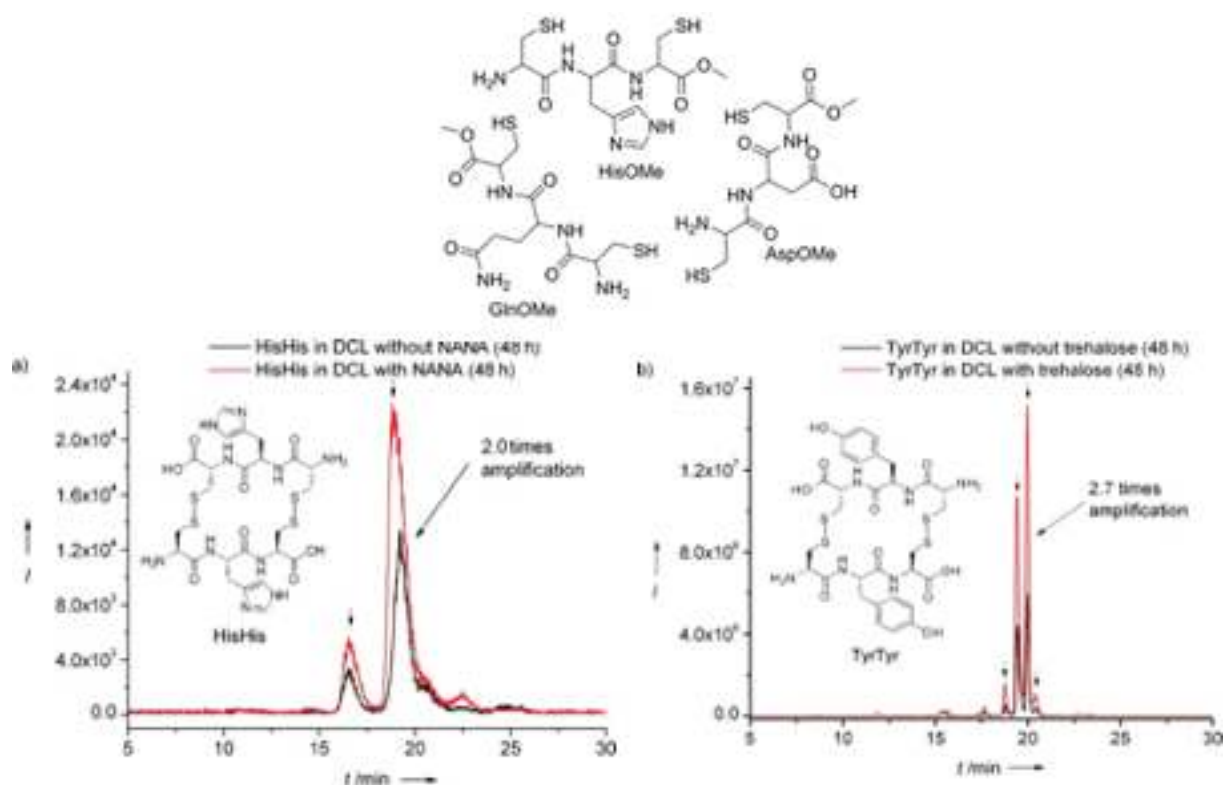


Fig. 18 : Structures de tripeptides utilisés par Rauschenberg. Ces séquences peuvent s'assembler par formation de ponts disulfures. Certains homodimères se sont révélés amplifiés lors de l'introduction de certains oses comme l'acide N-Acétyl NeurAminique ou le trehalose comme le montre les chromatogrammes représentés ci-dessus⁷³.

Le groupe de Miller a développé une méthodologie de chimie combinatoire dynamique sur support solide dans laquelle l'échange thiols disulfure permet la formation réversible d'un lien covalent entre des séquences supportées sur résine et des séquences en solution⁷⁴. Non seulement cette approche permet d'atteindre des tailles de bibliothèques considérables en s'affranchissant des problèmes liés à la solubilité mais cela leur a permis d'identifier une structure capable de lier une séquence d'ARN responsable de la dystrophie myotonique de type 1 (Fig. 19).

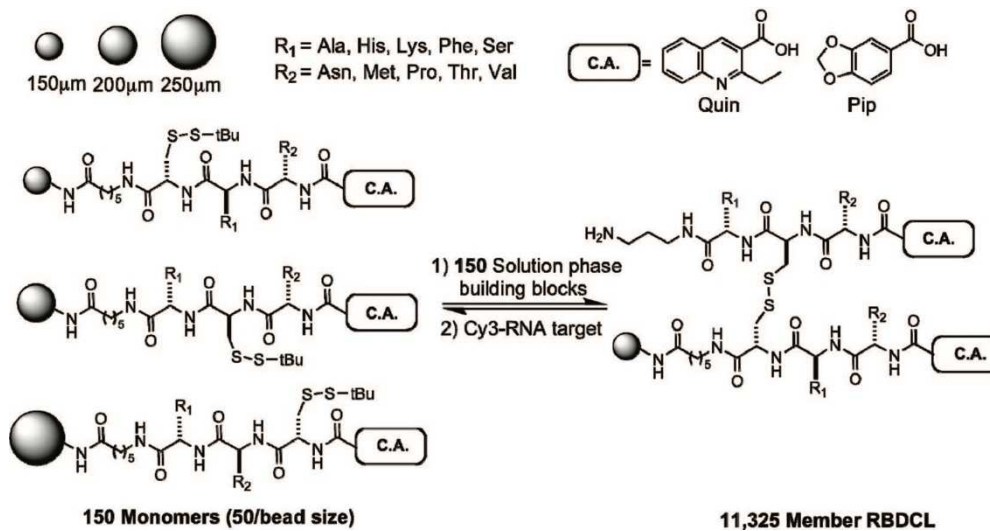


Fig. 19 : Librairie combinatoire dynamique supportée générée par échange thiols disulfure⁷⁴.

Un défi important ici serait de réussir à mettre au point des conditions dans lesquelles la formation des liaisons amides serait réversible dans l'eau. Une telle stratégie est décrite par Ulijn *et al.* et repose sur l'utilisation d'une enzyme protéolytique, la thermolysine, capable de catalyser à la fois l'hydrolyse et la formation des liaisons amides⁷⁵. Une librairie dynamique de dipeptides a ainsi été formée, capable de répondre à l'ajout de glycosaminoglycanes (GAG) permettant d'identifier pour différents GAG la séquence dipeptidique la plus apte à lier le polysaccharide (Fig. 20).

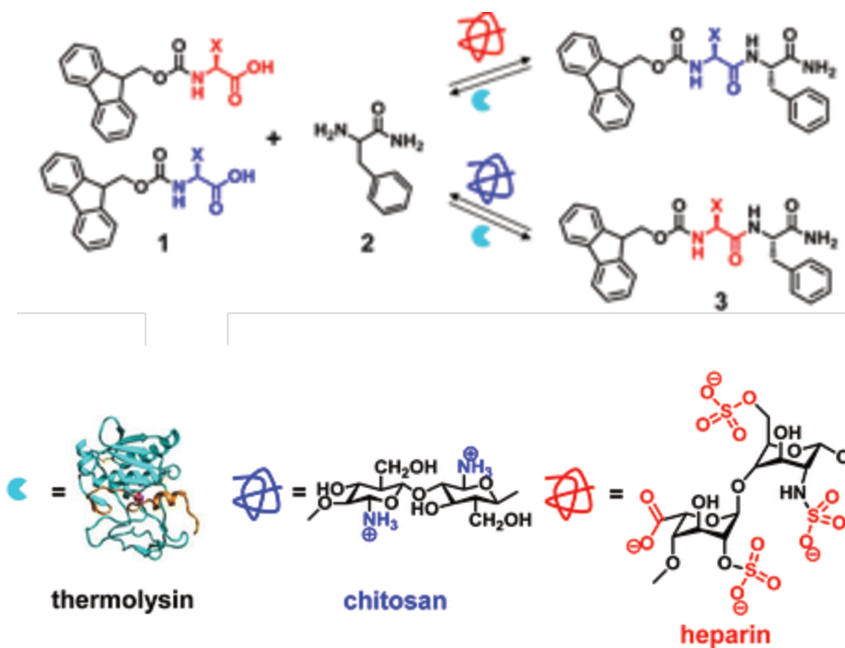


Fig. 20 : Système combinatoire dynamique développé par Ulijn dans lequel un acide aminé Fmoc protégé sur la fonction amine alpha peut former un dipeptide avec une phénylalanine. L'utilisation de l'enzyme thermolysine rend cette réaction réversible. Le criblage par des polysaccharides comme le chitosan et l'héparine conduit à la formation de la séquence la plus affine pour le polymère en question⁷⁵.

Giussepone *et. al.* ont mis au point une approche différente pour générer des séquences présentant des liaisons peptidiques natives en s'appuyant sur la réaction de ligation chimique native et en l'incorporant dans un système réversible par des réactions de transthioesterification (Fig. 21). Ainsi, les séquences encadrant ces sites réactionnels d'échange peuvent être interverties entre différents peptides de départ et permettre l'accès à une diversité de structures importante de manière très simple de mise en place⁷⁶. Aucun phénomène d'amplification n'a cependant été démontré dans ce cas à ce jour.

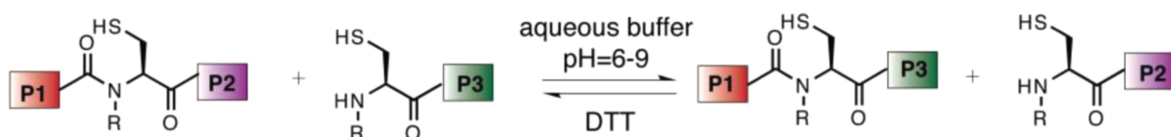


Fig. 21 : Librairie combinatoire dynamique générée par ligation chimique native réversible⁶⁹.

La dernière approche pour générer des mimes de protéines par chimie combinatoire dynamique repose sur le greffage de groupements prenant la place des chaînes latérales des acides aminés sur un squelette peptidique préexistant (Fig. 22).

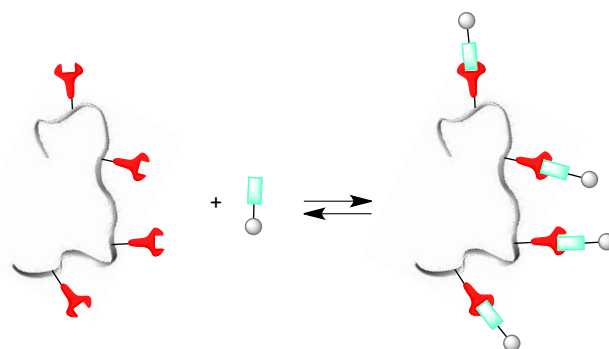


Fig. 22 : Greffage de groupements au niveau des chaînes latérales sur un squelette préexistant. Les fonctions rouges portées par le squelette et les fonction bleues peuvent s'associer de manière réversible en formant une nouvelle fonction et ainsi permettre le greffage dynamique de fonction imitant les chaîne latérales des acides aminés sur une structure imitant quant à elle le squelette.

Le groupe de Fulton a par exemple fonctionnalisé un polymère grâce à la formation réversible d'acylhydrazones et a également montré le déplacement de l'équilibre de fonctionnalisation induit par la présence de différentes cibles greffées sur un support. Le déplacement de l'équilibre se faisant en faveur des fonctionnalités permettant des interactions favorables avec la cible introduite (Fig. 23)^{77,78}.

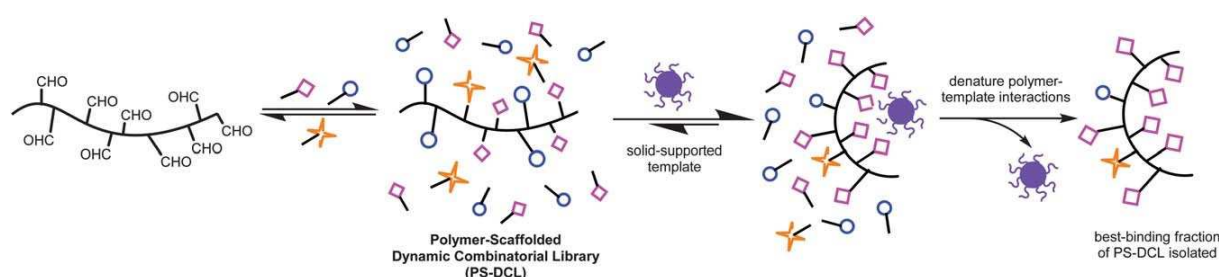


Fig. 23 : Fonctionnalisation d'un polymère par combinatoire dynamique reposant sur la formation d'acylhydrazones entre un polyaldéhyde et des blocs hydrazides. L'équilibre est déplacée par la présence d'un template sur support solide⁷⁷.

Le groupe de Ghadiri s'est appuyé sur la transthioesterification pour fonctionnaliser une séquence polycystéine grâce à des thioesters des différentes bases azotées rentrant dans la composition des acides nucléiques et ainsi générer des PNA (*peptide nucleic acid*) de manière combinatoire dynamique⁷⁹. L'introduction d'un oligonucléotide leur a permis de montrer l'effet induit par la reconnaissance moléculaire puisque la séquence complémentaire de l'oligonucléotide a été amplifiée sur la contrepartie PNA par déplacement de l'équilibre de fonctionnalisation (Fig. 24). Une stratégie très similaire a été rapportée simultanément par le groupe de Liu, employant quant à elle la formation réversible d'imines⁸⁰.

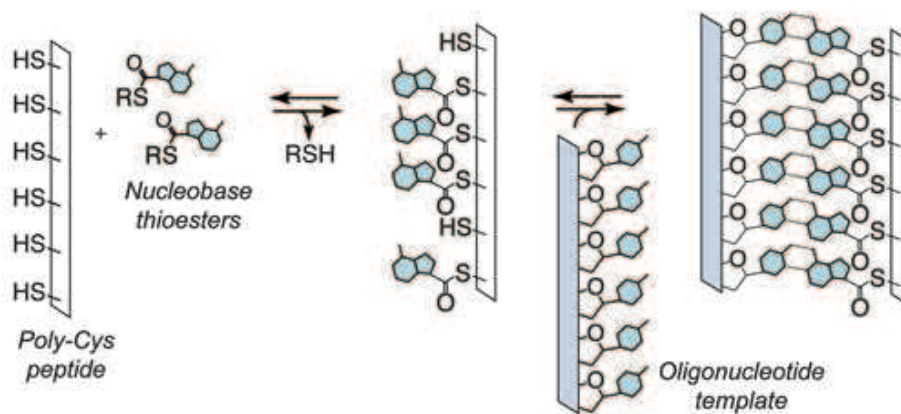


Fig. 24 : Fonctionnalisation combinatoire dynamique d'un polycystéine par transthioesterification et criblage par un oligonucléotide permettant l'accès à des PNA dynamiques⁷⁹.

Des travaux récents d'Ulrich *et al.* s'appuient sur la fonctionnalisation d'un squelette peptidique structuré par formation d'acylhydrazones afin d'y greffer des groupements guanidines permettant des interactions multivalentes avec l'ADN⁸¹. Les structures ainsi générées constituent des systèmes capables de réagir à leur environnement tout en étant capables de se lier et de libérer l'ADN selon les conditions appliquées. Ces travaux restent à ce jour les seuls s'appuyant sur un squelette peptidique possédant une conformation stable. Cependant aucun aspect combinatoire n'est traité dans ce travail.

Cet état de fait pourrait expliquer pourquoi à l'heure actuelle aucune méthodologie générale n'a vu le jour pour l'imitation de protéine par DCC en tant que cibles thérapeutiques de choix. En effet, la question de la structure tridimensionnelle des structures générées par ces approches visant des protéines d'intérêt n'est que très rarement abordée alors que ce paramètre est central pour l'affinité et la sélectivité.

III. Objectifs

L'objectif du projet développé au cours de ce travail de thèse est de mettre au point une nouvelle méthode pour la conception de peptides bioactifs, combinant un squelette peptidique de structure définie, conçu de manière rationnelle, avec l'aspect aléatoire de la DCC pour le greffage de chaînes latérales sur la surface de ce *scaffold* par auto-assemblage directement en présence de la cible (Fig. 25). Ainsi un *scaffold* peptidique serait couvert sur une des faces par des fonctions capables de réagir de manière réversible avec une série de blocs complémentaires, porteurs des chaînes latérales des acides aminés naturels dans des conditions physiologiques. La réaction d'échange conduit alors à une librairie dynamique de peptides de structures définies, possédant virtuellement à leur surface toutes les combinaisons possibles de chaînes

latérales et qui sont en équilibre thermodynamique les unes avec les autres. L'ajout d'une cible biologique d'intérêt va entraîner un déplacement d'équilibre en faveur des espèces les plus aptes à interagir avec cette cible au dépend des composés non affins, et permettre ainsi l'identification des meilleurs ligands.

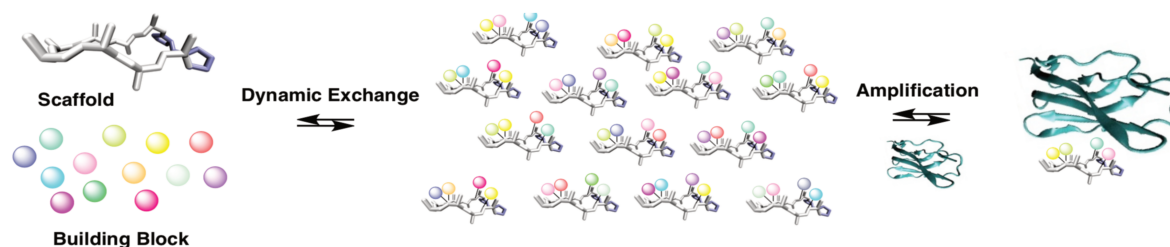


Fig. 25 : Greffage de groupements à la surface d'un peptide structuré de manière combinatoire dynamique. Les différentes combinaisons de groupements greffés sont en équilibre dynamique au sein d'une librairie combinatoire dynamique qui en présence d'une cible peut se réarranger vers la formation de partenaires affins pour cette cible.

Ce manuscrit développera plusieurs aspects liés au développement de cette méthodologie. Un premier chapitre sera consacré à la synthèse de blocs qui pourront être greffés de manière combinatoire dynamique sur un *scaffold* peptidique structuré. La synthèse de différentes structures peptidiques sera par la suite abordée : des peptides à agrafes structurés en hélice α et des peptides cycliques structurés en épingle à cheveux β . Différents aspects de la conception rationnelle de ces *scaffolds* seront abordés. Le second chapitre abordera le développement méthodologique pour la réaction de fonctionnalisation de peptides cycliques structurés en épingle à cheveux β , tout d'abord sur des systèmes modèles simples dans une optique de mise au point. Le chapitre abordera par la suite des systèmes avec une complexité de plus en plus importante. Le chapitre 3 sera consacré à l'application de la méthodologie développée au chapitre 2 pour l'identification de partenaire d'une cible biologique d'intérêt : l'héparine, un glycosaminoglycane modèle. Le chapitre 4 portera sur le développement d'une méthodologie de linéarisation des peptides β -*hairpin* pour la résolution de structures par spectrométrie de masse. Enfin, le dernier chapitre soulèvera des questions relatives à la structure de ces peptides β -*hairpin* grâce à des études préliminaires de modélisation moléculaire et de RMN. Notamment, s'ils adoptent une bonne structuration comme prévue par leur conception rationnelle, mais aussi si elle n'est pas affectée fondamentalement par la fonctionnalisation. A terme, ce projet visant à identifier des structures biologiquement actives, il sera nécessaire de synthétiser des analogues non dynamiques des structures identifiées par l'approche de DCC. L'influence de cette transposition y sera également abordée.

Chapitre 1 : Conception et synthèse de blocs et *scaffolds* peptidiques pour la chimie combinatoire dynamique

I. Choix de la réaction de DCC

Toute stratégie utilisant la chimie combinatoire dynamique repose sur une transformation réversible et équilibrée qui sera le support de la génération des bibliothèques combinatoires dynamiques et des possibles amplifications par des cibles d'intérêt lors du réarrangement de ces bibliothèques. Toutes les réactions réversibles et équilibrées sont potentiellement utilisables dans une telle approche. Cependant, l'ambition d'utiliser des molécules biologiques comme cibles restreint considérablement le nombre de ces réactions. En effet, ces molécules, et les protéines en particulier, sont sensibles et peuvent donc être dégradées rapidement, mais elles peuvent également perdre leur conformation native, dès lors que les conditions physico-chimiques dans lesquelles elles sont placées s'éloignent trop des conditions physiologiques.

En prenant en compte ces considérations, les caractéristiques de la réaction à choisir doivent être les suivantes : conductible en milieux aqueux tamponné proche du pH physiologique, à température raisonnable (d'ambiante à physiologique) et avec une cinétique rapide. Différentes réactions réversibles répondant plus ou moins à ces critères peuvent être utilisées (Fig. 26).

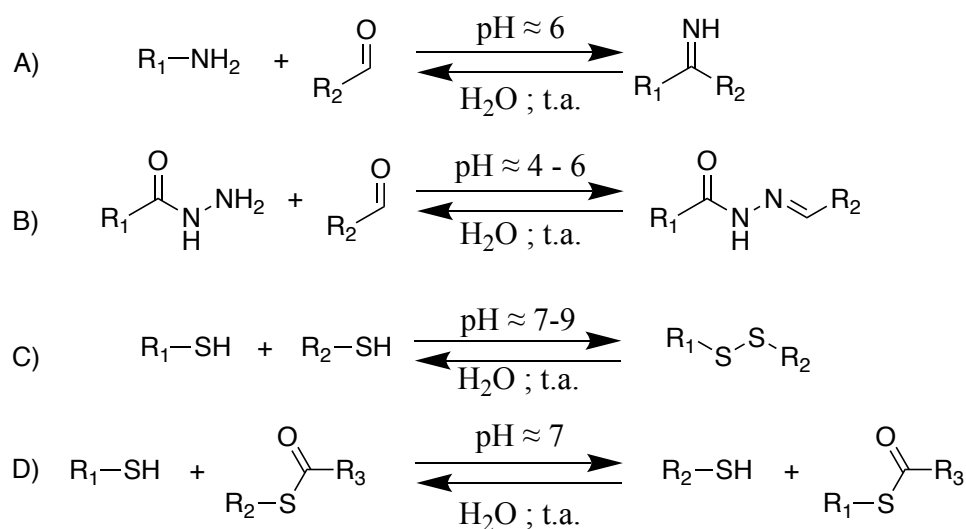


Fig. 26 : Schémas généraux de réactions covalentes réversibles couramment utilisées pour la génération de bibliothèques combinatoires dynamiques appliquées à des cibles biologiques. A) Formation d'imines. B) Formation d'acylhydrazones. C) Formation de disulfures. D) Transthioestérification

En premier lieu, on peut citer la formation réversible d'imines à partir d'amines et d'aldéhydes^{82, 83, 84}. Cette réaction est conductible en milieu aqueux à un pH légèrement acide, compatible avec la plupart des cibles biologiques. Cette dynamique peut être arrêtée par une étape de réduction, typiquement au NaBH₃CN en transformant les imines formées en amines inertes, avant analyse de la composition de la librairie. En revanche, le milieu aqueux favorisant l'équilibre du côté gauche, il est souvent nécessaire d'utiliser un excès d'amine afin de déplacer l'équilibre en faveur de la formation de l'imine, l'utilisation de cette chimie en présence de peptides ou de protéines peut être problématique, du fait de la présence de nombreuses amines sur les chaînes latérales des résidus lysines de ces biomolécules, qui peuvent entrer en compétition avec les amines des building blocks et réagir avec les aldéhydes utilisés pour générer la librairie combinatoire dynamique, conduisant ainsi à des modifications covalentes de la cible.

L'échange hydrazide-acylhydrazone a également été utilisé dans des systèmes de DCC. Pour cette réaction, la cinétique peut être arrêtée par alcalinisation du milieu réactionnel. Cependant cette réaction est réversible à pH sensiblement plus acide (pH 4 à 7)^{85, 86, 87, 88}, ce qui limite l'utilisation de cette chimie aux cibles compatibles avec de tels pH. Contrairement à la formation d'imine, les amines de la cible ne sont pas réactives aux pH davantage acides utilisés dans ce cas.

L'échange thiols-disulfure en milieu aqueux réalisable à pH neutre ou légèrement basique, est très utilisé en DCC^{89,90,91,92,93,94}. Cette réaction peut être arrêtée par acidification du milieu. Cette chimie spécifique du soufre présente l'avantage d'être orthogonale aux autres groupements potentiellement présents sur les cibles mis à part les atomes de soufre portés par les résidus cystéines. Dans ce cas, il est raisonnable de penser que la faible proportion de cystéines dans les protéines ainsi que leur enfouissement au cœur des structures tridimensionnelles peut minimiser l'altération des cibles utilisées par cette chimie. Cependant, un inconvénient de cette approche est qu'elle conduit à des combinaisons de blocs liées par ponts disulfures compliquant l'analyse des mélanges. Enfin, le disulfure n'est pas stable en milieu biologique et particulièrement intracellulaire. Le système de régulation de l'état rédox de la cellule médié par le glutathion notamment vient rapidement réagir sur ces ponts disulfures. Par conséquent, les composés identifiés par cette méthode de DCC doivent donc être modulés au niveau de ce lien s'il doivent être utilisés dans cette situation.

Une autre réaction impliquant l'atome de soufre utilisée en DCC est la transthioestérification, réversible à pH 7-8 et conductible en milieu aqueux. La cinétique de cet échange peut être stoppée là encore par diminution du pH afin de procéder à l'analyse des librairies générées sans risque de changement dans leur composition durant cette analyse^{95,96,97}.

Tout comme pour l'échange thiols-disulfure, elle devrait présenter le même comportement vis à vis des cystéines des cibles ne donnant que peu ou pas de réaction parasite entre les blocs choisis pour générer la librairie et la cible. De plus dans ce cas, une réaction parasite des thioesters avec les amines des chaînes latérales lysine de la cibles conduirait irréversiblement à des amides et à la rupture de l'équilibre réactionnel. Le paramètre du pH est donc primordial pour contrôler l'état de protonation de ces amines et donc leur réactivité. Nous avons choisi d'utiliser cette réaction pour le développement de notre système DCC.

II. Synthèse des blocs thioesters

1) Considérations structurales

Notre approche consiste à générer des librairies combinatoires dynamiques grâce à une réaction de transthioestérification entre des thiols portés par un peptide structuré et des blocs thioesters qu'il faut synthétiser. Ces thiols peuvent être introduits facilement sur le peptide via des résidus naturels, les cystéines. Les blocs thioesters peuvent théoriquement avoir n'importe quelle structure mais afin de garantir l'établissement de l'équilibre, les réactivités des thiols/thioesters portés par le peptide et par les blocs doivent être le plus similaires possibles afin de ne pas trop favoriser un côté de l'équilibre plutôt qu'un autre. Nous avons choisi pour cela d'utiliser la cystéine elle-même comme support des chaînes acyles qui seront échangées lors de la transthioestérification. Un groupement protecteur *N*-Acétyl est utilisé pour masquer la réactivité de l'amine vis à vis des thioesters et se rapprocher au plus près de la réactivité d'un thiol de cystéine impliquée dans une séquence peptidique. La fonction acide est maintenue libre afin de garantir une solubilité satisfaisante dans l'eau non seulement des thioesters de départ mais également du thiol libéré lors de la réaction d'échange. En effet, la présence de composés peu soluble ou volatile, constituerait un biais pour l'équilibre de transthioestérification (Fig. 27).

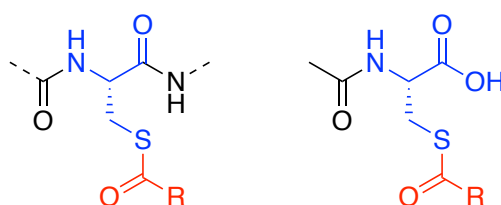


Fig. 27 : Comparaison des structures entre un résidu cystéine portant un thioester au sein d'une séquence peptidique et un bloc thioester. A gauche, structure d'un résidu cystéine impliquée dans une séquence peptidique. A droite, structure générique d'un bloc thioester.

2) Schéma général de synthèse

La synthèse des thioesters de *N*-Acétylcystéine a été envisagée selon le schéma suivant en adaptant une méthodologie développée par Katritzky, basée sur l'activation de la fonction acide sous forme d'acylbenzotriazole^{98,99} (Fig. 28).

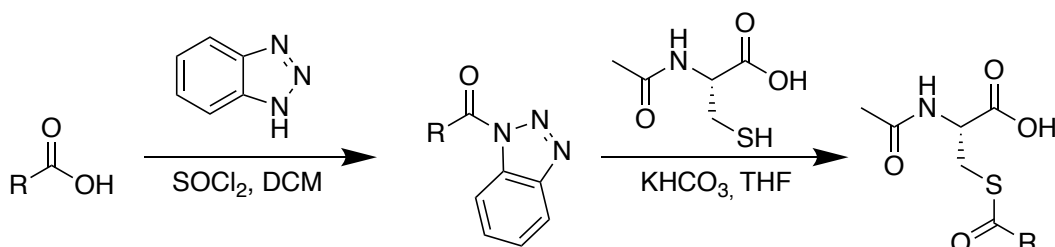


Fig. 28 : Schéma général de synthèse des blocs thioesters de *N*-Acétylcystéine.

Un choix judicieux des acides de départ permet d'accéder à une bibliothèque de building blocks **BB** dont les chaînes R sont représentatives de celles des acides aminés naturels ou d'analogues. D'une manière générale, afin d'obtenir un **BB** qui imitera un acide aminé possédant une chaîne R, l'acide carboxylique correspondant est l'analogue désaminé de l'acide aminé en question. (Fig. 29).

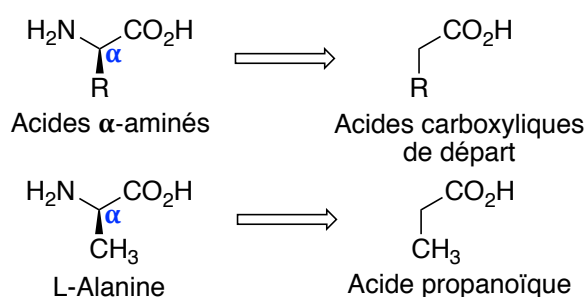


Fig. 29 : Structures illustrant le passage d'un acide aminé donné à l'acide carboxylique qui conduira au **BB** thioester correspondant. En haut, structure générique. En bas, exemple de l'alanine.

3) Blocs à chaînes aliphatiques et aromatiques

A l'exception du **BBGly**, tous les **BB** ont été synthétisés en suivant le schéma général décrit ci-dessus et ont pu être obtenus avec des rendements allant de 24 % jusqu'à 96 %. Malgré des rendements parfois faibles, ces synthèses n'ont pas été optimisées, les quantités de produits finaux obtenus étant suffisantes pour les expérimentations que nous souhaitons mener. Les

blocs correspondants aux acides aminés naturels à chaîne aliphatiques et aromatiques ont été obtenus : **BBAla**, **BBVal**, **BBLeu**, **BBPhe** et **BBTrp** mimant respectivement les chaînes de l’alanine, la valine, la leucine, la phénylalanine et du tryptophane. Des blocs correspondants à des acides aminés non naturels ont aussi été préparés d’une part afin d’exemplifier la faisabilité de cette synthèse sur différents substrats mais aussi d’obtenir des blocs variés pour étudier la réaction réversible de transthioestérification. Les blocs **BBhAla**, **BBAib**, **BBTle** ainsi que **BBPhg** ont été obtenus afin d’imiter respectivement les chaînes de l’homoalanine, l’acide 2-aminobutyrique, la *tert*-leucine et de la phénylglycine (Fig. 30).

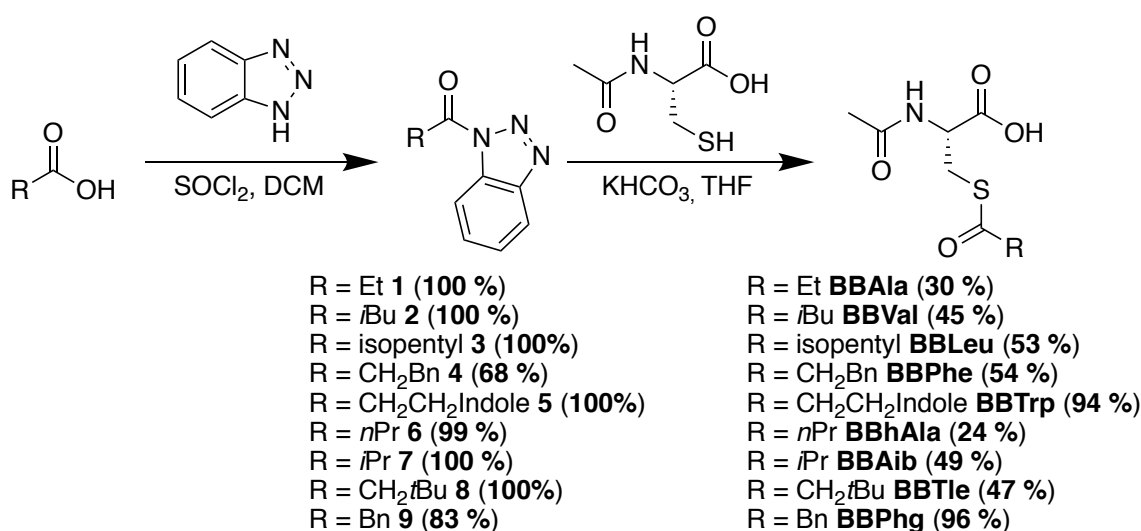


Fig. 30 : Synthèses des **BB** à chaînes aliphatiques et aromatiques

Le **BBGly** imitant la glycine a quant à lui été préparé en une seule étape par réaction de la L-Cystéine avec l’anhydride acétique et a pu être obtenu avec un rendement de 84 %¹⁰⁰ (Fig. 31).

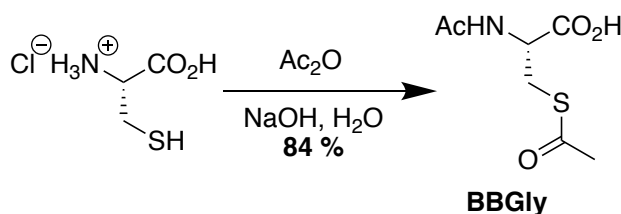


Fig. 31 : Synthèse du **BBGly**

4) Blocs à chaînes fonctionnalisées et potentiellement réactives

La même voie de synthèse a été exploitée afin d’obtenir les blocs correspondants aux chaînes des acides aminés porteurs de fonctions réactives en bout de chaîne. Cependant, des

étapes supplémentaires de synthèse ont dû être mises en place soit pour installer la fonctionnalité désirée, l'acide carboxylique correspondant n'étant pas commerciale, soit pour introduire et enlever des groupements protecteurs de ces fonctionnalités. La fonction thioester visée dans ces différentes synthèses impose des conditions de déprotection finale qui lui sont orthogonales. Ainsi, toutes protections susceptibles d'être supprimées en conditions basiques sont à proscrire afin d'éviter la destruction du thioester. De même, les déprotections par hydrogénolyse médiées par des catalyseurs métalliques ne sont pas envisageables : le soufre étant connu pour être un poison de ce type de catalyseurs. Enfin, la déprotection par un nucléophile est également à proscrire : ce dernier réagirait en premier lieu sur le thioester électrophile.

Ainsi, la stratégie la plus simple envisagée a été d'utiliser des groupements protecteurs sensibles au milieu acide : groupement Boc pour la protection des fonctions azotées (amine, guanidine, imidazole) et ester *tert*-butylique pour les fonctions acides carboxyliques pouvant tous deux être clivés par un traitement final acide au TFA auquel le thioester est résistant. Les fonctions alcools ont été protégées par un groupement éther silylé, déprotégé en fin de synthèse par des ions fluorures, par exemple par traitement au TBAF.

Le **BBLys** imitant la chaîne de la lysine a ainsi été obtenu à partir de l'acide 6-aminohexanoïque commercial avec un rendement global de 58 % sur quatre étapes (Fig. 32).

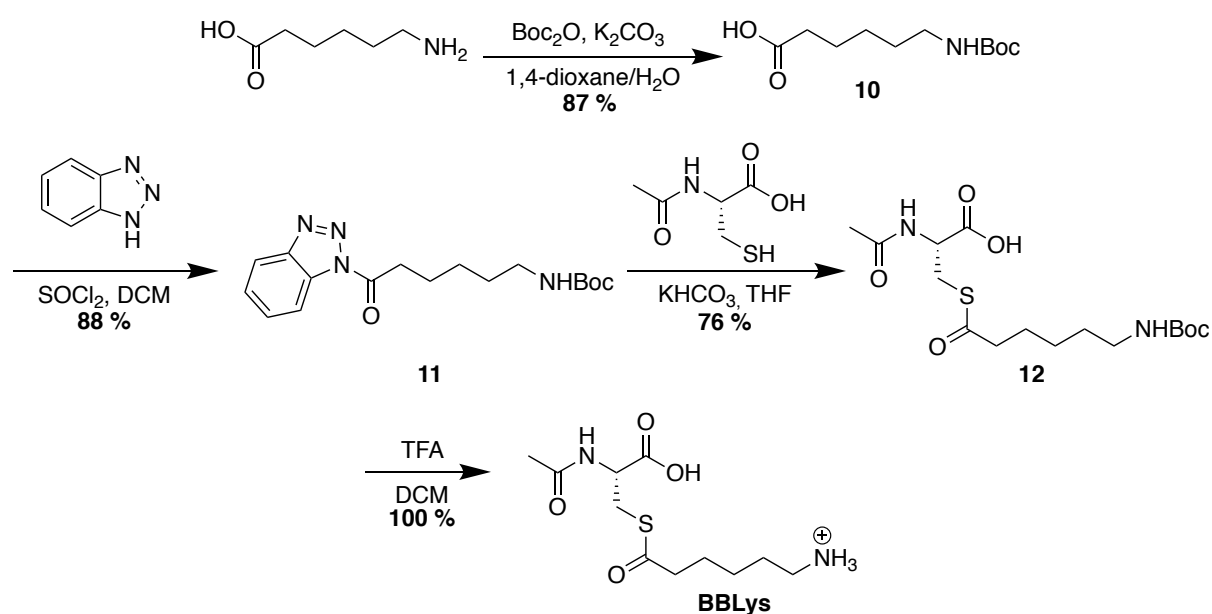
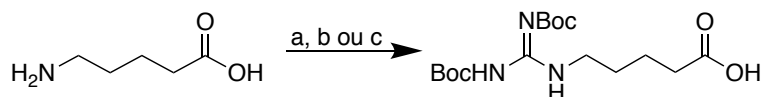


Fig. 32 : Synthèse du **BBLys**

Les **BBArg** et **BBhArg** mimant respectivement les chaînes de l'arginine et de l'homoarginine ont nécessité une étape pour l'introduction du groupement guanidine protégé. Divers réactifs de guanidylation ont été envisagés afin de réaliser cette transformation à partir de l'acide 5-aminopentanoïque pour le **BBArg** et 6-aminohexanoïque pour le **BBhArg** : le réactif de Goodman et le *N,N'*-di-Boc-1*H*-pyrazole-1-carboxamide commerciaux, ainsi que le 1,3-di-Boc-2-méthylisothiourée **13** obtenu après Boc protection de l'isothiourée correspondant (Fig. 33).



- a) Réactif de Goodman ((*N,N'*-Bis(*tert*-butoxycarbonyl)-*N''*-triflylguanidine)) (1,0 eq.), Et₃N (1,5 eq.), DMF, **49 %**.
 b) **13** (*N,N'*-Bis(*tert*-butoxycarbonyl)-*S*-méthylisothiourée) (1,2 eq.), HgCl₂ (1,7 eq.), Et₃N (2,5 eq.), DMF, **65 %**.
 c) *N,N'*-Di-Boc-1*H*-pyrazole-1-carboxamide (1,1 eq.), pyridine, **86 %**.

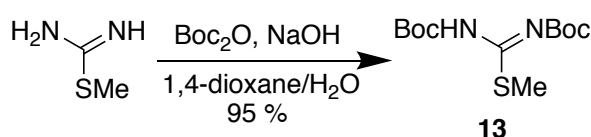
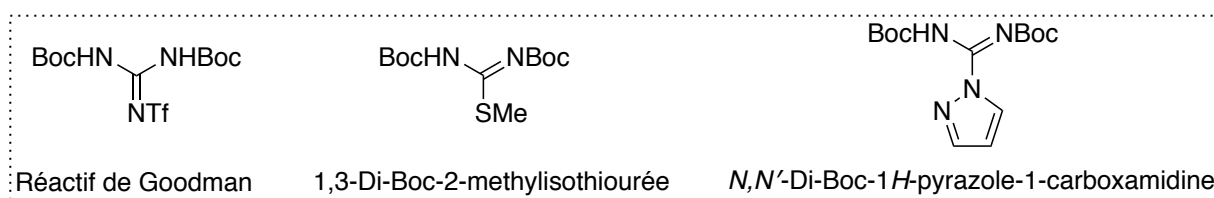


Fig. 33 : En haut, structures des différents réactifs de guanidylation utilisés. Au milieu, synthèse du réactif **13** *N,N'*-Bis(*tert*-butoxycarbonyl)-*S*-méthylisothiourée. En bas, synthèses de l'acide carboxylique **14** précurseur du **BBArg**.

Dans les trois cas, le produit est obtenu avec un rendement satisfaisant. Nous avons choisi la *N,N'*-Di-Boc-1*H*-pyrazole-1-carboxamide comme réactif à privilégier. En effet celui-ci est à la fois commercial et bon marché et ne requiert pas l'utilisation de composés dangereux tels que les sels de mercure¹⁰¹. Cette stratégie d'installation de la fonction guanidine a été utilisée pour obtenir les composés **14** et **17** avec des rendements satisfaisants. Ces derniers ont pu alors être utilisés pour obtenir les précurseurs **15** et **18** suivant le même schéma réactionnel, qui après déprotection finale des guanidines par un traitement au TFA permettent d'accéder respectivement aux deux blocs **BBArg** et **BBhArg** avec des rendements globaux respectifs de 20 % et 4 % (Fig. 34).

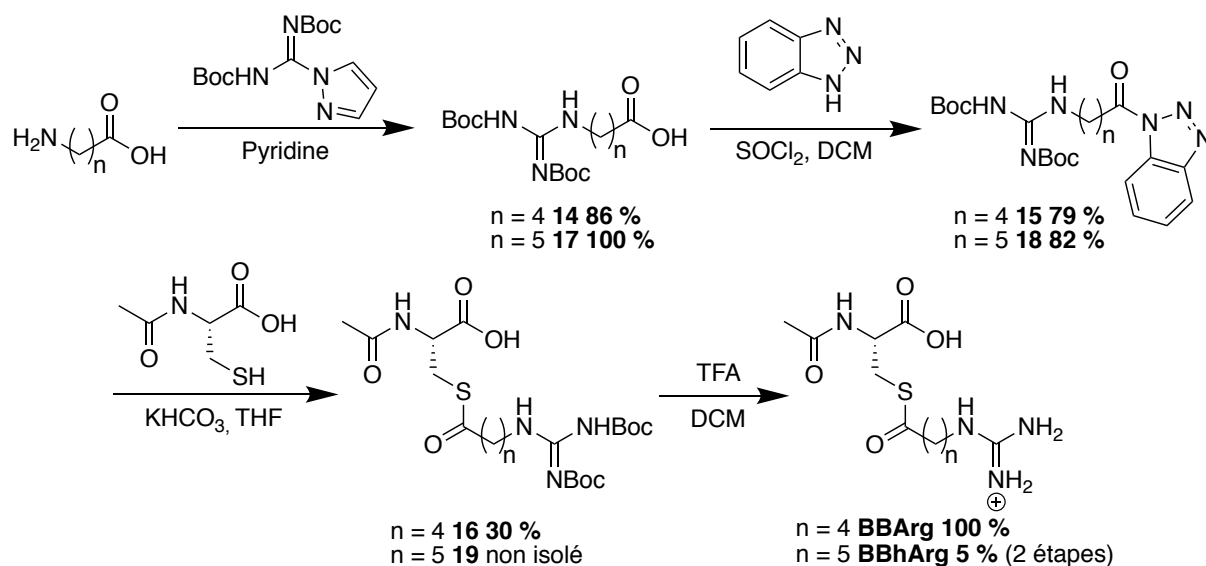


Fig. 34 : Synthèses des **BBArg** et **BBhArg** par une voie impliquant la *N,N'*-Di-Boc-1*H*-pyrazole-1-carboxamidine.

Comme introduit précédemment, les blocs à chaînes acides, **BBAsp** et **BBGlu**, porteurs respectivement des chaînes latérales de l'aspartate et du glutamate, ont été synthétisés *via* une protection intermédiaire sous forme d'ester *tert*-butylique. Afin d'accéder aux diacides monoprotégés sur le groupement de la chaîne R utilisable dans l'étape d'acylation, l'ouverture d'une anhydride cyclique a été envisagée.

A partir de l'anhydride succinique, deux voies de synthèse ont été explorées afin d'obtenir l'intermédiaire **20**. La première en une étape consiste en l'ouverture de l'anhydride succinique catalysée par le chlorure de zinc dans le *tert*-butanol à chaud, conduisant au produit souhaité avec un rendement modeste de 15 %¹⁰². Bien que facilement réalisable, le faible rendement de cette synthèse associé à un temps de réaction très long (9-10 jours) comme décrit dans la littérature nous ont conduit à tester une voie alternative. Cette dernière joue sur l'orthogonalité existant entre les esters méthyliques et *tert*-butyliques : en effet, après ouverture de l'anhydride cette fois-ci dans le méthanol à reflux, le monoester méthylique est obtenu. La fonction acide est alors transformée en chlorure d'acyle par traitement au chlorure de thionyle à reflux¹⁰³. Le chlorure d'acyle est ensuite transformé en ester *tert*-butylique par réaction avec le *tert*-butanol en présence de pyridine dans le chloroforme à reflux également. Pour finir, l'ester méthylique est saponifié en présence de soude afin de conduire au dérivé souhaité **20** avec 76 % de rendement sur 4 étapes. Ce composé a pu ensuite être utilisé dans les étapes générales de la synthèse des thioesters. L'ester *tert*-butylique **22** est déprotégé en dernière étape par un traitement au TFA afin de conduire au **BBAsp** avec un rendement global de 26 % (Fig. 35).

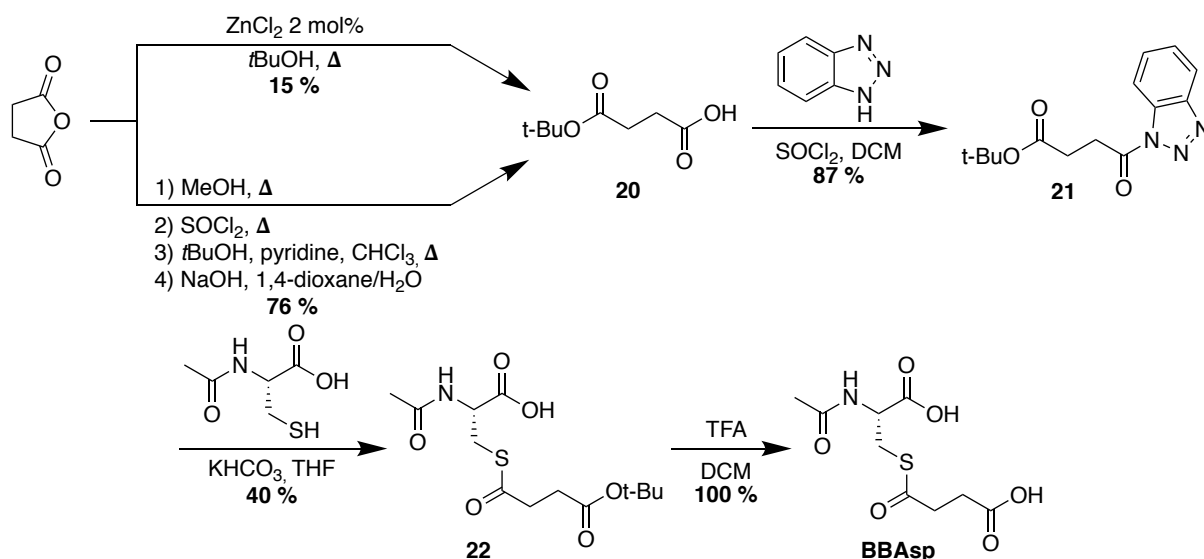


Fig. 35 : Synthèse du **BBAsp**

De manière similaire le **BBGlu** mimant l'acide glutamique a été obtenu à partir de l'anhydride glutarique. L'ouverture du cycle catalysée par le chlorure de zinc dans le *tert*-butanol à chaud permet dans ce cas d'accéder au monoester *tert*-butylique **23** correspondant avec un rendement satisfaisant de 45 %¹⁰⁴. Le composé **23** a été utilisé dans la synthèse des thioesters en suivant la méthodologie habituelle pour conduire au composé **25** qui a pu être déprotégé par traitement au TFA pour donner le **BBGlu**, avec un rendement global de 22 % (Fig. 36).

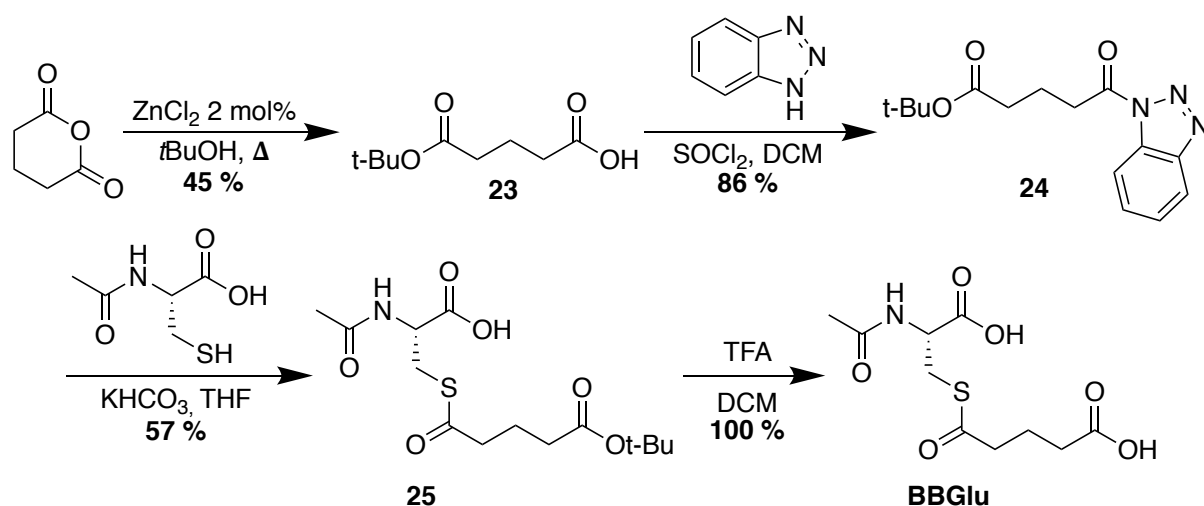


Fig. 36 : Synthèse du **BBGlu**

Le **BBTyr** mimant la chaîne de la tyrosine a été obtenu à partir de l'acide 3-(4-hydroxyphenyl)propanoïque préalablement protégé sous forme d'éther de *tert*-

butyldimethylsilyl sur la fonction phénol avant d'être engagé dans les étapes de synthèse des thioesters précédemment décrites¹⁰⁵. Après activation et thioesterification, l'intermédiaire **28** est traité au TBAF afin de déprotéger la fonction phénol. Le **BBTyr** est alors obtenu avec un rendement global de 19 % (Fig. 37).

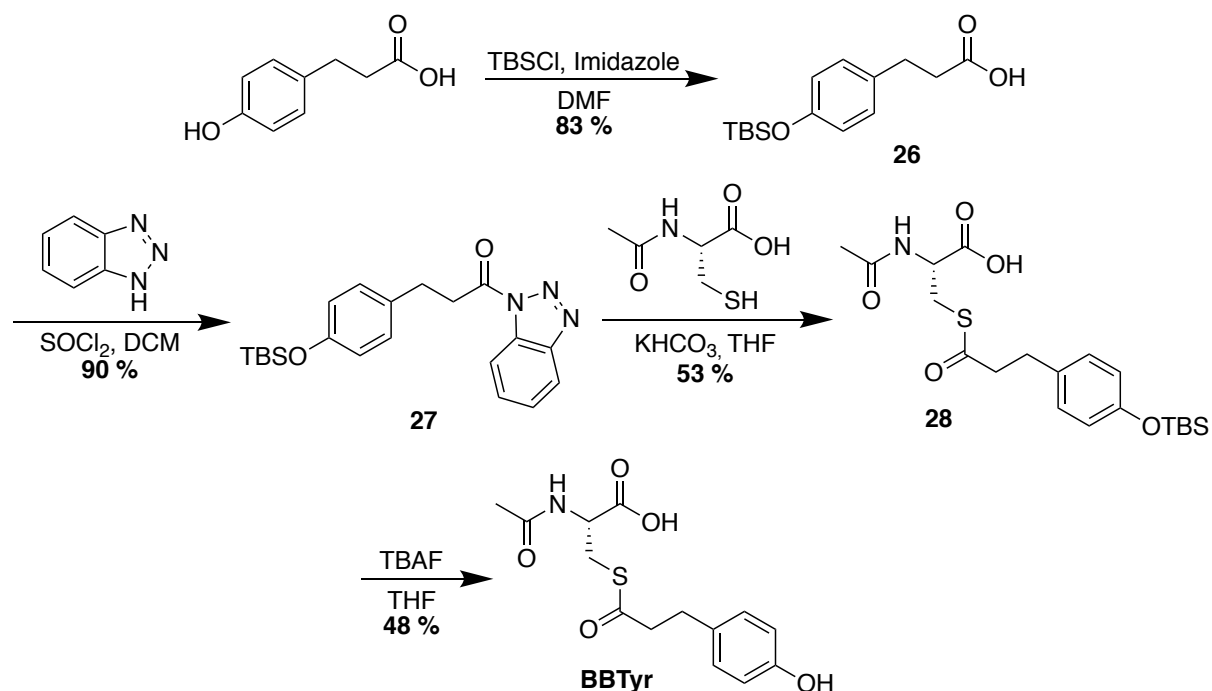


Fig. 37 : Synthèse du **BBTyr**

Avec une approche similaire, la synthèse du **BBSer** mimant la sérine a été débutée en utilisant une protection intermédiaire sous forme d'éther silylé. Le 1,3-propanediol est tout d'abord monoprotégé par le TBSCl après déprotonation par l'hydrure de sodium¹⁰⁶. Le composé **29** est ensuite oxydé par le BAIB en présence de TEMPO, pour obtenir l'acide carboxylique correspondant **30**¹⁰⁷. Après activation de cet acide sous forme d'acylbenzotriazole le composé **31** a été obtenu. L'accès au **BBSer** pourra être réalisé par thioesterification de la *N*-acétylcystéine puis déprotection de l'alcool par traitement au TBAF (Fig. 38).

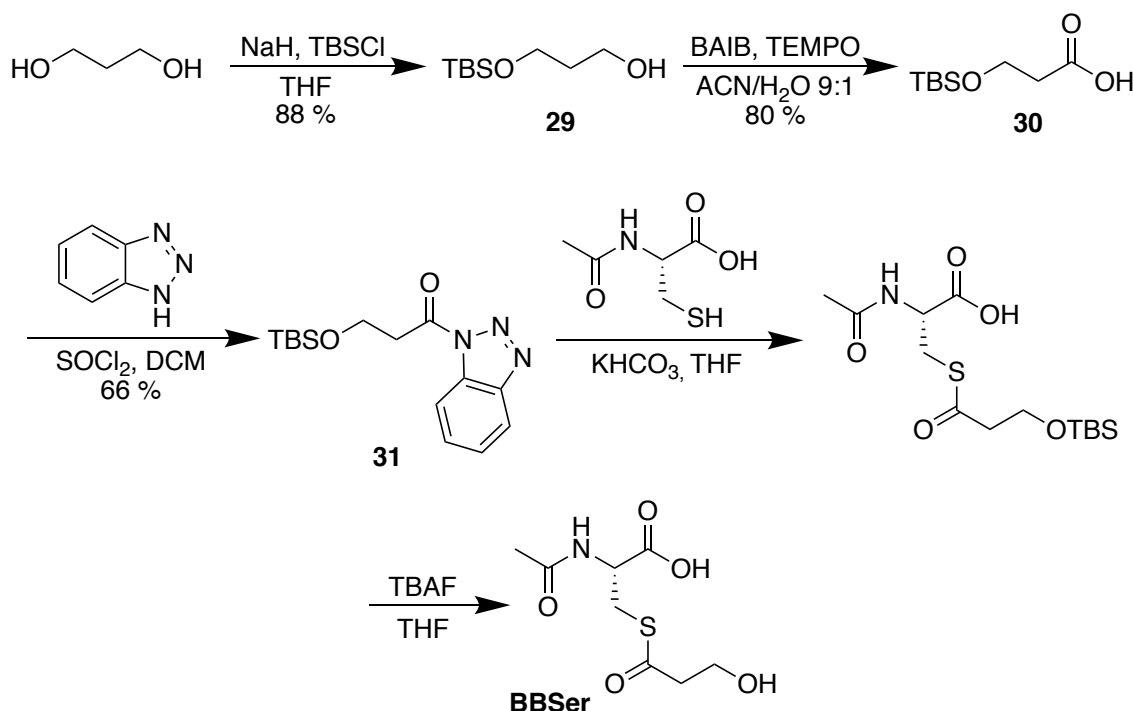


Fig. 38 : Synthèse du **BBSer**

5) Perspectives

Les blocs **BBHis**, **BBAsn**, **BBGln** mimant respectivement les chaînes de l'histidine, l'asparagine et la glutamine n'ont pas été synthétisés mais des voies de synthèse peuvent être proposées pour de futures investigations. Le **BBHis** pourrait être obtenu à partir de l'acide carboxylique correspondant en passant par une protection de l'imidazole par un groupement Boc. Les **BBAsn** et **BBGln** pourraient être obtenus par des synthèses similaires à celles des **BBAsp** et **BBGlu** : les mêmes anhydrides de départ pourraient être transformés en mono amides protégées tritylées. L'obtention de la majorité des blocs correspondants aux chaînes des acides aminés naturels a permis de mettre en évidence la grande polyvalence de cette voie de synthèse. Il est raisonnable de penser que cette approche peut être appliquée à la synthèse des autres blocs naturels, des acides aminés portant des équivalents de modifications post traductionnelles comme des groupements phosphates, des oses ou des lipides ou encore des fonctions qui n'existent pas dans le vivant.

III. Synthèses de *scaffolds* peptidiques

1) Structures tridimensionnelles

Afin de développer notre méthodologie de combinatoire dynamique, le choix préalable d'une plateforme peptidique structurée est nécessaire. Deux *scaffolds* basés sur les deux grands types de structures secondaires les plus représentées sur les interfaces d'interactions protéiques sont envisagés dans ce projet : l'hélice α et le feuillet β , représenté par sa forme la plus simple de structure en épingle à cheveux (*β -hairpin*) (Fig. 39).

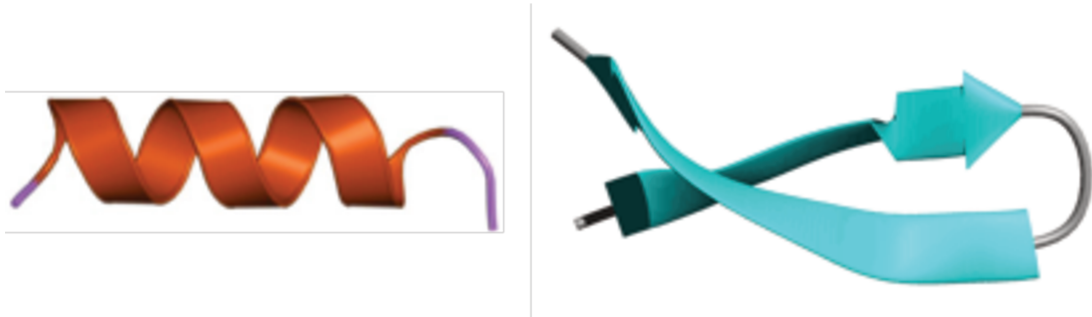


Fig. 39 : Représentation en ruban d'une hélice α (gauche) et d'un *β -hairpin* (droite).

a) *Scaffold en hélice α*

L'hélice α étant l'élément structural le plus abondant dans la Nature, un intérêt important est accordé au développement de structure mimant cet élément. Elle consiste en un enroulement régulier et périodique du squelette peptidique^{108,109}. Elle présente quelques caractéristiques clés comme une stabilisation par des liaisons hydrogène entre les CO et NH des résidus respectivement en position i et $i+4$ et ce pour tous les résidus constituant le segment hélicoïdal. Son pas est de 5,4 Å soit 3,6 résidus par tour d'hélice. Enfin, les chaînes latérales des résidus la composant sont orientées vers l'extérieur de l'hélice de manière régulière. Il est cependant difficile d'obtenir un simple brin peptidique structuré en hélice α dans l'eau, et il est souvent nécessaire d'introduire des contraintes structurales sur le squelette peptidique de manière à forcer ce repliement. Une méthode très largement utilisée pour cela est l'introduction de liens covalents agrafes (staple) entre deux chaînes latérales qui pointent dans la même direction, en position relative i et $i+4$. La nature de ce lien, sa longueur et la position à laquelle l'introduire au sein d'une séquence que l'on souhaite structurer sont des paramètres qui vont influencer la qualité de la structuration mais qui dépendent de la séquence elle-même : au final, il n'existe pas de lien universel pour contraindre des peptides à se replier en hélice α ^{110,111,112,113}.

La méthodologie historiquement la plus utilisée mais aussi la plus simple à mettre en place, ne nécessitant pas de résidu non naturels, consiste à créer un lien lactame entre une lysine et un acide aspartique^{114, 115, 116, 117, 118, 119, 120}. En particulier, le groupe de Fairlie a développé un *scaffold* robuste constitué par un module pentapeptidique *cyclo(1→5)Ac-[KARAD]-NH₂* où la chaîne latérale d'une lysine en position *i* est lié *via* une liaison amide à un résidu aspartate en position *i+4*. Plusieurs modules peuvent être combinés de manière à obtenir des peptides de 1 à 4 tours d'hélice très stable (Fig. 40).

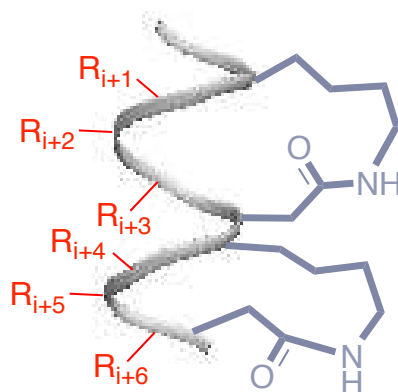


Fig. 40 : Schéma d'un peptide structuré en hélice α par des contraintes apportées par deux agrafes lactames entre les chaînes latérales de résidus lysine et acide aspartique en position relative *i* et *i+4*.

Ce *scaffold* a été largement exploité pour miniaturiser de nombreuses protéines : la protéine de fusion du virus respiratoire syncytial (RSV), la protéine Rev de HIV-1¹²¹, le récepteur ORL-1¹²², l'interaction protéine-protéine cJun-cFos^{28,123}, le domaine BAD de BH3¹²⁴, la protéine m-conotoxin KIIIA¹²⁵ et le complexe protéique p53-MDM2¹²⁶. Nous avons choisi d'utiliser cette structure pour le développement d'un *scaffold* fonctionnalisable par chimie combinatoire dynamique. Théoriquement, toutes les positions non impliquées dans les liens lactames c'est à dire les positions *i+1*, *i+2* et *i+3* du motif pentapeptidique peuvent être occupées par n'importe quel résidu.

b) *Scaffold* en β -hairpin

En ce qui concerne la structure feuillet β , notre choix s'est porté sur des décapeptides cycliques structurés en épingle à cheveux β que l'on peut décrire très succinctement comme un enchaînement brin β – coude ou boucle – brin β dans lequel les deux brins sont antiparallèles l'un par rapport à l'autre et établissent des liaisons hydrogènes entre eux¹²⁷. Ce choix a été

motivée par la simplicité des synthèses permettant l'accès à ces structures ainsi que le savoir-faire, déjà acquis au sein du laboratoire, pour les obtenir.

L'accès à ce type de structure est généralement obtenu par utilisation d'un fragment inducteur de coude β qui va forcer les séquences l'encadrant à adopter une structuration en brins β , recréant ainsi l'enchaînement caractéristique des épingles à cheveux β . De nombreux inducteurs de coude β ont été étudiés et décrits dans la littérature (Fig. 41)¹²⁸.

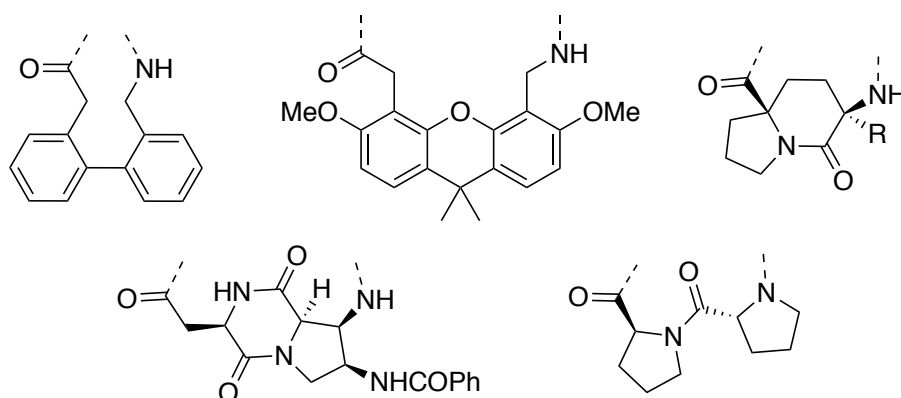


Fig. 41 : Structures de différents inducteurs de coude β .

Ces inducteurs présentent tous la capacité de stabiliser une structuration en épingle à cheveux β mais l'enchaînement D-Proline-L-Proline, représenté en bas à droite dans la figure précédente, reste le plus pratique à utiliser puisqu'il est directement possible de l'inclure dans une séquence peptidique en cours de synthèse sur support solide. Il a été montré que cet enchaînement induisait une conformation en coude β -II^{129, 130, 131}.

La structuration induite par le template confère une grande stabilité structurale au peptide et cela avec une grande tolérance vis à vis de la composition de la séquence tant que les positions où les liaisons hydrogènes inter-brins ne sont pas occupées par des résidus incapables de former ces liaisons H. Grâce à cela, ces plateformes peuvent être synthétisées avec une approche modulaire afin de concevoir une architecture répondant à nos besoins sans pour autant déstabiliser la structure.

L'arrangement des chaînes latérales des résidus formant les brins de part et d'autre de l'inducteur présente une alternance périodique de résidus en résidus. Les premiers résidus liés en N-ter et en C-ter de l'inducteur présenteront leur chaîne latérale sur la même face de la structure. Les résidus suivant ces derniers sur chacun des brins, sur la face opposée et ainsi de suite. Cette propriété intrinsèque de ces architectures nous a conduit à déterminer des positions sur lesquelles introduire les résidus cystéines qui seront engagées dans la DCC afin qu'elles

présentent leur chaîne sur la même face de la plateforme, imitant ainsi une surface de protéine susceptible d'interagir avec une cible par une région structurée en épingle à cheveux β (Fig. 42).

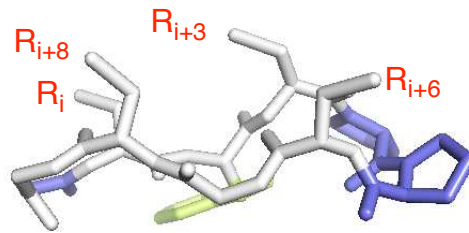


Fig. 42 : Schéma d'un peptide cyclique structuré en épingle à cheveux β par le motif D-Pro-L-Pro inducteur de coude β .

Deux *scaffolds* structurés respectivement en hélice α et en β -*hairpin* ont donc été développés. Dans chaque cas, seuls quelques résidus sont nécessaires au repliement du peptides, les résidus lysine/aspartate reliés par une liaison lactame dans le *scaffold* hélice α et le template D-Pro-L-Pro, ainsi que la glycine situé dans la boucle opposée pour ce qui est du β -*hairpin*. Les autres positions, et en particulier les résidus de surface peuvent être modifiés sans affecter la structure 3D de ces objets.

Les *scaffolds* β -*hairpin* synthétisés sont nommés HairpX-Y^Z, où X est le nombre de résidus cystéines incorporés, Y une caractéristique mise en avant pour l'exposé de ce travail et Z la position des chaînes thioester éventuellement présentes. De même, les *scaffold* hélicoïdaux sont nommés HelixX-Y où X désigne le nombre de répétition du motif cyclique pentapeptidique et Y le nombre de cystéines incorporées.

2) Séquence des *scaffolds* peptidiques

La figure 43 résume et schématise la structure de ces *scaffolds* et introduit également une numérotation qui sera celle employée tout au long de ce travail.

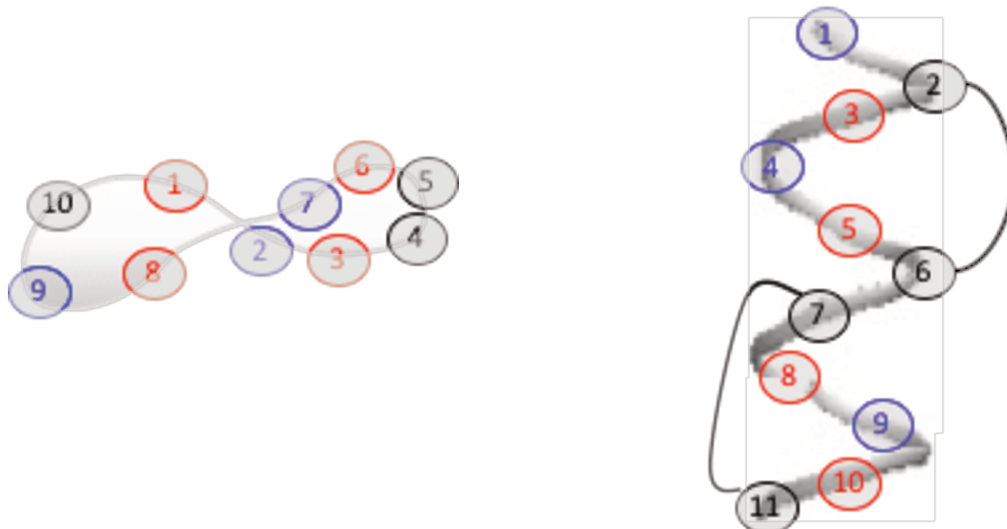


Fig. 43 : Schéma de la structure générale des *scaffolds* en épingle à cheveux β visés (gauche) et en hélice α (droite). Le squelette peptidique est représenté par le trait gris. Les résidus sont représentés par les cercles rouges (cystéines), noirs (inducteur structural), bleus et gris (résidus variables pour ajuster les propriétés) et sont numérotés d'après leur position dans le squelette. La position numéro une a été choisie arbitrairement pour que les positions des cystéines susceptibles d'être modifiées par transthioestérification aient les indices les plus petits possibles. Le sens de la numérotation suit l'enchaînement du squelette de N-ter vers C-ter.

Comme évoqué précédemment chacun des *scaffolds* contient un certains nombres de résidus structuraux qu'il n'est pas possible de modifier, les positions 4 et 5 qui constituent l'enchaînement D-Proline-L-Proline, inducteur de coude β et la position 10 occupée par une glycine pour le *scaffold* β -hairpin, et les positions 2/6 et 7/11 occupées par les paires Lys/Asp pour le *scaffold* hélice α . Pour ce qui des autres positions, une des faces du *scaffold* est recouverte par les résidus dont les fonctions présentes sur leurs chaînes latérales participeront à la réaction de chimie combinatoire dynamique, il s'agit des positions 1, 3, 6 et 8 pour le *scaffold* β -hairpin et 3, 5, 8 et 10 pour le *scaffold* hélice α . Les positions restantes peuvent quant à elles être exploitées comme nous le verrons par la suite pour moduler les propriétés physico-chimiques du peptide (solubilité, absorbance UV...).

a) Résidus impliqués dans la DCC

Dans le but d'avancer pas à pas dans le développement de notre méthodologie de chimie combinatoire dynamique nous avons dans un premier temps choisi de travailler sur des systèmes simples, ne présentant qu'une ou deux fonctions thiols réactives en DCC afin d'accéder à des bibliothèques de faibles complexités, facilement analysables. Si dans un premier

temps la nature des autres résidus placés dans ces positions avait été choisie arbitrairement (**Hairp-1-pNA**), nous avons par la suite décidé de remplacer les cystéines sur ces autres positions par l'acide L-2,3-diaminopropionique acétylé sur l'amine de la chaîne latérale (Dap(Ac)), de manière à avoir une chaîne iso-structurale d'une chaîne thioester mais non échangeable dans les conditions de la réaction puisque la liaison amide est stable en condition physiologique. (Fig. 44).

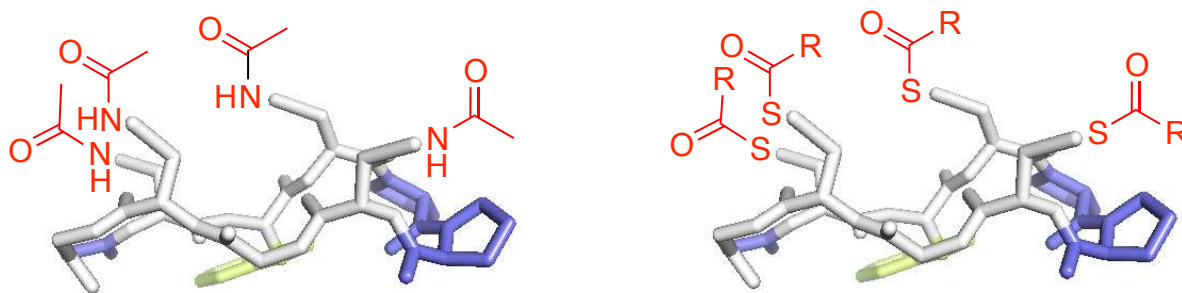


Fig. 44 : Structures comparées d'une plateforme à 4 résidus Dap(Ac) et 4 thioesters.

L'acide aminé correspondant Fmoc-L-Dap(Ac)-OH **32** a été préparé à partir du Fmoc-L-Dap(Boc)-OH commercial par un traitement au TFA dans le DCM puis acétylation de l'amine dans l'anhydride acétique en conditions basiques (Fig. 45). Ce résidu a été introduit dans les séquences peptidiques dans des conditions de couplage classiques.

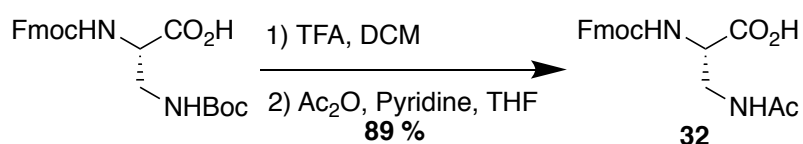


Fig. 45 : Synthèse du Fmoc-L-Dap(Ac)-OH.

b) Sonde spectroscopique

L'approche modulaire de la synthèse de ces molécules permet également d'introduire un résidu aux propriétés spectroscopiques intéressantes. En effet, le suivi de l'évolution de la composition des bibliothèques combinatoires dynamiques se faisant par RP-HPLC avec une détection UV-visible, avoir un chromophore permettant de ne détecter que les espèces peptidiques impliquées dans ces bibliothèques peut s'avérer un choix judicieux afin de simplifier l'analyse. En particulier, s'affranchir des signaux correspondants aux **BB** thioesters qui seront utilisés en excès, serait un avantage non négligeable. Dans un premier temps, un résidu

tryptophane a été utilisé pour cela, possédant un maximum d'absorption à 280 nm, orthogonal à celui des building blocks qui quant à eux absorbent pour la plupart uniquement à 220 nm grâce à leurs multiple fonctions carbonyles. Celui-ci a été utilisé dans la synthèse de différents *scaffolds* **Hairp0**, **Hairp1**, **Hairp4**, **Helix1-1**, **Helix2-2** et **Helix2-4** (Fig. 46).

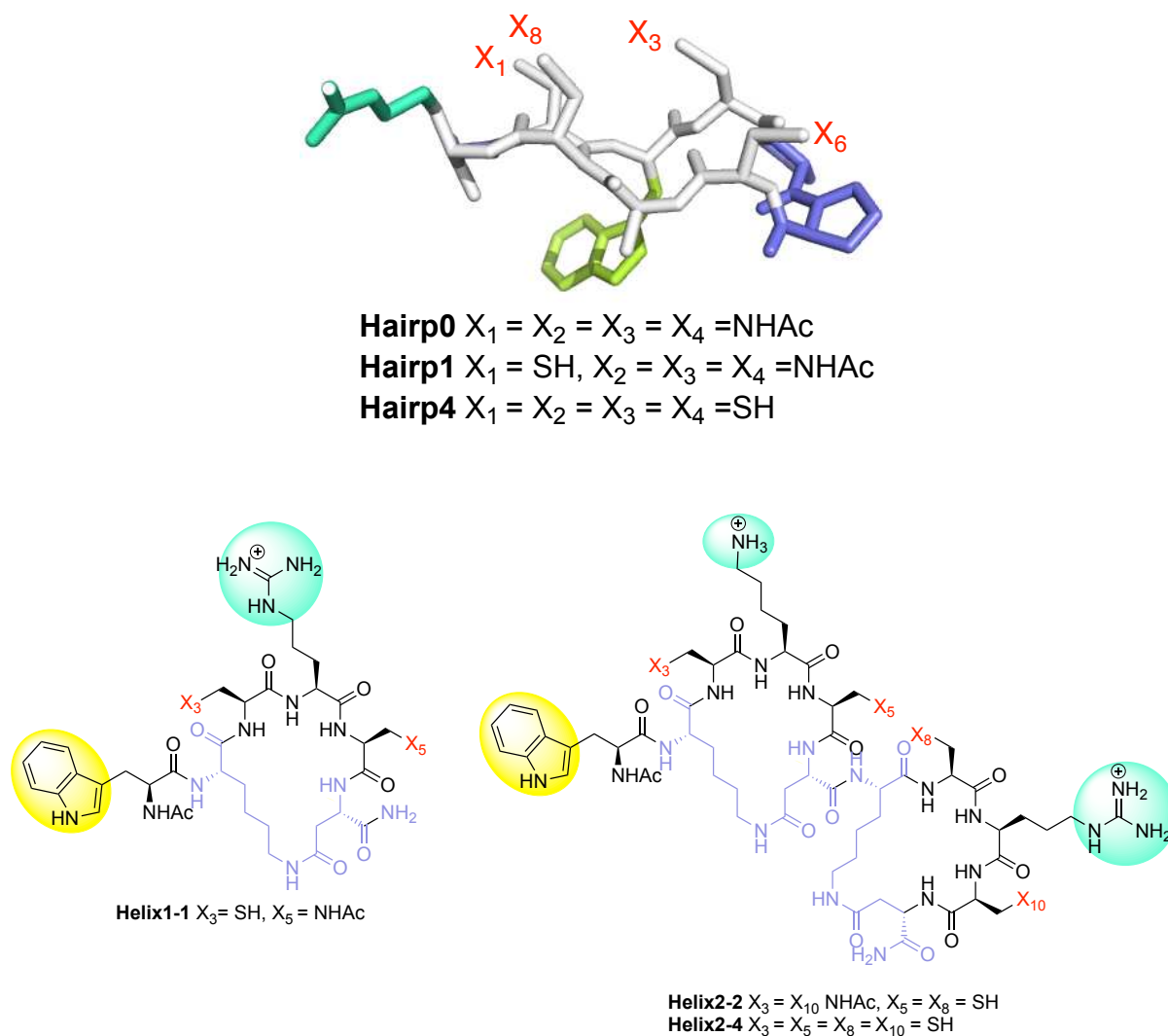


Fig. 46 : Structures génériques de différents *scaffolds* peptidiques synthétisés.

Cependant, cette sonde est loin d'être idéale puisque certains **BB** portant les chaînes latérales de résidus aromatiques (**BBTrp**, **BBTyr**) absorbent eux aussi fortement à cette longueur d'onde. Par conséquent, différents chromophores, possédant une longueur d'onde maximale différente de celle de ces résidus ont également été envisagés : le groupement *para*-nitroanilide, porté par la chaîne latérale du glutamate ($\lambda_{\text{max}} = 350 \text{ nm}$), le groupement naphthoyl, introduit sur la chaîne latérale d'un résidu lysine ($\lambda_{\text{max}} = 318 \text{ nm}$) et le groupement 4-

diméthylaminophthalimide porté par un résidu acide diaminopropionique ($\lambda_{\text{max}} = 325 \text{ nm}$ et 416 nm) (Fig. 47).

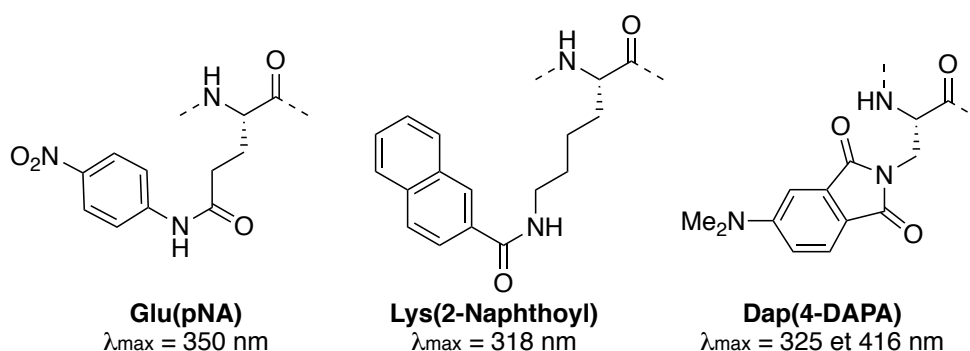


Fig. 47 : Structures des différents résidus chromophores utilisés.

La synthèse des résidus requis correctement protégés en vue de leur utilisation en SPPS et leur introduction dans la séquence peptidique du *scaffold* β -hairpin a été entreprise et leurs propriétés UV étudiées.

Le Glu(pNA) commercial a été protégé par un groupement Fmoc sur l'amine α (Fig. 48).

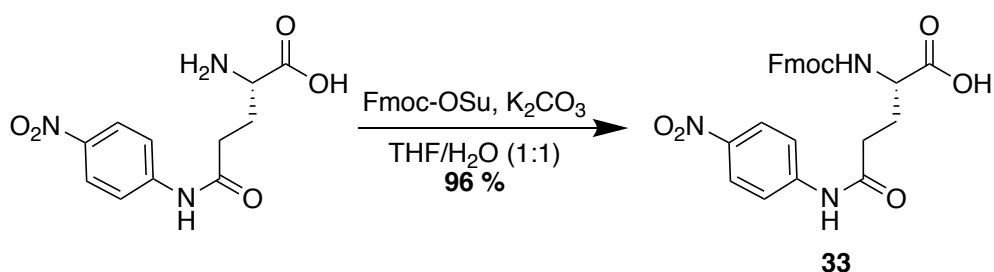


Fig. 48 : Synthèse du Fmoc-L-Glu(pNA)-OH

Ce résidu a été incorporé dans la séquence peptidique **Hairp1-pNA** afin d'en étudier les propriétés spectroscopiques.

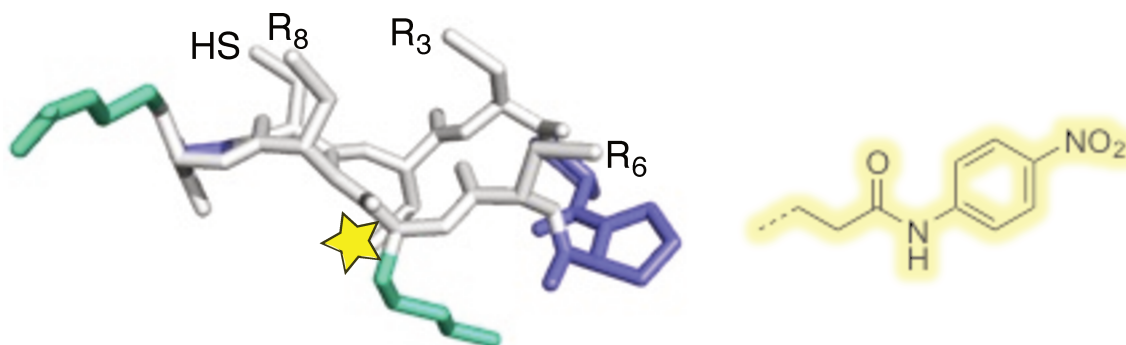


Fig. 49 : Schéma des structures de **Hairp1-pNA** et du chromophore associé.

Le Glu(pNA) présente une absorbance importante aux alentours de 350 nm, zone du spectre où les blocs thioesters n'absorbent plus du tout. Du point de vue de la spectroscopie UV-visible et donc de l'analyse des librairies dynamiques par RP-HPLC avec détection UV-visible, ce chromophore s'avère donc satisfaisant. Malheureusement, le motif *para*-nitroanilide s'est révélé instable en spectrométrie de masse MALDI-ToF et MALDI-ToF-ToF. En effet, ce motif subit une fragmentation après irradiation par le laser du MALDI selon le schéma décrit en figure 50¹³².

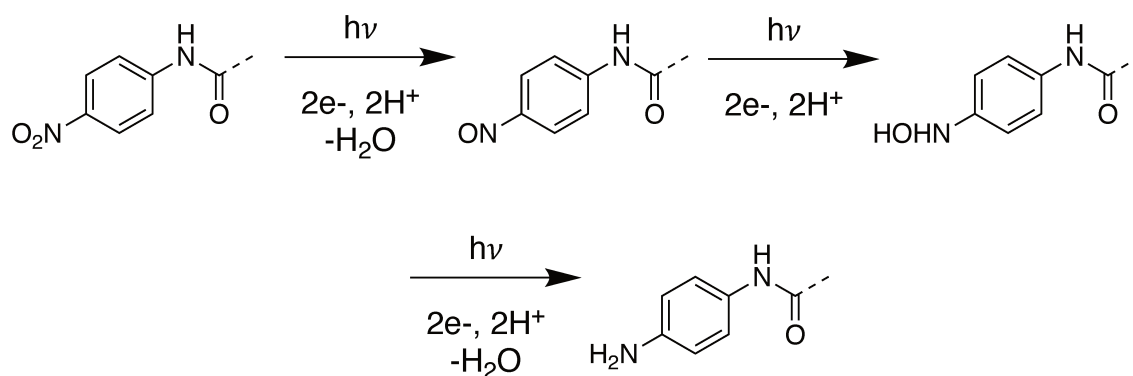


Fig. 50 : Schéma réactionnel de fragmentation du motif *para*-nitroanilide.

Ainsi, il est possible d'observer les espèces [MH-16]⁺, [MH-14]⁺, [MH-30]⁺. Ce mécanisme de fragmentation déjà présent en MALDI-ToF se révèle extrêmement problématique en spectrométrie de masse en tandem car il est très favorable énergétiquement et empêche la bonne fragmentation du squelette pour son séquençage. De plus, la multiplication du nombre de pics en spectrométrie de masse vient complexifier de manière importante l'analyse.

Le Fmoc-Lys(2-Naphthoyl) a été préparé à partir de la Fmoc-L-Lys(Boc)-OH commerciale, par clivage du groupement Boc au TFA et couplage de l'amine libre en présence de chlorure de 2-naphtoyl en présence de carbonate de potassium (Fig. 51).

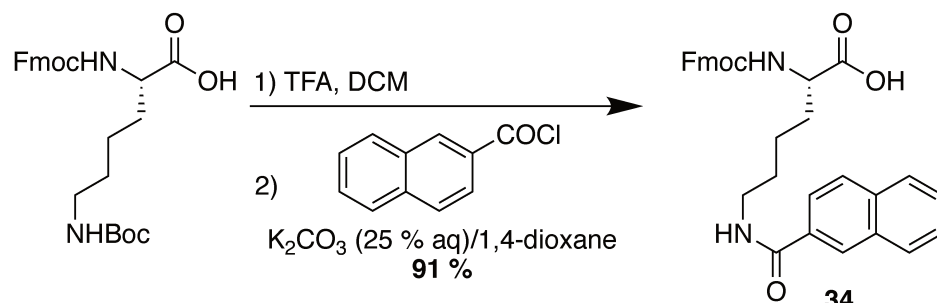


Fig. 51 : Synthèse de la Fmoc-L-Lys(2-Naphthoyl)-OH

Le chromophore 2-naphtoyl introduit sur la chaîne latérale d'une lysine a été utilisé pour la préparation des séquences peptidiques **Hairp1-Naphthoyl-loop**, **Hairp1-Naphthoyl-strand** et **Hairp2-Naphthoyl-strand** (Fig. 52).

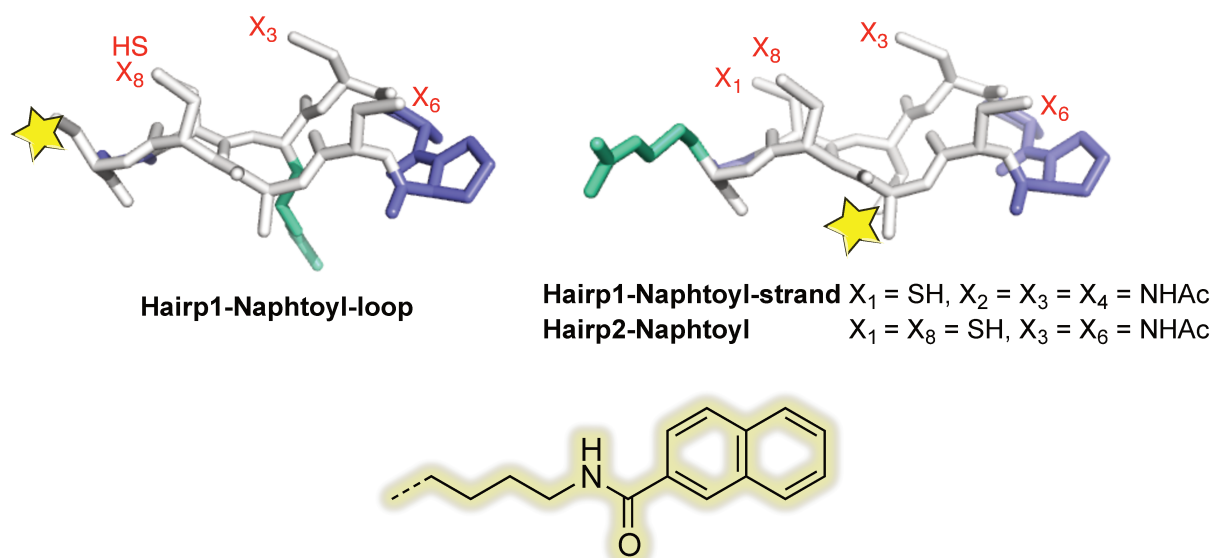


Fig. 52 : Schéma des structures de **Hairp1-Naphthoyl-loop**, **Hairp1-Naphthoyl-strand** et **Hairp2-Naphthoyl** et du chromophore associé.

En ce qui concerne la position du résidu chromophore au sein de la plateforme, celle-ci a été amenée à évoluer au cours de ce travail. Initialement localisé en position 9 (**Hairp1-Naphthoyl-loop**), celui-ci a été inversé avec le résidu arginine localisé en position 2 (**Hairp1-Naphthoyl-strand** et **Hairp2-Naphthoyl-strand**) afin de placer le résidu basique arginine, site de linéarisation par digestion enzymatique dans la boucle du peptide cyclique (cf chapitre 4).

Pour toutes les séquences étudiées, aucune fragmentation majeure du chromophore n'est observé en spectrométrie de masse à l'inverse du Glu(pNA). Cependant, au sein des séquences peptidiques préparées, une modification de ses propriétés UV-visible a été observée. En effet, la bande d'absorption initialement attendue aux alentours de 330 nm subit un décalage à 318 nm, conduisant à un début de recouvrement avec l'absorption à 280 nm des blocs thioesters. Surtout, le coefficient d'extinction molaire est largement réduit, diminuant significativement la sensibilité de la détection et rendant nécessaire l'utilisation de concentrations élevées en peptides pour l'analyse des bibliothèques. Enfin ce résidu confère aux peptides une grande hydrophobie ce qui diminue sensiblement leur solubilité.

Le Fmoc-Dap(4-DAPA) a été synthétisé en quelques étapes à partir du Fmoc-L-Dap(Boc)-OH, suivant la méthodologie développée par Imperiali *et. al.*^{133,134} L'acide est d'abord protégé sous forme d'ester d'allyle, avant clivage du groupement Boc. L'anhydride **35** est préparé par amination réductrice de l'acide 4-aminophtalique commercial et déshydratation en présence d'anhydride acétique. Celui-ci est couplé à l'amine **36** en présence de HBTU/HOBt. La Fmoc-Dap(4-DAPA) est alors obtenu après déprotection du groupement allyle par le phenylsilane catalysée par le palladium tetrakis triphenylphosphine (Fig. 53).

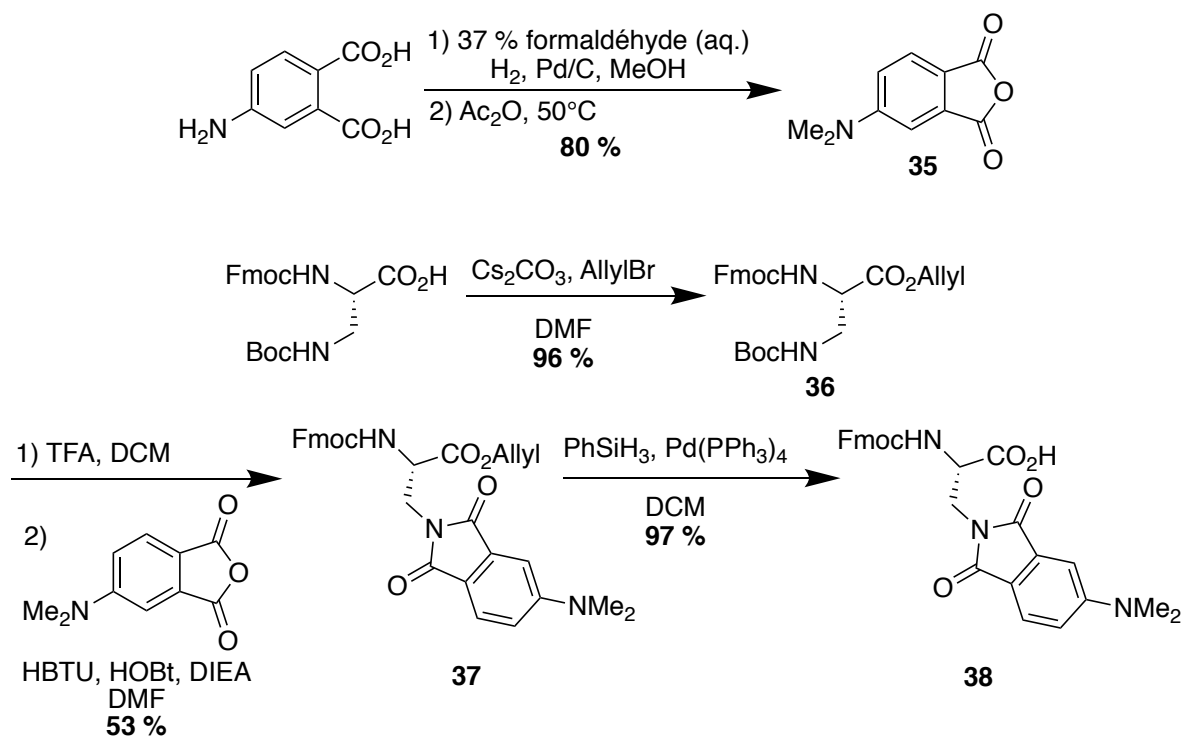


Fig. 53 : Synthèse du Fmoc-L-Dap(4-DAPA)-OH.

Le Dap(4-DAPA) a été utilisé pour la préparation des peptides **Hairp1-DAPA-Ala**, **Hairp1-DAPA-Lys** et **Hairp4-DAPA-Lys** (Fig. 54).

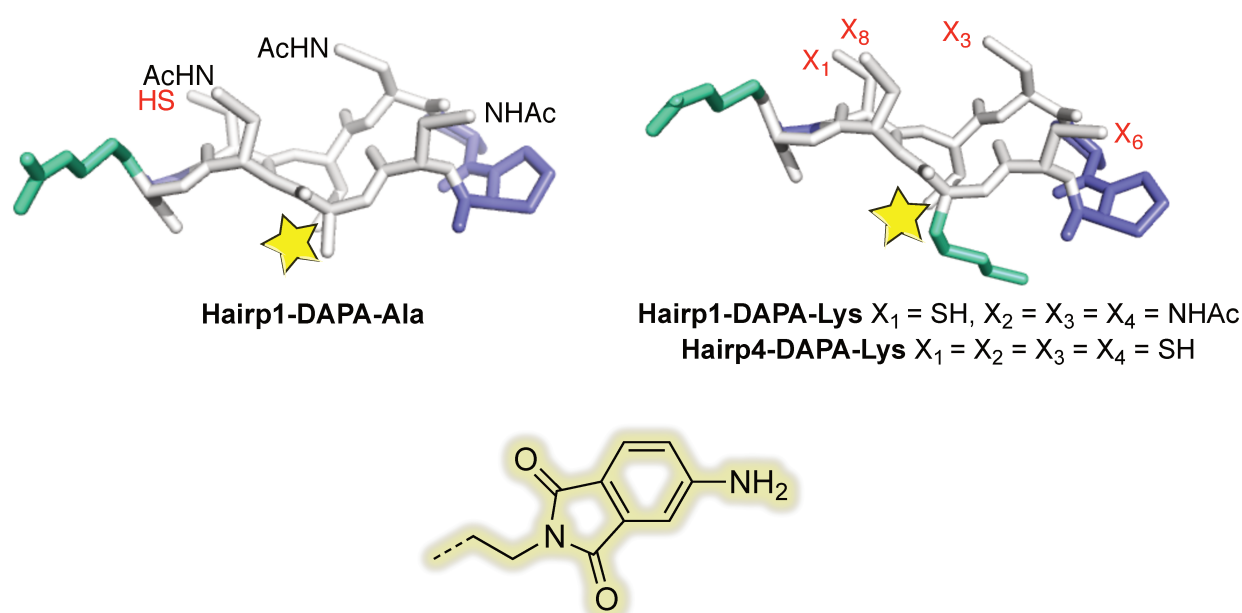


Fig. 54 : Schéma des structures de **Hairp1-DAPA-Ala**, **Hairp1-DAPA-Lys** et **Hairp4-DAPA-Lys**.

Ces peptides se sont révélés compatibles avec la méthode d'analyse par spectrométrie de masse puisqu'aucune fragmentation du chromophore n'est observée ici, en MALDI ainsi qu'en MS-MS. La bande d'excitation de fluorescence du chromophore à 416 nm, suffisamment éloignée de l'absorption des thioesters et très intense peut être exploitée pour l'analyse des librairies. Ce groupement étant également fluorophore, il est possible d'envisager un changement de système de détection en passant de l'absorption UV-visible à l'émission de fluorescence. De plus, il pourrait servir de sonde intéressante pour la caractérisation d'interactions intermoléculaires entre une plateforme fonctionnalisée et une cible.

c) Solubilité

Dans tous les peptides synthétisés, une arginine, résidu cationique à pH physiologique, a été introduite en position 2 ou en position 9 (cf chapitre 4), afin d'assurer une bonne solubilité du peptide dans l'eau. Pour tous les peptides ne comportant qu'une seule cystéine, cette arginine s'est révélée en effet suffisante pour garantir une excellente solubilité des peptides aux concentrations utilisées avant et après acylation par DCC. Cependant lorsque le nombre de cystéine augmente (**Hairp2-Naphthoyl** et **Hairp4-Ala**) les peptides deviennent globalement moins solubles à ces mêmes concentrations, en particulier lorsque des fonctions thioesters sont

introduites en DCC. Cela s'explique par le fait que les acétamides, groupements polaires et à la fois donneur et accepteur de liaison H participent grandement à la solubilité en milieu aqueux. A contrario, les fonctions thiols/thioesters sont peu polaires. Cette solubilité décroissante avec le nombre de cystéines présentes a donc conduit à modifier le résidu alanine en position 7 par un résidu lysine. Cette modification a été faite dans le but d'obtenir des librairies dynamiques comportant des membres à la solubilité équivalente. En effet, une disparité trop importante de ce paramètre pourrait entraîner un biais dans la composition de la librairie si certaines espèces s'avéraient trop peu solubles puisque ces dernières pourraient précipiter et ainsi ne plus être en équilibre avec le reste des membres de la librairie. Ce résidu basique et polaire a permis de retrouver une solubilité acceptable pour les plateformes **Hairp4-Lys** et **Hairp4-DAPA-Lys** (cf Fig. 46 et Fig. 54).

3) Synthèse des *scaffolds*

a) *Scaffolds en hélice- α*

L'accès au motif *cyclo*(1 \rightarrow 5)-Ac-[KARAD]-NH₂ se fait par synthèse sur support solide en stratégie Fmoc/*t*Bu en choisissant des groupements protecteurs pour les chaînes latérales de la lysine et de l'acide aspartique orthogonaux aux conditions de clivage. Typiquement, des groupement protecteur type alloc/allyle peuvent être utilisés pour les chaînes latérales qui seront impliquées dans l'agrafe, orthogonaux au groupement Fmoc/*t*Bu puisqu'ils peuvent être clivés par traitement au palladium tetrakis(triphénylphosphine) sur résine permettant ainsi la cyclisation des deux chaînes latérales avant clivage/déprotection finale. Une seconde stratégie consiste à jouer sur la différence de sensibilité de groupements protecteurs acido-labiles, en utilisant pour les chaînes latérales lysine/aspartate des groupements sensibles à des conditions acides très douces qui n'auront pas d'effets sur les autres groupements protecteurs ou le lien à la résine : le groupement 4-méthyltrityl (Mtt) pour la lysine et le groupement 2-phénylisopropyl (O-2-PhiPr) pour l'acide aspartique qui sont clivés tous les deux à l'aide d'une solution de TFA à 1 % dans le DCM par exemple (Fig. 55).

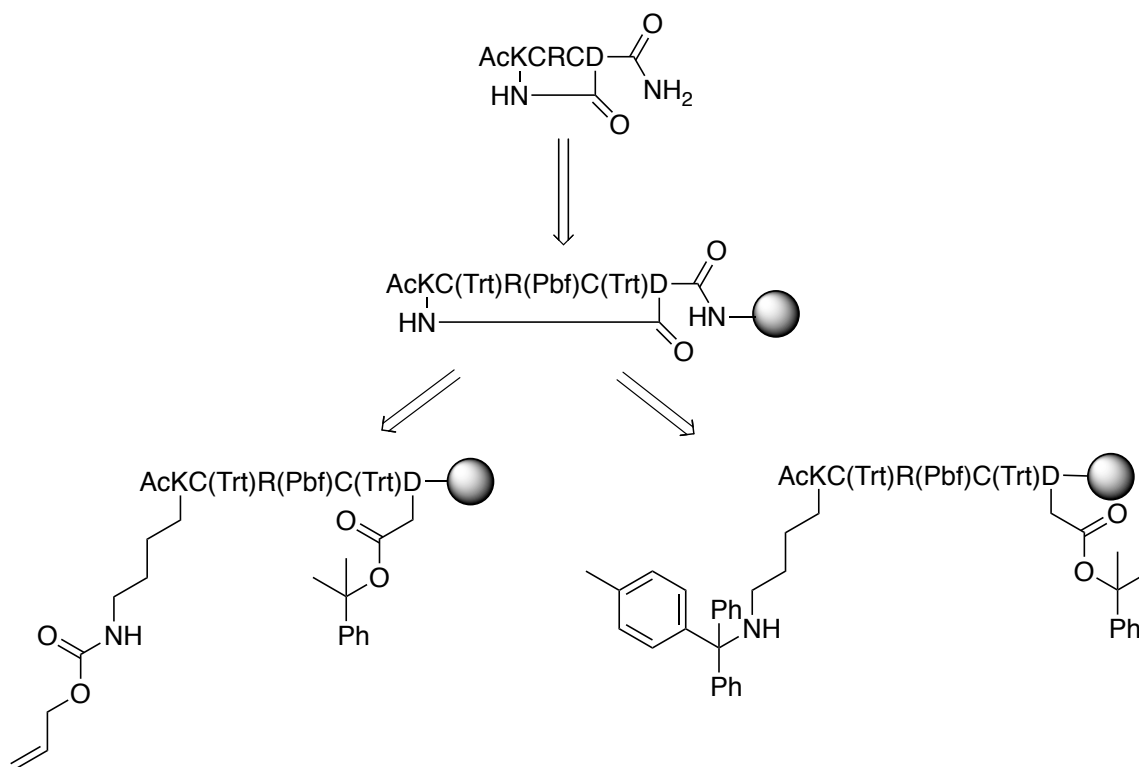
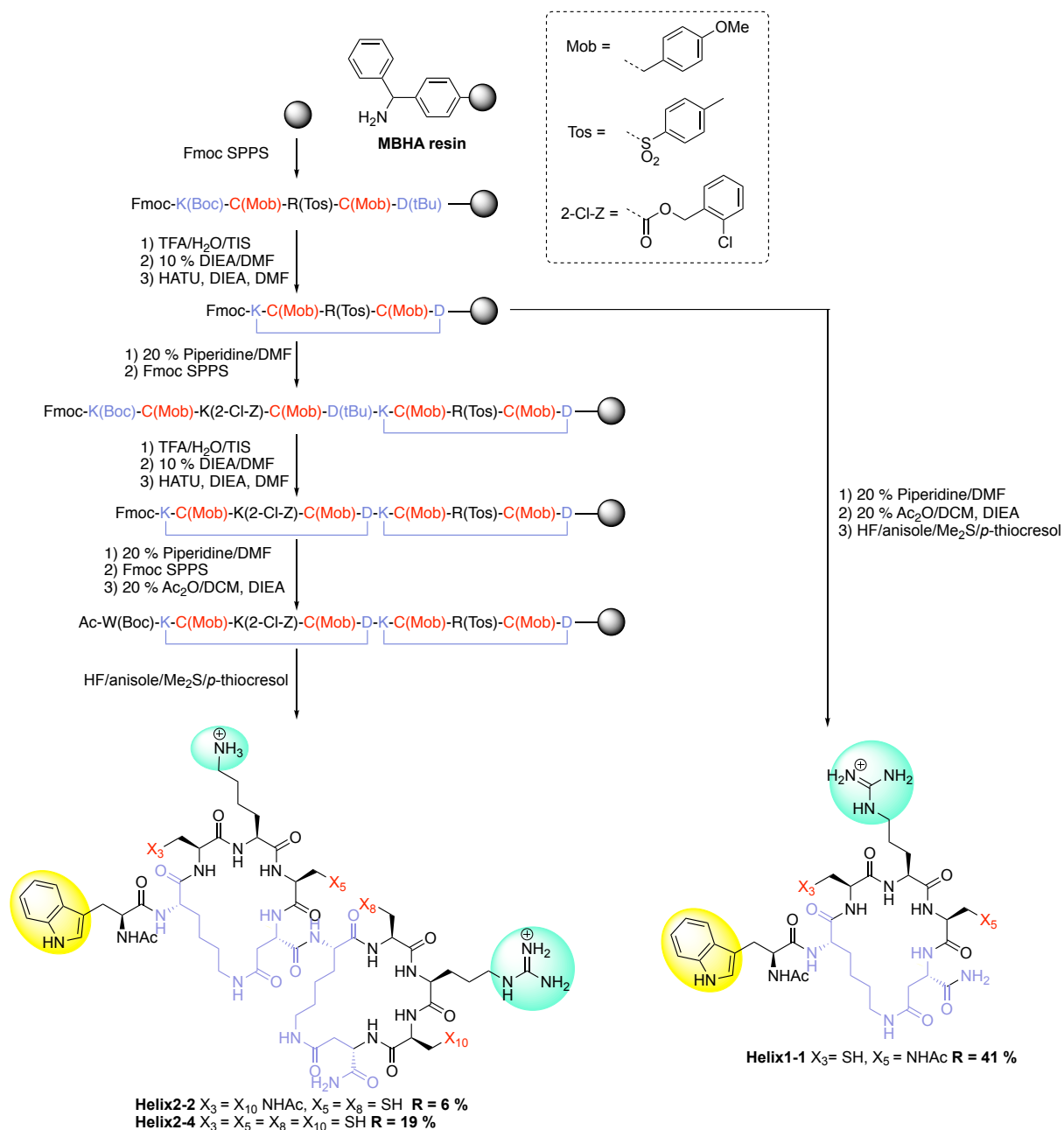


Fig. 55 : Schéma rétrosynthétique pour l'accès aux pentapeptides avec groupements protecteurs couramment utilisés pour la synthèse des motifs à agrafes lactames.

Nous avons choisi d'exploiter cette seconde stratégie, et la séquence a été ainsi préparée par SPPS, incluant une Lys(Mtt) en position 1 et un Asp(O-2-PhiPr) en position 5. Malheureusement lors du clivage de ces groupements par 1% de TFA, un pourcentage important des groupements trityles portés par les cystéines est également clivé. Les différentes tentatives consistant à jouer sur la concentration de TFA et le temps d'exposition n'ont malheureusement pas permis d'obtenir une orthogonalité totale entre Mtt/O-2-PhiPr d'une part et Trt d'autre part.

Nous avons donc établi une nouvelle stratégie de synthèse reposant sur l'orthogonalité entre les groupements protecteurs des chaînes latérales des lysines et acides aspartiques impliquées dans les agrafes d'une part, les groupements protecteurs des chaînes latérales des autres résidus du peptide et le lien à la résine d'autre part. Pour ce faire nous avons utilisé une résine MBHA d'ordinaire destinée à la stratégie Boc, résistante au TFA et clivable donc par traitement au HF. La lysine et l'acide aspartique qui seront impliqués dans les agrafes peuvent donc être protégés respectivement par un groupement Boc et *t*Bu afin de pouvoir les déprotéger par traitement au TFA sans impacter le reste du peptide et pouvoir ensuite réaliser le couplage chaîne latérale à chaîne latérale sur support solide. Les autres chaînes latérales sont quant à elles protégées par des groupements orthogonaux, qui seront clivés uniquement en fin de synthèse

par le HF : tosylé (Tos) pour l'arginine, méthyloxybenzyl (Mob) pour la cystéine et 2-chlorobenzoylé (2-Cl-Cbz) pour la lysine (Fig. 56).

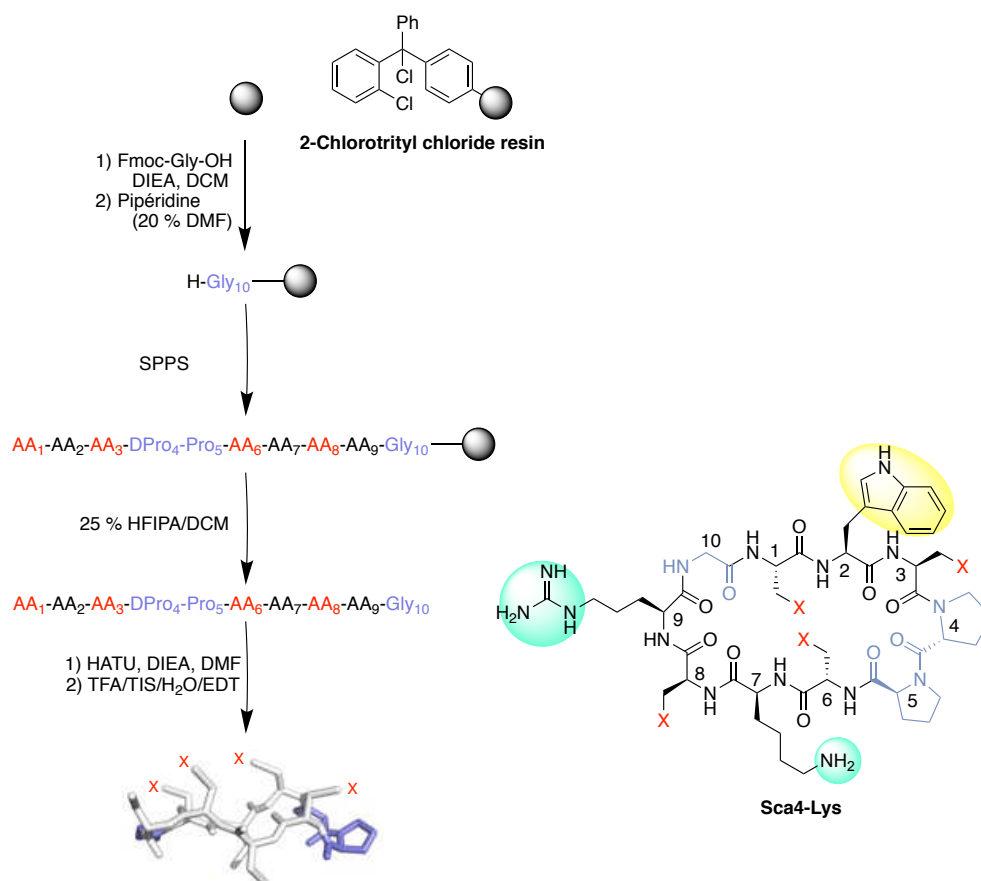


Peptide	Temps de rétention HPLC (min)	MS MALDI-ToF MH+	R
Helix1-1	7.0	858.35	41%
Helix2-2	9.3	1142.12	6%
Helix2-4	5.4	1392.13	19%

Fig. 56 : Schéma de synthèse des *scaffolds* hélice α et structures des différents peptides synthétisés. Les structures développées représentées ici ne rendent pas compte de la structure tridimensionnelle hélicoïdale attendue. Le tableau résume les caractérisations et rendements de synthèse obtenus.

b) Scaffold β -hairpin

L'accès aux *scaffolds* en épingle à cheveux β s'est fait par synthèse peptidique sur support solide en stratégie Fmoc/*t*Bu à partir d'une résine chlorure de 2-chlorotriyle, suivant la stratégie décrite par le groupe de Robinson¹³⁵. Après avoir accroché l'acide aminé qui sera en C-terminal de la séquence linéaire (donc la glycine en position 10 dans nos structures), des cycles de déprotection du groupement Fmoc et couplage peptidique sont répétés afin d'assembler la séquence complète, résidu après résidu. De manière classique, les étapes de déprotections ont été réalisées par traitement avec une solution de pipéridine à 20 % dans le DMF et les couplages obtenus après activation de l'acide carboxylique à coupler par le HBTU et une base, la DIEA. La résine utilisée étant très sensible aux conditions acides, elle permet une fois la séquence assemblée d'utiliser des conditions très douces par traitement à l'HFIPA, afin de libérer le fragment peptidique linéaire sans toucher aux groupements protecteurs des chaînes latérales et obtenir une extrémité C-terminale sous forme d'acide carboxylique. Ainsi, on peut procéder à la cyclisation tête à queue entre l'amine N-terminale et l'acide carboxylique C-terminal par couplage peptidique en solution diluée, afin de favoriser la réaction intramoléculaire et d'éviter les problèmes d'oligomérisation. Dans nos structures cette cyclisation se fera donc entre le carboxyle de la glycine en position 10 et l'amine du résidu en position 1. Le choix de la glycine comme résidu C-terminal est judicieux ici en tant qu'acide aminé à activer pour ce couplage puisqu'elle permet d'éviter les risques d'épimérisation souvent observés lors de l'activation puisque la glycine ne possède pas de centre stéréogène. Après cyclisation, la déprotection finale des chaînes latérales peut-être réalisée par traitement avec un mélange de TFA et de scavengers avant de passer à l'étape de purification par RP-HPLC préparative (Fig. 57).



Peptide	Temps de rétention HPLC (min)	MS MALDI-ToF MH+	R
Hairp0	4.7	1234.84	24%
Hairp1	6.9	1152.24	14%
Hairp1-pNA	8.2	1222.42	15%
Hairp1-Naphthoyl-loop	6.5	1248.70	7%
Hairp1-Naphthoyl-strand	7.2	1248.70	4%
Hairp1-DAPA-Ala	5.3	1225.50	6%
Hairp1-DAPA-Lys	7.8	1282.73	8%
Hairp2-Naphthoyl	7.4	1223.77	7%
Hairp4-Ala	7.3	1077.52	13%
Hairp4-Lys	6.5	1134.43	19%
Hairp4-DAPA-Lys	7.0	1207.70	8%

Fig. 57 : A gauche, schéma de synthèse générale pour l'accès aux décapeptides cycliques en épingle à cheveux β . La structure représentée en fin de synthèse est une représentation construite grâce au logiciel PyMOL (sur structure X = SH ou NHAc). Elle n'a qu'une vocation illustrative afin de simplifier la lecture par rapport aux structures développées car elle rend mieux compte de la disposition tridimensionnelle des chaînes latérales. Elle ne correspond cependant pas à une structure qui aurait été déterminé par RMN. A droite, représentation à plat d'un exemple de peptide obtenu par cette méthode : **Sca4-Lys**, portant 4 cystéines, la stéréochimie est fidèle dans cette représentation mais pas la structuration tridimensionnelle. Le tableau résume les caractérisations et rendements de synthèse obtenus.

4) Bilan et perspectives

L'approche modulaire de ces synthèses a permis une évolution pas à pas des structures des plateformes en épingle à cheveux β . Ainsi, les différentes propriétés et paramètres liés à ces structures ont pu être affinés progressivement afin de satisfaire aux besoins imposés par la chimie combinatoire dynamique et les stratégies de linéarisation en spectrométrie de masse. Cela a également permis de construire rapidement des structures modèles simples afin de travailler sur la mise au point des différents aspects du projet. La structure idéale n'est à l'heure actuelle pas encore obtenue mais il s'agit d'un compromis à faire entre les différentes propriétés de la plateforme. Celles-ci devant être optimales pour les différents aspects de l'utilisation de la plateforme : la chimie combinatoire dynamique avec la réactivité, la détection par spectroscopie UV-Visible ou fluorescence et la solubilité, la linéarisation pour la spectrométrie de masse et enfin la structuration.

Chapitre 2 : Chimie combinatoire dynamique par transthioestérification sur *scaffolds* en épingle à cheveux β

I. Généralités

La méthodologie de chimie combinatoire dynamique a été mise au point sur la plateforme structurée en épingle à cheveux β présentant donc au plus quatre sites de fonctionnalisation cystéines sur une même face. Cette fonctionnalisation réversible est réalisée par transthioestérification à partir des **BB** thioesters décrits dans le chapitre 1 (Fig. 58).

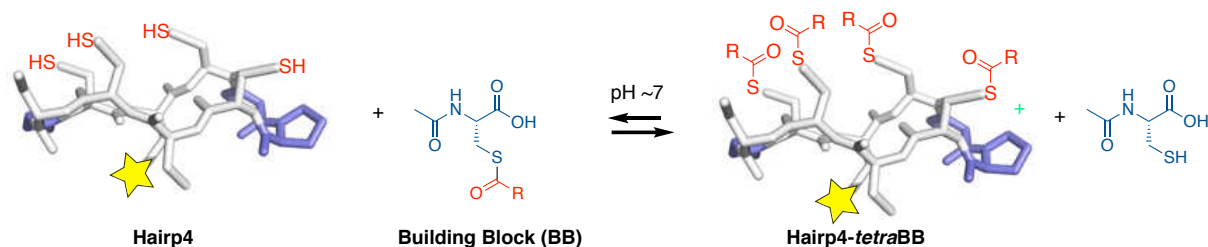


Fig. 58 : Équilibre de transthioestérification à mettre en place sur un système cible à quatre cystéines.

Les structures *tétra* fonctionnalisées visées par cette approche sont en équilibre entre elles et avec leurs analogues *mono*, *di* et *tri* fonctionnalisées également. Dans un système comme le nôtre, le nombre de membre d'une DCL va suivre une loi polynomiale de la forme $(n+1)^x$ où n est le nombre de sites fonctionnalisables et x le nombre de blocs utilisés pour introduire ces fonctionnalités. Si l'on se place dans un système à 4 sites comme c'est le cas ici, le nombre de membres d'une librairie suit l'évolution suivante en fonction du nombre de blocs (Fig. 59).

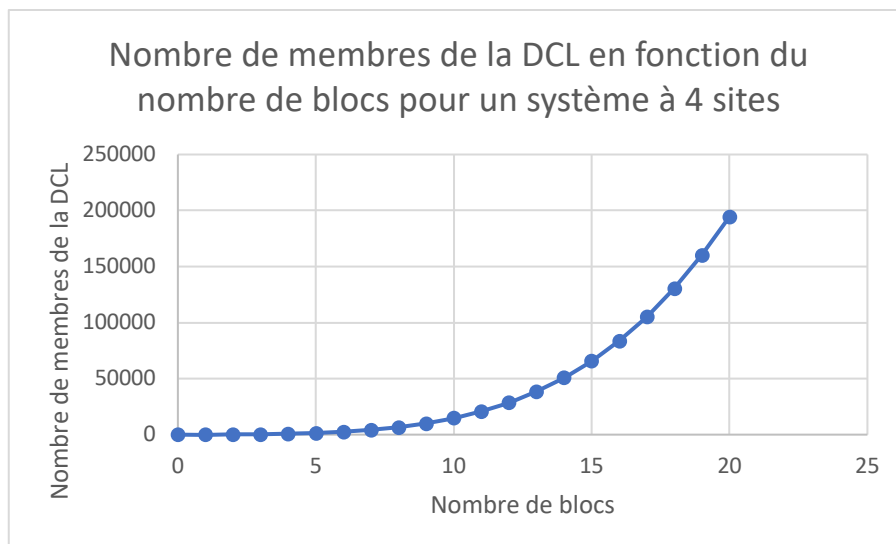


Fig. 59 : Graphique représentant l'évolution du nombre de blocs de la DCL en fonction du nombre de **BB** utilisés pour un système à 4 sites.

On constate que la taille de la librairie augmente très rapidement avec le nombre de blocs puisqu'avec un seul bloc la librairie comporte déjà 16 structures peptidiques : en plus du peptide de départ ne comportant aucun thioester, 4 composés portant une seule chaîne latérale thioester, 6 composés di-acylés, 4 tri-acylés, et un composé où toutes les fonctions thiols sont fonctionnalisées. Avec deux **BB**, la librairie contient théoriquement 81 membres, avec trois blocs, ce nombre passe à 256... Cette augmentation très rapide est importante à prendre en compte pour deux points. D'une part cela permet de souligner la puissance de l'approche par DCC pour cribler des nombres importants de composés en « one-pot ». D'autre part, l'évolution rapide de la taille de ces librairies va se confronter à la capacité de les analyser tant d'un point de vue de la détection que de l'interprétation des criblages. Il est donc nécessaire de travailler initialement sur des systèmes relativement simples et dont le maximum de paramètres sont ajustables.

II. Fonctionnalisation d'un peptide modèle à une cystéine

1) Mise au point des conditions expérimentales

Dans le but de mettre au point les conditions de cette réaction sur des systèmes modèles simples, facilement analysables, les peptides **Hairp1**, **Hairp1-Glu(PNA)**, **Hairp1-Naphtoyl-strand** et **Hairp1-DAPA**, possédant un seul site de fonctionnalisation ont tout d'abord été évalués dans la réaction de transthioesterification.

Le peptide est ainsi traité par un unique bloc, soit en conditions stœchiométriques, soit en excès (5 ou 10 équiv.) de blocs en milieu aqueux tamponné à pH physiologique (Fig. 60).

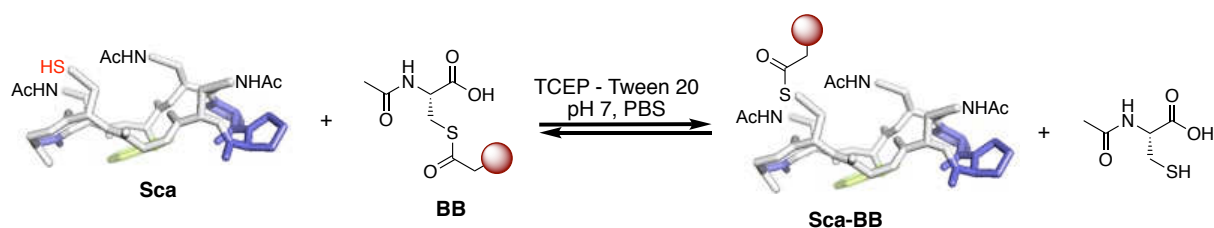


Fig. 60 : Schéma réactionnel de la transthioestérification entre une plateforme à une position fonctionnalisable et un **BB** générique.

Un agent réducteur, la TCEP (Tris(2-CarboxyEthyl)Phosphine), est introduit dans le tampon afin d'empêcher la réaction compétitive de formation de disulfures par oxydation des thiols qui compliquerait le système. Dans les réactions impliquant les peptides **Hairp1** et **Hairp1-DAPA-Ala**, un détergent doux, le TWEEN-20 est également introduit à raison de 1 % (v/v) afin d'éviter l'agrégation des peptides qui peuvent devenir particulièrement hydrophobes une fois fonctionnalisés par des **BB** à chaînes aliphatiques ou aromatiques. Nous avons choisi de travailler à une concentration en peptide de 1 mM dans le milieu réactionnel. Ce choix représente un compromis entre une concentration suffisamment élevée pour permettre une bonne détection des composés en HPLC et compatible avec l'utilisation future d'une cible biologique, qui quant à elle serait utilisée à une concentration 10 à 100 fois plus faible que le peptide. La réaction est suivie par HPLC après prélèvement d'un échantillon du milieu réactionnel et arrêt de l'échange par acidification. Les constantes d'équilibre ainsi que les temps nécessaires pour atteindre cet équilibre sont donnés dans le tableau 1. Les chromatogrammes obtenus pour les réactions entre **Hairp1** et **BBPhe** ou **BBTrp**, à différents ratios, sont donnés à titre d'exemples dans la figure 61.

Tableau 1 : Paramètres réactionnels des réactions entre plateformes à une cystéine et un **BB** ainsi que les constantes d'équilibres apparentes calculées à partir des intégration des pics HPLC.

Entrée	Sca	BB	Ratio	pH	Temps	K'
1	Hairp1	BBPhe	1:1	7	6-8h	non atteint à 24h
2			1:5			0,2
3			1:10			0,1
4		BBTrp	1:1	7	8h	0,2
5			1:5			0,5
6			1:10			0,2

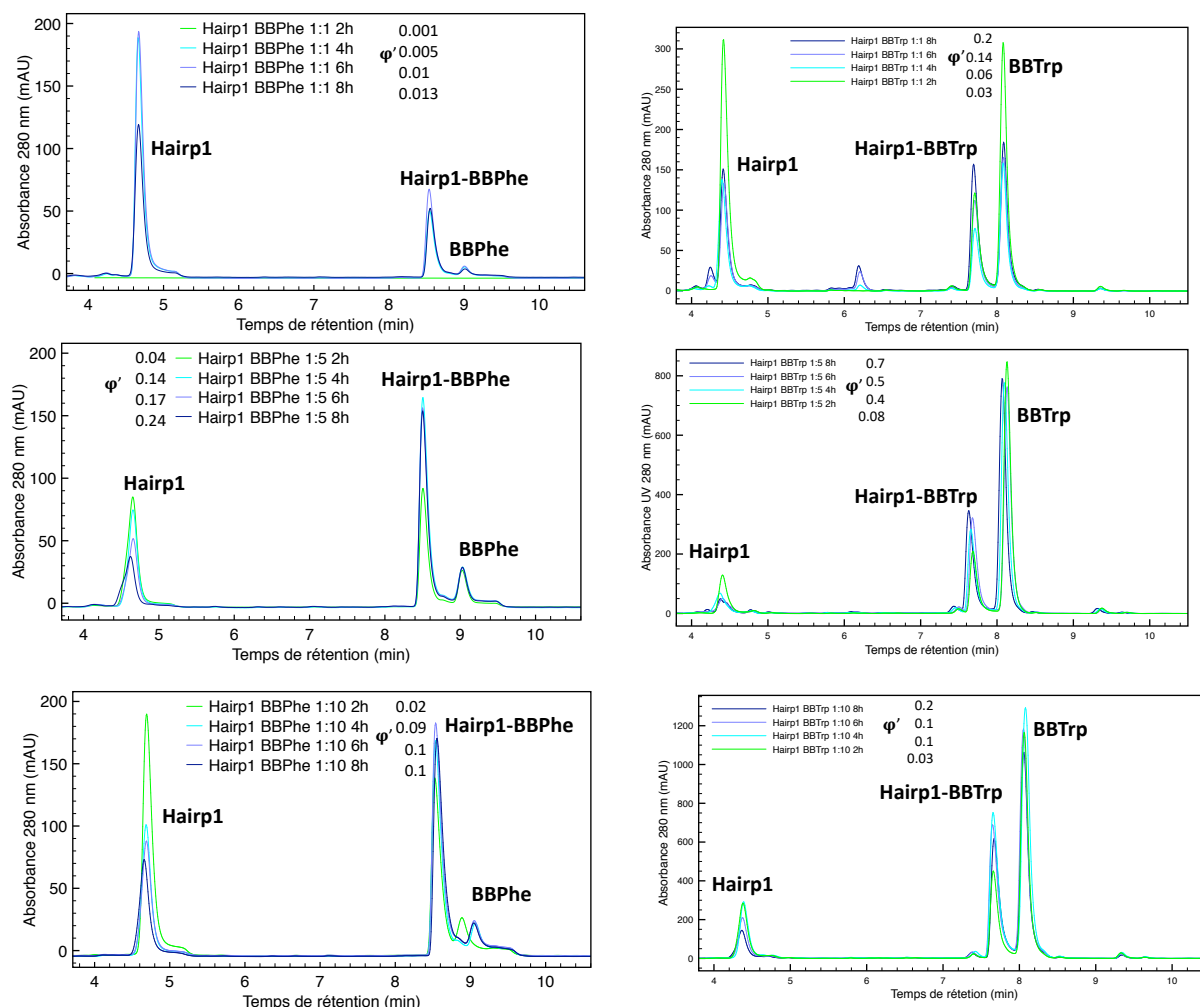


Fig. 61 : Chromatogrammes HPLC obtenus pour les réaction entre **Hairp1** et **BBPhe** (colonne de gauche) et **BBTrp** (colonne de droite) aux différents ratios : 1:1 en haut, 1:5 au milieu, 1:10 en bas. Les quotients de réaction apparents ϕ' sont indiqués pour chaque temps du suivi de la réaction.

On observe dans chaque cas la formation d'un unique nouveau produit correspondant au peptide acylé, confirmé après collecte du pic correspondant et analyse MS-MALDI-ToF. La proportion de ce nouveau composé augmente au cours du temps, jusqu'à atteindre un équilibre. En conditions stœchiométriques, la réaction est relativement lente et l'équilibre est en général

atteint seulement après 24h. La réaction peut être significativement accélérée lorsque les blocs sont utilisés en excès, avec un équilibre atteint en 6h environ avec un ratio de 1:5 et 4h environ avec un ratio de 1:10.

Le mécanisme de la réaction de transthioestérification décrit dans la figure 62, fait intervenir dans l'étape cinétiquement déterminante l'attaque nucléophile des ions thiolates sur le thioester. (Fig. 62).

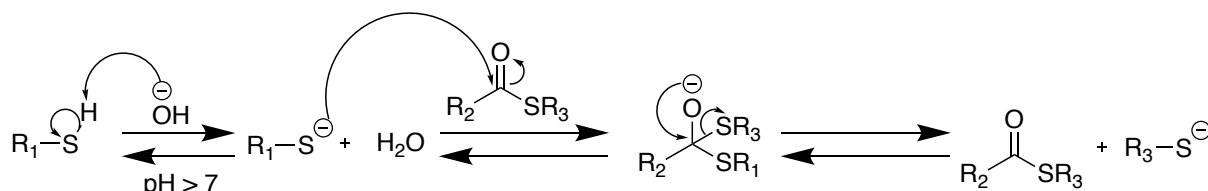


Fig. 62 : Réaction de transthioestérification en milieux aqueux légèrement basique.

Le pKa du couple (thiol/thiolate) d'une cystéine dans un peptide est décrit comme étant d'environ 8,5¹³⁶. À pH neutre, la population de thiols présents dans le milieu réactionnel est donc partiellement déprotonnée. Cette forme déprotonnée des thiols étant l'espèce réactive, le pH peut avoir une influence significative sur la cinétique de la réaction. En effet plus la concentration en thiolate sera importante et plus la cinétique globale de la transthioestérification devrait s'en trouver augmentée. Un compromis est cependant à faire car à pH trop élevé, l'hydrolyse basique des thioesters, irréversible, deviendra favorable venant briser l'équilibre d'échange¹³⁷. Afin d'étudier l'influence du pH sur la cinétique de la réaction, le peptide **Hairp1** (1 mM) a été traité par **BBPhe** ou **BBTrp** (5 équiv.), à pH 7, 7,3 et 7,5, dans les mêmes conditions que décrites précédemment. De la même manière, l'évolution de la composition du mélange a été suivie par HPLC (Tableau 2 et Fig. 63).

Tableau 2 : Paramètres réactionnels des réactions conduites aux pH 7, 7,3 et 7,5 entre plateformes à une cystéine et un **BB** ainsi que les constantes d'équilibres apparentes calculées à partir des intégration des pics HPLC.

Entrée	Sca	BB	Ratio	pH	Temps	K'
2	Hairp1	BBPhe	1:5	7	6-8h	0,2
8				7,3	4-5h	1
9				7,5	2-3h	0,8
5		BBTrp	1:5	7	4-6h	0,5
10				7,3	4-5h	2
11				7,5	2-4h	1

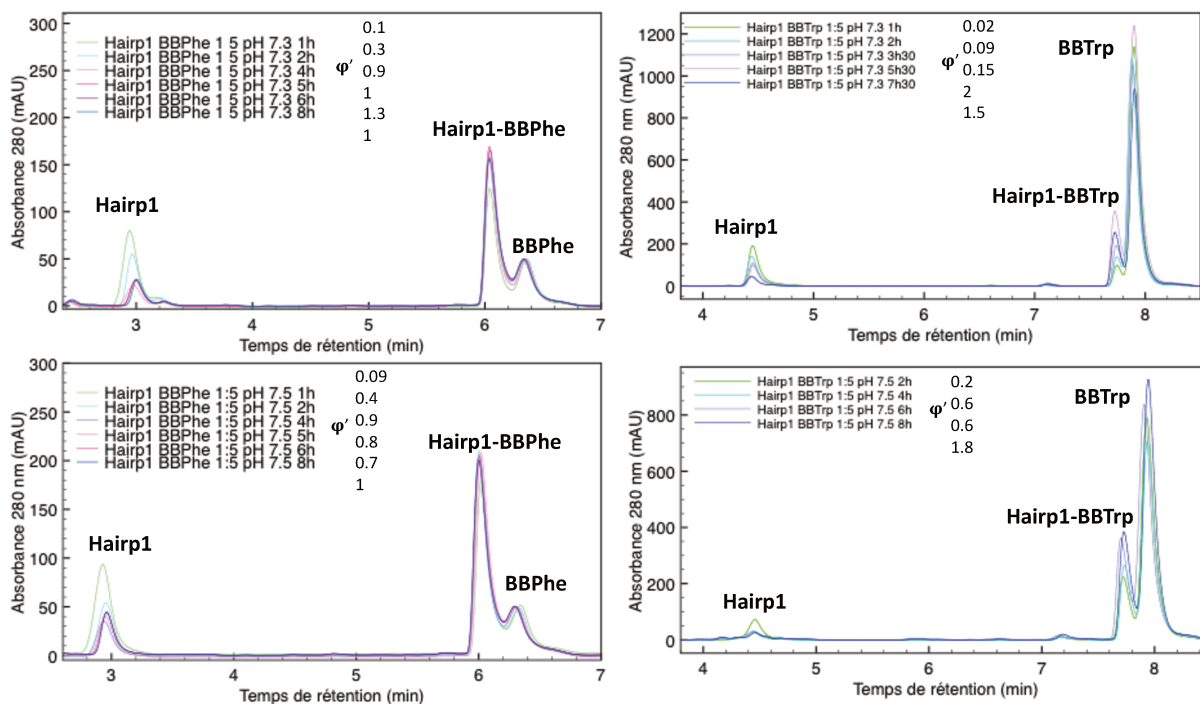


Fig. 63 : Chromatogrammes HPLC obtenus pour la réaction entre **Hairp1** et **BBPhe** (colonne de gauche) ou **BBTrp** (colonne de droite) en ratio 1:5 à pH 7.3 en haut et 7.5 en bas.

Ces résultats montrent en effet que l'augmentation du pH permet d'accélérer sensiblement la réaction avec un équilibre atteint en 4h à pH 7.3 et en environ 2h à pH 7.5.

Nous avons également évalué le potentiel d'une catalyse pour notre système. Le domaine de la ligation chimique native emploie lui-aussi la réaction de transthioestérification dans l'eau sur des peptides afin de coupler des fragments peptidiques entre eux¹³⁸. Sommairement, un fragment présente un thioester en C-terminal tandis qu'un second présente une cystéine N-terminale. La NCL consiste à faire réagir le thiol de la cystéine sur le thioester par transthioestérification puis l'amine α libre de la cystéine réagit par transfert d'acyl irréversible sur le nouveau thioester donnant un amide constituant une nouvelle liaison peptidique reliant les deux fragments initiaux. Il a été montré que l'utilisation de l'imidazole permettait d'accélérer ce processus. Le mécanisme proposé dans la littérature passe par l'activation du thioester de départ par l'imidazole conduisant à un intermédiaire acyl-imidazolium activé plus réactif vis-à-vis du thiol de la cystéine (Fig. 64).

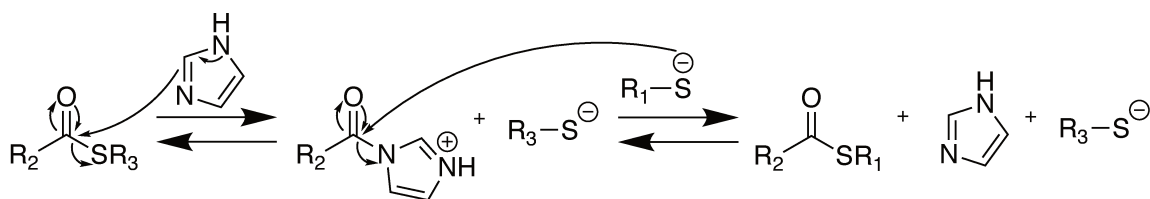


Fig. 64 : Mécanisme attendue de la transthioestérification catalysée à l'imidazole.

En s'inspirant de ces résultats, la réaction entre **Hairp1** (1 mM) et **BBPhe** ou **BBTrp** (5 équiv.) à pH 7 a été réalisée également en présence d'imidazole (5 mol% par rapport au **BB**). Le suivi dans le temps a été conduit de manière identique (Fig. 65).

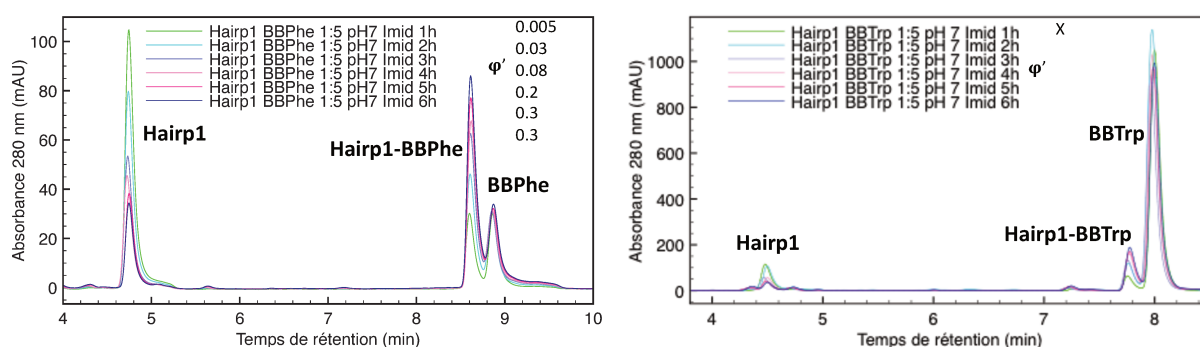


Fig. 65 : Chromatogrammes obtenus pour le suivi de la réaction entre **Hairp1** et **BBPhe** (à gauche) ou **BBTrp** (à droite) en présence de 5 mol% d'imidazole par rapport au **BB**.

Cependant, les résultats montrent que la vitesse de la réaction n'est pas significativement différente puisqu'ici l'équilibre n'est pas atteint avant 6h. Aucun effet sur la constante d'équilibre apparente K' n'est observé.

2) Influence de la nature des blocs sur la constante d'équilibre

Lors du développement d'un système combinatoire dynamique que l'on souhaite à terme soumettre à un criblage, il est important de confirmer que la librairie générée ne soit pas biaisée, c'est à dire qu'un des membres la composant soit thermodynamiquement plus stable et donc majoritairement représenté ou bien qu'un des membres se forme à une vitesse beaucoup plus grande ou plus faible que le reste de la librairie ou autrement dit, il faut que les membres de la librairie soient iso-énergétiques et iso-cinétiques. En effet, le biais cinétique peut entraîner une interprétation faussée des résultats. Le biais thermodynamique quant à lui conduit à une surreprésentation des membres de la librairie les plus stables énergétiquement. Cela peut

s'avérer problématique lors de l'étape de criblage de la librairie pour deux raisons : premièrement si la cible sélectionne un des membres de la librairie les moins représentés, l'amplification associée peut ne pas s'avérer suffisante pour le rendre majoritaire devant le membre initialement plus stable, conduisant donc à une interprétation faussée. Deuxièmement, si la librairie est biaisée énergétiquement en faveur d'un membre de manière trop importante, il est possible que l'énergie nécessaire pour déplacer l'équilibre lors du criblage soit trop importante pour sortir de ce puit énergétique lié au biais.

Afin de vérifier qu'un tel biais n'existe pas dans le système développé, nous avons évalué les constantes d'équilibre de transthioestérification entre les peptides à une cystéine et les différents **BB** synthétisés représentatifs des chaînes latérales des aminoacides et possédant donc des chaînes variées en terme de longueur, ramification et nature des fonctions chimiques. Pour ce faire, les peptides **Hairp1-pNA**, **Hairp1**, **Hairp1-DAPA** et **Hairp1-Naphthoyl-strand** (1 mM) ont été traités chacun par un unique **BB** (1, 5 ou 10 équiv.) dans les conditions décrites précédemment. La constante d'équilibre apparente a été déterminée à partir de l'intégration des pics en HPLC (Tableau 3).

Tableau 3 : Paramètres réactionnels des réactions entre plateformes à une cystéine et un **BB** ainsi que les constantes d'équilibres apparentes calculées à partir des intégration des pics HPLC.

Entrée	Sca	BB	Ratio	pH	Temps	K'	
9	Hairp1	BBPhe	1:10	7,5	2-3h	0,8	
11		BBTrp			2-4h	1	
12		BBLeu			6h	0,4	
13		BBVal			6h	non atteint à 6h	
14		BBLys			6h	0,2	
15		BBTyr			6h	0,7	
16		BBAla			6h	0,4	
17		BBArg			6h	n.d. (chevauchement)	
18		BBAsp			6h	n.d. (chevauchement)	
19	BBGlu	6h	0,4				
20	Hairp1-DAPA-Ala	BBPhe	1:5	7,5	3h	0,5	
21		BBAla				0,4	
22		BBArg				0,1	
23		BBTrp				1,5	
24	Hairp1-Naphthoyl-strand	BBLeu	1:5	7	7h	0,5	
25		BBPhe				5h	1
26		BBLys				7h	0,4
27	Hairp1-pNA	BBAib	1:1	7	24h	non atteint à 24h	
28		BBPhg				0,08	
29		BBTrp				1	
30		BBLys				0,8	
31		BBhAla				1,8	
32		BBVal				0,08	
33		BBLeu				1,2	
34		BBTle				pas de réaction	

Dans la plupart des cas, ces valeurs de constantes d'équilibre sont satisfaisantes pour un système combinatoire dynamique puisqu'elles sont aux alentours de 1, valeur qui traduit l'équilibre réparti équitablement entre les réactifs et les produits de la réaction. En effet, si la valeur était bien supérieure à 1, la réaction ne serait plus équilibrée mais s'approcherait d'une réaction totale et donc irréversible. De manière inverse, une valeur proche de zéro traduirait une réaction totalement déplacée vers les réactifs ce qui rendrait difficile la détection de phénomène d'amplification.

Une seconde observation intéressante est la faible variation de ces constantes en fonction du bloc considéré ou encore du *scaffold*. En effet, qu'il s'agisse d'une chaîne courte (**BBAla**), plus longue et ramifiée (**BBLeu**), portant une charge positive (**BBLys**) ou négative (**BBGlu**) ou encore un cycle aromatique relativement encombrant (**BBPhe**, **BBTyr**, **BBTrp**), la valeur de la constante d'équilibre apparente ne varie que peu traduisant une réactivité similaire, propriété nécessaire pour la combinatoire dynamique comme discuté plus haut.

En revanche, pour les **BB** possédant un fort encombrement stérique **BBVal**, **BBAib**, **BBPhg** ou encore **BBTle**, la réaction est moins favorable. Deux scénarios peuvent être distingués. Dans certains cas, l'encombrement stérique ralentit la réaction et l'équilibre n'est atteint qu'à un temps beaucoup plus long. A titre d'exemple, la réaction de **Hairp1-pNA** avec le **BBAib** n'a pas atteint son état d'équilibre après 24h et c'est seulement après 6 jours que celui-ci est atteint avec néanmoins une constante K' d'environ 1, similaire aux autres **BB**. Dans d'autres cas, l'équilibre favorise les réactifs plutôt que les produits du fait d'une certaine instabilité de ces derniers. Par exemple, la réaction du **Hairp1-pNA** avec le **BBPhg** atteint l'équilibre dans le même temps que les autres blocs mais présente une constante d'équilibre apparente plus faible. Ceci a été mis en évidence par analyse du mélange après 6 jours : aucune évolution n'est observée. Il en va de même pour le **BBVal**. Le cas extrême est celui du **BBTle** puisque dans son cas, la formation du produit attendu n'est jamais observée. Dans les deux cas, les produits issus des réactions avec ces blocs ne sont donc pas isoénergétiques ou isocinétiques par rapport aux produits des réactions avec les autres **BB**. Leur utilisation dans des DCL peut donc engendrer des biais.

3) Librairie issue d'une plateforme à un site fonctionnalisable

Des petites librairies ont été générées à partir des peptides comportant un site fonctionnalisable et plusieurs **BB** simultanément dans les conditions décrites précédemment (Fig. 66).

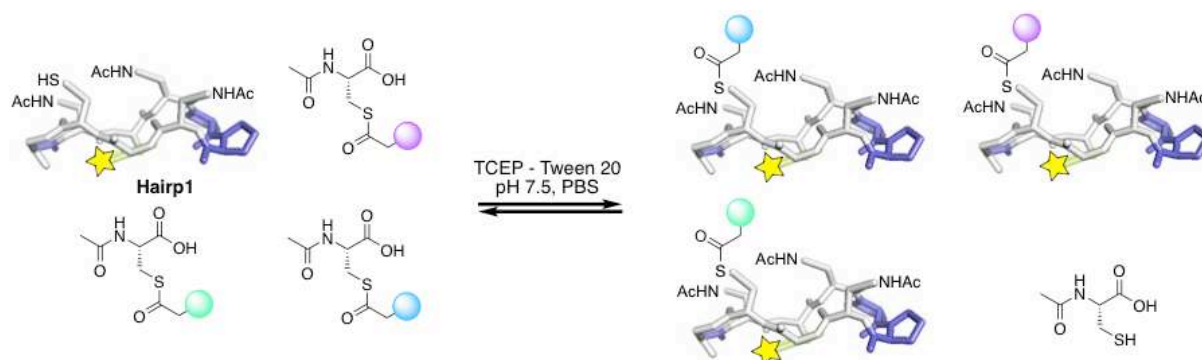


Fig. 66 : Équilibre réactionnel d'une DCL formée par la réaction d'une plateforme à une cystéine et 3 **BB**.

Trois DCL ont ainsi été générées à partir de **Hairp1-pNA** et **Hairp1**. Les résultats sont regroupés dans le tableau 4 et les chromatogrammes HPLC représentés en figure 67.

Tableaux 4 : Paramètres et proportions à l'équilibre des espèces peptidiques composant les DCL issues de plateformes à une cystéine et plusieurs blocs.

Sca	BB	Ratio	pH	Temps (h)	Proportion à l'équilibre
Hairp1-pNA	BBAib	1 eq de chaque BB	7	8	Hairp1-pNA-BBAib 7%
	BBPhg				Hairp1-pNA-BBPhg 5%
	BBTrp				Hairp1-pNA-BBTrp 34%
	BBLys				Hairp1-pNA-BBLys 24%
					Hairp1-pNA 29%
Sca	BB	Ratio	pH	Temps (h)	Proportion à l'équilibre
Hairp1-pNA	BBAib	1 eq de chaque BB	7	8	Hairp1-pNA-BBhAla + Hairp1-pNA-BBAib 27%
	BBhAla				
	BBVal				Hairp1-pNA-BBVal 3%
	BBLeu				Hairp1-pNA-BBLeu 22%
					Hairp1-pNA-BBTle 0%
					Hairp1-pNA 45%
Sca	BB	Ratio	pH	Temps (h)	Proportion à l'équilibre
Hairp1	BBPhe	5 eq de chaque BB	7,5	24	Hairp1-BBPhe 46%
	BBAla				Hairp1-BBAla 35%
	BBLys				Hairp1-BBLys 6%
					Hairp1 10%

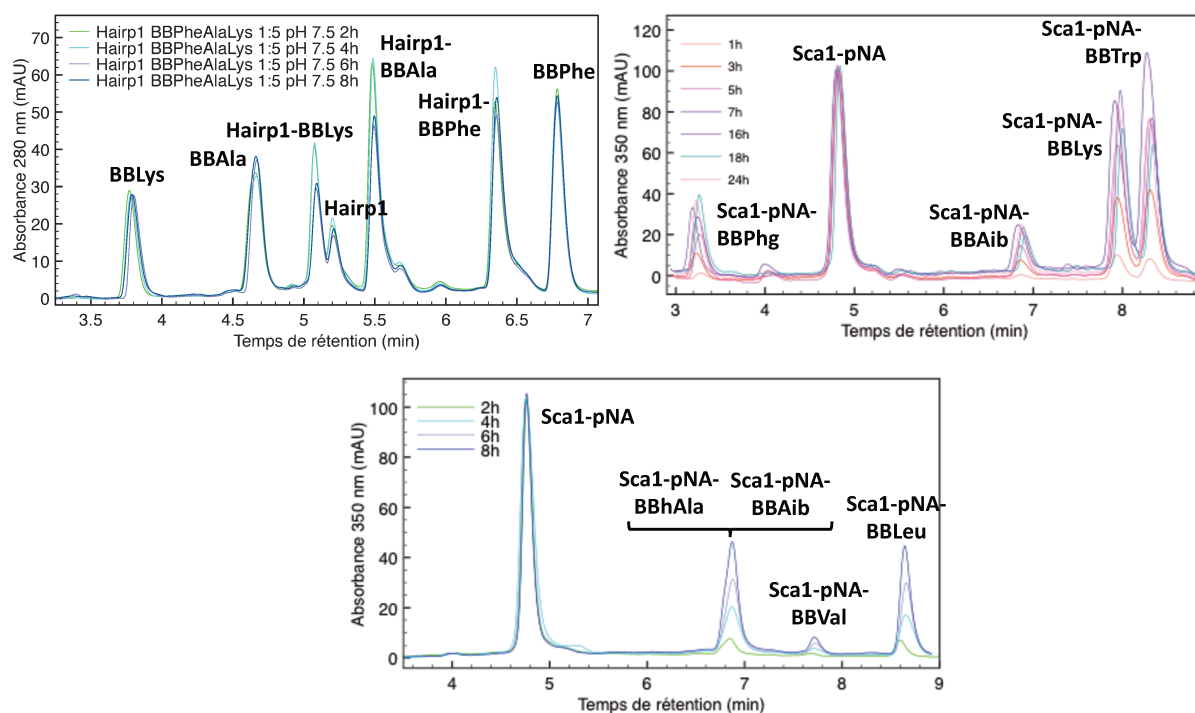


Fig. 67 : Chromatogrammes obtenus pour la librairie combinatoire dynamique générée à partir de **Hairp1** et des **BBPhe**, **BBLys** et **BBAla** (en haut à gauche), **Sca1-pNA** et des **BBPhg**, **BBAib**, **BBLys** et **BBTrp** (en haut à droite) et **Sca1-pNA** et des **BBhAla**, **BBAib**, **BBVal**, **BBTle** et **BBLeu** (en dessous).

Les chromatogrammes obtenus à l'équilibre (8h) montrent la formation des nouveaux adduits **Sca-BB** attendus dans des proportions qui sont en bonne adéquation avec leurs constantes d'équilibre respectives indiquant que la réaction entre un bloc et la plateforme n'a pas d'influence sur les réactions de cette même plateforme avec d'autres blocs.

4) Réversibilité du système

Une propriété qu'il est impératif de vérifier pour un système combinatoire dynamique est sa réversibilité et donc sa capacité à répondre à un stimulus extérieur. Afin de mettre en évidence expérimentalement la réversibilité de notre système, des expériences visant à étudier l'influence de l'ordre d'introduction des composants sur la composition du milieu réactionnel à l'équilibre ont été menées. Pour cela, des expériences ont été réalisées où cette fois l'introduction des différents blocs a été réalisée de manière séquentielle (Fig. 68). Les peptides **Hairp1** et **Hairp1-DAPA-Ala** (1 mM) ont ainsi été traités chacun en parallèle par 5 équiv. de **BBAla** (Exp1) et 5 équiv. de **BBPhe** (Exp2) durant 3 heures avant analyse. Le **BBLys** (5 équiv.) a alors été ajouté à chacun des milieux réactionnels et l'agitation a été maintenue durant 3h afin de permettre la rééquilibration du système. Après ajout de 5 équiv. du troisième bloc dans chacun des milieux réactionnels, respectivement **BBPhe** dans les Exp1 et **BBAla** dans les Exp2 et trois nouvelles heures de réaction, la composition de chacune des librairies finales est analysée par HPLC (Fig. 69).

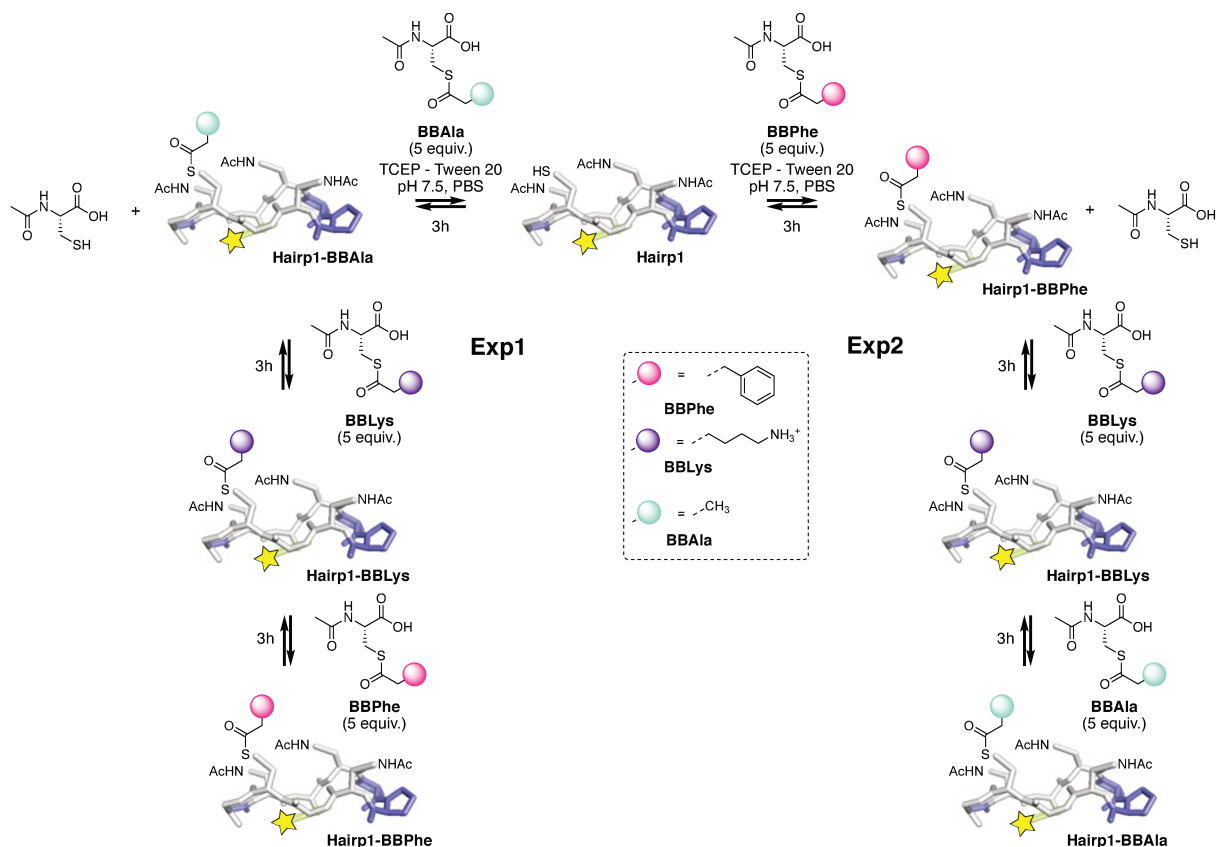


Fig. 68 : Équilibres réactionnels successifs permettant d'accéder à la librairie issue de **Hairp1** et des **BBPhe**, **BBLys** et **BBAIa** par ajouts consécutifs de ces derniers dans l'ordre de l'Exp1 **BBAIa**, **BBLys** puis **BBPhe** à gauche et dans l'ordre de l'Exp2 **BBPhe**, **BBLys** puis **BBAIa** à droite. Le schéma est transposable au peptide **Hairp1-DAPA-Ala**.

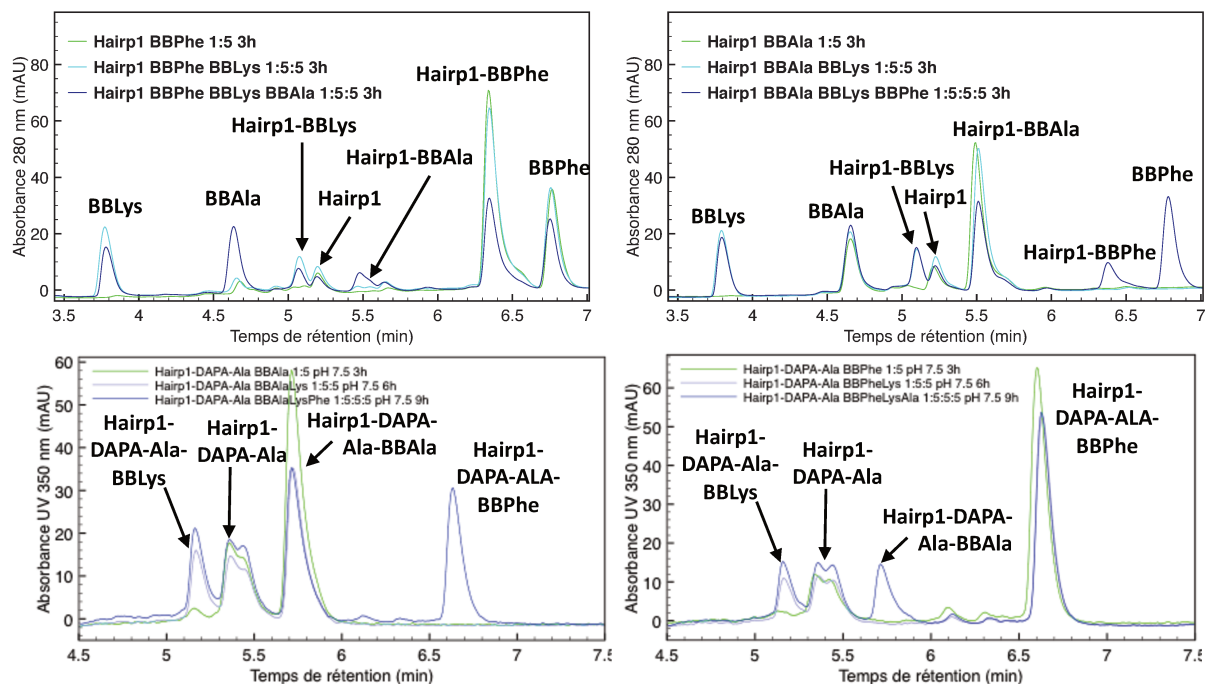


Fig. 69 : Chromatogrammes HPLC du suivi des réactions précédentes à 3 heures après l'ajout d'un des blocs. Exp1 à gauche, Exp2 à droite, pour **Hairp1** en haut et **Hairp1-DAPA-Ala** en bas

Les chromatogrammes obtenus révèlent bien la présence des trois adduits **Hairp1-BBPhe**, **Hairp1-BBAla** et **Hairp1-BBLys** dans les deux expériences, avec une diminution de la proportion du premier puis du second adduit formé respectivement lors du second puis du troisième ajout, au profit des nouvelles espèces formées, ce qui confirme la réversibilité de notre système combinatoire dynamique. Il est à noter cependant que les proportions des différentes espèces composant la DCL ne sont pas les mêmes dans les deux librairies obtenues par ajouts successifs et également dans la librairie où tous les blocs sont ajoutés simultanément, avec les adduits formés lors du premier ajout surreprésentés par rapport à leur constante d'équilibre. Cela pourrait être dû au fait que le système n'a pas eu le temps à chaque fois d'atteindre le nouvel état d'équilibre. En effet, après la première équilibration la concentration en peptide thiol dans le milieu réactionnel se trouve fortement réduite, et de ce fait la vitesse d'attaque de ce thiol sur le second puis le troisième **BB** ralentit corrélativement. Cette constatation pourrait avoir des conséquences quant au choix de la méthode qu'il conviendra d'adopter lors de criblage de cibles biologiques. En effet, deux approches peuvent être envisagées : l'établissement dans un premier temps d'un équilibre qui est ensuite réorganisé lors de l'ajout de la cible ou la méthode adaptative où l'équilibration se fait directement en présence de la cible. Dans notre cas, la rééquilibration semblant être un processus plus lent que l'équilibration, il est probable que l'emploi de librairies adaptatives soit la plus adéquate.

5) Combinatoire dynamique sur un système à deux cystéines

Afin d'étudier l'influence que peut avoir une première chaîne latérale thioester sur l'acylation d'un thiol voisin, un peptide contenant cette fois deux cystéines, en position 1 et 8, **Hairp2-Naphthoyl**, a ensuite été testé dans les expériences de DCC. Nous avons choisi ici de disposer les deux sites fonctionnalisables face à face sur chacun des brins de l'épingle à cheveux β de manière à avoir le maximum de proximité entre les deux thiols, les rendant ainsi le plus sensible l'un vis à vis de l'autre et également les plus promptes à former un pont disulfure. Le peptide **Hairp2-Naphthoyl** (1 mM) a ainsi été traité par un seul **BB**, **BBLeu**, **BBPhe** ou **BBLys** (5 équiv.), dans un tampon phosphate à pH 7. (Fig. 70 et Tableau 5).

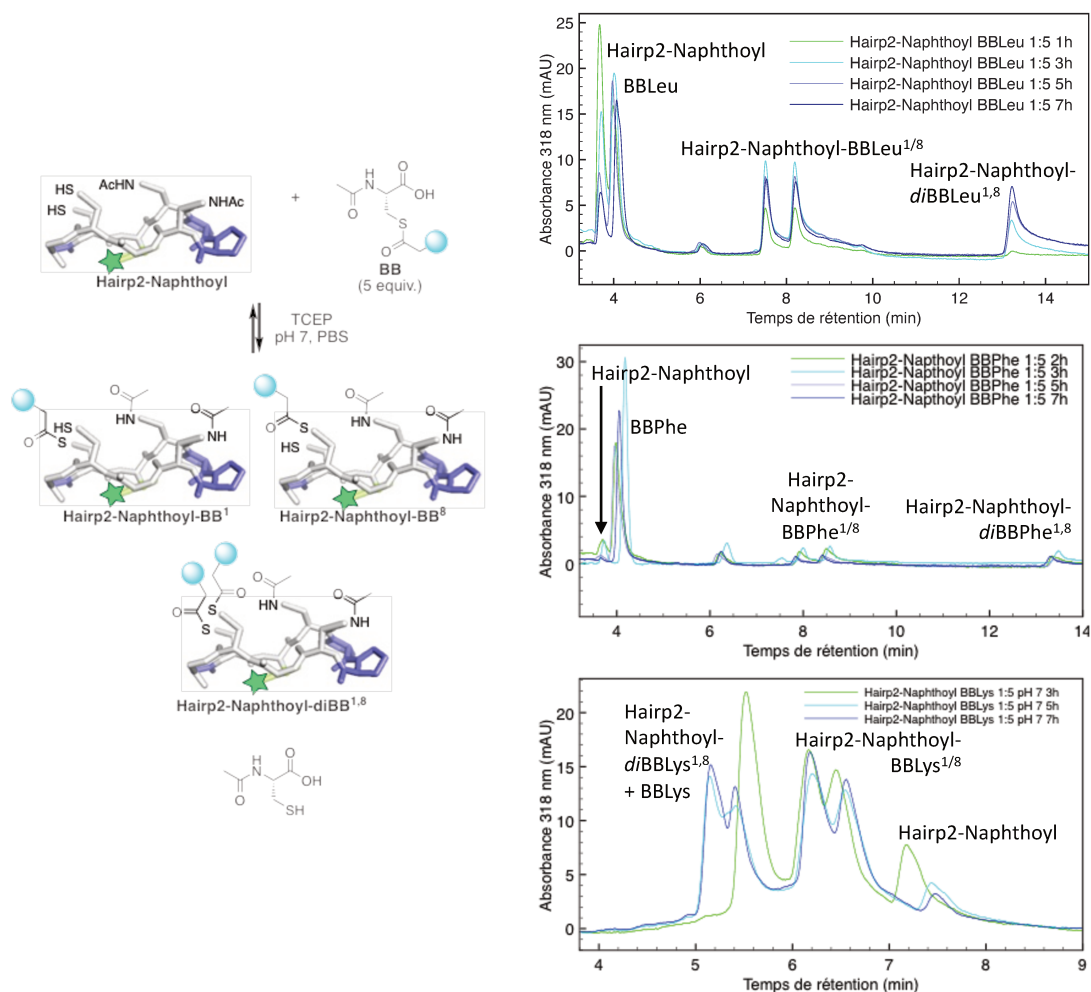


Fig. 70 : Équilibre réactionnel obtenu pour les bibliothèques dynamiques générées à partir de **Hairp2-Naphthoyl** et d'un **BB** et chromatogrammes HPLC du suivi des bibliothèques obtenues (**BBLeu** (en haut), **BBPhe** (au milieu) et **BBLys** (en bas)).

Tableau 5 : Paramètres et proportions à l'équilibre des espèces peptidiques composant les DCL issues de **Hairp2-Naphthoyl** et **BBLeu**, **BBLys** ou **BBPhe**.

Sca	BB	Ratio	pH	Temps (h)	Proportion à l'équilibre
Hairp2-Naphthoyl	BBLeu	1:5	7	7h	Di 38%
					Mono 25% et 23%
					Sca 14%
	BBLys				Di 31%
					Mono 62%
					Sca 7%
	BBPhe				Di 46%
					Mono 23% et 23%
					Sca 7%

Le suivi de la réaction par HPLC révèle bien la formation dans chaque cas des trois peptides fonctionnalisés attendus : deux ne portant qu'une seule chaîne acyle **Hairp2-monoBB**¹ et **Hairp2-monoBB**⁸ et un qui en porte deux que l'on appelle **Hairp2-diBB**^{1,8}. Il est à noter que dans tous les cas les deux isomères monofonctionnalisés se forment à la même vitesse et dans des proportions identiques. Cela indique qu'il n'y pas de différence de réactivité entre les deux cystéines situées en face à face sur la plateforme. De plus, on peut observer que la formation du produit difonctionnalisé se fait bien à partir des produits monofonctionnalisés de façon identique pour ces deux produits confirmant une réactivité similaire pour les deux sites mais aussi une absence d'influence de la fonctionnalisation d'un site sur la réactivité de l'autre site. A l'équilibre, les différents composés non fonctionnalisés, mono et difonctionnalisé sont tous présents et dans des proportions qui sont en accord avec les ratios utilisés.

Cela confirme que le système est sous contrôle thermodynamique et non cinétique et qu'en travaillant avec un plus fort excès de bloc il sera possible d'augmenter la proportion de produit possédant un nombre plus important de fonctionnalités et sa vitesse de formation également. Notons que la résolution obtenue pour les deux composés isomères possédant une unique chaîne acyle est relativement bonne bien que des recouvrements soient obtenus pour les produits portant la chaîne du **BBLys**.

Enfin, il a été constaté lors de ces expériences que dans le cas de la fonctionnalisation par le **BBPhe** relativement hydrophobe, des phénomènes d'agrégation de faible ampleur des produits fonctionnalisés et en particulier difonctionnalisés avaient lieu. Ceci pourrait être à l'origine de la forte proportion de produit diacylé obtenu dans ce cas. Du fait de ce type de phénomène, un résidu cationique supplémentaire a été introduit par la suite dans la séquence peptidique comme discuté dans le chapitre 1.

6) Combinatoire dynamique sur un systèmes à quatre cystéines : **Hairp4-Lys**

a) *Exemple d'une librairie simple à un bloc*

Afin d'illustrer la complexité importante que de tels systèmes peuvent générer très rapidement et avant de s'intéresser à un cas d'application particulière, une librairie générée à partir de **Hairp4-Lys** et un unique bloc **BBLeu** (10 équiv.) est présentée ici (Fig. 71).

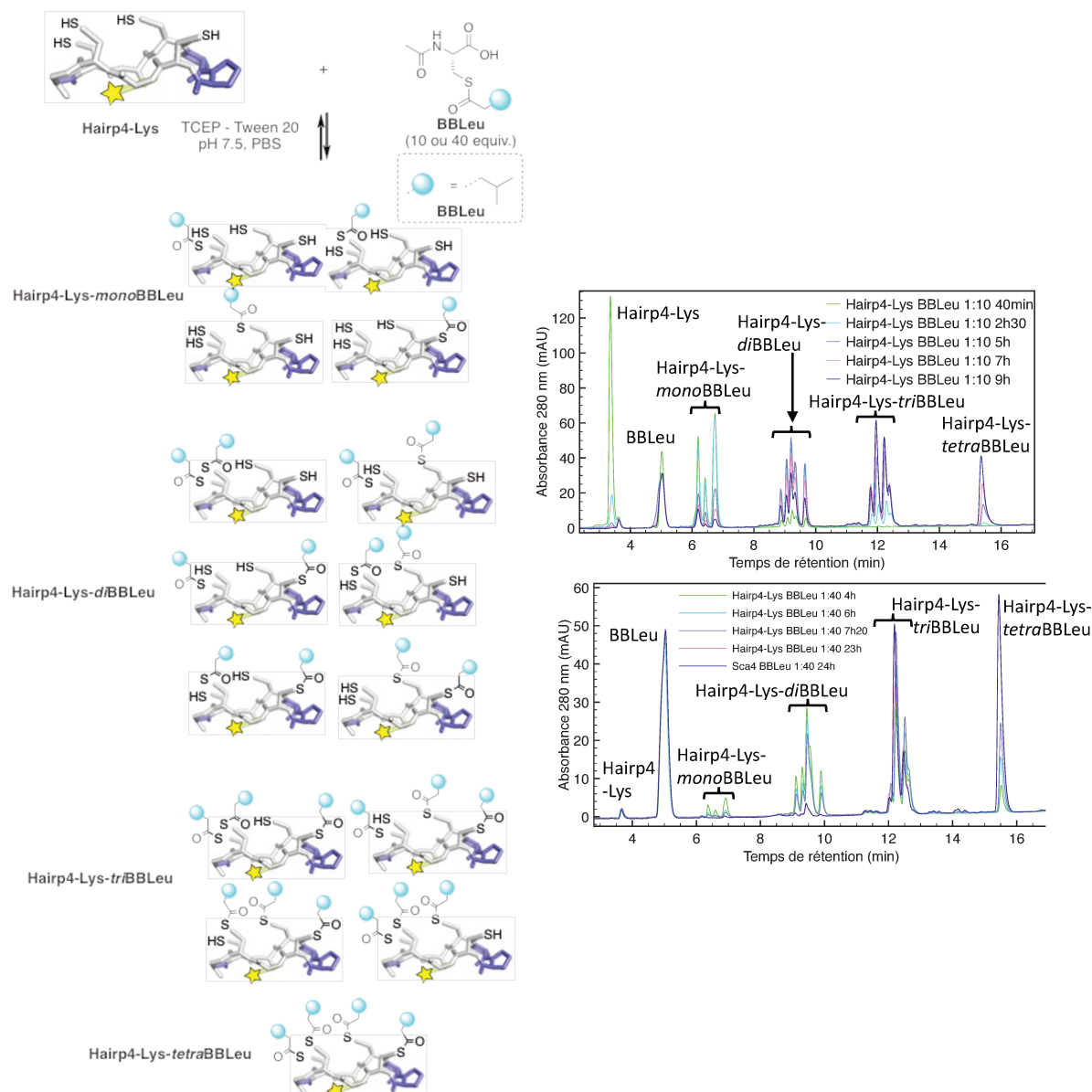


Fig. 71 : Équilibre réactionnel obtenu pour la génération d'une librairie dynamique à partir de **Hairp4-Lys** et de **BBLeu**, à droite et chromatogrammes HPLC (à droite) du suivi de l'évolution de cette librairie lors de l'utilisation de 10 équivalents de **BBLeu** (au-dessus) ou 40 équivalents de **BBLeu** (en dessous).

Le suivi de la réaction par HPLC, dans les premières heures de la réaction, révèle la formation majoritaire des espèces mono-acylés sur une seule des 4 fonctions thiols sous la forme de 3 pics, et 5 signaux correspondant aux 6 espèces di-fonctionnalisées. Au cours du temps, la formation des espèces poly-fonctionnalisées (tri et tétra) devient prédominantes avec 4 pics observés pour les 4 espèces tri-fonctionnalisées et 1 seul pour l'unique composé tétra-fonctionnalisée. Dans les deux cas, les différents pics ont été isolés par HPLC après 60 minutes de réaction et analysés par spectrométrie de masse MALDI-TOF permettant d'identifier les espèces constitutives de chaque librairie. À l'équilibre toutes les espèces mono, di, tri et tétra-acylées sont présentes dans des proportions variables, les produits tri-fonctionnalisés étant dans

ces conditions formés majoritairement. Il est à noter par ailleurs que le temps nécessaire pour atteindre l'équilibre ici est beaucoup plus long puisque 24h sont nécessaires pour qu'aucune évolution du mélange ne soit observable. Cela peut s'expliquer par l'augmentation significative d'encombrement stérique à la surface du peptide au fur et à mesure que les chaîne thioester sont introduites, ralentissant ainsi les réactions d'acylation suivantes. La cinétique pour atteindre la molécule tétra-fonctionnalisée et l'équilibre global s'en trouve diminuée. Il est à noter également ici que bien que la résolution des différents signaux ne soit pas parfaite, il existe peu de chevauchement des signaux dans le chromatogramme puisque 14 signaux distinct sont observés pour les 16 produits attendus et la plupart des composés ont pu être séparés par simple HPLC sur une colonne courte utilisée en routine. Cette séparation pourrait être largement améliorée par l'utilisation de conditions plus performantes. Il est possible de modifier les proportions des différentes espèces en jouant sur le ratio peptide/**BB**. Par exemple, avec un excès de **BB** de 40 équiv. (10 équiv. par fonction thiol) le produit tétrafonctionnalisé, qui *a priori* pourrait être le plus intéressant du point de vue de la reconnaissance moléculaire lors du criblage car possédant le plus de fonctionnalités susceptibles d'établir des interactions, devient majoritaire. De plus ces conditions permettent de s'affranchir des nombreux signaux correspondant aux produits mono et difonctionnalisés qui compliquent l'analyse et qui ne sont dans ce cas quasiment plus détectables devenant ainsi des espèces virtuelles de la DCL.

Chapitre 3 : Application à l'identification de nouvelles structures CPP

I. Contexte et choix de la cible

Les *cell-penetrating peptides* (CPP) aussi appelés peptides pénétrants, découverts à la fin des années 1980, sont des peptides de taille modeste (moins de 30 résidus), généralement cationiques, ayant la capacité de traverser les membranes cellulaires de manière non invasive, sans interaction préalable avec un récepteur membranaire¹³⁹. Ces peptides peuvent également être conjugués à des agents thérapeutiques de manière covalente ou non, et jouer le rôle de cargos pour un transport à l'intérieur des cellules. C'est pourquoi ils sont également qualifiés de peptides vecteurs. Depuis leur découverte, ils sont un objet d'étude important tant sur le plan fondamental pour comprendre les mécanismes de leur internalisation cellulaire qu'applicatif pour créer des conjugués capables de délivrer des molécules d'intérêt à l'intérieur des cellules. De très nombreuses séquences cationiques capables de pénétrer les cellules ont été reportées à ce jour, leurs efficacités ainsi que leurs stabilités *in vivo* restent cependant insuffisantes pour des applications thérapeutiques. Bien que les CPPs possèdent à ce jour des séquences très différentes, ils présentent néanmoins souvent des propriétés physico-chimiques communes. La plupart d'entre eux contiennent plusieurs résidus arginine et lysine et sont donc polycationiques à pH physiologique. Le rôle crucial de ces charges positives dans le processus d'internalisation a été mis en évidence. Des études sur des oligomères de différents acides aminés cationiques (arginine, lysine, histidine, ornithine) ont montré que les homopolymères d'arginine sont internalisés de manière plus efficace. Ensemble et avec des études complémentaires, ces résultats définissent le groupement guanidinium de l'arginine comme motif central, étant capable de former des liaisons hydrogènes bidentées de manière stable avec les anions de la membrane tels que les phosphates ou les sulfates. Par ailleurs, de nombreux CPP comportent des résidus hydrophobes, qui lorsqu'ils sont présents, jouent un rôle clé dans le processus d'internalisation. Enfin, bien que ce ne soit pas toujours le cas, de nombreux CPP présentent un caractère amphiphile. La plupart du temps les CPP flexibles en solution adoptent des structures secondaire bien définies, telles que des hélices ou des *β -hairpin*, en présence d'un environnement lipidique, permettant ainsi la ségrégation des chaînes latérales cationiques sur une face et des groupement hydrophobes sur une autre face. Cependant le lien entre ce caractère amphiphile et la perméabilité du peptide reste controversé. Ces travaux ont également montré que lors de l'interaction, la structuration en brin *β* de ces séquences était un moteur pour leur

auto-assemblage en agrégats stables à la surface de la membrane cellulaire avant d'être internalisés.

Par ailleurs, bien que le mécanisme de l'internalisation soit toujours le sujet de nombreuses études, il est communément admis que celui-ci procède par une première phase commune à toutes les voies décrites, qui correspond à la fixation du peptide sur des glycoaminoglycanes (GAG) à la surface des cellules et en particulier ceux présentant des groupements sulfates comme l'héparine notamment. En effet, les CPP cationiques sont capables d'interagir avec les GAG par interactions électrostatiques avec les groupements sulfates anioniques mais également par liaisons H et interactions hydrophobes avec les sucres de ces polymères.

Avec ces considérations, l'objectif de ce travail est de concevoir de nouvelles séquences peptidiques de taille réduite, capables de lier les GAG et d'entraîner leur internalisation à l'intérieur des cellules. Pour ce faire, des bibliothèques dynamiques possédant différentes combinaisons de résidus cationiques et hydrophobes à leur surface ont été générées. Nous avons choisi pour cela les chaînes latérales des résidus Arg et Trp, connues pour leur rôle clé dans les processus d'internalisation. Notre plateforme étant déjà structurée en épingle à cheveux β , il se peut que cette structuration préalable favorise l'interaction.

II. Bibliothèques à un bloc

Afin de faciliter la déconvolution de la bibliothèque générée à partir des deux **BB** simultanément, deux bibliothèques ont préalablement été étudiées en combinant individuellement soit le **BBTrp** soit le **BBArg** avec le peptide **Hairp4-Lys** avec 10 équivalents de **BB**, à pH 7.5 en présence de TCEP et de TWEEN 20 (Fig. 72).

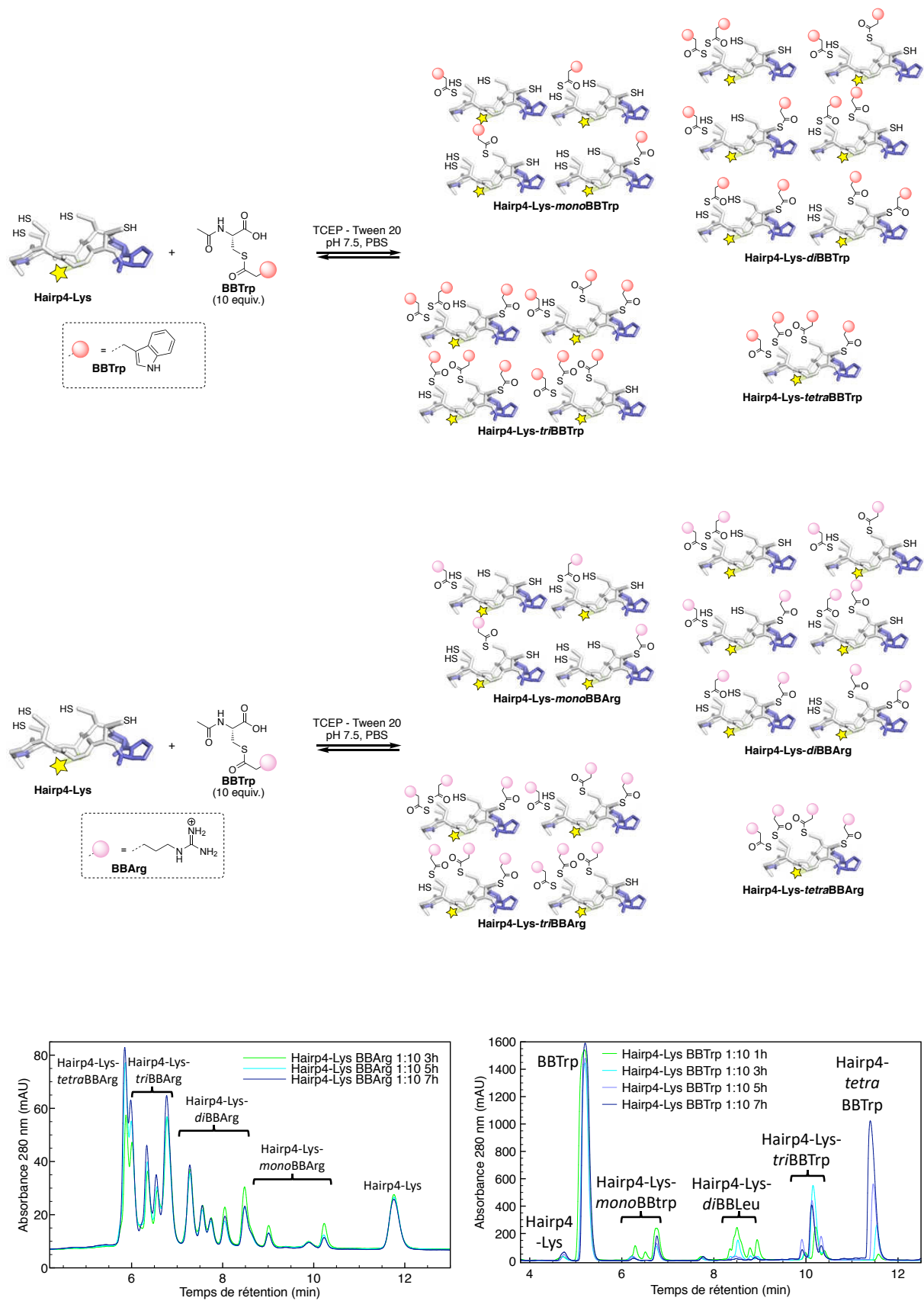


Fig. 72 : Équilibres réactionnels des bibliothèques dynamiques générées par BBTrp ou BBArg et Hairp4-Lys et chromatogrammes obtenus pour le suivi de leur évolution.

Le suivi des réactions par HPLC a permis de repérer les pics correspondants aux produits fonctionnalisés à différents degrés par un des deux blocs. Par exemple dans la série **BBTrp**, le chromatogramme montre 3 signaux pour les 4 espèces mono-fonctionnalisées, six pour chacune des espèces di-fonctionnalisées, 3 pour les 4 tri-fonctionnalisées, ainsi que le pic correspondant au peptide totalement acylé. Au cours du temps, l'intensité des signaux correspondant aux espèces tri- et tétra-fonctionnalisées augmente, l'espèce tétra-fonctionnalisée étant majoritairement formée à l'équilibre de la réaction, ce qui est en accord avec le large excès de **BB** utilisé ici. En revanche, de manière surprenante, un des pics correspondant à une ou deux espèces mono-fonctionnalisés est toujours présent en quantité significative à l'équilibre, ce qui suggérerait que l'une des espèces mono-fonctionnalisées serait particulièrement plus stable d'un point de thermodynamique.

En ce qui concerne les expériences menées avec le **BBArg**, là encore le nombre de pics observés ne correspond pas au nombre de produit attendus ce qui suggère de la même manière des chevauchements, notamment pour les espèces mono et di fonctionnalisées, avec 2 signaux pour les 4 peptides mono-acylés, 3 pour les 6 espèces di-fonctionnalisées, 4 pour chacune des espèces tri-acylés ainsi que le pic correspondant à l'espèce tétra-acylée. En revanche si les produits de la librairie générée à partir du **BBTrp** ont tous pu être formellement identifiés par MS-MALDI-ToF, ceux issus de la librairie obtenue à partir du **BBArg** semblent se fragmenter lors de leur analyse par MS-MALDI-ToF. En effet, le spectre de masse obtenu pour un pic correspondant à un degré de fonctionnalisation n présente un mélange de signaux aux rapports m/z correspondant aux produits de degré de fonctionnalisation n mais tous ceux de degrés inférieurs également. Cela laisse penser que ces produits tendent à se fragmenter en leurs analogues possédant moins de fonctionnalités **BBArg** lors de l'analyse en masse. Des conditions plus douces devront donc être envisagées dans le but de s'affranchir de ce problème.

III. Librairie mixte à deux blocs

La librairie mixte générée à partir des deux blocs **BBTrp** et **BBArg** a été obtenue par traitement similaire du **Hairp4-Lys**, 81 composés peuvent être en principe formés dont 16 espèces tétrafonctionnalisées attendues de manière majoritaire dans ces conditions (Fig. 73).

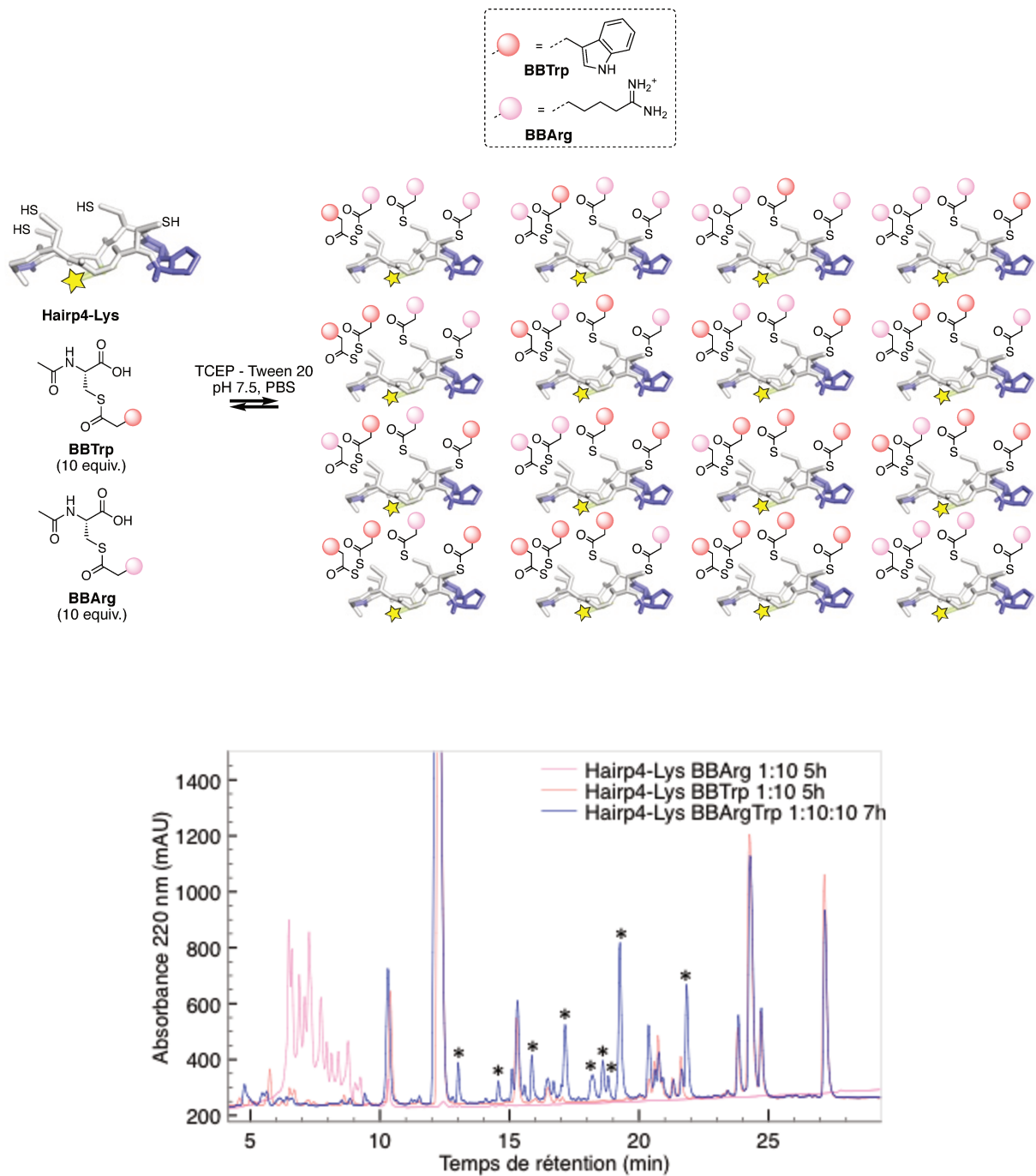


Fig. 73 : Equilibre réactionnel obtenu de la librairie dynamique générée par les **BBTrp**, **BBArg** et **Hairp4-Lys** (seules les structures tétra fonctionnalisées sont représentées) et chromatogrammes superposés des trois librairies analysées à 5h pour les librairies à bloc unique et 7h pour la librairie à deux blocs . Les pics spécifiques de la librairie mixte sont signalés par une *.

La superposition des chromatogrammes obtenus avec les 2 chromatogrammes des expériences précédentes nous a tout d'abord permis d'identifier tous les composés portant la même chaîne latérale sur les différentes positions. Les pics restants pourraient donc correspondre à de nouvelles espèces mixtes Arg/Trp. Certaines de ces espèces ont pu être caractérisées et ne correspondent exclusivement qu'à des espèces comportant plusieurs chaînes du **BBTrp**. Les

espèces comportant majoritairement les chaînes latérales de l'arginine subissent probablement le même phénomène de fragmentation évoqué précédemment.

IV. Discussion

Comme l'apparition de nouveaux pics par rapport aux expériences à un bloc le prouve, des produits portant des chaînes provenant de chacun des deux blocs se sont bien formés. Certains d'entre eux ont pu être caractérisés comme correspondant exclusivement à des espèces comportant plusieurs fonctionnalités **BBTrp**. Les espèces présentant majoritairement des chaînes issues du **BBArg** sont certainement sujettes au même comportement de fragmentation en spectrométrie de masse que les produits issus de la librairie générée à partir du **BBArg** uniquement. Ce problème rencontré uniquement avec ce **BB** pourrait être lié à la nucléophilie du groupement guanidinium qui pourrait réagir avec la fonction thioester. Le fait que ce phénomène ne soit pas observé avec le **BBLys** nous a conduit à l'hypothèse que c'est la longueur de la chaîne qui permettrait ce réarrangement dans un cas et pas dans l'autre. Pour le vérifier, la synthèse du **BBhArg** possédant une chaîne plus longue d'un carbone et du **BBOrn** possédant une chaîne plus courte que le **BBLys** ont été entreprises afin d'évaluer si celui-ci conduirait à des thioesters peptidiques plus stables en spectrométrie de masse. Les nouveaux **BB** seront évalués dans les réactions de DCC prochainement.

Un problème qui peut être également soulevé avec ces expériences vient de la détection des produits de la librairie. En effet, d'une part, plus le nombre de produits générés augmente plus le nombre de pics correspondant augmente et globalement le signal correspondant à la plateforme initiale sera réparti en autant de signaux qu'il y aura de produits. Cela peut rapidement conduire à une détection trop faible des produits de la librairie si sa taille est trop importante. D'autre part, une détection à 280 nm met en évidence le problème lié à l'observation du pic de **BBTrp**, très intense car en fort excès mais aussi le problème d'une absorption différente des produits de la librairie en fonction de leur degré de fonctionnalisation par le **BBTrp**. Ainsi les composés possédant moins de chaînes **BBTrp** sont moins faciles à détecter. Ces derniers obstacles ont quant à eux motivé la recherche d'une dernière sonde : le Dap(4-DAPA) afin de les contourner. Celle-ci reste encore à évaluer.

V. Conclusion et perspectives

Le travail accompli sur l'aspect combinatoire dynamique de ce projet a permis de motiver tous les autres aspects du projet allant de la synthèse de nouveaux blocs au raffinement d'une structure de plateforme le plus idéale possible. Il a aussi permis d'établir des conditions expérimentales pour la génération de bibliothèques dynamiques qui conduisent à la fois à des équilibres relativement rapides mais aussi suffisamment déplacés vers les produits désirés. Une approche a également été mise au point afin d'évaluer le potentiel d'un nouveau bloc avant de l'impliquer dans des expériences de DCC ce qui permettra dans le futur de sélectionner ces blocs pour s'assurer des conditions de DCC au plus proche de l'optimal.

La variété des bibliothèques qui ont pu être générées a par ailleurs permis d'exemplifier la tolérance de ce système tant vis à vis du type de bloc impliqué que de la structure variable de la plateforme utilisée. Des tentatives préliminaires de conception de bibliothèques intéressantes à cribler ont permis également de montrer le potentiel de cette approche pour l'identification de partenaires de molécules d'intérêts.

Les étapes suivantes seront d'évaluer les nouvelles structures récemment synthétisées afin de répondre aux derniers problèmes soulevés par les dernières expériences. Le **BBhArg** devra être testé comme alternative au **BBArg** afin de vérifier si celui-ci permet d'accéder à des thioesters peptidiques stables en spectrométrie de masse. S'il ne répondait pas à ces attentes, il serait possible d'aller vers d'autres analogues de ce bloc (à chaîne plus courte ou plus longue) ou bien encore d'aller vers des conditions d'analyse différentes en spectrométrie de masse (ESI par exemple). Notons tout de même que les produits issus de la réaction par le **BBArg** sont les seuls à ce jour dont la formation ait pu être observée mais dont les analyses en MALDI-ToF soient problématiques, cela montre tout de même un comportement tout à fait singulier de la fonctionnalité apportée par ce bloc.

L'utilisation d'un nouveau chromophore, le Dap(4-DAPA), pourra également participer à la résolution des problèmes liés à la détection par spectroscopie UV-Visible. Ce dernier a pu être incorporé dans différentes versions de la plateforme et devra être évalué lors d'expériences de génération de bibliothèques. En corollaire à l'utilisation de ce chromophore en absorption UV-Visible, l'utilisation d'une détection par fluorescence possible avec ce chromophore/fluorophore pourrait s'avérer avantageuse. Cela permettrait d'avoir un seuil de détection beaucoup plus bas et donc d'abaisser davantage la concentration des échantillons

analytiques pour les suivis de l'évolution des librairies tout en étant *a priori* aussi efficace, peu importe le ou les types de fonctionnalités introduites sur la plateforme.

Enfin, en ce qui concerne le criblage des DCL, celui-ci n'a pas été réalisé durant ce travail suites aux obstacles qui se sont dressés sur le chemin et qu'il était nécessaire de surmonter avant toute tentative. Il sera nécessaire de procéder à des essais méthodologiques lors de l'introduction d'une cible et des protocoles et conditions spécifiques à chacune d'elles seront certainement à mettre au point. Pour terminer, il sera également intéressant et nécessaire d'évaluer le comportement de nouvelles structures comme les *scaffolds* hélices α déjà synthétisés ici ou de nouvelles qui seront obtenues dans le futur afin de montrer les possibilités de transfert de ce système de DCC.

Chapitre 4 : Méthodologies de linéarisation de peptides cycliques β -*hairpin* pour leur analyse en spectrométrie de masse

I. Généralités et problématiques

Le développement d'une stratégie de chimie combinatoire dynamique s'accompagne obligatoirement de la mise en place d'une méthodologie analytique robuste permettant l'analyse de mélanges complexes qui constituent les librairies dynamiques. Celle-ci doit aussi être suffisamment souple pour permettre la caractérisation de molécules d'autant plus variées que les librairies générées contiendront un grand nombre de combinaisons. Il est à noter qu'en pratique il n'est pas nécessaire d'identifier et caractériser tous les membres d'une librairie dynamique¹⁴⁰. Ludlow *et al.* ont par exemple travaillé sur des librairies dynamiques reposant sur l'association de fragments poly thiols par formation de disulfure comportant plus de 9000 composés. Après criblage de leurs librairies par l'éphédrine, ils ont pu identifier par LC-MS le composé majoritairement amplifié par le criblage parmi tous ceux potentiellement présents dans la librairie. La philosophie de la DCC ne repose donc pas sur la caractérisation formelle de tous les potentiels candidats au criblage mais davantage sur la capacité de pouvoir identifier ceux que le criblage révèle comme affins. Le problème est qu'en augmentant le nombre de combinaisons possibles que ce soit en termes de nature ou d'agencement, la méthode analytique envisagée se doit d'être de plus en plus tolérante vis à vis de nouvelles structures moléculaires mais tout aussi efficace dans leur analyse.

Dans le cadre de la mise en place d'une stratégie de DCC pour l'assemblage de chaînes latérales d'acides aminés sur un peptide, nous avons choisi d'utiliser une approche de couplage « off-line » entre la HPLC et la MS-MALDI-ToF. La RP-HPLC est une méthode largement établie pour la séparation de peptides, le MALDI-ToF permettant quant à lui l'identification de macromolécules par leurs masses moléculaires.

La figure 74 présente une superposition de chromatogrammes obtenus lors du suivi de l'évolution d'une librairie issue de **Hairp4-Lys** et du **BBLeu** sur une colonne HPLC de routine (cf. chapitre 2). Cette méthode analytique présente une bonne séparation des familles de composés en fonction de leur degré de fonctionnalisation. Bien que la résolution des isomères au sein d'une même famille ne soit pas totale, l'emploi d'une colonne RP-HPLC plus performante permettrait de l'améliorer. L'identification de ces composés par MS-MALDI-ToF a permis d'identifier formellement ces différentes familles de composés. Les conditions

d'analyse du MALDI-ToF permettent d'une manière générale une bonne compatibilité des thioesters avec cette méthode analytique.

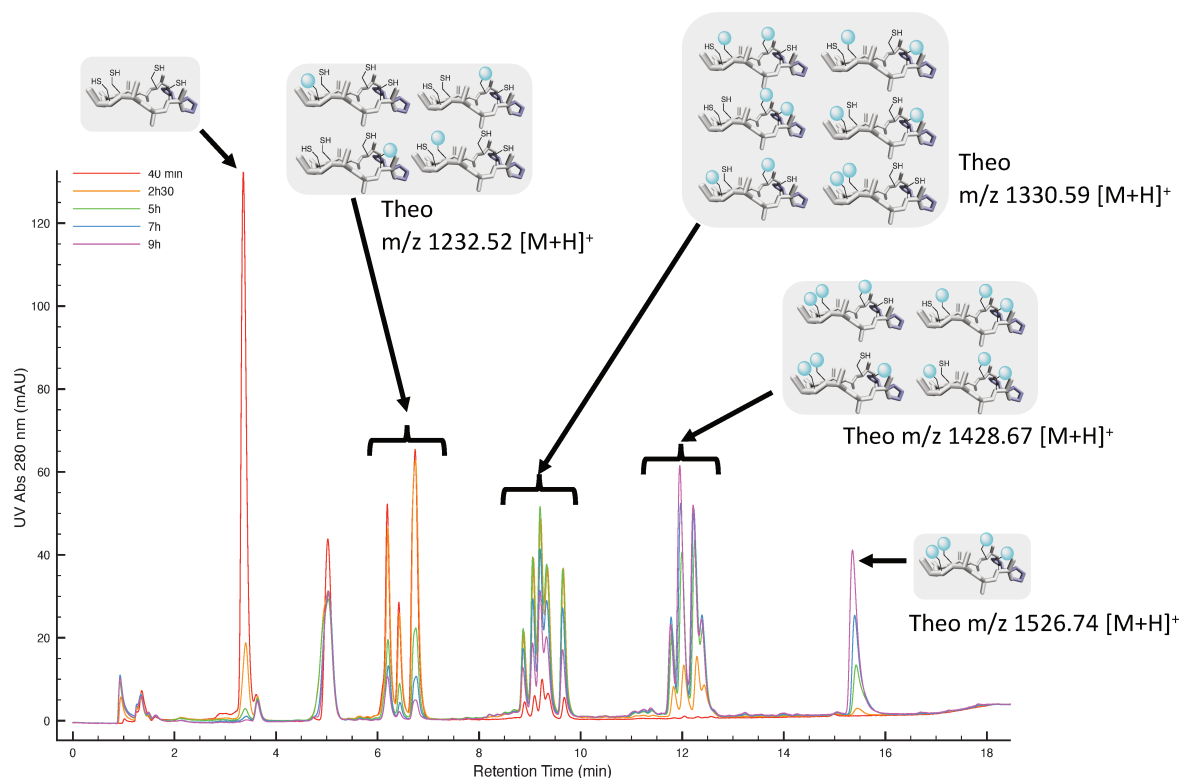


Fig. 74 : Chromatogrammes HPLC du suivi de la librairie issue de **Hairp4** et **BBLeu**.

Le véritable défi analytique imposé par notre approche de combinatoire dynamique est lié à la caractérisation d'isomères de position qui seront générés intrinsèquement avec notre stratégie. En effet, ces isomères, fonctionnalisés par les mêmes blocs mais avec différents agencements mèneront à des ions isobares et donc indistinguables en MS simple. Par exemple, dans la figure 74 si la MS permet d'identifier facilement le degré de substitution de chaque peptide, l'identification de la séquence exacte de chacun des peptides au sein d'une même famille mono-, di-, tri-, tétra-fonctionnalisée, n'est quant à elle pas possible. Ce problème est d'autant plus prépondérant lorsque des librairies impliquant plus d'un **BB** sont générées.

Afin de résoudre ce problème, l'emploi de la spectrométrie de masse en tandem ou MS/MS (ici, MALDI-ToF/ToF) est une solution de choix. En effet, après avoir été ionisées, les molécules à analyser peuvent être fragmentées en ions fragments qui permettent de remonter à la structure de la molécule de départ et donc à ou aux cystéines fonctionnalisées par un **BB**. Dans le cas des peptides et des protéines, cette approche MS/MS permet de remonter à la séquence primaire de la molécule. La fragmentation des peptides cycliques est plus complexe que celle des peptides linéaires correspondants et l'élucidation des spectres MS/MS est bien

plus difficile voire parfois impossible. En effet, la première fragmentation d'une des liaisons peptidiques conduit à la linéarisation du peptide. Théoriquement, cette linéarisation peut avoir lieu sur l'une ou l'autre des liaisons peptidiques, conduisant à une population d'ions isobares de différentes séquences, qui peuvent, si l'énergie interne est suffisante, conduire à une fragmentation supplémentaire de la chaîne peptidique en deux ions fragments. L'attribution des ions fragments à telle ou telle séquence sera rendue très difficile si on a un mélange préalable d'ions linéarisés au niveau de liaisons peptidiques différentes (Fig. 75). Afin de faciliter la caractérisation des bibliothèques de peptides cycliques, différentes stratégies conduisant à l'ouverture du cycle à une position contrôlée avant analyse MS et MS/MS ont été proposées dans la littérature : ouverture chimique au bromure de cyanogène (CNBr), spécifique des résidus méthionine^{141,142} ou des liaisons non peptidiques introduites dans les séquences et qui sont sensibles par exemple aux UV¹⁴³ ou à une aminolyse réalisée en phase gazeuse¹⁴⁴.

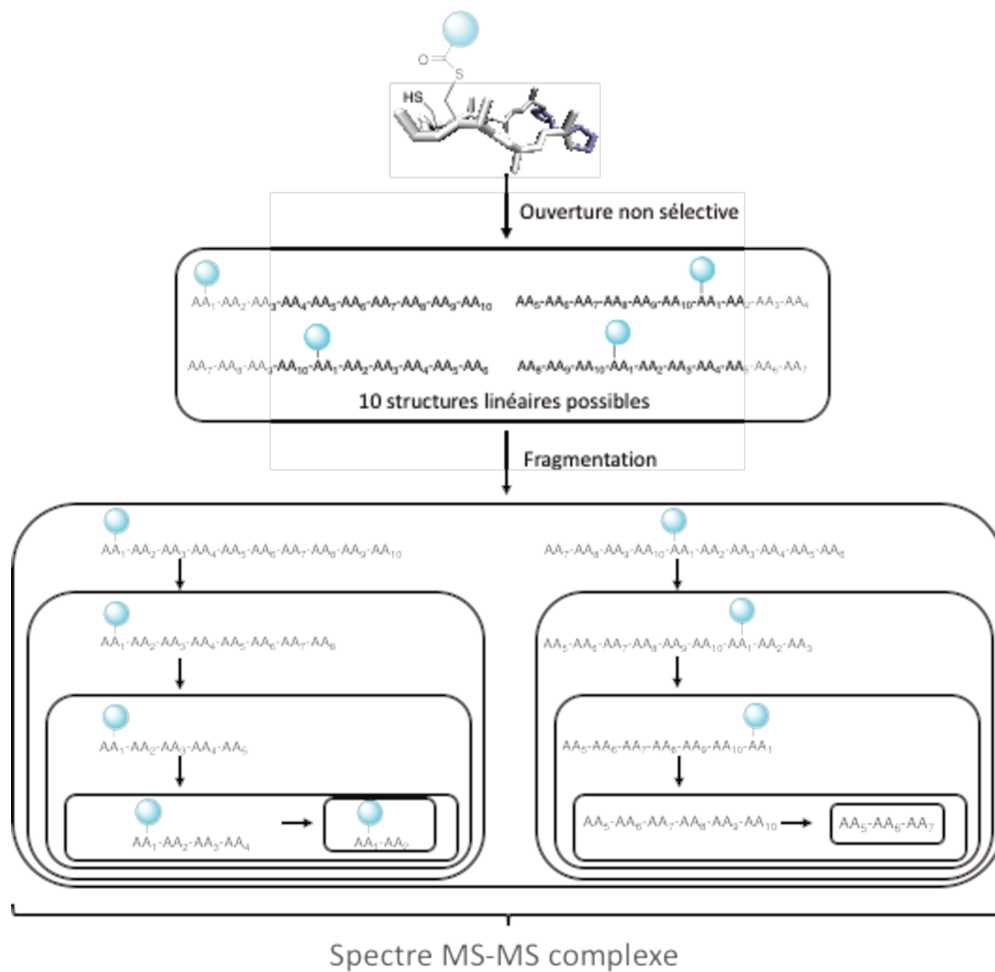


Fig. 75 : L'ouverture d'un peptide cyclique non sélective conduit à la formation d'une population de structures linéaires. Chacune de ces structures va suivre une voie de fragmentation qui vont se retrouver superposées dans le spectre rendant l'interprétation difficile. Dans ce schéma, seules 4 structures linéaires sur les 10 possibles à partir d'un décapeptide cyclique sont illustrées ainsi que des voies de fragmentations possibles à partir de deux d'entre elles, ceci dans un soucis de clarté.

A partir de nos peptides cycliques, plusieurs voies d'ouverture ont été explorées. La structure particulière de la plateforme peptidique imposant la présence de résidus Cys et Pro nous a conduit à évaluer la possibilité d'ouverture par ISD (*In-Source Decay*) qui semble facilitée en N-terminal de ces résidus, nous avons testé l'ouverture chimique au bromure de cyanogène et enfin l'ouverture par hydrolyse enzymatique.

II. Ouvertures sélectives du cycle peptidique

Afin de rappeler au lecteur la nomenclature des ions fragments issus de la fragmentation d'une séquence peptidique et de lui faciliter la lecture des paragraphes à venir, la figure 76 est à sa disposition¹⁴⁵. Elle reprend les bases de la nomenclature proposée par Biemann, la plus largement utilisée à l'heure actuelle.

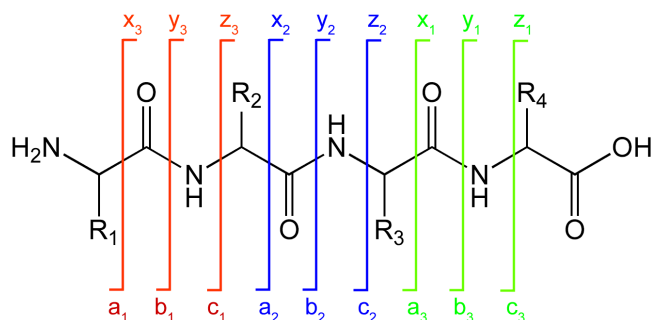


Fig. 76 : Nomenclature des ions fragments issus d'une séquence peptidique protonée selon le système de Biemann.

1) Ouverture par ISD

Une approche d'ouverture par ISD paraissait appropriée et élégante pour ouvrir et fragmenter les peptides linéaires directement dans l'appareil de spectrométrie de masse sans préparation d'échantillon préalable. L'ISD est un processus de fragmentation radicalaire rapide pouvant survenir juste après le tir laser du MALDI mais avant l'accélération des ions formés vers l'analyseur et le détecteur. Ce mécanisme est induit par certaines matrices MALDI comme le 2,5-DHB (acide 2,5-dihydroxybenzoïque) et le 1,5-DAN (1,5-Diaminonaphtalène) par exemple. L'ion moléculaire protoné $[M+H]^+$ généré par ablation laser peut être réduit par un radical hydrogène ($H\cdot$) produit à partir de la matrice. Cet ion moléculaire réduit peut ensuite conduire à différents ions fragments *via* différentes voies de fragmentations par rupture homolytique induite par le radical. Des ions de type *c*- et *z*- sont obtenus majoritairement mais d'autres ions peuvent également être obtenus (Fig. 77)¹⁴⁶. Les peptides cycliques ouverts par ISD dans la source d'ions peuvent être sélectionnés et fragmentés à nouveau par CID (*Collision Induced Dissociation*) par exemple, on parle alors de pseudo MS. Les mécanismes mis en jeu lors de la fragmentation ISD peuvent être comparés aux fragmentations ETD (*Electron-Transfer Dissociation*) ou ECD (*Electron-Capture Dissociation*). Ces méthodes de fragmentation douces

sont connues pour laisser intactes les modifications post-traductionnelles des peptides et protéines¹⁴⁷. Dans notre cas, on peut supposer que les liaisons thioesters obtenues après chimie combinatoire dynamique sont maintenues en phase gazeuse, par analogie avec des modifications post-traductionnelles.

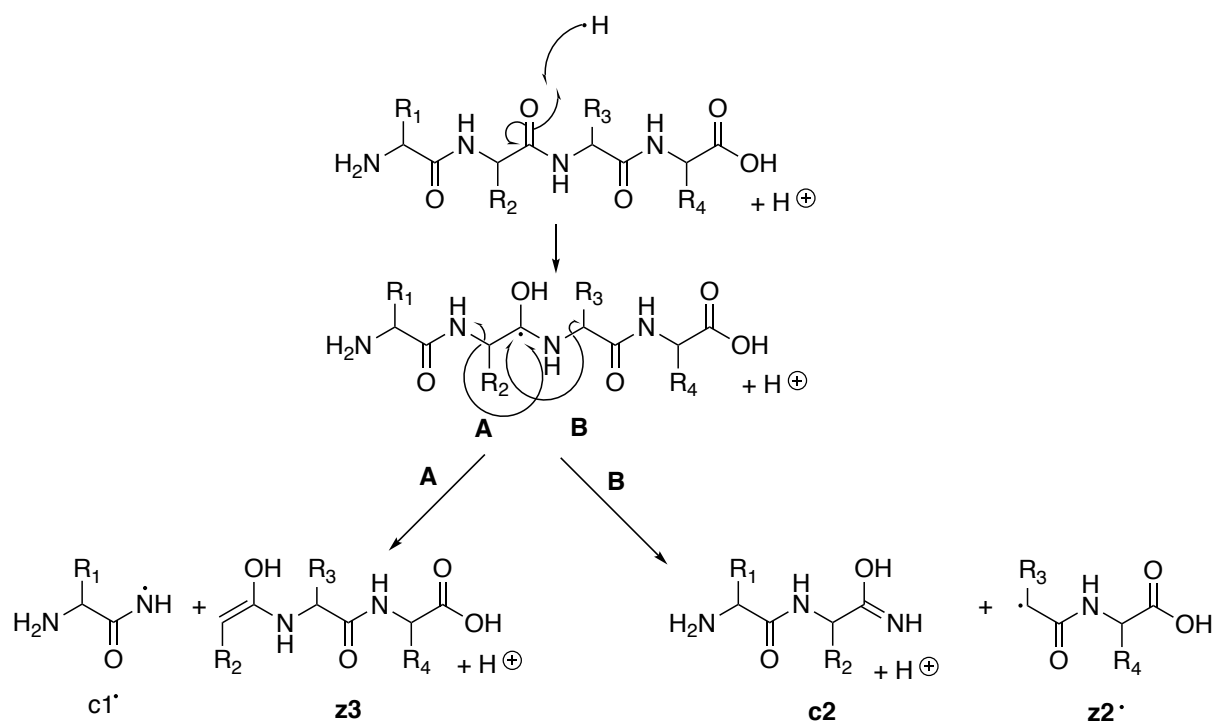


Fig. 77 : Mécanisme de fragmentation ISD à partir d'un peptide protoné en utilisant une matrice MALDI capable de donner un radical hydrogène H· (matrice réductrice de type DHB ou 1,5-DAN par exemple).

D'une manière générale, l'ISD peut se produire au niveau de toutes les liaisons peptidiques d'une même séquence. Néanmoins, des études visant à déterminer si certaines séquences sont plus promptes à se fragmenter par ISD que d'autres sont publiées et notamment il a été montré par Asakawa *et al.* que les enchaînements Xxx-Cys ou Xxx-Xxx-Pro sont plus favorables à ce type de fragmentation¹⁴⁸. La présence dans les peptides étudiés de telles séquences laisse à penser qu'une ouverture pourrait être favorisée par ISD et éventuellement qu'elle pourrait être plus ou moins sélective d'une liaison peptidique particulière soit au niveau d'une cystéine, soit au niveau d'une proline (Fig. 78).

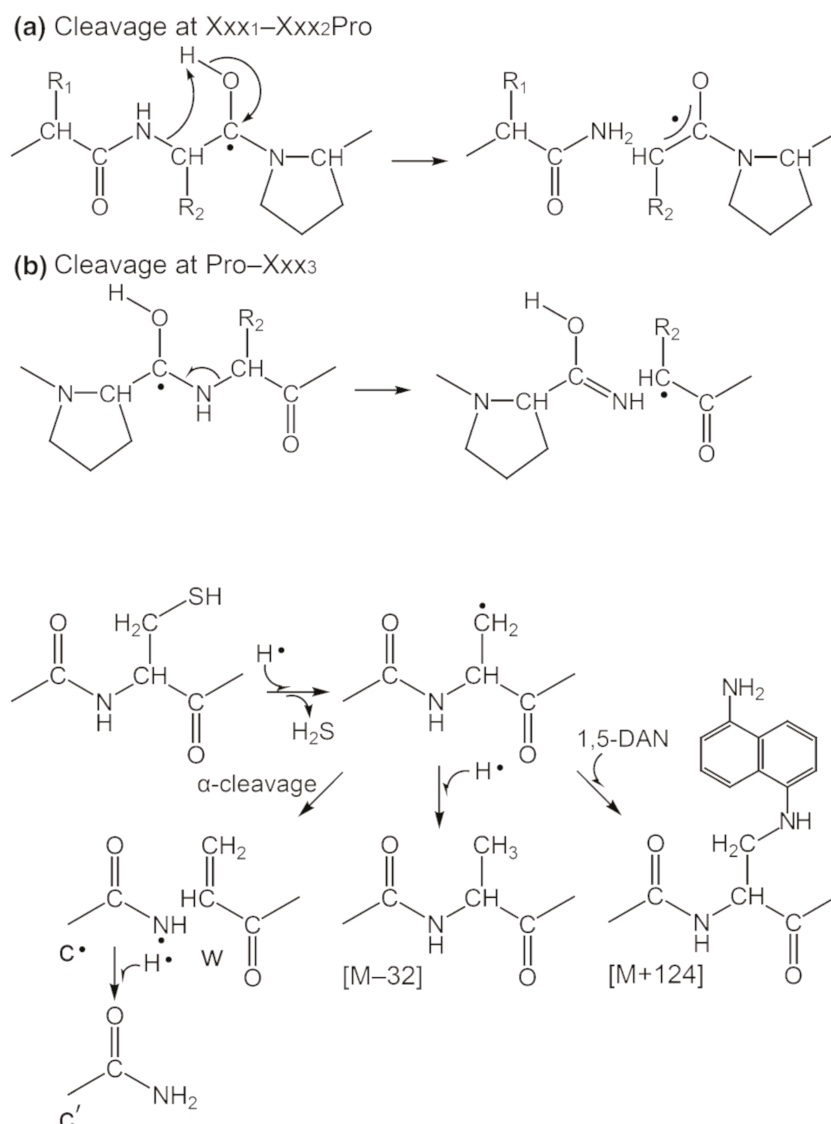


Fig. 78 : Mécanismes de fragmentation au niveau des résidus prolines et cystéines suggérés par Asakawa *et al.*

Les essais d'ouverture par ISD ont été réalisés sur les peptides **Hairp1-pNA** et **Hairp1-pNA-BBLeu** en présence de matrices décrites pour favoriser l'ISD : le 2,5-DHB ou le 1,5-DAN. Ces matrices sont capables de fournir facilement le radical H^\bullet à l'analyte cyclique et conduisent à l'ouverture des peptides cycliques par fragmentations dans la source MALDI, ce qui produit des ions linéaires, mono chargés, de forme $[M+3H]^+$. Dans le cas de peptides cycliques on obtient donc une version linéaire du cycle (Fig. 79).

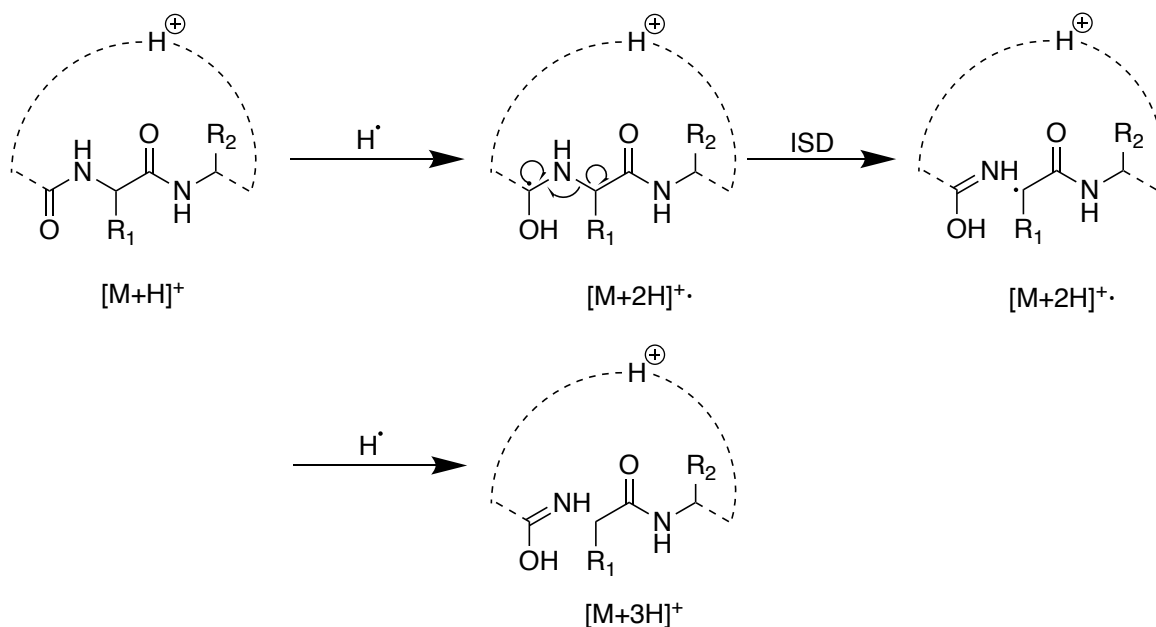
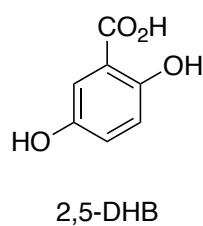
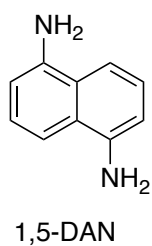
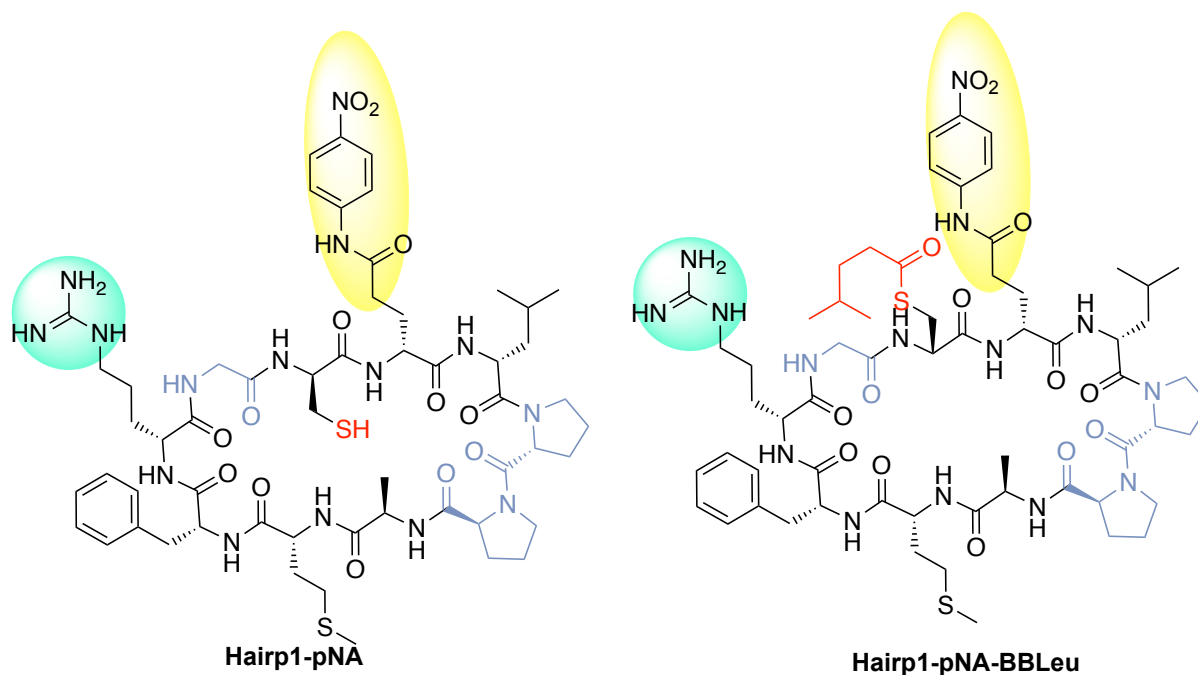


Fig. 79 : En haut, structures des peptides **Hairp1-pNA** et **Hairp1-pNA-BBLeu** ainsi que des matrices 1,5-DAN et DHB favorisant la fragmentation par ISD. En bas, mécanisme de fragmentation par ISD d'un peptide cyclique protoné conduisant à un ion moléculaire linéaire mono chargé de forme $[M+3H]^+$.

Le massif isotopique correspondant au peptide cyclique mono protoné obtenu dans le spectre MALDI-ToF est alors affecté selon l'efficacité du mécanisme de fragmentation ISD. Si on obtient 100 % d'ISD, le massif isotopique est déplacé de deux unités de masses vers les plus hauts m/z. Si le mécanisme d'ISD n'est pas complet, le massif est complexe et contient deux populations d'ions séparés de deux unités de masse, les ions de type $[M+H]^+$ et $[M+3H]^+$. En comparant la distribution isotopique de ce massif par rapport au massif isotopique théorique, simulé à partir de la formule brute de l'ion et à un contrôle en matrice non inductrice d'ISD, HCCA par exemple (acide α -cyano-4-hydroxycinnamique), on peut donc évaluer le rapport ion cyclique/ion linéaire. Les différents spectres obtenus pour le peptide **Hairp1-pNA** sont présentés dans la figure 80.

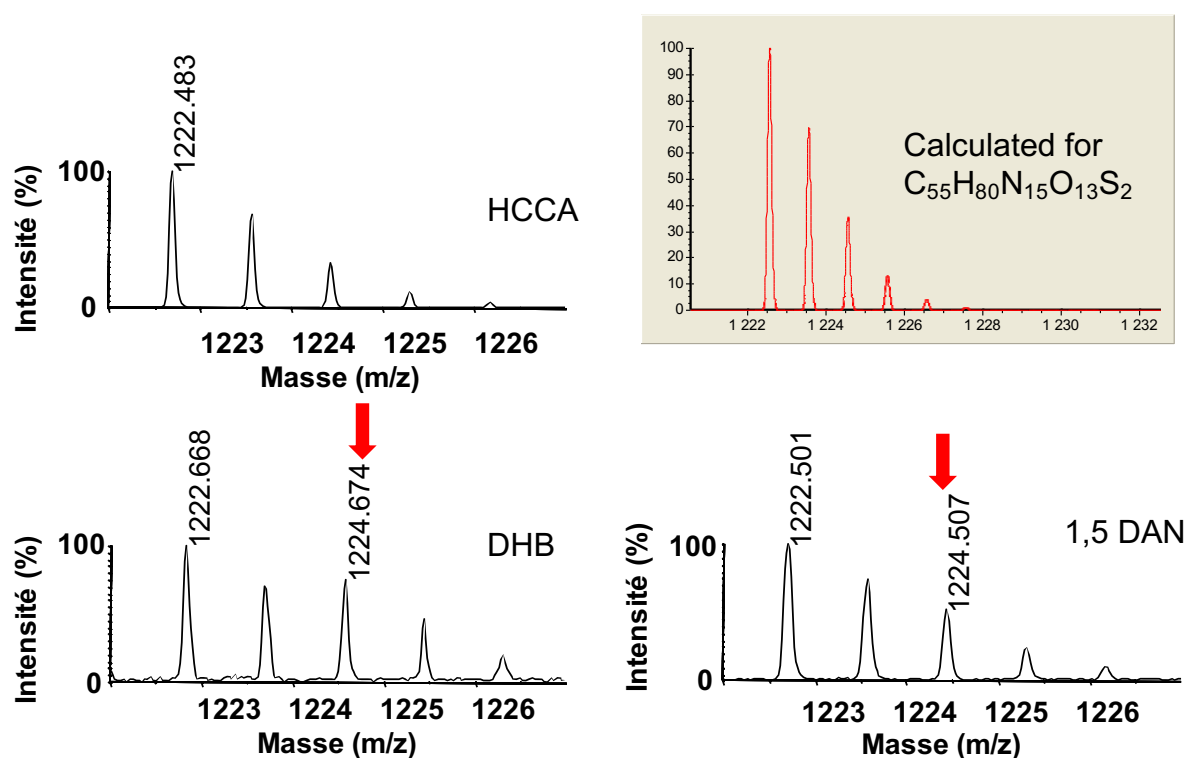


Fig. 80 : Spectres de masse MALDI-ToF du peptide **Hairp1-pNA** acquis en utilisant différentes matrices en mode réflecteur ions positifs. En haut à gauche, matrice HCCA. En bas à gauche, matrice DHB. En bas à droite, matrice 1,5-DAN. Le spectre en haut à droite est une simulation de la distribution isotopique attendue pour ce même peptide protoné.

La distribution isotopique obtenue en matrice HCCA correspond à celle attendue pour un ion monochargé de cette masse. Le massif isotopique observé est superposable au massif théorique calculé pour cette composition élémentaire. Les matrices 2,5-DHB ou 1,5-DAN, favorisant l'ISD conduisent à une modification de cette distribution isotopique : l'intensité relative du troisième isotope à m/z 1224 (indiqué par une flèche rouge sur la figure 80)

augmente. En effet, cet isotope est superposé avec l'espèce monochargée $[M+3H]^+$ issu de l'ISD. Dans les deux matrices, le taux de conversion a été évalué à environ 10 %. Des résultats similaires ont été obtenus pour l'analogue de ce peptide fonctionnalisé par le **BBLeu**. Comme anticipé, la fonction thioester est stable à ces conditions. En revanche la faible conversion du peptide cyclique en linéaire par ce mécanisme n'a pas permis d'isoler un ion parent pur correspondant au peptide linéarisé par ISD, ce qui rend l'interprétation des spectres MS/MS difficile. Différentes tentatives pour induire des mécanismes radicalaires et donc augmenter l'ISD, en ajoutant des initiateurs de radicaux comme l'APS (Ammonium PerSulfate) ont été réalisées mais sans succès. Ainsi, malgré une simplicité de mise en œuvre initiale, cette méthodologie n'a pas permis de mettre en place un protocole utilisable en routine pour l'analyse de nos DCL. Une approche d'ouverture du cycle préalable à l'analyse par spectrométrie de masse a alors été envisagée.

2) Ouverture chimique au bromure de cyanogène (CNBr)

Les séquences protéiques et peptidiques peuvent être découpées de manière sélective au niveau des résidus méthionine grâce à un traitement au CNBr. Cette approche est généralement utilisée en biochimie afin d'obtenir des peptides plus courts avant leur identification par des techniques de séparation classiques (électrophorèse SDS PAGE, HPLC) ou en spectrométrie de masse.

Le traitement d'une séquence contenant une méthionine par le CNBr conduit spécifiquement à l'hydrolyse de la liaison peptidique du côté C-terminale de la méthionine. Cette réaction est conduite à pH acide et procède selon le mécanisme décrit ci-dessous (Fig. 81).

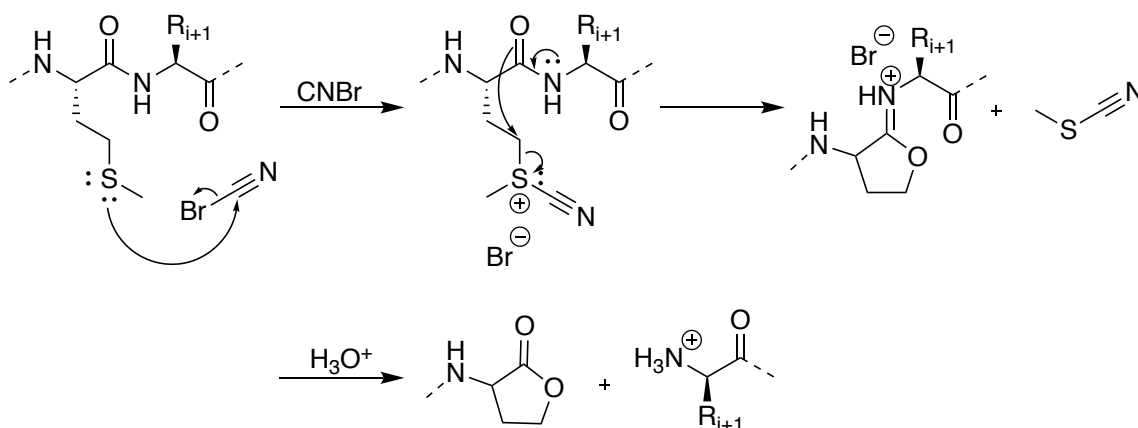


Fig. 81 : Mécanisme de coupure d'une séquence peptidique au niveau d'un résidu méthionine par action du bromure de cyanogène à pH acide.

Après addition du soufre de la méthionine sur le CNBr et élimination du brome, l'intermédiaire sulfonium subit une cyclisation conduisant à la libération de thiocyanate de méthyle. Cet intermédiaire cyclique subit ensuite une hydrolyse acide pour conduire à deux fragments peptidiques : un possédant une nouvelle extrémité C-terminale homosérine lactone à la place de la méthionine et un possédant une extrémité N-terminale libre constituée du résidu précédemment en C-terminal de la méthionine.

Dans notre cas, l'introduction d'une méthionine sur le *scaffold* **Hairp1-pNA** a permis d'explorer cette solution pour l'ouverture du cycle peptidique et d'évaluer la mise en place d'un protocole de routine reposant sur cette chimie (Fig. 82).

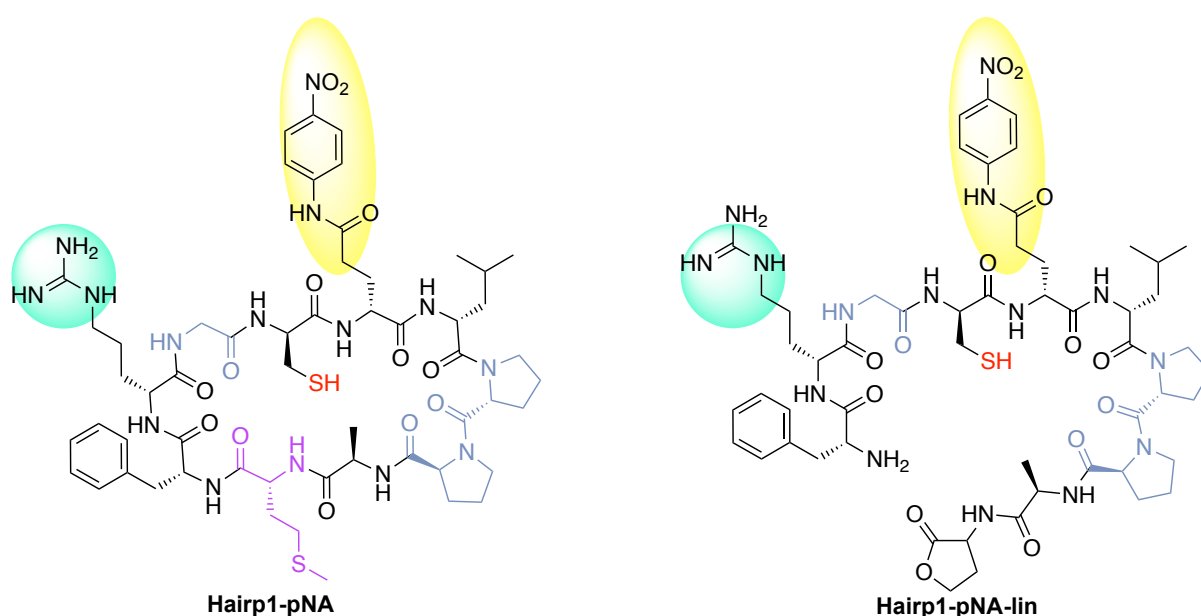


Fig. 82 : Structures de **Hairp1-pNA** et du composé linéaire correspondant après traitement au CNBr, **Hairp1-pNA-lin**. Le résidu méthionine est indiqué en violet.

Le traitement de **Hairp1-pNA** par un excès de CNBr à température ambiante dans une solution aqueuse à 70 % (v/v) en acide formique conduit bien à l'ouverture du cycle avec la formation du composé linéaire **Hairp1-pNA-lin** confirmé en MS-MALDI-ToF. Cependant des produits secondaires issus de l'oxydation de la méthionine en sulfoxyde majoritairement et sulfone dans une moindre mesure sont également observés. De plus, la forme sulfoxyde du peptide peut se fragmenter dans le tube de vol du MALDI-ToF (fragmentation PSD, Post-Source Decay) libérant de l'acide méthanesulfénique (CH₃SOH). Cela conduit à la présence de pics supplémentaires dans le spectre compliquant davantage l'analyse.

Cette oxydation difficilement contrôlable dans un protocole de routine associée aux faibles rendements de clivage et aux risques liés à l'utilisation du CNBr, nous ont conduit à explorer une nouvelle stratégie.

3) Ouverture enzymatique par hydrolyse trypsique

L'expertise en protéomique au sein du Laboratoire des Biomolécules a naturellement conduit à l'utilisation d'une approche courante dans ce domaine : l'emploi d'une enzyme pouvant hydrolyser de manière spécifique une liaison peptidique. Différentes endopeptidases avec des sélectivités variées peuvent être utilisées mais l'enzyme la plus couramment utilisée en protéomique est la trypsine. Elle hydrolyse spécifiquement les liaisons peptidiques dans lesquelles le groupement carboxyle des résidus basiques lysine et arginine est impliqué sauf s'il est suivi d'une proline dans la séquence. Le rendement est aussi abaissé dans le cas de la présence de doublet de résidus basiques ou de résidus acides côté N-ter ou C-ter du résidu basique. Cette spécificité ainsi que les rendements d'hydrolyse élevés généralement obtenus ont conduit à la choisir pour notre méthodologie.

Un avantage non négligeable de l'emploi de la trypsine est qu'elle crée un peptide linéaire dont le résidu C-terminal sera un acide aminé basique, aisément protonable donc ionisable en MS. Les peptides **Hairp1**, **Hairp1-pNA**, **Hairp1-Naphthoyl-loop**, **Hairp1-Naphthoyl-strand**, **Hairp2-Naphthoyl**, **Hairp4-Ala** et **Hairp4-Lys** contenant tous une arginine ont pu être testés pour la mise au point de cette méthode.

a) *Mise au point des conditions*

Comme présenté dans le chapitre sur la structure générale des *scaffolds*, mis à part les positions dédiées à la structuration par l'enchaînement D-Pro-L-Pro et la glycine ainsi que les positions dédiées à la fonctionnalisation occupées par les cystéines ou leurs analogues non réactifs, les trois positions restantes peuvent être choisies *a priori* pour introduire l'arginine qui servira de site de clivage par la trypsine, soit les positions 2 et 7 dans chacun des deux brins de l'épingle β soit en position 9 située dans la boucle refermant l'épingle du côté opposé aux prolines.

Initialement, la structure du *scaffold* qui avait été envisagée comportait une arginine en position 2 comme pour **Hairp1-Naphthoyl-loop** par exemple. Les premiers essais de digestion trypsique ont montré qu'aucune digestion n'est observée quelles que soient les conditions. Nous avons émis l'hypothèse que ce manque de réactivité était dû à la position de ce résidu, au sein du brin de l'épingle, fortement structuré et donc peu accessible à l'enzyme. Afin de tester cette hypothèse, un nouveau peptide, de même séquence, mais où les résidus 2 et 9 ont été inversés a été préparé, **Hairp1-Naphthoyl-strand**, contenant ainsi le site de clivage cette fois dans la boucle *a priori* plus flexible (Fig. 83).

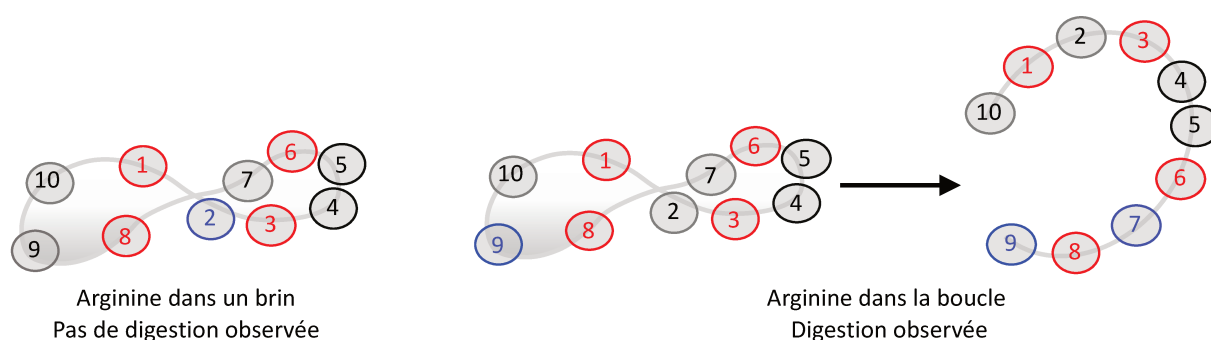


Fig. 83 : Schémas de la structure du *scaffold* où l'arginine est en position 2, dans le brin, à gauche, ou en position 9, dans la boucle, à droite. Les résidus sont symbolisés par des cercles, la nature du résidu est indiquée par la couleur du cercle : en bleu l'arginine, en rouge les cystéines et en noir les prolines. Les autres cercles gris correspondent aux autres résidus du *scaffold*.

En effet dans ce cas, lors du traitement du peptide par la trypsine, à pH 5 dans de l'eau bidistillée (compatibilité avec la stabilité des liaisons thioester), à 37°C, pendant 24h une digestion partielle est observée avec une conversion de 60 %.

Il est donc possible de parvenir à une ouverture sélective et satisfaisante grâce à l'emploi de la trypsine. Pour ce faire, nous avons optimisé nos protocoles en jouant sur différents paramètres. Nous avons augmenté le temps d'hydrolyse de quelques heures à 24h, la quantité d'enzyme, le pH tout en restant compatible avec les fonctions thioesters (pH 4 et pH 5). Nous avons aussi essayé de travailler dans des conditions réductrices en TCEP pour limiter la présence de ponts disulfures intramoléculaires qui pourraient limiter l'accessibilité de la trypsine.

Néanmoins les rendements d'hydrolyse enzymatique sont loin d'être quantitatifs. Cela peut être expliqué par le fait que nos conditions d'hydrolyse s'éloignent significativement des conditions optimales de fonctionnement de cette enzyme, notamment le pH. En effet, le pH optimal de fonctionnement de la trypsine est aux alentours de 8¹⁴⁹. Les thioesters utilisés dans notre approche de combinatoire dynamique n'étant pas stables à ce pH, il nous a fallu utiliser la trypsine à pH 5 mais dans ce cas, son activité est fortement diminuée.

De plus, les peptides fonctionnalisés issus de la chimie combinatoire dynamique peuvent s'avérer très hydrophobes et leur faible solubilité dans l'eau nous a aussi conduit à utiliser la trypsine dans un mélange à 50 % en acétonitrile. Des tests ont été réalisés et la trypsine fonctionne correctement dans ces conditions de solvant organique. Pour certaines séquences peptidiques particulièrement insolubles, nous avons également inclus une étape de dénaturation

des peptides par traitement à l'urée 8M puis dilution à 2M lors de la digestion enzymatique dans l'optique que leur possible déstructuration dans ces conditions permettrait une meilleure hydrolyse.

Au fil de l'analyse des nombreux peptides fonctionnalisés générés lors de ce travail, nous avons tout de même pu observer une tendance de comportement vis à vis de la digestion enzymatique. Tout d'abord plus le nombre de cystéine versus Dap(Ac) augmente, plus les peptides sont sensibles à l'hydrolyse trypsique. Ceci pourrait être dû à la présence des fonctions amides des résidus Dap(Ac) et leur caractère donneur de liaisons H qui peut perturber la reconnaissance moléculaire avec la trypsine. Les peptides ne comportant que des cystéines représentent donc de bien meilleurs modèles pour la mise au point de ces protocoles d'analyse. D'autre part, plus les peptides portent un nombre important de thioesters, plus l'efficacité de l'hydrolyse enzymatique est importante.

Pour illustrer ce propos, la figure 84 montre une fraction HPLC contenant trois versions du peptide **Hairp2-Naphthoyl** : non fonctionnalisé, mono et di-fonctionnalisé, digérés par la trypsine à 37°C, pH 5, pendant 24h. Les peptides linéarisés sont observés à une masse plus élevée de 18 unités de masse correspondant à l'ajout d'une molécule d'eau lors de la réaction d'hydrolyse enzymatique catalysée par la trypsine au niveau de la liaison peptidique en C-terminal de l'arginine.

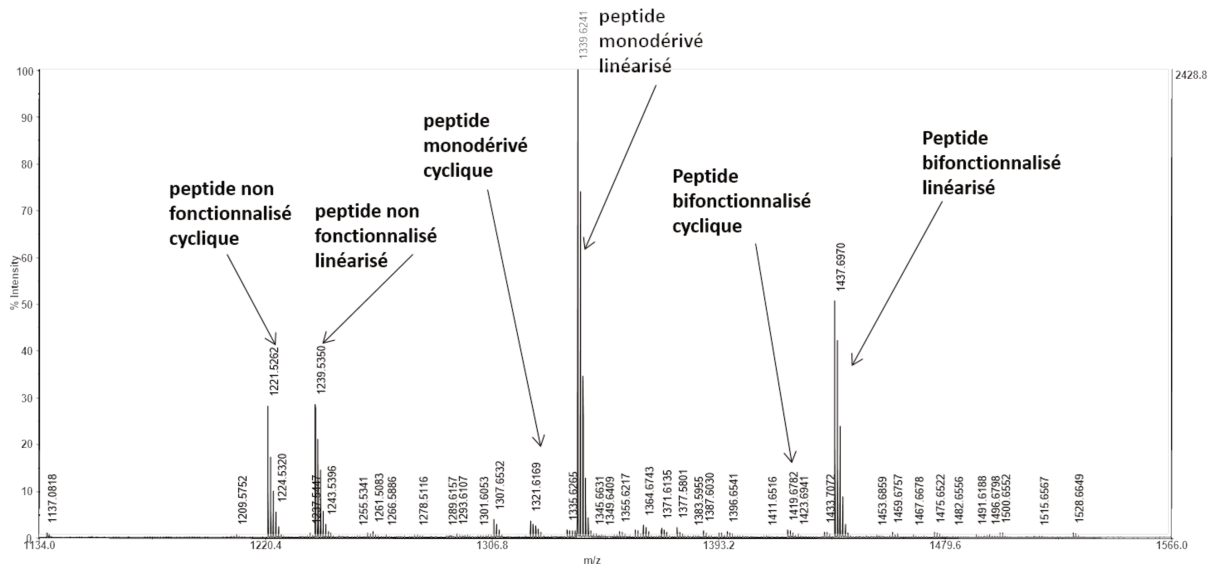


Fig. 84 : Spectre MALDI-TOF d'un mélange contenant **Hairp2-Naphthoyl** ainsi que les peptides correspondants mono et di fonctionnalisés par le **BBLeu**, après digestion trypsique.

On observe bien que le rendement d'hydrolyse par la trypsine augmente avec le nombre de chaînes thioesters en passant de 50 % pour le peptide non fonctionnalisé à quasiment 100 %

pour le diacylé. Bien que les rendements ne soient pas toujours aussi élevés, cette tendance est générale, quelle que soit la séquence peptidique : les peptides *tetra*-fonctionnalisés sont hydrolysés de manière plus efficace que les peptides *tri*-fonctionnalisés eux-mêmes davantage hydrolysés que les peptides *di*-fonctionnalisés eux-mêmes davantage hydrolysés que les peptides *mono*-fonctionnalisés eux-mêmes davantage hydrolysés que les peptides ne portant que des fonctions thiols libres. Ceci pourrait être dû au fait qu'en augmentant le nombre de chaînes latérales thioesters à la surface du peptide, l'encombrement stérique devient important, rendant le cycle peptidique de plus en plus tendu et contraint. L'hydrolyse par l'enzyme permettant la linéarisation et donc le relâchement de ces tensions et contraintes serait d'autant plus favorisée que le peptide porterait un grand nombre de fonctions thioesters.

Dans tous les cas, l'efficacité de digestion reste toujours inférieure à ce que l'on pourrait attendre pour une utilisation plus habituelle de la trypsine. Cela dénote une certaine résistance de ces peptides vis-à-vis de la dégradation protéolytique, ce qui est un point crucial lors de l'application de peptide pour des visées thérapeutiques. Ces observations appuient donc notre démarche de développer ces plateformes peptidiques pour de telles applications. Dans l'optique d'employer cette stratégie de combinatoire dynamique pour du criblage avec des molécules biologiques, il est raisonnable de penser que les espèces les plus affines seront probablement celles portant un maximum de fonctions pouvant former un pharmacophore. Les espèces susceptibles d'être identifiées comme affines pour une cible d'intérêt seront donc aussi celles qui seront linéarisées le plus efficacement par cette approche enzymatique.

Une fois linéarisés, les peptides peuvent être analysés par spectrométrie de masse en tandem avec fragmentation par CID haute énergie MALDI-ToF-ToF et des spectres de masse montrant des fragmentations informatives ont pu être obtenus. La figure 85 montre le spectre obtenu après fragmentation par CID du **Hairp2-Naphthoyl-diBBLeu**^{1,8} fonctionnalisé deux fois par le **BBLeu** et les masses des ions fragments b et y associés. Les ions en rouges sont des ions diagnostiques permettant de caractériser les cystéines qui sont fonctionnalisées.

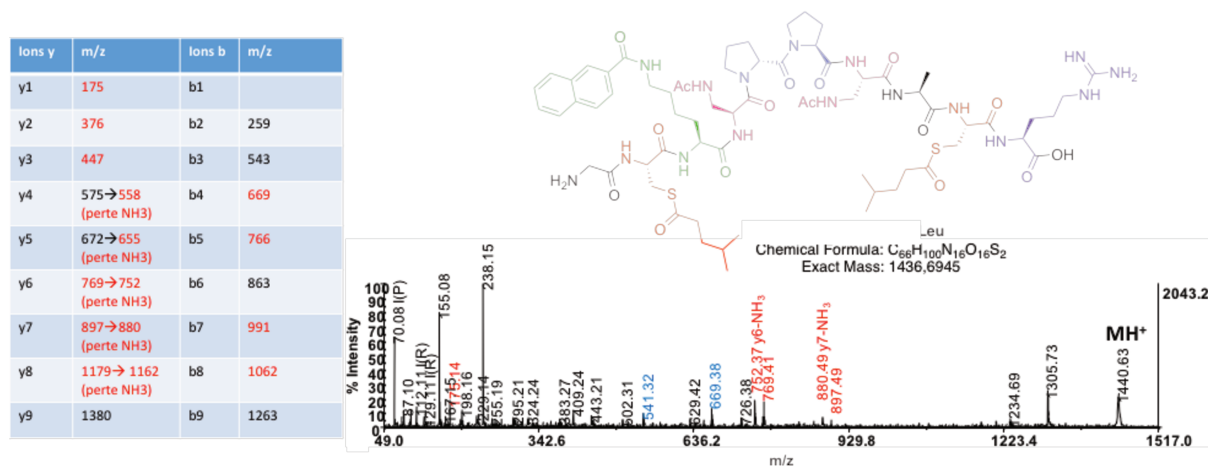


Fig. 85 : Structure du **Hairp2-Naphthoyl-diBBLeu¹⁻⁸** et spectre MS/MS obtenu par fragmentation CID haute énergie à partir de l'ion moléculaire mono protoné correspondant. Les ions b sont marqués en bleu et les ions y sont marqués en rouge.

Les fragmentations principales observées correspondent à deux clivages consécutifs au niveau de la sonde UV à m/z 155 et 238 puis à la perte d'une des deux fonctions thioesters à partir de l'ion parent à m/z 1305. Un mécanisme primaire pour sa formation est proposé en figure 86.

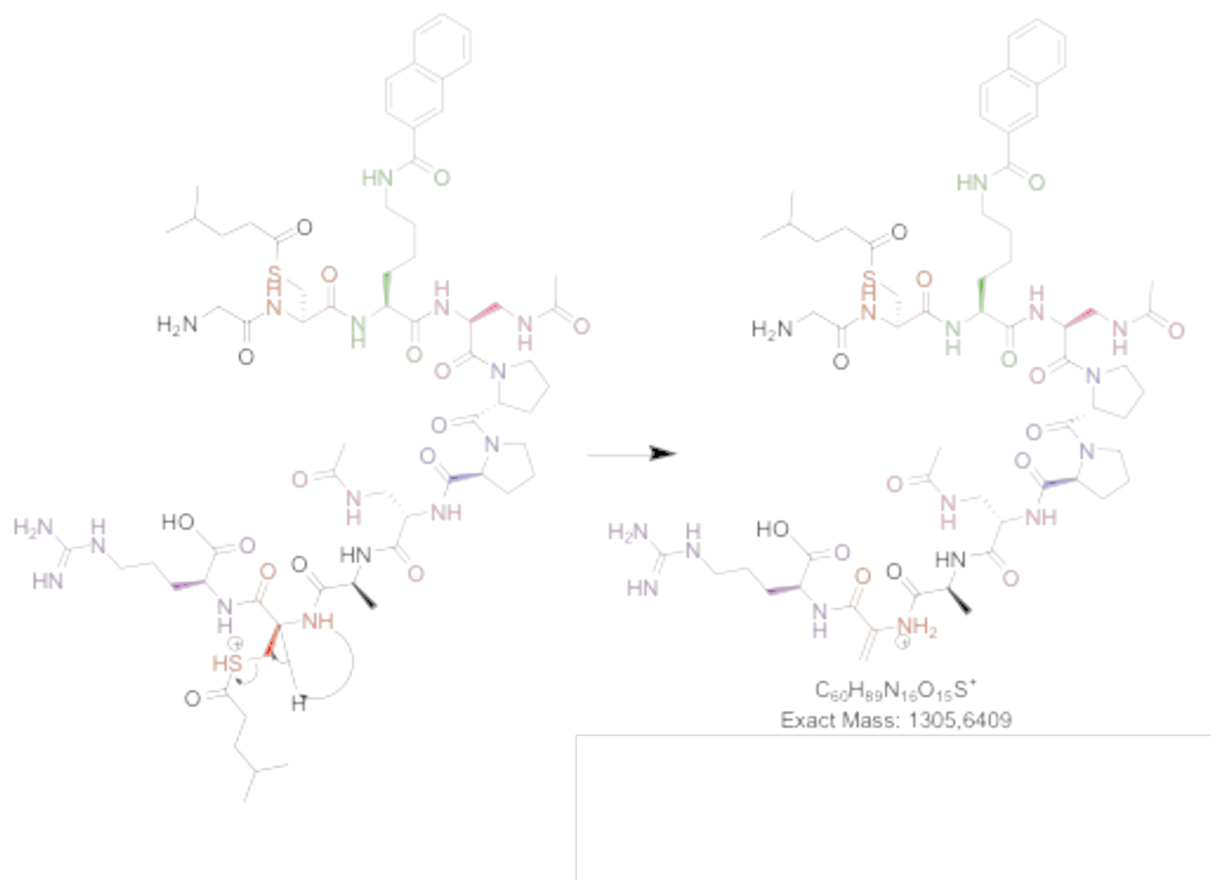


Fig. 86 : Proposition de mécanisme pour la formation de l'ion à m/z 1305 par perte d'un des thioesters

Bien que ces voies de fragmentation soient privilégiées, il est important de noter que des fragmentations se produisent aussi le long de la séquence et les ions observés permettent de remonter à la structure et donc à la position des fonctions thioesters. On peut noter des pertes de NH₃ quasiment systématiques sur les ions y, probablement provenant de la chaîne latérale de l'arginine. Les ions b sont globalement de faible intensité.

Malgré la fragmentation importante des fonctions thioesters en CID, une partie de l'énergie conduit à des voies de fragmentation informatives le long de la séquence peptidique, produisant des ions qui permettent de remonter à la structure et donc aux positions des fonctions thioesters. D'autres méthodes de fragmentation comme l'ETD pourront être employées dans l'éventualité où les substrats montreraient une trop forte perte des groupements thioesters en phase gazeuse.

4) Évolution de la structure du *scaffold* et nouveaux défis

Au cours de ce travail, la structure du *scaffold* a évolué petit à petit afin de moduler ses propriétés générales et permettre de se rapprocher d'un système satisfaisant sur tous les fronts : propriétés physico-chimiques et spectroscopiques, réactivité en DCC, digestion enzymatique et fragmentation MS/MS. La solubilité des *scaffolds* a été un des problèmes majeurs durant ce travail et nous avons donc choisi d'introduire un second résidu susceptible de l'augmenter. Ce résidu ne peut être placé qu'en position 2 ou 7 du *scaffold*, seules encore disponibles. Dans notre cas, nous avons arbitrairement choisi de placer ce nouveau résidu en position 7 et l'acide aminé chromophore en position 2.

La lysine a été choisie au dépend de l'arginine afin de favoriser une diversité maximale des résidus et faciliter l'identification des séquences. En revanche, tout comme l'arginine, la lysine est également reconnue par la trypsine. Au vu des résultats obtenus précédemment, nous avons pensé que le placement de ce résidu sur l'un des brins de l'épingle devrait lui conférer une plus grande résistance vis à vis de la trypsine, et permettre ainsi de conserver un unique site d'hydrolyse ou du moins un site d'hydrolyse préférentiel au niveau de l'arginine dans la boucle de l'épingle.

Des études de digestion tryptique ont donc été réalisées sur le peptide **Hairp4-Lys** et ses dérivés mono, di, tri et tétra acylés afin d'étudier les conséquences de la présence de ce second site basique sur la réactivité du peptide. Après digestion tryptique de ces nouveaux peptides, nous avons observé majoritairement une double digestion conduisant à un mélange des différents fragments de digestion A, B, C, D comme montré sur la figure 87 ci-dessous.

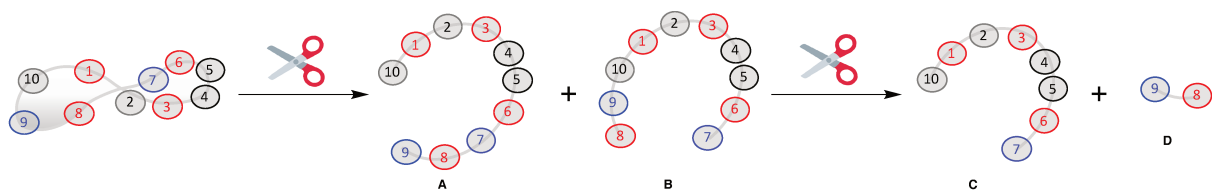


Fig. 87 : Schéma représentant les deux digestions successives par la trypsine ici symbolisée par une paire de ciseaux, à partir d'un peptide portant à la fois une arginine en position 9 et une lysine en position 7.

Des essais de digestion ménagée ont été réalisés sur deux séries de peptides issus du **Hairp4-Lys** fonctionnalisés par le **BBLeu** ou le **BBLys** en faisant varier soit le temps de digestion de seize à quatre à deux heures soit la quantité d'enzyme utilisée de 1 à 0,5 à 0,1 μg . Les résultats montrent qu'aucune digestion ménagée n'a pu être atteinte au travers des conditions explorées peu importe le nombre ou la nature des fonctions introduites sur les cystéines.

Le plus long des deux peptides **C** issus de ces digestions porte les cystéines 1, 3 et 6. Dans ce cas, une analyse MS simple permet de savoir s'il porte une, deux ou trois fonctions thioesters et la nature de celles-ci mais identifier leur position respective nécessite la spectrométrie de masse en tandem. Le plus court des deux peptides issus de la digestion trypsique est un dipeptide **D** ne comptant que l'arginine 9 et la cystéine 8. Dans ce cas, on peut savoir en ayant la masse du fragment **D** si oui ou non la cystéine 8 est fonctionnalisée et par quel **BB** sans avoir recours à la MS/MS.

De façon moins fréquente, on a pu observer la première ouverture sur la lysine qui conduit au peptide linéaire **B** qui peut alors être hydrolysé en deux peptides **C** et **D**. Cette ouverture a été confirmée par MS/MS du peptide linéaire **B** puisque dans ce cas le résidu en C-terminal est une lysine, ce qui conduit à l'ion y_1 à m/z 147 et pas 175 comme dans le cas de l'arginine.

III. Conclusion

Nous avons donc pu mettre en place une approche enzymatique permettant d'accéder à des peptides linéarisés. L'emploi de la trypsine dans des conditions assez éloignées de ses conditions optimales de fonctionnement et orthogonales aux fonctions thioesters a certes conduit à des rendements d'hydrolyse limités mais suffisants pour permettre des analyses en spectrométrie de masse en tandem afin de caractériser nos peptides. A défaut de permettre de distinguer formellement des isomères de position dans le dernier cas présenté, les résultats obtenus et les observations faites durant ce travail ont permis de mettre en évidence les points encore perfectibles de cette approche.

La double digestion semble être substrats dépendante et comme mentionné précédemment, la fonctionnalisation au voisinage d'un site pourrait diminuer son accessibilité pour la trypsine, accessibilité qui est déjà faible pour le résidu basique localisé dans un des brins en position 7. Cependant, cette double digestion a l'avantage de permettre d'avoir accès à l'état de fonctionnalisation de la cystéine 8 sans avoir recours à la MS-MS. Puisque les structures *a priori* visées par la stratégie de DCC employée seront des molécules *tétra* fonctionnalisées et donc portant une modification sur la cystéine 8, cette double digestion peut alors s'avérer avantageuse. Le peptide le plus long obtenu après digestion enzymatique peut quant à lui être analysé par MS-MS. Cela permet alors d'avoir accès à l'état de fonctionnalisation des cystéines restantes.

La méthodologie de couplage « off-line » a permis d'explorer trois approches différentes pour la linéarisation de nos peptides cycliques et leur potentiel en tant que méthode de routine. L'ouverture par ISD et le clivage au bromure de cyanogène se sont révélés trop peu efficaces pour une application en routine. L'hydrolyse enzymatique à la trypsine en revanche s'est avérée compatible. L'utilisation de l'enzyme en conditions non optimales mais compatibles avec la stabilité des thioesters (pH 5) et la solubilité parfois faible des molécules à analyser (solvant organique nécessaire) conduit à limiter les rendements de digestion enzymatique mais permet d'arriver jusqu'à la caractérisation des peptides.

Un point important qui a été mis en évidence lors de ce travail est la difficulté voire l'impossibilité de digestion enzymatique de peptides monofonctionnalisés portant des chaînes amides en lieu et place des cystéines ou thioesters, dans les conditions utilisées. Il est donc raisonnable de considérer que les peptides ayant un maximum de cystéines initialement sont donc de meilleurs objets d'études pour la mise au point de notre stratégie analytique en spectrométrie de masse alors que la mise au point de la méthodologie de combinatoire dynamique peut être mise au point sur des peptides avec davantage de fonctions acétamides pour limiter la complexité des bibliothèques générées. Ceci est un avantage en ce qui concerne l'application finale. En effet, les architectures affines identifiées suite au criblage des DCLs par une cible sont destinées à être transposées de la série thioester à la série amide avant toutes possibilités d'études biologiques. Ainsi, la résistance à la digestion enzymatique présentée par les peptides présentant des groupements acétamides est un argument permettant de soutenir la stabilité biologique des potentielles structures identifiées par notre stratégie et leur potentiel en tant que molécules à visée thérapeutique dans le futur.

Une autre possibilité de linéarisation sélective qui pourrait être envisageable ici est l'ouverture par irradiation UV d'un groupement photoclivable au sein du *scaffold* peptidique.

Plusieurs groupements photoclivables peuvent être envisagés d'après la littérature comme la 2-nitrophenylglycine (2-Npg)¹⁵⁰ ou la 2-nitrophenylalanine (2-Npa)¹⁵¹ par exemple (Fig. 88).

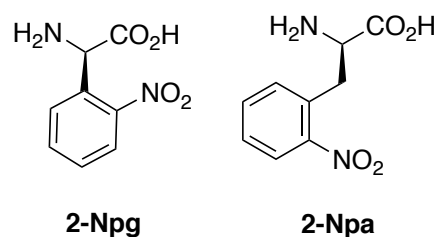


Fig. 88 : Structures de la 2-nitrophenylglycine (2-Npg) et de la 2-nitrophenylalanine (2-Npa).

Soumis à une irradiation UV à 365 nm, ces résidus peuvent conduire à la coupure spécifique du squelette peptidique au niveau de leur site d'incorporation. Dans le cas de peptides cycliques, cela conduirait à des séquences linéaires. Il serait bien sûr nécessaire de vérifier si la réaction de clivage photochimique est réalisable sur nos architectures tant au niveau de la stabilité des thioesters en condition de clivage photochimique que celui des propriétés spectroscopiques des résidus photoclivables en question. En effet, comme discuté au chapitre 1, l'expérience a montré que les chromophores incorporés dans nos séquences possèdent des propriétés différentes de celles décrites dans la littérature. A priori, les fonctions thioesters semblent stables en conditions MALDI, avec une irradiation UV par un laser 337 ou 355 nm. Cette approche d'ouverture de cycle pourrait trouver son application dans la transposition de cette approche analytique à des structures polycycliques qui ne sauraient être linéarisées totalement par une enzyme comme les *scaffolds* en hélice α par exemple.

Sur le plan de la technique d'analyse en spectrométrie de masse, la méthode mise au point en HPLC et MALDI-ToF avec couplage « off-line » sera transférée sur un système couplé nano-LC ESI MS/MS pour permettre une analyse de type « haut débit ». Pour cela, il serait nécessaire de considérer nos DCL comme des mélanges biologiques complexes et les analyser avec une approche de type protéomique/peptidomique. Les problèmes de coélution plus ou moins importants rencontrés en HPLC seraient aussi présents et ce problème restera toujours présent au fur et à mesure que nous augmenterons la complexité des bibliothèques. Dans cette approche, la complexité de la séparation HPLC serait accrue du fait que les peptides linéaires et cycliques obtenus par digestion trypsique incomplète seraient contenus dans le mélange à séparer. Afin de décomplexifier le système il serait judicieux d'ajouter une dimension de séparation supplémentaire à la chromatographie en phase inverse. Une voie envisageable serait l'utilisation d'un système de type UPLC/ESI/IMMS (Fig. 89). La première séparation se ferait

en UPLC selon la polarité des composés de la librairie comme le fait la HPLC, mais la plus grande performance de la UPLC pourrait permettre la séparation des peptides cycliques et linéaires. Après ionisation par ESI, une étape de séparation supplémentaire par mobilité ionique permettrait de séparer les ions isomères de position en phase gazeuse selon leur mobilité au travers d'un gaz inerte avant d'être analysés par MS et MS-MS. En effet, il est probable que deux isomères de position ne possèdent pas la même section efficace et ne doivent pas migrer de la même façon dans la cellule de mobilité ionique.



Fig. 89 : Schéma en blocs d'un système UPLC/ESI/IMMS à deux dimensions de séparation et permettant une approche à « haut débit » pour l'analyse des bibliothèques combinatoires dynamiques.

Enfin, il serait intéressant de développer un outil bio-informatique pour rendre l'analyse de plus grandes bibliothèques rapide et automatique, analogue à ceux utilisés en protéomique classiques. Il s'agirait de simuler la librairie dynamique entière en la comparant à une protéine fictive dont la séquence correspondrait aux séquences linéaires des peptides théoriquement obtenus à partir de la librairie, mises bout à bout. Il serait alors possible de digérer *in silico* par la trypsine cette protéine fictive pour obtenir la liste des masses et rapports m/z des ions linéaires issus de ces digestions, en tenant compte des éventuelles erreurs de clivage de la trypsine (1 ou 2 sites d'hydrolyse avec des rendements plus ou moins complets) mais aussi les m/z des ions fragments potentiels issus de la fragmentation MS-MS de ces peptides. Ainsi, une base de données des ions attendus pour une librairie générée avec un *scaffold* et des building blocks donnés pourrait être générée. Par comparaison de cette banque de données avec les spectres MS/MS enregistrés, il serait possible d'identifier et caractériser les peptides fonctionnalisés présents dans les bibliothèques.

Pour finir, le développement de nouvelles architectures peptidiques pour élargir la stratégie de combinatoire dynamique à d'autres structures, comme des hélices par exemple, entraînera le besoin d'un nouvel effort analytique pour l'adaptation de cette méthode d'étude. Il y a fort à parier que chaque nouvelle structure utilisée en DCC nécessitera une méthode spécifique d'analyse en spectrométrie de masse.

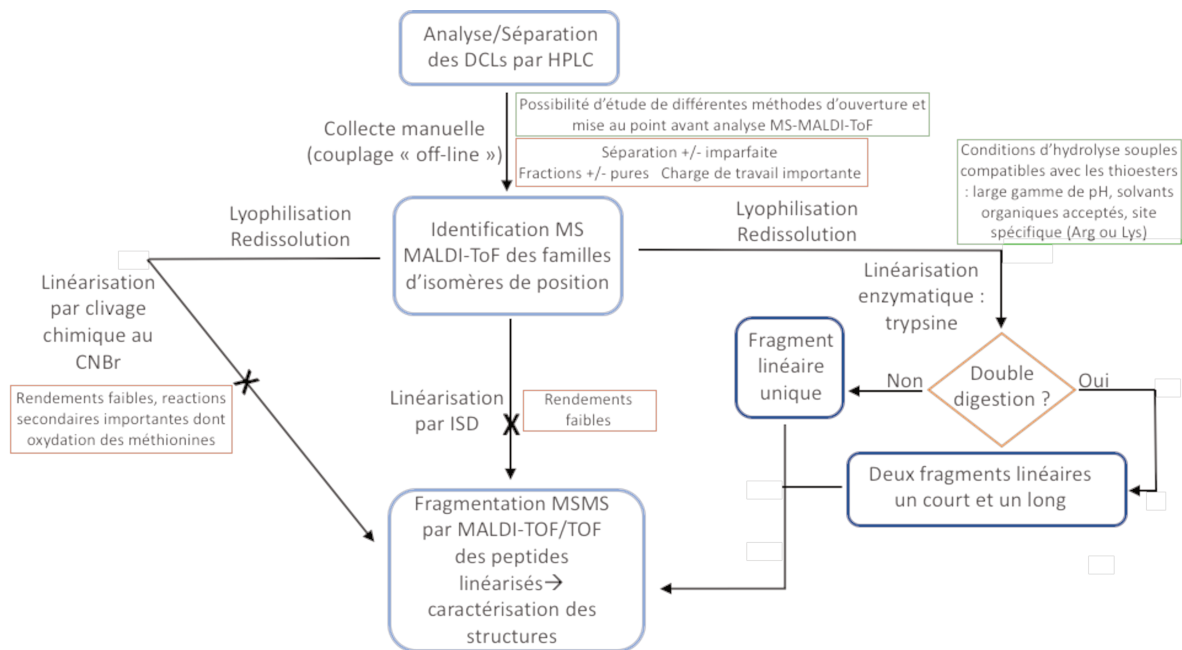


Fig. 90 : Diagramme synoptique résumant les différentes voies explorées pour la caractérisation des bibliothèques combinatoires dynamiques. Les cadres verts soulignent les points positifs et améliorables, les cadres rouges montrent les inconvénients. Les voies possédant une croix sur une des flèches n'ont pas été retenues.

Chapitre 5 : Études Structurales

I. Problématiques

La structuration des plateformes peptidiques représente un paramètre important dans la conception et la conduite de ce projet puisque l'idée centrale est de générer des combinaisons variées de fonctionnalités à la surface d'un peptide par DCC sans affecter la géométrie spatiale de l'échafaudage moléculaire qui les porte. Plusieurs questions relatives à cette structuration sont et seront à soulever au cours de ce projet. Tout d'abord, il est nécessaire de s'assurer que la plateforme utilisée respecte la structure secondaire imaginée lors de sa conception rationnelle. De plus, il faut savoir si la fonctionnalisation lors de l'étape de DCC et en particulier par quatre fonctions thioesters modifie de façon significative la structuration de la plateforme. En effet, c'est lors de cette étape que la reconnaissance moléculaire avec une cible interviendra et que la structure tridimensionnelle de la plateforme et l'arrangement spatial des chaînes des thioesters seront cruciaux pour que le criblage ait tout son sens. Enfin, à plus long terme, si une structure est identifiée après criblage d'une librairie dynamique comme partenaire privilégié pour une cible, il sera nécessaire de produire son analogue où les thioesters seront remplacés par des amides afin de pouvoir procéder à des évaluations biologiques, les thioesters n'étant pas assez stables en milieu biologique. Là encore, la vérification de la structuration après ce passage en version amide sera nécessaire pour complètement valider la stratégie.

Des études structurales par modélisation moléculaire et RMN ont été entamées sur les peptides **Hairp4-Lys**, **Hairp0**, **Hairp4-tetra(BBGly)**^{1,3,5,8} afin d'offrir des pistes de réponses à ces questions (Fig. 91).

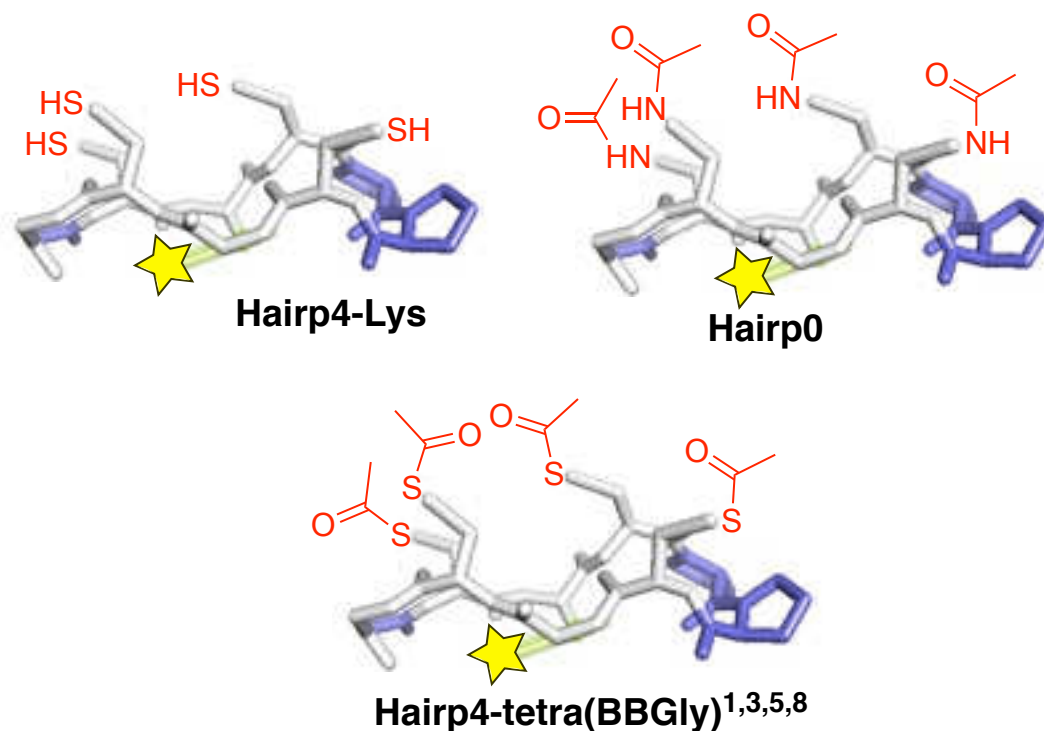


Fig. 91 : Structures schématiques des peptides **Hairp4-Lys** et ses analogues acetamide **Hairp0** et acyl thioesters **Hairp4-tetra(BBGly)^{1,3,5,8}**.

II. Simulation de dynamique moléculaire sur **Hairp4-Lys**

Une approche reposant sur la modélisation par informatique a été entreprise afin d'apporter plus d'éléments de réponse quant à la structure de ce peptide. Le logiciel Amber^{152,153} a été utilisé pour générer une dynamique moléculaire de 10 ns en solvant implicite à 300 K dans le champ de force ff14SB. La structure initiale a été générée par une technique de recuit simulé sans contrainte, puis minimisée. Il est important de noter qu'à ce stade, l'ensemble des calculs a été effectué en absence de contrainte expérimentale.

Au cours de cette simulation, l'évolution de différents paramètres géométriques clés de la molécule peut être suivie comme des distances entre atomes et les angles dièdres des liaisons peptidiques. On peut ainsi visualiser si la molécule présente différents ensembles de valeurs pour ces paramètres structuraux ce qui indique le nombre de conformations qu'elle adopte. Tout d'abord, l'évolution de la valeur quadratique moyenne de la position des atomes de carbone et d'azote du squelette peptidique par rapport à une structure de référence a été observée (Fig. 92).

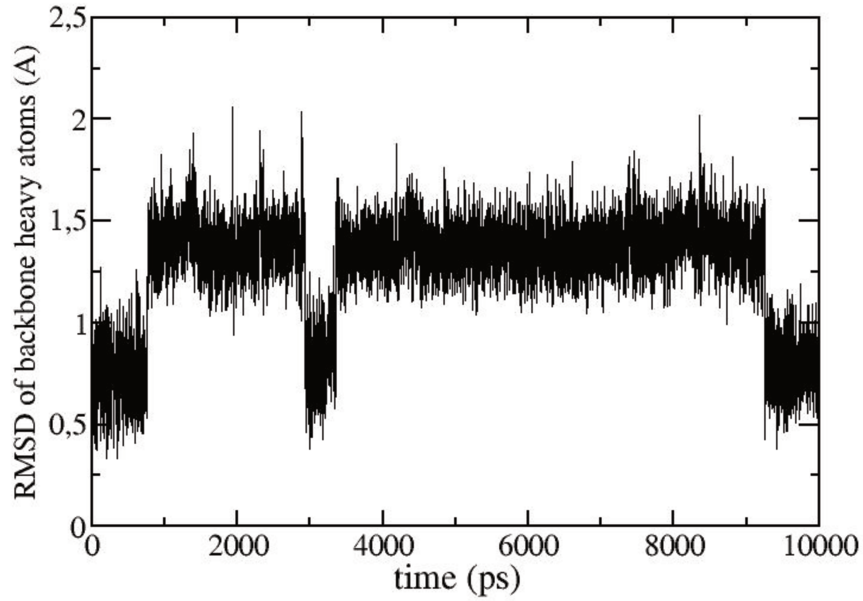


Fig. 92 : Évolution de la RMSD des atomes d'azote et de carbone du squelette peptidique au cours de la dynamique moléculaire par rapport à la structure initiale de la dynamique prise comme référence.

On peut observer que les écarts de RMSD sont très faibles (au maximum de 2 Å) ce qui indique que la molécule explore un espace conformationnel relativement restreint. Cependant, les valeurs de RMSD se répartissent en deux populations séparées par des sauts répétés au cours de la dynamique indiquant deux conformations apparentes. Afin d'obtenir plus de détails, ce même suivi a été effectué pour les angles dièdres Φ et Ψ de l'ensemble des résidus du peptide (Fig. 93).

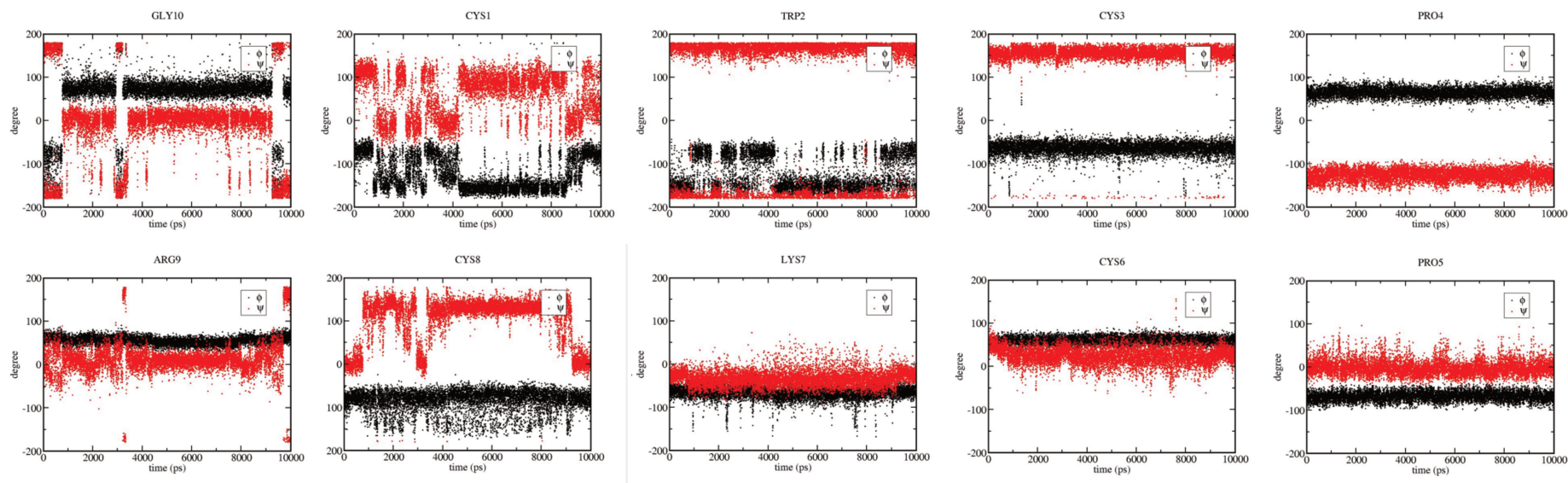


Fig. 93 : Évolution des angles dièdres Φ (noir) et Ψ (rouge) des résidus de la plateforme **Hairp4-Lys** au cours de la dynamique moléculaire. Les graphiques sont disposés de manière à reproduire le cycle peptidique avec l'inducteur de coude β D-Pro-L-Pro ici placé sur la droite de la page.

On observe que les valeurs d'angles dièdres des résidus 3, 4, 5, 6 et 7 sont extrêmement contraintes traduisant une forte rigidité de cette partie de la molécule sur le temps de dynamique exploré. On remarque en revanche que les résidus Cys1 (angles Φ (-156° et -74°) et Ψ (100° et -8°), Trp2 (angle Φ (-150° et -70°), Cys8 (angle Ψ (132° et 4°) et Gly10 (angles Φ (75° et -75°) et Ψ (0° et 180°)) explorent deux régions du diagramme de Ramachandran, ce qui se traduit par de nombreux sauts sur leur trajectoire respective. De manière intéressante, ces événements conformationnels semblent se produire de manière concertée et correspondent aux transitions observées pour la RMSD du squelette. Ceci montre que ce sont ces résidus qui sont impliqués dans le changement conformationnel. Cela traduit également une flexibilité plus importante de la boucle contenant la Gly10 opposée aux prolines. Cela est en bonne adéquation avec les travaux réalisés en digestion enzymatique pour la spectrométrie de masse.

Enfin, les distances entre certains atomes ont été suivies au cours de la simulation afin de vérifier l'implication de certains protons amide dans une liaison H (Fig. 94). La distance entre le NH de la cystéine 6 et le CO de la cystéine 3 reste inférieure à 2.5 Å sur 50 % de la trajectoire ce qui indique une très bonne stabilité de la liaison H. La distance entre leur carbone α reste inférieure à 7 Å sur toute la durée de la trajectoire ce qui atteste de la présence constante d'un coude β II', structure validée également par la valeurs des angles dièdres de la D-Pro4 ($\Phi \approx 60^\circ$, $\Psi \approx -120^\circ$) et de la Pro5 ($\Phi \approx 80^\circ$, $\Psi \approx 0^\circ$).

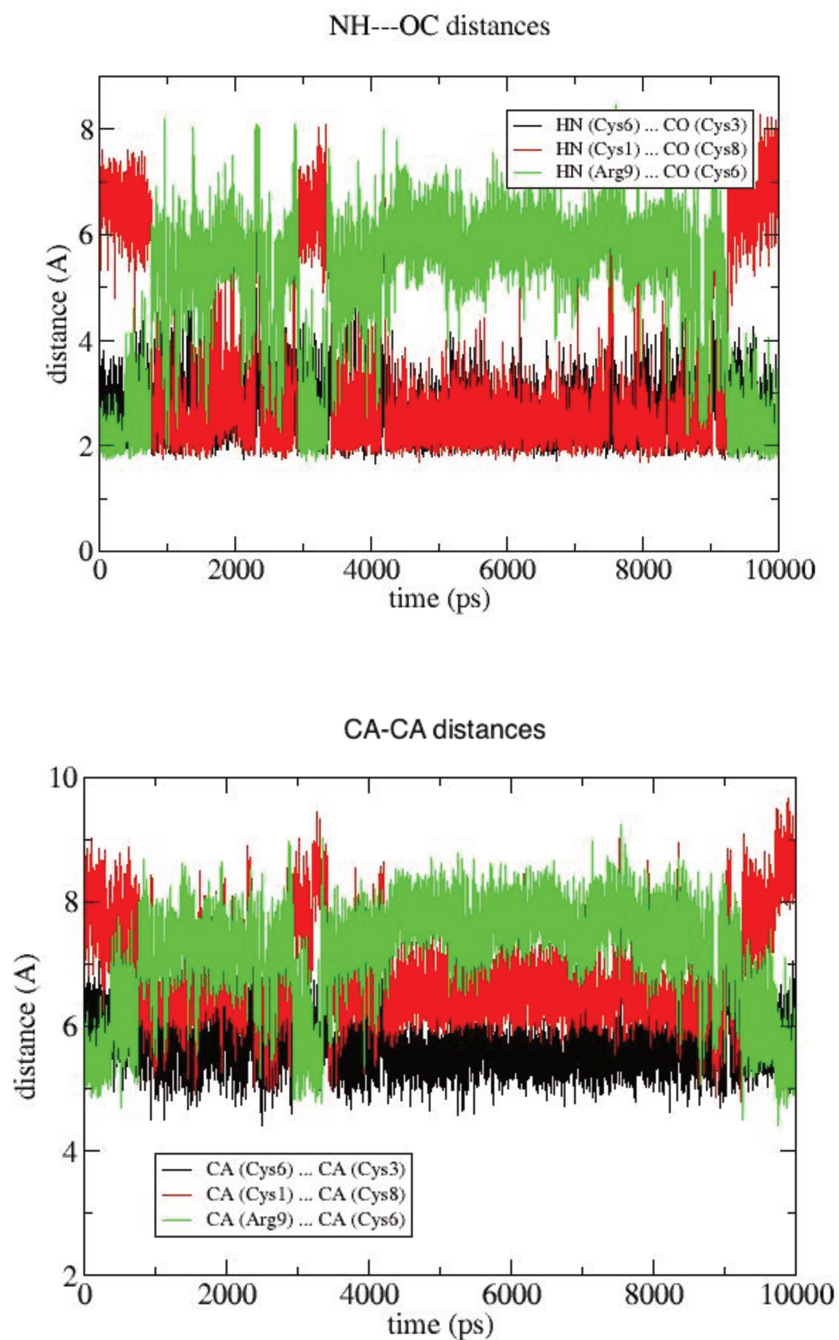


Fig. 94 : Évolutions des distances NH-OC au cours de la simulation (en haut). Distance NH(Cys6)-OC(Cys3) en noir, distance NH(Cys1)-OC(Cys8) en rouge, distance NH(Arg9)-OC(Cys6) en vert. Évolutions des distances entre carbonés α au cours de la simulation (en bas). Distance C_{α} (Cys6)- C_{α} (Cys3) en noir, distance C_{α} (Cys1)- C_{α} (Cys8) en rouge, distance C_{α} (Arg9)- C_{α} (Cys6) en vert.

L'analyse des autres distances HN-OC et C α -C α a été faite sur l'ensemble de la molécule et révèle que les résidus Cys6-Lys7-Cys8-Arg9 pourraient également former un coude β de manière transitoire au cours de la dynamique. On peut en effet observer que sur environ 15 % de la durée de la simulation, le NH de l'arginine 9 et le CO de la cystéine 6 se rapprochent à moins de 2.5 Å ce qui est caractéristique d'une liaison H. Cette proximité se retrouve également entre les carbones α correspondants. Ce phénomène étant simultané aux changements conformationnels mentionnés plus tôt, il est raisonnable de penser que ce conformère minoritaire soit stabilisé par une liaison H entre l'arginine 9 et la cystéine 6.

Les structures correspondantes à chacune de ces deux populations de conformères ont été extraites de la dynamique moléculaire (respectivement à 4,5 ns et 9,9 ns). La figure 95 en présente une superposition afin de mettre en évidence les particularités structurales de ces deux conformères. On peut remarquer que la différence majeure entre les deux structures vient de la formation d'un repli du squelette au niveau de la partie flexible identifiée plus tôt. L'autre conformation semble en revanche plus étendue.

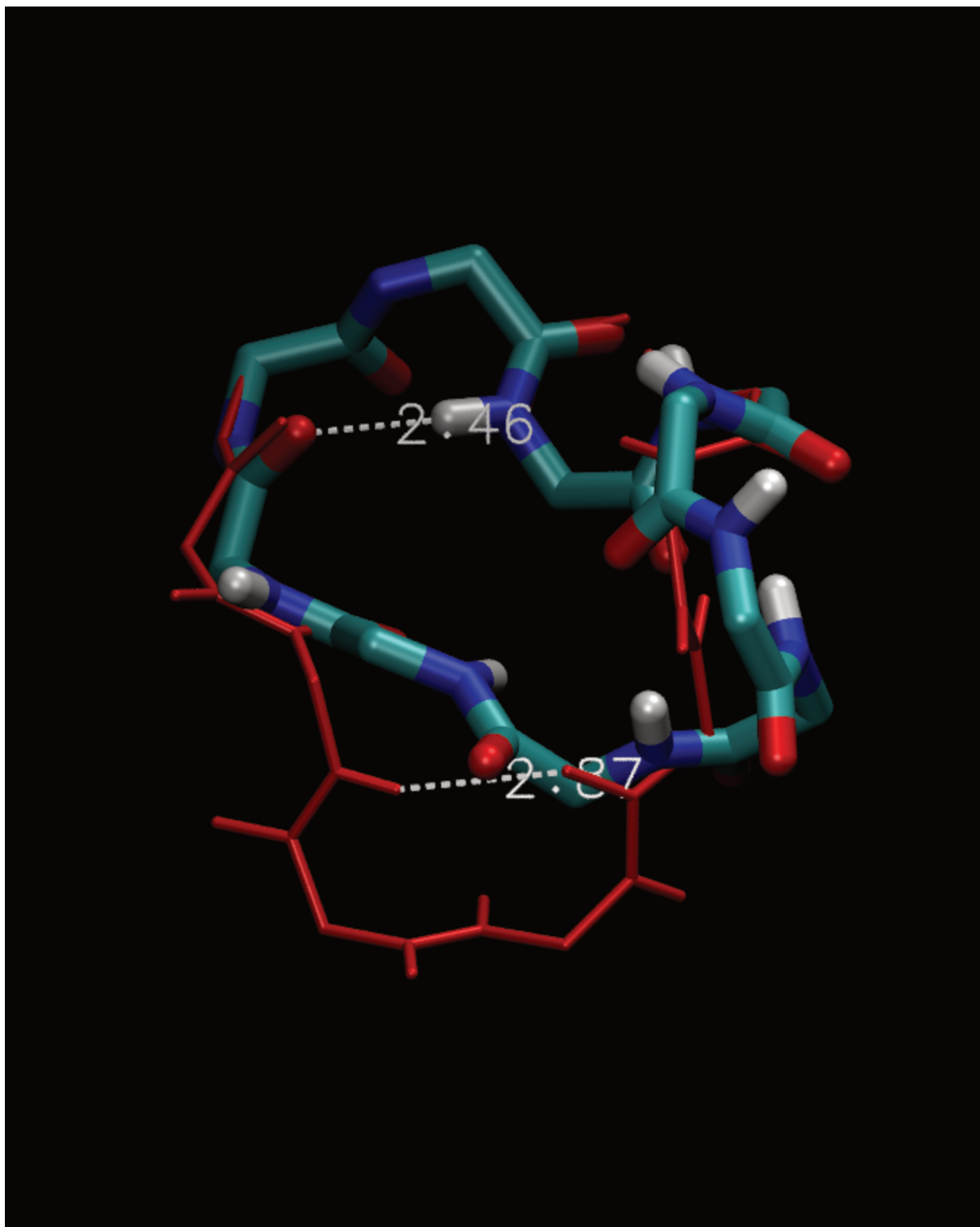


Fig. 95 : Superposition des structures extraites de la simulation de dynamique moléculaire à 4,5 ns (atomes représentés avec différentes couleurs) et 9,9 ns (rouge). Seuls les squelettes peptidiques sont représentés et ils ont été superposés au niveau du coude β formé par l'enchaînement D-Pro-L-Pro. La liaison H mesurée à 2,46 Å est celle formée entre la Cys3 et la Cys6 au niveau du coude β II' contenant les prolines. La liaison H mesurée à 2,37 Å est celle formée entre la Cys1 et la Cys8.

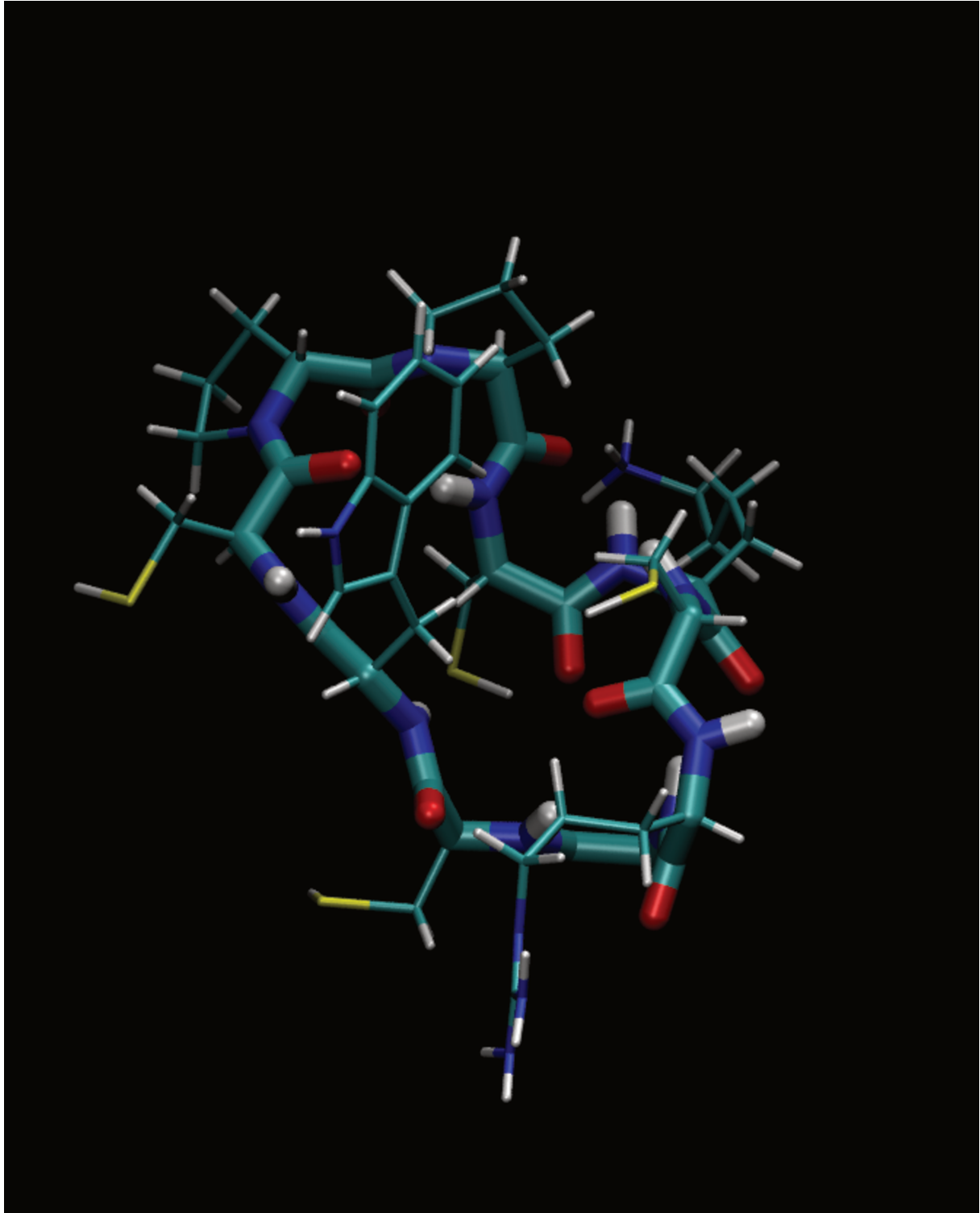


Fig. 96 : Structure extraite de la simulation de dynamique moléculaire à 4,5 ns. On peut remarquer que le repliement de la boucle flexible autour de la glycine conduit une cystéine à présenter sa chaîne latérale sur la face opposée aux trois autres.

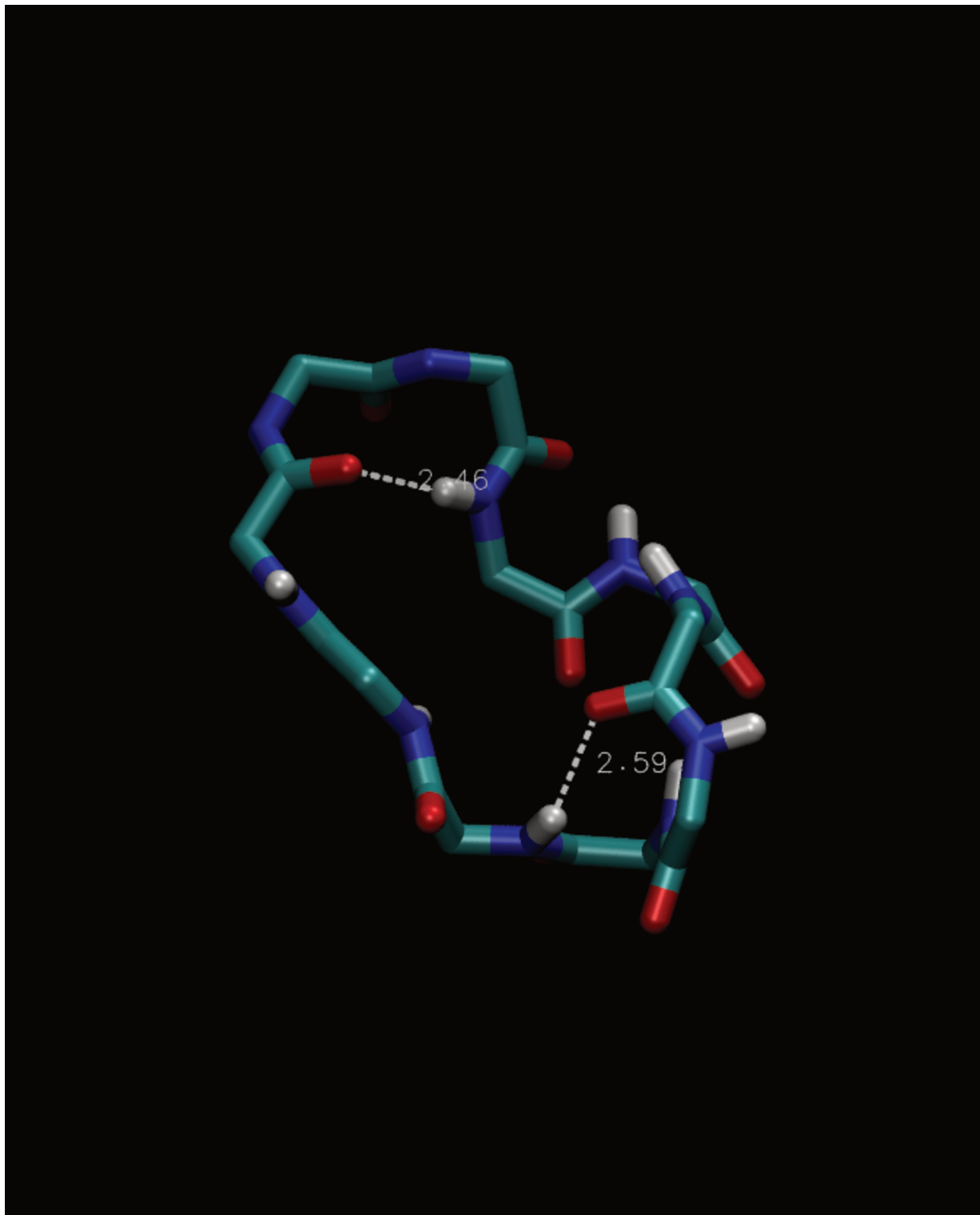


Fig. 97 : Structure extraite de la simulation de dynamique moléculaire à 4,5 ns, seul le squelette peptidique est représenté. La liaison H mesurée à 2,46 Å est celle formée entre la Cys3 et la Cys6 au niveau du coude β II' contenant les prolines. La liaison H mesurée à 2,59 Å est celle formée entre la Cys6 et l'Arg9 stabilisant le conformère minoritaire.

III. Étude RMN sur la plateforme non fonctionnalisée **Hairp4-Lys**

Afin d'avoir une base de travail solide et un point de comparaison futur pour les plateformes fonctionnalisées, le peptide **Hairp4-Lys** à quatre sites fonctionnalisables a été choisi afin d'évaluer sa bonne structuration en épingle à cheveux β . Sa meilleure solubilité par rapport à son analogue **Hairp4-Ala** a été le paramètre décisif pour le choix entre ces deux plateformes.

Afin de pouvoir apporter des éléments de réponses cohérents quant à la structuration de **Hairp4-Lys** en épingle à cheveux β , l'étude a été conduite dans des conditions similaires à celles utilisées en DCC : un tampon phosphate aqueux (10 % D₂O) à pH 7 et contenant du TCEP afin de prévenir la formation de disulfures durant le temps des expériences RMN. Différentes expériences ont pu donc être menées pour conduire à l'attribution : des mesures de spectres ¹H, ¹³C, TOCSY et HSQC ont permis d'attribuer les signaux correspondant à chacun des atomes, mais également des expériences NOESY/ROESY afin d'extraire des contraintes de distances et d'angles qui permettront de réaliser des modélisations informatiques de cette structure sous contraintes RMN. Toutes ces expériences ont été réalisées avec des séquences dites « watergate » permettant une suppression du signal de l'eau.

Dans un premier temps, des expériences ¹H ont été réalisées à différentes températures sur un spectromètre 500 MHz équipé d'une cryosonde. La figure 98 présente une superposition de la région correspondant aux protons aromatiques et aux NH amides des différents spectres obtenus.

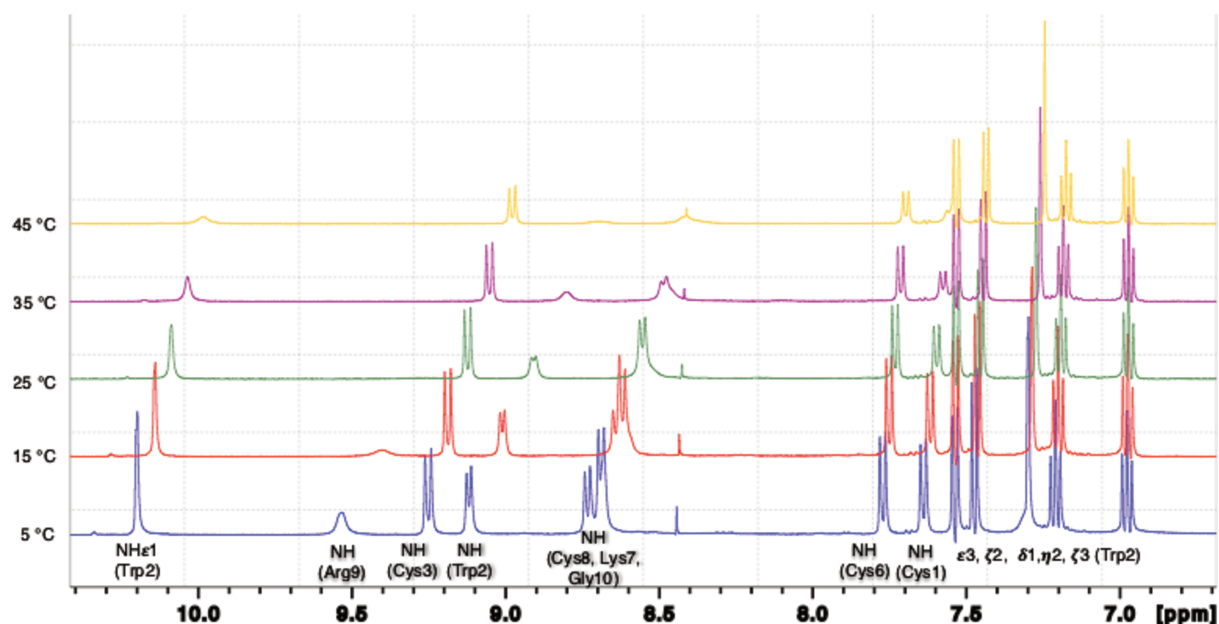


Fig. 98 : Spectres RMN ^1H de **Hairp4-Lys** acquis à différentes températures (5 à 45 °C). Seule la région des protons aromatiques et des NH amides est affichée. L'attribution est affichée en dessous des tracés et reprend d'une part la numérotation des résidus du *scaffold* et d'autre part la nomenclature recommandée pour les acides aminés¹⁵⁴.

Il est tout d'abord important de noter qu'un seul jeu de signaux est observable. Cela indique donc qu'une seule population de conformère est présente. D'autre part, l'éclatement important (~2 ppm) des signaux correspondants aux protons des amides du squelette peptidique indique une structuration importante. Enfin, ces expériences ont permis de choisir la température de 15 °C comme température de travail pour les expériences à deux dimensions qui ont permis en parallèle l'attribution totale des signaux. Cette température a été choisie car à cette température les signaux des protons amides présentaient un minimum de recouvrement entre eux tout en étant tous distinguables de manière satisfaisante.

Ces expériences ont également permis de déterminer les coefficients de température des protons amides du squelette peptidique qui correspondent à $\left| \frac{\Delta\delta(\text{ppm})}{\Delta T (\text{K})} \right|$. Ceux-ci apportent une information à propos de l'implication de ces protons amides dans une liaison hydrogène intramoléculaire. En effet, on estime que si ce coefficient est inférieur à 4.6 ppb/K, le proton amide correspondant est impliqué dans une liaison hydrogène intramoléculaire avec une probabilité de 85% et 93% s'il est inférieur à 4 ppb/K¹⁵⁵. La valeur seuil de 6 ppb/K est également retrouvé dans la littérature.

Les droites obtenues pour deux des résidus du squelette sont présentées ici, les autres sont disponibles en annexes (Fig. 99). Le tableau 6 qui les accompagne résume également les coefficients obtenus pour tous les NH du peptide.

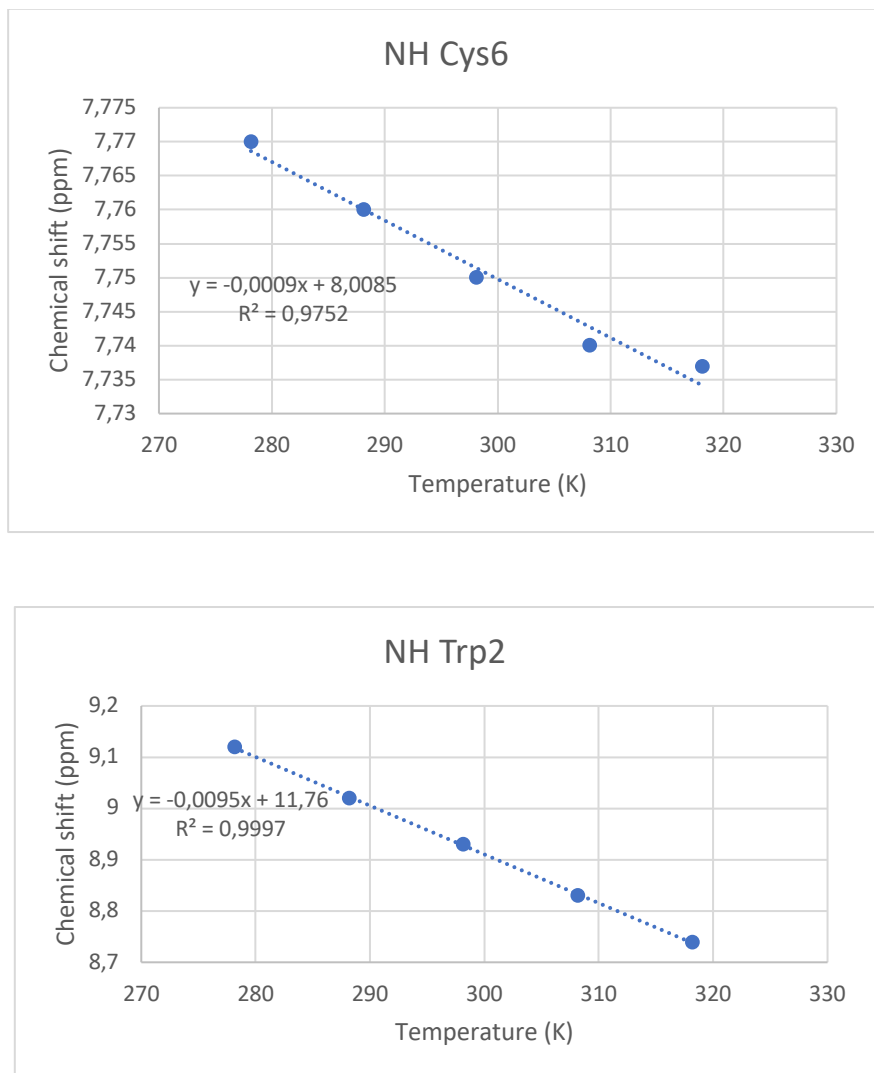


Fig. 99 : Graphiques représentant l'évolution des déplacements chimiques des NH amides des résidus Trp2 et Cys6 en fonction de la température. Une régression linéaire peut être faite sur ces données, la pente de la droite obtenue correspondant alors au coefficient de température du proton considéré.

Tableau 6 : Coefficients de température des NH de la plateforme **Hairp4-Lys** en ppb/K (partie par milliard par Kelvin).

	NH ε1 Trp2	NH Arg9	NH Cys3	NH Trp2	NH Lys7	NH Cys8	NH Gly10	NH Cys6	NH Cys1
$\Delta\delta/\Delta T$ (ppb/K)	4,5	12,5	5,8	9,5	6,8	5,8	6,1	0,9	1

Grâce à ces résultats il est raisonnable de penser que les protons amides des résidus Cys1 et Cys6 établissent des liaisons hydrogène. Il est intéressant de noter que ce sont également les protons amides les moins sensibles à la température en terme de largeur de raie, ce qui signifie une vitesse d'échange plus lente avec les protons de l'eau. Les HN des résidus Cys3 et Cys8 semblent par ailleurs présenter une légère propension à interagir avec des groupements carbonyles voisins. Comme décrit dans la littérature^{54,127,128,135}, le mode de liaisons H canonique pour ce type de structure serait que les Cystéines 3 et 6 d'une part, et que les Cystéines 1 et 8 d'autre part, interagissent entre elles deux à deux. Cela apporterait un élément de réponse important quant à la structuration de cette plateforme qui diffère sensiblement de l'étude préliminaire menée par MD.

La MD a permis de mettre en évidence la liaison HN Cys6-CO Cys 3. Aucune liaison H n'a été trouvée pour la Cys 1 qui présente un coefficient très bas. De plus, le résidu Arg9 est celui qui présente le coefficient de température le plus élevé il est donc très peu probable qu'il soit impliqué dans une liaison H, même fugace, comme indiqué par les simulations de MD. Ces résultats mettent donc en lumière la prudence dont il faut faire preuve en analysant les résultats de MD. La trajectoire réalisée est sans doute trop courte pour pouvoir explorer la totalité de l'espace conformationnel et une trajectoire approchant la microseconde devra être réalisée. Par ailleurs, l'implémentation de contraintes expérimentales (RMN) sera une prochaine étape afin d'avoir un modèle plus réaliste.

IV. Étude RMN comparative sur **Hairp4-tetra(BBGly)** et **Hairp0**

Des premières analyses sur les analogues fonctionnalisés par une chaîne acyl **Hairp0** et **Hairp4-tétraGly** ont également été entamées. Tout aussi préliminaires soient-elles, elles participent malgré tout à répondre aux questions initiales posées. Des expériences RMN ¹H menées sur ces peptides ont été réalisées dans les mêmes conditions de tampon, de concentration et de température (15 °C) que celles précédemment présentées pour **Hairp4-Lys**. La figure 100 présente une superposition des spectres protons obtenus pour les trois plateformes, centrés sur la région des NH amides.

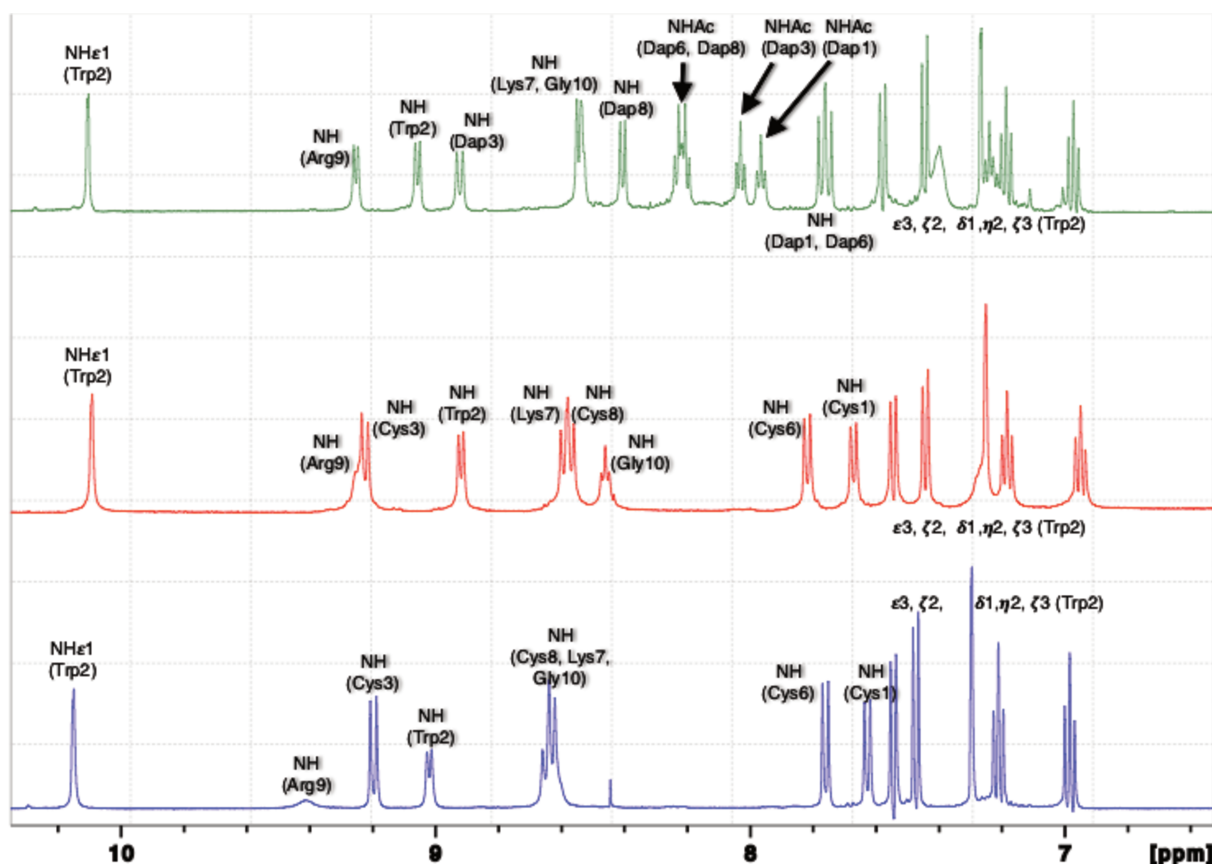


Fig. 100 : Spectres RMN ^1H de **Hairp4-Lys** (bleu), **Hairp4-tetra(BBGly) 1,3,5,8** (rouge) et **Hairp0** (vert) à 15 °C. Seule la région des protons aromatiques et des NH amides est affichée.

D'après ces résultats, il est d'ores et déjà possible de confirmer une bonne structuration des plateformes fonctionnalisées à la fois en série thioester et en série amide. En effet, le passage à l'une ou à l'autre des séries n'engendre pas de changement majeur dans le spectre en terme de variation de déplacement chimique des signaux amides. Cela tend à indiquer que ces protons sont dans un environnement similaire quelle que soit la fonctionnalité portée par la plateforme (thioester ou amide), et donc que le squelette peptidique qui les porte adopte une structuration similaire également.

V. Discussion et perspectives

Le travail réalisé et présenté au sein de ce chapitre bien que préliminaire démontre un certain effort quant aux questions liées à la structure des plateformes développées pour ce projet. Même si ces résultats restent préliminaires, des premiers éléments de réponses ont pu être mis en lumière afin de répondre aux questions initiales ayant motivées ce travail. La structuration de la plateforme **Hairp4-Lys** a pu être démontrée d'après les données

expérimentales obtenues par RMN relatives à l'existence de liaisons hydrogènes inter-brins indiquant une structuration canonique. Le développement d'un modèle virtuel a pu être entamé afin de soutenir davantage les preuves empiriques. Les premiers résultats liés à cette modélisation vont effectivement dans le même sens que les résultats expérimentaux à savoir une structuration rigide et stable de la plateforme mais présentant une portion plus flexible malgré tout, ce que d'autres travaux, par exemple les études de digestion enzymatique, ont également mis en lumière.

Les premières études réalisées sur les peptides **Hairp0** et **Hairp4-tétraBBGly**^{1,3,6,8} soutiennent l'hypothèse d'une plateforme rigide et résistante à la fonctionnalisation sur le plan de la stabilité structurale. Cela apporte des premiers éléments de réponses positifs quant à la validité de notre stratégie. Il sera bien sûr nécessaire d'approfondir ces résultats par des expériences RMN à deux dimensions non seulement pour les caractériser totalement mais également pour pouvoir les comparer à la référence **Hairp4-Lys** avec plus de précision et de détails. Il sera également intéressant de reproduire cette étude avec un analogue portant un autre type de fonctionnalité comme une chaîne plus encombrée ou portant un groupement chargé par exemple.

Enfin et à plus long terme, ce type d'étude devra être étendu à toutes structures fonctionnalisées identifiées comme partenaire d'une cible par criblage d'une librairie combinatoire dynamique après transposition en série amide, non seulement afin de la caractériser mais de confirmer sa bonne structuration et valider cette stratégie quelle que soit la structure générée et sélectionnée. Des analyses similaires cette sur la série de peptides en hélices α seront également à réaliser par la suite.

Conclusion et perspectives

Le travail réalisé au cours de cette thèse a permis d'atteindre les objectifs fixés au départ et ainsi de passer les étapes initiales importantes pour la mise en place de la stratégie de combinatoire dynamique imaginée. La synthèse de blocs thioesters a été réalisée sur seize substrats différents représentatifs des acides aminés naturels mais aussi d'analogues. La polyvalence de cette voie de synthèse présage la possible extension de cette méthode aux blocs correspondants au reste des acides aminés naturels de manière simple mais aussi à des blocs à chaînes osidiques ou lipidiques par exemple mais aussi à des fonctions autres que celles trouvées dans le Vivant. Sur le plan des *scaffolds* peptidiques, ce travail s'est fortement concentré sur les plateformes β -*hairpin* : onze plateformes avec cette architecture ont été synthétisés au cours de ce travail tout en faisant varier les éléments de cette structure afin d'ajuster petit à petit les différentes propriétés du *scaffold*. En parallèle, des efforts ont été fournis pour l'obtention de plateformes hélice α grâce à une stratégie originale de synthèse de peptides à agrafes lactames : trois plateformes ont ainsi été obtenues en faisant varier dans ce cas la taille et le nombre de sites cystéines.

Le développement de la méthodologie de fonctionnalisation réversible de ces plateformes a été conduit sur les architectures β -*hairpin* dans une approche pas à pas en partant de systèmes modèles simples. Ainsi, différents paramètres physico-chimiques (nombre d'équivalents, pH, agent réducteur, détergent) ont été ajustés afin d'obtenir un équilibre de DCC rapide et déplacé de manière satisfaisante vers les produits fonctionnalisés. La complexification graduelle du système tant en nombre de sites fonctionnalisables que de blocs a également permis un ajustement de ces paramètres au fur et à mesure du travail. Le parcours suivi pour la mise en place de cette méthodologie a permis d'acquérir une expertise importante pour l'analyse de ces bibliothèques par HPLC, expertise qui sera valorisée pour toutes applications futures. Au final, ce travail a permis d'arriver à une plateforme à quatre sites fonctionnalisables comme imaginée initialement. Le niveau de complexité atteint s'est limité à deux blocs introduits simultanément lors de la tentative d'application pour le criblage à l'héparine, soit une DCL d'une taille théorique de 81 composés peptidiques. Les difficultés rencontrées pour l'analyse par MS des composés issus de cette expérience ont toutefois bloqué l'avancée du travail vers cette application qui reste la perspective majeure à poursuivre.

Sur les aspects analytiques, un travail important de méthodologie a été réalisé en spectrométrie de masse MALDI-ToF afin de caractériser les membres des DCL générés tout au long de cette thèse. Les obstacles liés à la méthode de DCC et aux objets utilisés, et qui avaient été identifiés en amont, ont conduit au développement d'une méthode de choix pour la

linéarisation de peptides cycliques par digestion enzymatique à la trypsine. Cette méthode a permis d'ouvrir la voie vers des analyses par spectrométrie de masse en tandem pour la caractérisation des isomères de position générés ce qui constitue la prochaine étape clé dans cette axe du projet.

Enfin, le démarrage des analyses structurales par simulation de dynamique moléculaire d'une part et par RMN d'autre part ont permis d'apporter des éléments de réponse quant aux aspects structuraux de ce projet. Ainsi, la plateforme modèle **Hairp4-Lys** a pu être complètement caractérisée par RMN et sa bonne structuration par RMN mais aussi par MD a pu être mise en évidence. Ce travail reste à poursuivre afin de faire converger ces deux approches notamment par l'implémentation des contraintes RMN issues de l'expérience au sein du modèle informatique. Pour finir, des études comparatives préliminaires confirment la tolérance de cette plateforme à la fonctionnalisation par des thioesters et au passage en série amide. Là encore, ce travail sera à poursuivre en l'étendant à des fonctionnalités plus encombrantes ou chargées non seulement avec davantage de détails en RMN mais aussi en MD.

A court terme, il sera intéressant de continuer à augmenter la complexité du système en augmentant le nombre de blocs introduits au sein des DCL en vue d'une application de criblage de nombreux composés en « one-pot ». Des ajustements méthodologiques seront peut-être à envisager pour résoudre les problèmes liés à la détection lors de l'analyse HPLC de ces mélanges. Par ailleurs, la conception d'un modèle théorique a déjà démarré afin de modéliser les équilibres thermodynamiques de ces mélanges. Le but étant de pouvoir prédire le comportement de nouveaux blocs synthétisés d'une part, mais aussi de prédire le paysage énergétique d'une DCL d'après le comportement des blocs la composant afin de s'assurer d'une distribution relativement équitable entre les différents produits. Enfin, ce modèle sera également utile en prévision des étapes applicatives de criblage afin d'évaluer la gamme d'énergie d'interaction et donc d'affinité acceptable afin d'avoir une amplification observable.

Les étapes de criblages avec des cibles variées constituent la perspective majeure de ce travail puisqu'elles sont l'application directe du travail méthodologique réalisé. La poursuite du travail démarré pour le criblage à l'héparine et l'identification de nouvelles structures CPP reste intéressante et s'inscrit parfaitement dans les thématiques du laboratoire. D'autres cibles sont également envisagées comme le VEGF et son récepteur dans le but d'identifier des molécules anti-angiogéniques, des enzymes de type protéases à sérine dans le but d'identifier des molécules anti-inflammatoires ou encore d'autres biomolécules comme des ARN par exemple.

Pour terminer, le transfert de cette méthodologie de DCC a de nouvelles architectures est une perspective importante non seulement pour démontrer la généralité de cette stratégie mais aussi pour permettre l'identification d'inhibiteurs de PPI avec des structures toujours plus variées et complexes. Ainsi, les hélices α déjà obtenues sont un bon point de départ pour l'extension de cette stratégie à de nouvelles structures mais d'autres peuvent être imaginées comme des β -hairpin de taille plus importante, des mini-protéines utilisés en *protein grafting* comme l'APP par exemple ou encore des foldamères. Les différents aspects développés pour les β -hairpin au cours de ce travail seront donc à transférer à ces nouvelles architectures tant sur le plan de la chimie que de l'analyse en spectrométrie de masse et en analyses structurales par RMN et MD ce qui ouvre autant de perspectives pour du travail future sur ces nouveaux objets.

Experimental section

I. Building Blocks Synthesis

- **General Considerations**

About products preparation:

All reagents and solvents were purchased from commercial suppliers and used without further purification. Flash chromatography was performed using silica gel Merck 60 with 0.040-0.063 μm on a Grace Reveleris X2TM apparatus. Analytical thin-layer chromatography (TLC) was performed using silica gel Merck 60 on alumina, visualised by UV fluorescence at 254 nm and revealed with one of the following staining solutions:

- 1) Phosphomolybdic acid (10 g in 100 mL absolute EtOH)
- 2) Bromocresol green (40 mg in 100 mL absolute EtOH, slightly alkalised till color change with 0.1 M aqueous NaOH)
- 3) Ninhydrine (1.5 g in 100 mL nBuOH and 3 mL AcOH)
- 4) p-anisaldehyde (3.7 mL in 135 mL absolute EtOH, 5 mL conc. H_2SO_4 and 1.5 mL glacial AcOH)

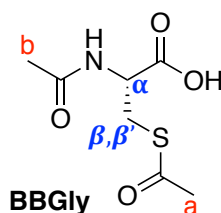
About products characterisation:

NMR spectra were recorded on Bruker Avance III nanobay 300 MHz or 400 MHz spectrometers. Proton chemical shifts are reported in parts per million (ppm) downfield from tetramethylsilane (TMS) unless noted otherwise. Coupling constants (J) are reported in Hertz (Hz). Carbon chemical shifts are reported in parts per million (ppm) with reference to internal solvent CDCl_3 (77.0 ppm) or CD_3OD (49.00 ppm). Multiplicities are abbreviated as follows: singlet (s), doublet (d), triplet (t), quartet (q), quintuplet (quint), sextuplet (sextet), septuplet, (septet), octuplet (octet), multiplet (m) and broad (br).

HRMS analyses were performed by FT-MS at the HRMS core of the Institut Parisien de Chimie Moléculaire.

Analytical HPLC was performed on a Dionex system consisted of an analytical automated LC system (Ultimate 3000) equipped with an autosampler, a pump block composed of two ternary gradient pumps and a dual wavelength detector, piloted by Chromeleon software. HPLC analysis were performed on a *Proto 200* (RP C18, 100 x 4.6 mm, 3 μm , 300 Å) from *Higgins-Analytical Inc.* with a $\text{H}_2\text{O}/\text{ACN}/0.1\%$ TFA mixture (linear gradient).

BBGly



L-Cystein hydrochloride (9.0 g, 57.1 mmol) was added to a solution of NaOH (6.4 g, 160 mmol, 2.8 eq.) in 40 mL of water at 0 °C, followed by the addition of acetic anhydride (11.3 mL, 120 mmol, 2.1 eq.). The reaction mixture was stirred for 30 minutes at 0 °C, TLC monitoring with a mixture of *n*-Butanol/Acetic Acid/Water 4:1:5 (upper phase) as the mobile phase showed complete transformation of the starting material (staining was performed with staining solution 3). The reaction mixture was acidified to pH 2 with aqueous 1M HCl then extracted with EtOAc (4 x 40 mL). The combined organic phase was dried over MgSO₄ and filtered before the solvent was removed under reduced pressure to give pure **BBGly** as a white powder in 84 % yield. ¹H-NMR (300 MHz; MeOD): δ 4.62 (dd, *J* = 8.0, 4.7 Hz, 1H, α), 3.54 (dd, *J* = 13.9, 4.7 Hz, 1H, β), 3.15 (dd, *J* = 13.9, 8.0 Hz, 1H, β'), 2.35 (s, 3H, *a*), 1.99 (s, 3H, *b*). ¹³C-NMR (75 MHz; MeOD): δ 196.3 (C_{IV}, CO-S), 173.3 (C_{IV}, CO-OH), 173.0 (C_{IV}, CO-NH), 53.2 (CH, α), 31.3 (CH₂, β), 30.3 (CH₃, *a*), 22.3 (CH₃, *b*).

RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): *t_r* = 3.8 min, > 99 % purity

HRMS *m/z* calculated for C₇H₁₁NNaO₄S [MNa⁺] Calc. *m/z* 228.0301 Obs. *m/z* 228.0303

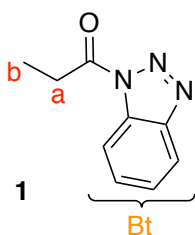
- **General Procedure A for the preparation of the N-acylbenzotriazole intermediates**

To a solution of 1H-Benzotriazole (4 eq) in DCM (1.25 mL/mmol) was added SOCl₂ (1.1-1.3 eq) at room temperature with stirring. After 30 minutes, the desired carboxylic acid or acyl chloride substrate (1.0 eq) was added in one portion and stirring was continued until complete transformation of the carboxylic acid according to TLC monitoring (eluent: Cyclohexane/AcoEt 6:4 to 8:2). Staining was performed with staining solution 1) or 2). The white precipitate was filtered off and washed with DCM. The combined organic layer was washed with aq 1M NaOH, dried over MgSO₄ and the solvent was removed under reduced pressure. The residue was washed with pentane when necessary and was used in the next step without further purification.

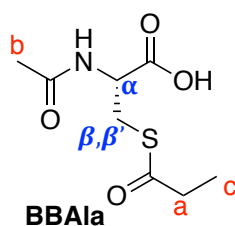
- **General Procedure B for the preparation of N-Acetyl-Cysteine thioesters from acylbenzotriazole intermediates**

To a solution of *N*-acylbenzotriazole (1.0 eq) in THF (3.33 mL/mmol) were added KHCO_3 (1.1-1.3 eq) and *N*-Acetyl-Cysteine (1.0 eq). The mixture was stirred at room temperature until complete transformation of the *N*-Acetyl-Cysteine according to TLC monitoring of the reaction (DCM/MeOH/AcOH 93:7:0.1). Staining was performed with staining solutions 3) or 4). For the thermostable substrates, heating until refluxing THF was performed to speed up the reaction rate. The solvent was removed under reduced pressure; the residue was dissolved in water. The mixture was cooled to 0°C and acidified to pH 2 with aq. HCl 1M. The aqueous phase was then extracted with DCM. The combined organic phases were dried over MgSO_4 and the solvent was removed under reduced pressure. The crude product was purified by flash chromatography on silica gel.

BBA1a



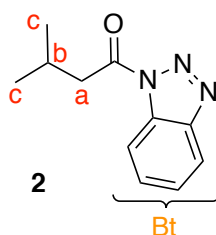
Compound **1** was obtained as an off-white solid in quantitative yield according to general procedure A starting from propanoic acid (750 μL , 10 mmol). $^1\text{H-NMR}$ (300 MHz; CDCl_3): δ 8.29 (d, $J = 8.3$ Hz, 1H, Bt), 8.11 (d, $J = 8.3$ Hz, 1H, Bt), 7.67-7.62 (m, 1H, Bt), 7.52-7.47 (m, 1H, Bt), 3.46 (q, $J = 7.4$ Hz, 2H, a), 1.42 (t, $J = 7.4$ Hz, 3H, b). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 173.3 (C_{IV} , CO-N), 146.2 (C_{IV} , Bt), 131.1 (C_{IV} , Bt), 130.3 (CH, Bt), 126.1 (CH, Bt), 120.1 (CH, Bt), 114.4 (CH, Bt), 29.1 (CH_2 , a), 8.4 (CH_3 , b).



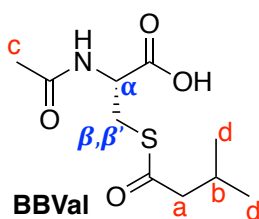
Treatment of **1** (520 mg, 3 mmol) according to the general procedure B lead to **BBAIa**. The crude product was purified by flash chromatography (20:80:0.5 Cyclohexane/EtOAc/AcOH) to afford the pure product in 30 % yield. ¹H-NMR (400 MHz; MeOD) : δ 4.58 (dd, *J* = 8.0, 4.8 Hz, 1H, **α**), 3.52 (dd, *J* = 13.9, 4.8 Hz, 1H, **β**), 3.14 (dd, *J* = 13.9, 8.0 Hz, 1H, **β'**), 2.60 (q, *J* = 7.5 Hz, 2H, **a**), 1.96 (s, 3H, **b**), 1.15 (t, *J* = 7.5 Hz, 3H, **c**). ¹³C-NMR (75 MHz; MeOD): δ 200.7 (C_{IV}, CO-S), 173.3 (C_{IV}, CO-NH), 173.2 (C_{IV}, CO-OH), 53.5 (CH, **α**), 38.0 (CH₂, **β**), 31.1 (CH₂, **a**), 22.4 (CH₃, **b**), 9.9 (CH₃, **c**).

RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): t_r = 4.7 min, 98 % purity
 HRMS m/z calculated for C₈H₁₃NNaO₄S [MNa⁺] Calc. m/z 242.0457 Obs. m/z 242.0458

BBVal



Compound **2** was obtained as a white solid, in quantitative yield according to the general procedure A, starting from isovaleryl chloride (1.2 mL, 10 mmol). ¹H-NMR (300MHz, CDCl₃) δ 8.26 (dt, *J* = 8.1, 0.9, 1H, **Bt**), 8.07 (dt, *J* = 8.1, 0.9, 1H, **Bt**), 7.61 (ddd, *J* = 8.1, 7.2, 0.9, 1H, **Bt**), 7.46 (ddd, *J* = 8.1, 7.2, 0.9, 1H, **Bt**), 3.27 (d, *J* = 6.9, 2H, **a**), 2.42 (octet, *J* = 6.9, 1H, **b**), 1.08 (d, *J* = 6.9, 6H, **c**). ¹³C-NMR (75MHz, CDCl₃) δ 172 (C_{IV}, CO-N), 146.2 (C_{IV}, **Bt**), 131.2 (C_{IV}, **Bt**), 130.3 (CH, **Bt**), 126.1 (CH, **Bt**), 120.2 (CH, **Bt**), 114.6 (CH, **Bt**), 44.2 (CH₂, **a**), 25.7 (CH₂, **b**), 22.6 (CH₃, **c**).

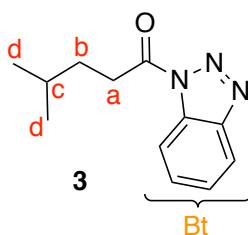


Treatment of **2** (609 mg, 3 mmol) according to the general procedure B lead to **BBVal**. The crude product was purified by flash chromatography (97:3:0.2 DCM/MeOH/AcOH) to afford the pure product in 45 % yield. ¹H-NMR (300MHz, CD₃OD) δ 4.58 (dd, *J* = 8.2, 4.7, 1H, **α**), 3.51 (dd, *J* = 13.9, 4.7, 1H, **β**), 3.15 (dd, *J* = 13.9, 8.2, 1H, **β'**), 2.46 (d, *J* = 6.8, 2H, **a**), 2.12 (octet, *J* = 6.8, 1H, **b**), 1.97 (s, 3H, **c**), 0.95 (d, *J* = 6.8, 6H, **d**) ; ¹³C-NMR (75MHz, CD₃OD) δ 199.3 (C_{IV}, CO-S), 173.2 (C_{IV}, CO-NH), 173 (C_{IV}, CO-OH), 53.5 (CH₂, **a**), 53.3 (CH, **α**), 31.1 (CH₂, **β**), 27.6 (CH, **b**), 22.5 (CH₃, **c**), 22.3 (CH₃, **d**).

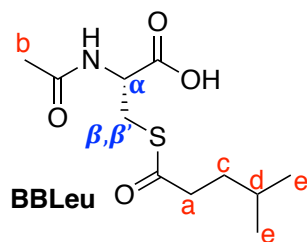
RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): t_r = 5.9 minutes, 97 % purity

HRMS m/z calculated for C₁₀H₁₇NNaO₄S [MNa⁺] Calc. m/z 270.0770 Obs. m/z 270.0772

BBLeu



Compound **3** was obtained as a white solid in quantitative yield according to the general procedure A, starting from 4-methylvaleric acid (1.3 mL, 10 mmol). ¹H-NMR (300MHz, CDCl₃) δ 8.27 (dt, *J* = 8.1, 0.9, 1H, **Bt**), 8.09 (dt, *J* = 8.1, 1.2, 1H, **Bt**), 7.62 (ddd, *J* = 8.1, 7.1, 0.9, 1H, **Bt**), 7.48 (ddd, *J* = 8.1, 7.1, 1.2, 1H, **Bt**), 3.41 (m, 2H, **a**), 1.74-1.83 (m, 3H, **b** and **c**), 0.99 (d, *J* = 4.2, 6H, **d**). ¹³C-NMR (75MHz, CDCl₃) δ 172.7 (C_{IV}, CO-N), 146.1 (C_{IV}, **Bt**), 131 (C_{IV}, **Bt**), 130.2 (CH, **Bt**), 125.9 (CH, **Bt**), 120 (CH, **Bt**), 114.3 (CH, **Bt**), 33.5 (CH₂, **a**), 33.1 (CH₂, **b**), 27.7 (CH, **c**), 22.1 (CH₃, **d**).

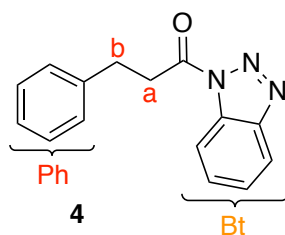


Treatment of **3** (651 mg, 3 mmol) according to the general procedure B lead to **BBLeu**. The crude product was purified by flash chromatography (97:3:0.2 DCM/MeOH/AcOH) to afford the pure product in 53 % yield. ¹H-NMR (300MHz, CD₃OD) δ 4.58 (dd, *J* = 8.1, 4.7, 1H, α), 3.51 (dd, *J* = 13.9, 4.7, 1H, β), 3.14 (dd, *J* = 13.9, 8.1, 1H, β'), 2.59 (m, 2H, a), 1.96 (s, 3H, b), 1.53-1.57 (m, 3H, c and d), 0.90 (d, *J* = 6.6, 6H, e). ¹³C-NMR (75MHz, CD₃OD) δ 200 (C_{IV}, CO-S), 173.2 (C_{IV}, CO-NH), 173 (C_{IV}, CO-OH), 53.3 (CH, α), 42.8 (CH₂, a), 35.5 (CH₂, c), 31.1 (CH₂, β), 28.7 (CH, d), 22.6 (CH₃, e), 22.3 (CH₃, b).

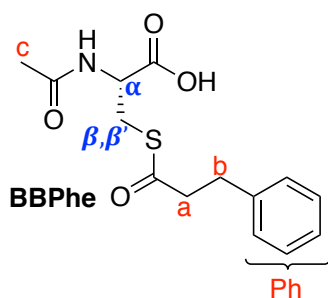
RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): *t_r* = 6.6 minutes, 95 % purity

HRMS *m/z* calculated for C₁₁H₁₉NNaO₄S [MNa⁺] Calc. *m/z* 284.0927 Obs. *m/z* 284.0929

BBPhe



Compound **4** was obtained as a white solid in 68 % yield according to the general procedure A, starting from hydrocinnamoic acid (1.5 g, 10 mmol). ¹H-NMR (300 MHz; CDCl₃): δ 8.28 (dt, *J* = 8.3, 0.9 Hz, 1H, Bt), 8.11 (dt, *J* = 8.3, 0.9 Hz, 1H, Bt), 7.64 (ddd, *J* = 8.3, 7.2, 0.9 Hz, 1H, Bt), 7.49 (ddd, *J* = 8.3, 7.2, 0.9 Hz, 1H, Bt), 7.33-7.28 (m, 4H, Ph), 7.25-7.19 (m, 1H, Ph), 3.77 (t, *J* = 7.7 Hz, 2H, a), 3.24 (t, *J* = 7.7 Hz, 2H, b). ¹³C-NMR (75 MHz; CDCl₃): δ 171.7 (C_{IV}, CO-N), 146.2 (C_{IV}, Bt), 139.9 (C_{IV}, Ph), 131.2 (C_{IV}, Bt), 130.5 (CH, Bt), 128.7 (CH, Ph), 128.5 (CH, Ph), 126.6 (CH, Ph), 126.2 (CH, Bt), 120.2 (CH, Bt), 114.5 (CH, Bt), 37.2 (CH₂, a), 30.3 (CH₂, b).

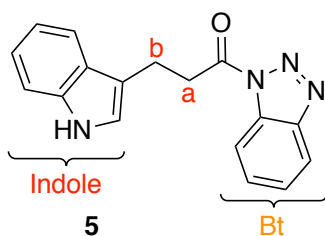


Treatment of **4** (753 mg, 3 mmol) according to the general procedure B lead to **BBPhe**. The crude product was purified by flash chromatography (98:2:0.2 DCM/MeOH/AcOH) to afford the pure product in 54 % yield. ¹H-NMR (300MHz, CD₃OD) δ 7.24-7.28 (m, 2H, Ph), 7.17 (td, *J* = 7.3, 4.3, 3H, Ph), 4.58 (dd, *J* = 8.1, 4.6, 1H, α), 3.52 (dd, *J* = 13.9, 4.6, 1H, β), 3.13 (dd, *J* = 13.9, 8.1, 1H, β'), 2.86-2.97 (m, 4H, a and b), 1.94 (s, 3H, c); ¹³C-NMR (75MHz, CD₃OD) δ 199.1 (C_{IV}, CO-S), 173.2 (C_{IV}, CO-NH), 173.1 (C_{IV}, CO-OH), 141.4 (C_{IV}, Ph), 129.5 (CH, Ph), 129.3 (CH, Ph), 127.3 (CH, Ph), 53.4 (CH, α), 46.2 (CH₂, a), 32.3 (CH₂, b), 31.1 (CH₂, β), 22.4 (CH₂, c).

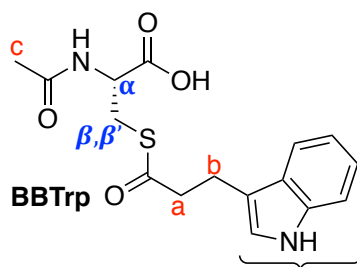
RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): t_r = 6.6 minutes, 97 % purity

HRMS m/z calculated for C₁₄H₁₇NNaO₄S [MNa⁺] Calc. m/z 318.0770 Obs. m/z 318.0772

BBTrp



3-indole propionic acid (1.89 g, 10 mmol) was treated according to general procedure A to give **5** as a dark yellow solid in quantitative yield. ¹H-NMR (400 MHz; CDCl₃): δ 8.31 (dt, *J* = 8.3, 0.9 Hz, 1H, Bt), 8.12 (dt, *J* = 8.3, 0.9 Hz, 1H, Bt), 8.03 (br, 1H, NH Indole), 7.70 (ddt, *J* = 7.8, 1.3, 0.7 Hz, 1H, Indole), 7.65 (ddd, *J* = 8.3, 7.2, 0.9 Hz, 1H, Bt), 7.50 (ddd, *J* = 8.3, 7.2, 0.9 Hz, 1H, Bt), 7.36 (dt, *J* = 8.0, 1.0 Hz, 1H, Indole), 7.21 (ddd, *J* = 8.0, 7.0, 1.0 Hz, 1H, Indole), 7.17-7.12 (m, 2H, Indole), 3.85 (t, *J* = 7.5 Hz, 2H, a), 3.41 (t, *J* = 7.5 Hz, 2H, b). ¹³C-NMR (75 MHz; CDCl₃): δ 172.2 (C_{IV}, CO-N), 146.3 (C_{IV}, Bt), 136.4 (C_{IV}, Indole), 131.2 (C_{IV}, Bt), 130.5 (C_{IV}, Indole), 127.2 (CH, Bt), 126.2 (CH, Bt), 122.3 (CH, Indole), 121.9 (CH, Indole), 120.2 (CH, Indole), 119.6 (CH, Bt), 118.8 (CH, Indole), 114.6 (C_{IV}, Indole), 114.4 (CH, Bt), 111.3 (CH, Indole), 36.4 (CH₂, a), 20.2 (CH₂, b).

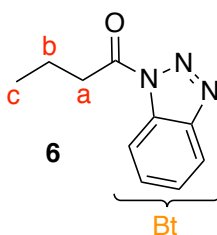


Treatment of **5** (870 mg, 3 mmol) according to the general procedure B lead to **BBTrp**. The crude product was purified by flash chromatography (20:80:0.5 Cyclohexane/EtOAc/AcOH) to afford the pure product as a light orange solid in 94 % yield. $^1\text{H-NMR}$ (300 MHz; MeOD): δ 7.48 (d, $J = 7.7$ Hz, 1H, **Indole**), 7.32 (d, $J = 8.0$ Hz, 1H, **Indole**), 7.11-6.98 (m, 3H, **Indole**), 4.61 (dd, $J = 8.0, 4.6$ Hz, 1H, α), 3.51 (dd, $J = 13.9, 4.6$ Hz, 1H, β), 3.15 (dd, $J = 13.9, 8.0$ Hz, 1H, β'), 3.06 (t, $J = 7.4$ Hz, 2H, **a**), 2.90 (t, $J = 7.2$ Hz, 2H, **b**), 1.90 (s, 3H, **c**). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 198.6 (C_{IV} , CO-S), 171.9 (C_{IV} , CO-NH and CO-OH), 136.7 (C_{IV} , **Indole**), 127.0 (C_{IV} , **Indole**), 121.8 (CH, **Indole**), 121.0 (CH, **Indole**), 118.3 (CH, **Indole**), 117.7 (CH, **Indole**), 112.9 (CH, **Indole**), 110.9 (C_{IV} , **Indole**), 52.3 (CH, α), 44.3 (CH_2 , **a**), 29.9 (CH_2 , β), 21.0 (CH_2 , **b**), 20.9 (CH_2 , **c**).

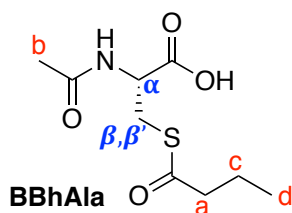
RP-HPLC ($\text{H}_2\text{O}/\text{MeCN}$, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): $t_{\text{r}} = 7.9$ min, 99 % purity.

HRMS m/z calculated for $\text{C}_{16}\text{H}_{18}\text{N}_2\text{NaO}_4\text{S}$ [MNa^+] Calc. m/z 357.0879 Obs. m/z 357.0880

BBhAla



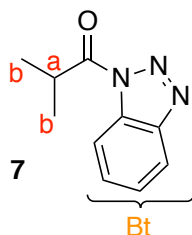
Compound **6** was obtained as a white solid in 99 % yield according to the general procedure A starting from butanoic acid (914 μL , 10 mmol). $^1\text{H-NMR}$ (400MHz, CDCl_3) δ 8.29 (dt, $J = 8.3, 0.9$, 1H, **Bt**), 8.11 (dt, $J = 8.3, 0.9$, 1H, **Bt**), 7.64 (ddd, $J = 8.3, 7.1, 0.9$, 1H, **Bt**), 7.49 (ddd, $J = 8.3, 7.1, 0.9$, 1H, **Bt**), 3.39 (t, $J = 5.8$, 2H, **a**), 1.95 (sextet, $J = 5.8$, 2H, **b**), 1.10 (t, $J = 5.4$, 3H, **c**). $^{13}\text{C-NMR}$ (75MHz, CDCl_3) δ 172.7 (C_{IV} , CO-N), 146.4 (C_{IV} , **Bt**), 131.3 (C_{IV} , **Bt**), 130.5 (CH, **Bt**), 126.2 (CH, **Bt**), 120.3 (CH, **Bt**), 114.6 (CH, **Bt**), 37.5 (CH_2 , **a**), 18.2 (CH_2 , **b**), 13.8 (CH_2 , **c**).



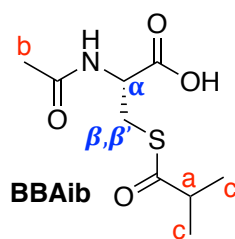
Treatment of **6** (572 mg, 3 mmol) according to the general procedure B lead to **BBhAla**. The crude product was purified by flash chromatography (9:1:0.1 DCM/MeOH/AcOH) to afford the pure product in 24 % yield. ¹H-NMR (300MHz, CD₃OD) δ 4.58 (dd, *J* = 8, 4.6, 1H, **α**), 3.52 (dd, *J* = 13.9, 4.7, 1H, **β**), 3.14 (dd, *J* = 13.9, 8.1, 1H, **β'**), 2.56 (t, *J* = 7.3, 2H, **a**), 1.96 (s, 3H, **b**), 1.67 (sextet, *J* = 7.3, 2H, **c**), 0.94 (t, *J* = 7.3, 3H, **d**). ¹³C-NMR (75MHz, CD₃OD) δ 200.2 (C_{IV}, CO-S), 176.1 (C_{IV}, CO-NH), 172.2 (C_{IV}, CO-OH), 54.3 (CH, **α**), 45.9 (CH₂, **a**), 30.6 (CH₂, **β**), 22.9 (CH₂, **c**), 19.2 (CH₃, **b**), 13.6 (CH₃, **d**).

RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): t_r = 5.3 min, 96 % purity
 HRMS m/z calculated for C₉H₁₅NNaO₄S [MNa⁺] Calc. m/z 256.0614 Obs. m/z 256.0615

BBAib



Compound **7** was obtained as a white solid in quantitative yield according to the general procedure A starting from isobutyric acid (926 μL, 10 mmol). ¹H-NMR (300 MHz; CDCl₃): δ 8.29 (dt, *J* = 8.3, 0.9 Hz, 1H, **Bt**), 8.11 (dt, *J* = 8.3, 1.0 Hz, 1H, **Bt**), 7.64 (ddd, *J* = 8.3, 7.1, 1.1 Hz, 1H, **Bt**), 7.49 (ddd, *J* = 8.2, 7.1, 1.1 Hz, 1H, **Bt**), 4.15 (septet, *J* = 6.9 Hz, 1H, **a**), 1.43 (d, *J* = 6.9 Hz, 6H, **b**). ¹³C-NMR (75 MHz; CDCl₃): δ 176.8 (C_{IV}, CO-N), 146.3 (C_{IV}, **Bt**), 131.4 (C_{IV}, **Bt**), 130.4 (CH, **Bt**), 126.2 (CH, **Bt**), 120.2 (CH, **Bt**), 114.7 (CH, **Bt**), 34.1 (CH₂, **a**), 19.2 (CH₃, **b**).

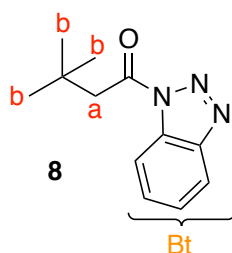


Treatment of **7** (572 mg, 3 mmol) according to the general procedure B lead to **BBAib**. The crude product was purified by silica gel column chromatography (95:5:0.2 CH₂Cl₂/MeOH/AcOH) to afford the pure product in 49 % yield. ¹H-NMR (300 MHz; MeOD): δ 4.56 (dd, *J* = 8.1, 4.6 Hz, 1H, **α**), 3.50 (dd, *J* = 13.8, 4.6 Hz, 1H, **β**), 3.14 (dd, *J* = 13.8, 8.1 Hz, 1H, **β'**), 2.76 (7, *J* = 6.9 Hz, 1H, **a**), 1.96 (s, 3H, **b**), 1.17 (d, *J* = 6.9 Hz, 6H, **c**). ¹³C-NMR (75 MHz; MeOD): δ 204.6 (C_{IV}, CO-S), 173.2 (C_{IV}, CO-NH and CO-OH), 53.6 (CH, **α**), 44.2 (CH₂, **a**), 31.0 (CH₂, **β**), 22.4 (CH₃, **b**), 19.7 and 19.6 (2CH₂, **c**).

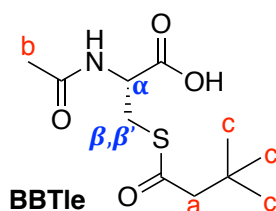
RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): *t_r* = 5.2 min , > 99 % purity

HRMS *m/z* calculated for C₉H₁₇NO₄S [MH⁺] Calc. *m/z* 234.0795 Obs. *m/z* 234.0793

BBTle



Compound **8** was obtained as a white solid in quantitative yield according to the general procedure A starting from *tert*-butylacetic acid (1.3 mL, 10 mmol). ¹H-NMR (300MHz, CDCl₃) δ 8.29 (d, *J* = 8.1, 1H, **Bt**), 8.07 (d, *J* = 8.1, 1H, **Bt**), 7.60 (td, *J* = 8.1, 0.6, 1H, **Bt**), 7.45 (td, *J* = 8.1, 0.6, 1H, **Bt**), 3.32 (s, 2H, **a**), 1.14 (s, 9H, **b**). ¹³C-NMR (75MHz, CDCl₃) δ 171.4 (C_{IV}, CO-N), 146.3 (C_{IV}, **Bt**), 131.2 (C_{IV}, **Bt**), 130.3 (CH, **Bt**), 126.1 (CH, **Bt**), 120.2 (CH, **Bt**), 114.7 (CH, **Bt**), 42.3 (CH₂, **a**), 32.1 (C_{IV}, *t*Bu), 29.7 (CH₃, **b**).

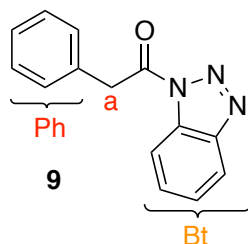


Treatment of **8** (615 mg, 3 mmol) according to the general procedure B lead to **BBTle**. The crude product was purified by flash chromatography (97:3:0.2 CH₂Cl₂/MeOH/AcOH) to afford the pure product in 47% yield. ¹H NMR (300MHz, CD₃OD) δ 4.58 (dd, *J* = 8.2, 4.7, 1H, **α**), 3.49 (dd, *J* = 13.9, 4.7, 1H, **β**), 3.14 (dd, *J* = 13.9, 8.3, 1H, **β'**), 2.47 (s, 2H, **a**), 1.96 (s, 3H, **b**), 1.02 (s, 9H, **c**). ¹³C-NMR (75MHz, CD₃OD) δ 198.3 (C_{IV}, CO-S), 173.1 (C_{IV}, CO-NH), 173 (C_{IV}, CO-OH), 57.4 (CH₂, **a**), 53.3 (CH, **α**), 32.3 (CH₂, **β**), 31.3 (C_{IV}, tBu), 30 (CH₃, **c**), 22.4 (CH₃, **b**).

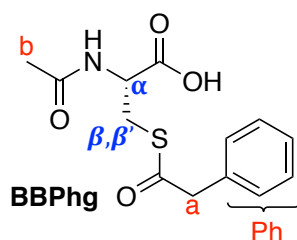
RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100% MeCN in 10 minutes): *t_r* = 6.6 minutes, > 99 % purity

HRMS *m/z* calculated for C₁₁H₁₉NNaO₄S [MNa⁺] Calc. *m/z* 284.0927 Obs. *m/z* 284.0926

BBPhg



Compound **9** was obtained as a white solid in 83 % yield according to the general procedure A starting from phenylacetic acid (3.6 g, 10 mmol). ¹H-NMR (300 MHz; CDCl₃): δ 8.30 (d, *J* = 8.2 Hz, 1H, **Bt**), 8.16 (d, *J* = 8.2 Hz, 1H, **Bt**), 7.67 (t, *J* = 7.7 Hz, 1H, **Bt**), 7.56-7.50 (m, 3H, **Bt** and **Ph**), 7.40-7.28 (m, 3H, **Ph**), 4.73 (s, 2H, **a**). ¹³C-NMR (75 MHz; CDCl₃): δ 170.4 (C_{IV}, CO-N), 146.4 (C_{IV}, **Bt**), 132.6 (C_{IV}, **Bt**), 130.6 (CH, **Bt**), 130.0, 129.5, 129.0, 128.8, 127.8, 127.4, 126.4 (CH, **Bt**), 120.3 (CH, **Bt**), 114.6 (CH, **Bt**), 42.2 (CH₂, **a**).

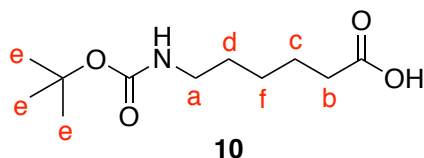


Treatment of **9** (711 mg, 3 mmol) according to the general procedure B lead to **BBPhg**. The crude product was purified by flash chromatography (95:5:0.3 DCM/MeOH/AcOH) to afford the pure product in 96% yield. ¹H-NMR (300 MHz; MeOD): δ 7.34-7.26 (m, 5H, **Ph**), 4.57 (dd, *J* = 8.2, 4.6 Hz, 1H, **α**), 3.87 (s, 2H, **a**), 3.49 (dd, *J* = 13.9, 4.6 Hz, 1H, **β**), 3.14 (dd, *J* = 13.9, 8.3 Hz, 1H, **β'**), 1.90 (s, 3H, **b**). ¹³C-NMR (75 MHz; MeOD): δ 196.8 (C_{IV}, CO-S), 171.88 (C_{IV}, CO-NH), 171.71 (C_{IV}, CO-OH), 133.7 (C_{IV}, **Ph**), 129.3 (CH, **Ph**), 128.2 (CH, **Ph**), 127.0 (CH, **Ph**), 51.9 (CH, **α**), 49.6 (CH₂, **a**), 30.0 (CH₂, **β**), 20.9 (CH₃, **b**).

RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100% MeCN in 10 minutes): *t_r* = 6.2 minutes, > 99 % purity

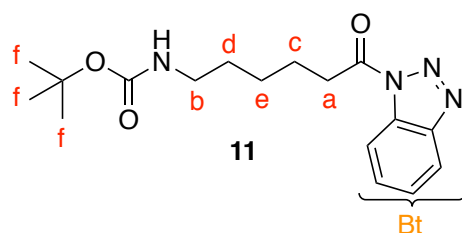
HRMS *m/z* calculated for C₁₃H₁₅NNaO₄S [MNa⁺] Calc. *m/z* 304.0614 Obs. *m/z* 304.0613

BBLys

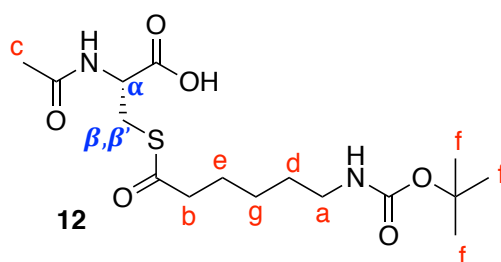


6-aminohexanoic acid (10 g, 76.3 mmol) was dissolved in 300 mL of a 1,4-dioxane/water (2:1). Potassium carbonate (21 g, 152.4 mmol, 2 eq.) and di-*tert*-butyl dicarbonate (18.3 g, 83.9 mmol, 1.1 eq.) were added and the mixture was stirred overnight at room temperature. TLC monitoring using DCM/MeOH/AcOH 98:2:0.1 as the mobile phase (staining performed with staining solution 3)) showed complete transformation of the starting material. Volatiles were removed under reduced pressure. The aqueous phase was washed with EtOAc (3 x 50 mL) then cooled to 0°C, acidified to pH 2 with 1M aqueous HCl and extracted with EtOAc (3 x 50 mL). The combined organic layer was dried over MgSO₄ and filtered before the solvent was removed under reduced pressure to afford the Boc-protected 6-aminohexanoic acid **10** in 87 % yield, as a colourless oil which solidified upon storage. ¹H-NMR (300 MHz; CDCl₃): δ 3.10 (m, 2H, **a**), 2.34 (t, *J* = 7.4 Hz, 2H, **b**), 1.64 (quintet, *J* = 7.5 Hz, 2H, **c**), 1.54-1.47 (m, 2H, **d**), 1.43 (s, 9H, **e**), 1.40-1.33 (m, 2H, **f**). ¹³C-NMR (75 MHz; CDCl₃): δ 179.0 (C_{IV}, CO-OH), 160.4 (C_{IV}, O-

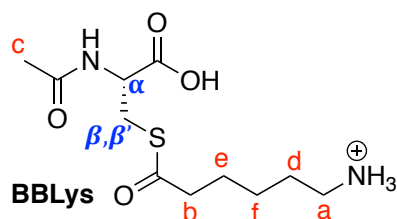
CO-NH), 80.3 (C_{IV}, tBu), 40.5 (CH₂, a), 34.0 (CH₂, b), 29.8 (CH₂, c), 28.6 (CH₃, e), 26.3 (CH₂, f) 24.5 (CH₂, d).



Compound **10** (2.31 g, 10 mmol) was treated according to general procedure A to give **11** as a white solid in 88 % yield. ¹H-NMR (300 MHz; CDCl₃): δ 8.28 (dt, *J* = 8.3, 0.8 Hz, 1H, Bt), 8.11 (dt, *J* = 8.3, 0.8 Hz, 1H, Bt), 7.65 (ddd, *J* = 8.3, 7.2, 0.8 Hz, 1H, Bt), 7.50 (ddd, *J* = 8.3, 7.2, 0.8 Hz, 1H, Bt), 3.42 (t, *J* = 7.4 Hz, 2H, a), 3.15 (m, 2H, b), 1.93 (quintet, *J* = 7.4 Hz, 2H, c), 1.61-1.47 (m, 4H, d and e), 1.43 (s, 9H, f). ¹³C NMR (75 MHz; CDCl₃): δ 172.5 (C_{IV}, CO-N), 156.1 (C_{IV}, O-CO-NH), 146.3 (C_{IV}, Bt), 131.2 (C_{IV}, Bt), 130.5 (CH, Bt), 126.2 (CH, Bt), 120.3 (CH, Bt), 114.6 (CH, Bt), 79.3 (C_{IV}, tBu), 40.5 (CH₂, a), 35.5 (CH₂, b), 30.0 (CH₂, c), 28.5 (CH₃, f), 26.4 (CH₂, d), 24.2 (CH₂, e).



Treatment of **11** (997 mg, 3 mmol) according to the general procedure B lead to the Boc-protected **BBLys 12**. The crude product was purified by flash chromatography (20:80:0.5 Cyclohexane/EtOAc/AcOH) to afford the pure product in 76 % yield. ¹H-NMR (300 MHz; MeOD): δ 4.59 (dd, *J* = 8.0, 4.7 Hz, 1H, α), 3.52 (dd, *J* = 13.9, 4.7 Hz, 1H, β), 3.14 (dd, *J* = 13.9, 8.1 Hz, 1H, β'), 3.02 (t, *J* = 6.9 Hz, 2H, a), 2.59 (t, *J* = 7.3 Hz, 2H, b), 1.97 (s, 3H, c), 1.66 (quintet, *J* = 7.5 Hz, 2H, d), 1.51-1.47 (m, 2H, e), 1.43 (s, 9H, f), 1.39-1.29 (m, 2H, g). ¹³C-NMR (75 MHz; MeOD): δ 199.9 (C_{IV}, CO-S), 175.2 (C_{IV}, CO-NH), 173.2 (C_{IV}, CO-OH), 157.0 (C_{IV}, O-CO-NH), 79.8 (C_{IV}, tBu), 53.4 (CH, α), 44.6 (CH₂, b), 41.1 (CH₂, a), 31.1 (CH₂, β), 28.8 (CH₂, b), 27.1 (CH₂, g), 26.3 (CH₂, e), 22.4 (CH₃, c).



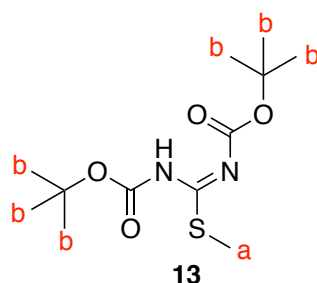
This compound (851 mg, 2.26 mmol) was dissolved in DCM (7 mL) and treated with TFA (5 mL). The mixture was stirred at room temperature for 2h30. The solvent was removed under reduced pressure. The residue was dissolved in water and freeze-dried to afford the desired product **BBLys** as a light yellow oil in quantitative yield. ¹H-NMR (300 MHz; MeOD): δ 4.59 (dd, *J* = 8.4, 4.9 Hz, 1H, α), 3.54 (dd, *J* = 13.9, 4.9 Hz, 1H, β), 3.09 (dd, *J* = 13.9, 8.4 Hz, 1H, β'), 2.92 (t, *J* = 7.6 Hz, 2H, a), 2.63 (t, *J* = 7.1 Hz, 2H, b), 1.97 (s, 3H, c), 1.73-1.60 (m, 4H, d), 1.42 (m, 2H, e). ¹³C-NMR (75 MHz; MeOD): δ 199.9 (C_{IV}, CO-S), 173.2 (C_{IV}, CO-OH and CO-NH), 53.2 (CH, α), 44.2 (CH₂, b), 40.5 (CH₂, a), 31.2 (CH₂, β), 28.1 (CH₂, d), 26.5 (CH₂, f), 25.9 (CH₂, e), 22.4 (CH₃, c).

RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100% MeCN in 10 minutes): *t_r* = 3.8 min, 99 % purity.

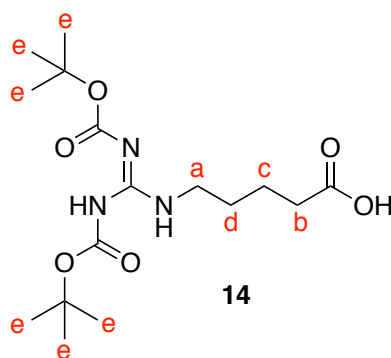
HRMS *m/z* calculated for C₁₁H₂₁N₂O₄S [M⁺] Calc. *m/z* 277.1217 Obs. *m/z* 277.1218

BBArg

1,3-Di-Boc-2-methylisothiourrea route:



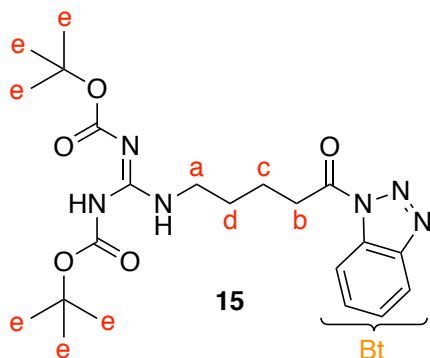
S-Methylisothiourrea hemisulfate salt (6.96 g, 25 mmol) was dissolved in a 1:1 water/1,4-dioxane mixture (120 mL) then 1M aqueous NaOH (50 mL) and di-*tert*-butyl dicarbonate (27 g, 125 mmol, 5 eq.) were added. The mixture was stirred overnight at room temperature. TLC monitoring showed complete transformation of the starting material (*n*-BuOH/AcOH/H₂O; 4:1:5). The precipitate was filtered and washed with water. The volume of the filtrate was reduced by half under reduced pressure then filtered one more time and the solid was washed with water. The solids were combined and suspended in water (500 mL). The suspension was stirred for 15 minutes at 50°C then filtered. The white solid was dried under vacuum to afford the desired product **13** in 95% yield. ¹H-NMR (300 MHz; CDCl₃): δ 11.58 (s, 1H, NH), 2.39 (s, 3H, a), 1.51 (s, 18H, b).



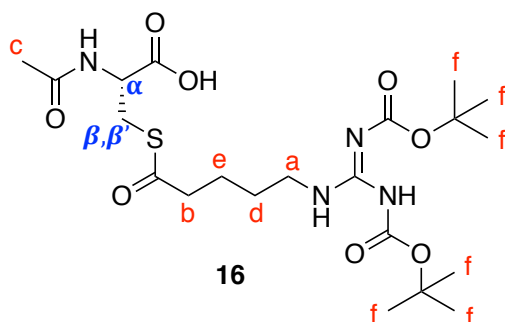
5-aminopentanoic acid (1.76 g, 15 mmol) was dissolved in dry DMF (100 mL). Compound **13** (5.23 g, 18 mmol, 1.2 eq.), triethylamine (5.3 mL, 38 mmol, 2.5 eq.) and mercury(II) chloride (7.06 g, 26 mmol, 1.7 eq.) were added. The mixture was stirred at room temperature for 3 days. TLC monitoring showed complete transformation of the starting material (*n*-BuOH/AcOH/H₂O ; 4:1:5, staining solution 3)). The reaction mixture was poured in water then extracted with EtOAc. The combined organic layers were dried over MgSO₄ and filtered before the solvent was removed under reduced pressure. The crude product was purified on silica gel column chromatography with DCM/MeOH (98:2) as the mobile phase to afford the desired product **14** as a white solid in 65% yield (see NMR characterization below).

N,N'-Di-Boc-1*H*-pyrazole-1-carboxamide route :

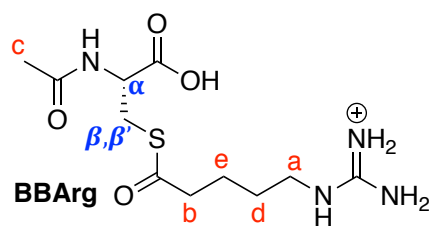
5-aminopentanoic acid (1.57 g, 13 mmol) and *N,N'*-Di-Boc-1*H*-pyrazole-1-carboxamide (4.44 g, 14.3 mmol, 1.1 eq) were suspended in pyridine (40 mL). The mixture was stirred at room temperature for 2 days, until the mixture became clear. TLC monitoring of the reaction showed complete transformation of the 5-aminopentanoic acid (*n*-BuOH/AcOH/H₂O; 4:1:5, staining with staining solution 3)). NaOH 1M (60 mL) was added to the reaction mixture, which was then stirred for 15 min. The aqueous phase was extracted with EtOAc (3x50 mL), acidified to pH 3 with aqueous 1M HCl then re-extracted with EtOAc (3x50 mL). The organic phase was dried over MgSO₄ and filtered before the solvent was removed under vacuum. Compound **14** was obtained as a white solid in 86% yield (4.02 g). H-NMR (300 MHz; CDCl₃): δ 8.34 (s, 1H, NH), 3.42 (q, *J* = 5.8 Hz, 2H, **a**), 2.38 (t, *J* = 7.0 Hz, 2H, **b**), 1.72-1.59 (m, 4H, **c** and **d**), 1.48 (s, 18H, **e**). ¹³C-NMR (75 MHz; CDCl₃) : δ 178.7 (C_{IV}, CO-OH), 163.6 (C_{IV}, HN-CN-NH), 156.3 (C_{IV}, O-CO-N), 153.4 (C_{IV}, O-CO-NH), 83.2 (C_{IV}, tBu), 79.5 (C_{IV}, tBu), 40.6 (CH₂, **a**), 33.7 (CH₂, **b**), 28.6 (CH₂, **d**), 28.4 (CH₃, **e**), 28.2 (CH₂, **e**), 22.1 (CH₂, **c**).



Compound **14** (3.65 g, 10 mmol) was treated according to general procedure A to give **15** as a white solid in 79% yield. $^1\text{H-NMR}$ (400 MHz; CDCl_3): δ 8.29 (dt, $J = 8.3, 0.9$ Hz, 1H, **Bt**), 8.12 (dt, $J = 8.3, 0.9$ Hz, 1H, **Bt**), 7.66 (ddd, $J = 8.3, 7.2, 0.9$ Hz, 1H, **Bt**), 7.51 (ddd, $J = 8.3, 7.2, 0.9$ Hz, 1H, **Bt**), 3.55-3.47 (m, 4H, **a** and **b**), 2.02-1.94 (m, 2H, **c**), 1.84-1.75 (m, 2H, **d**), 1.50-1.48 (m, 18H, **e**). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 172.0 (C_{IV} , CO-N), 163.5 (C_{IV} , HN-CN-NH), 156.2 (C_{IV} , O-CO-N), 153.2 (C_{IV} , O-CO-NH), 146.1 (C_{IV} , **Bt**), 131.0 (C_{IV} , **Bt**), 130.3 (CH, **Bt**), 126.0 (CH, **Bt**), 120.1 (CH, **Bt**), 114.3 (CH, **Bt**), 83.0 (C_{IV} , tBu), 79.1 (C_{IV} , tBu), 40.3 (CH_2 , **a**), 35.0 (CH_2 , **d**), 28.5 (CH_2 , **b**), 28.2 (CH_2 , **e**), 28.0 (CH_2 , **e**), 21.5 (CH_2 , **c**).



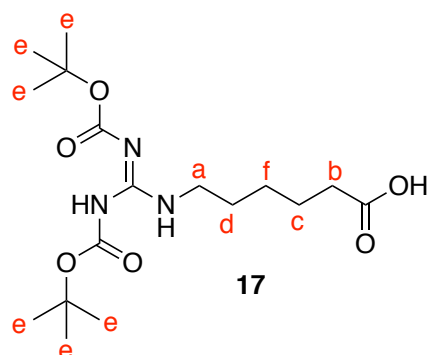
Treatment of **15** (1.38 g, 3 mmol) according to the general procedure B lead to the diBoc-protected **BBArg** **16**. The crude product was purified by flash chromatography (20:80:0.5 Cyclohexane/EtOAc/AcOH) to afford the pure product in 30% yield. $^1\text{H-NMR}$ (300 MHz; MeOD): δ 4.59 (dd, $J = 7.8, 4.6$ Hz, 1H, α), 3.53 (dd, $J = 13.9, 4.6$ Hz, 1H, β), 3.37 (t, $J = 6.8$ Hz, 2H, **a**), 3.16 (dd, $J = 13.9, 7.8$ Hz, 1H, β), 2.65 (t, $J = 7.1$ Hz, 2H, **b**), 1.97 (s, 3H, **c**), 1.73-1.58 (m, 4H, **d** and **e**), 1.52 (s, 9H, **f**), 1.47 (s, 9H, **f**). $^{13}\text{C-NMR}$ (101 MHz; MeOD): δ 199.7 (C_{IV} , CO-S), 173.1 (C_{IV} , CO-NH), 173.0 (C_{IV} , CO-OH), 164.2 (C_{IV} , HN-CN-NH), 157.4 (C_{IV} , O-CO-N), 154.1 (C_{IV} , O-CO-NH), 84.5 (C_{IV} , tBu), 80.4 (C_{IV} , tBu), 53.6 (CH, α), 44.1 (CH_2 , **b**), 41.3 (CH_2 , **a**), 31.3 (CH_2 , β), 29.3 (CH_2 , **d**), 28.6 (CH_3 , **f**), 28.3 (C CH_3 , **f**), 23.7 (CH_2 , **e**), 22.5 (CH_2 , **c**).



Compound **16** (580 mg, 1.15 mmol) was dissolved in DCM (7 mL) and treated with TFA (5 mL). The mixture was stirred at room temperature for 2h30. The solvent was removed under reduced pressure. The residue was dissolved in water and freeze-dried to afford the desired product **BBArg** as a light yellow oil in quantitative yield. ¹H-NMR (300 MHz; MeOD): δ 4.60 (dd, $J = 8.5, 4.7$ Hz, 1H, α), 3.55 (dd, $J = 13.9, 4.7$ Hz, 1H, β), 3.18 (t, $J = 13.7$ Hz, 2H, a), 3.09 (dd, $J = 13.9, 8.5$ Hz, 1H, β'), 2.65 (t, $J = 7.0$ Hz, 2H, b), 1.97 (s, 3H, c), 1.77-1.56 (m, 4H, d and e). ¹³C-NMR (75 MHz; MeOD): δ 199.6 (C_{IV}, CO-S), 173.3 (C_{IV}, CO-NH), 173.1 (C_{IV}, CO-OH), 157.2 (C_{IV}, HN-CN₂-NH₂), 53.2 (CH, α), 43.9 (CH₂, b), 42.0 (CH₂, a), 31.3 (CH₂, β), 29.0 (CH₂, d), 23.6 (CH₂, e), 22.4 (CH₂, c).

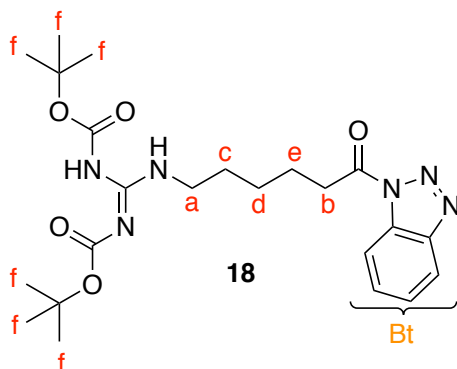
RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100% MeCN in 10 minutes): $t_r = 3.8$ min, 99% purity. HRMS m/z calculated for C₁₁H₂₁N₄O₄S [M⁺] Calc. m/z 305.1278 Obs. m/z 305.1280

BBhArg

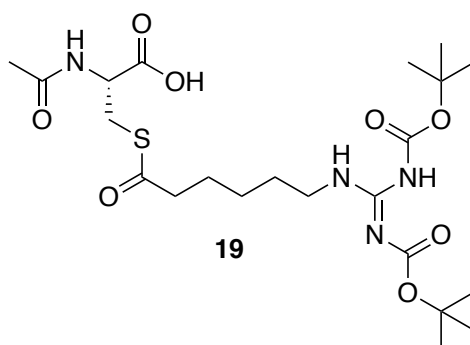


6-aminohexanoic acid (1.97 g, 15 mmol) and *N,N'*-Di-Boc-1*H*-pyrazole-1-carboxamide (5.12 g, 16.5 mmol, 1.1 eq) are suspended in pyridine (50 mL). The mixture was stirred at room temperature for 2 days, until the mixture became clear. TLC monitoring of the reaction showed complete transformation of the 5-aminopentanoic acid (*n*-BuOH/AcOH/H₂O; 4:1:5, staining with staining solution 3)). NaOH 1M (100 mL) was added to the reaction mixture, which was then stirred for 15 min. The aqueous phase was extracted with EtOAc (3x50 mL), acidified to pH 3 with aqueous 1M HCl then re-extracted with EtOAc (3x50 mL). The organic phase was dried over MgSO₄ and filtered before the solvent was removed under vacuum. Compound **17** was obtained as a white solid in quantitative yield. ¹H-NMR (300 MHz; CDCl₃): δ 8.30 (br, 1H, NH), 3.38 (q, $J = 6.0$ Hz, 2H, a), 2.33 (t, $J = 7.4$ Hz, 2H, b), 1.70-1.55 (m, 4H, c and d),

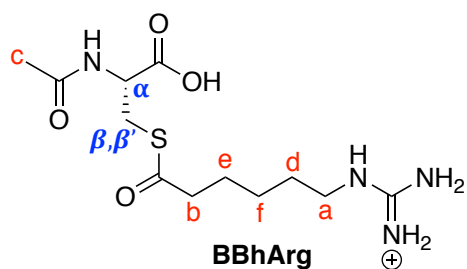
1.47 (s, 18H, e), 1.44-1.35 (m, 2H, f). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 178.8 (C_{IV} , CO-OH), 163.6 (C_{IV} , HN-CN-NH), 156.2 (C_{IV} , N-CO-O), 153.4 (C_{IV} , HN-CO-O), 83.2 (C_{IV} , tBu), 79.4 (C_{IV} , tBu), 40.8 (CH_2 , a), 34.0 (CH_2 , b), 28.8 (CH_2 , f), 28.4 (CH_3 , e), 28.2 (CH_3 , a), 26.4 (CH_2 , c), 24.4 (CH_2 , d).



Compound **17** (3.65 g, 10 mmol) was treated according to general procedure A to give **18** as a colourless oil in 82% yield. $^1\text{H-NMR}$ (300 MHz; CDCl_3): δ 8.34 (t, $J = 5.0$ Hz, 1H, NH), 8.29 (d, $J = 8.3$ Hz, 1H, Bt), 8.12 (d, $J = 8.3$ Hz, 1H, Bt), 7.65 (ddd, $J = 8.2, 7.2, 1.0$ Hz, 1H, Bt), 7.50 (ddd, $J = 8.2, 7.2, 1.0$ Hz, 1H, Bt), 3.48-3.41 (m, 4H, a and b), 1.94 (quintet, $J = 7.5$ Hz, 2H, c), 1.70-1.63 (m, 2H, d), 1.59-1.51 (m, 2H, e), 1.49 (s, 9H, f), 1.48 (s, 9H, f). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 172.5 (C_{IV} , CO-N), 163.8 (C_{IV} , HN-CN-NH), 156.3 (C_{IV} , N-CO-O), 153.5 (C_{IV} , HN-CO-O), 146.3 (C_{IV} , Bt), 131.3 (C_{IV} , Bt), 130.5 (CH, Bt), 126.3 (CH, Bt), 120.3 (CH, Bt), 114.6 (CH, Bt), 83.2 (C_{IV} , tBu), 79.4 (C_{IV} , tBu), 40.8 (CH_2 , a), 35.5 (CH_2 , b), 28.9 (CH_2 , d), 28.5 (CH_3 , f), 28.2 (CH_3 , f), 26.5 (CH_2 , e), 24.2 (CH_2 , c).



Treatment of **18** (475 mg, 3 mmol) according to the general procedure B lead to the diBoc-protected **BBhArg 19**. The crude product was purified by flash chromatography with (20:80:0.5 Cyclohexane/EtOAc/AcOH) as the mobile phase to afford a mixture of mono Boc-protected product and fully deprotected **BBhArg**, compound **19** was not isolated. The product was used in the next step without further purification.

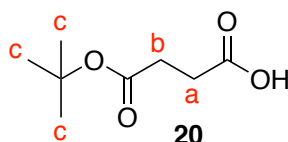


The mixture of products obtained in the last step was dissolved in DCM (7 mL) and treated with TFA (7 mL). The mixture was stirred at room temperature for 30 minutes. The solvent was removed under reduced pressure. The residue was dissolved in water and freeze-dried to afford the desired product **BBhArg** as a colorless oil in 5 % yield over two steps. ¹H-NMR (300 MHz; MeOD): δ 4.59 (dd, *J* = 8.5, 4.7 Hz, 1H, α), 3.54 (dd, *J* = 13.9, 4.7 Hz, 1H, β), 3.17 (t, *J* = 7.0 Hz, 2H, a), 3.09 (dd, *J* = 13.9, 8.5 Hz, 1H, β'), 2.62 (t, *J* = 7.2 Hz, 2H, b), 1.97 (s, 3H, c), 1.75-1.65 (m, 2H, d), 1.64-1.54 (m, 2H, e), 1.45-1.34 (m, 2H, f). ¹³C-NMR (75 MHz; MeOD): δ 199.9 (C_{IV}, CO-S), 173.3 (C_{IV}, CO-NH), 173.1 (C_{IV}, CO-OH), 158.7 (C_{IV}, HN-CN₂-NH₂), 53.3 (CH, α), 44.4 (CH₂, b), 42.2 (CH₂, a), 31.2 (CH₂, β), 29.5 (CH₂, e), 26.8 (CH₂, f), 26.1 (CH₂, d), 22.4 (CH₂, c).

HRMS *m/z* calculated for C₁₂H₂₃N₄O₄S [MNa⁺] Calc. *m/z* 319.1435 Obs. *m/z* 319.1435

RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100% MeCN in 10 minutes): *t_r* = 5.2 min, >99% purity.

BBAsp



One step protocol:

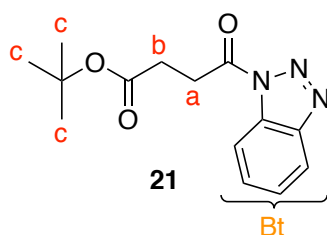
Succinic anhydride (3.0 g, 30 mmol) and ZnCl₂ (96.4 mg, 0.7 mmol, 2 mol%) were suspended in *tert*-butanol (17 mL, 180 mmol, 6 eq.). The mixture was stirred and heated at 60 °C for 10 days. 1M aqueous NaOH (15 mL) was added and stirring was maintained for 30 min. The aqueous mixture was extracted with diethylether (3 x 15 mL). The combined organic layer was washed with water (3 x 20 mL), dried over MgSO₄ then filtered before the solvent was removed under reduced pressure. Compound **20** was obtained as a pure white solid in 15 % yield.

Four steps protocol:

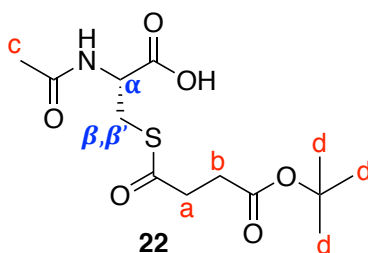
Succinic anhydride (2.00 g, 20 mmol) was suspended in anhydrous MeOH (15 mL). The mixture was stirred and refluxed for 3 hours. The solvent was removed under reduced pressure to obtain the mono methyl ester as a white solid, which was used without further purification.

The solid was suspended in SOCl_2 (3 mL, 41 mmol, 2.7 eq.). The mixture was stirred and refluxed for two hours. Excess thionyl chloride was removed under vacuum to yield a yellow oil which was re-dissolved in 5 mL of CHCl_3 . *Tert*-butanol (3.5 mL, 36 mmol, 2.6 eq.) and pyridine (1.8 mL, 23 mmol, 1.6 eq.) were slowly added while preventing the mixture from heating above 50 °C. The mixture was refluxed at 80 °C for 3h30 then let to cool before the content was poured in a beaker with 16 g of ice. The aqueous phase was acidified to pH 1-2 with aqueous 1M HCl then extracted with DCM (3 x 15 mL). The combined organic phase was washed with aqueous 1M HCl (3 x 10 mL), saturated aqueous NaHCO_3 (10 mL), water (10 mL) then dried over MgSO_4 and filtered before the solvent was removed under reduced pressure. The methyl *tert*-butyl diester was obtained as an orange oil and was used without further purification.

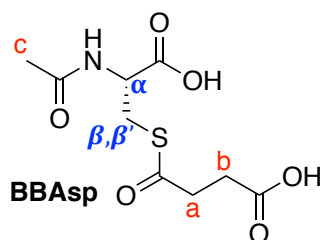
The crude compound (1.95 g, 10 mmol) and NaOH (0.4 g, 10 mmol, 1.0 eq.) were dissolved in 50 mL of a 1,4-dioxane/water (1:1). The mixture was stirred at room temperature for 12 hours. Volatiles were removed under reduced pressure. The aqueous phase was extracted with diethylether (15 mL), acidified to pH 1-2 with aqueous 1M HCl then re-extracted with diethylether (3 x 15 mL). The combined organic layer was dried over MgSO_4 and filtered before the solvent was removed under reduced pressure. Compound **20** was obtained as a dark orange oil which crystallized upon storage in 76 % yield over 4 steps and was used without further purification. $^1\text{H-NMR}$ (400 MHz; CDCl_3): δ 2.62 (m, 2H, **a**), 2.53 (m, 2H, **b**), 1.44 (s, 9H, **c**). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): 178.4 (C_{IV} , CO-OH), 171.5 (C_{IV} , CO-OtBu), 81.2 (C_{IV} , tBu), 30.3 (CH_2 , **b**), 29.3 (CH_2 , **a**), 28.2 (CH_3 , **c**).



Compound **20** (1.74 g, 10 mmol) was treated according to general procedure A to give **21** as an orange solid in 87% yield. $^1\text{H-NMR}$ (300 MHz; CDCl_3): δ 8.27-8.22 (m, 1H, **Bt**), 8.12 (m, 1H, **Bt**), 7.66-7.60 (m, 1H, **Bt**), 7.50 (m, 1H, **Bt**), 3.70 (t, $J = 6.5$ Hz, 2H, **a**), 2.83 (t, $J = 6.5$ Hz, 2H, **b**), 1.43 (s, 9H, **c**). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 171.5 (C_{IV} , CO), 171.1 (C_{IV} , CO), 146.3 (C_{IV} , **Bt**), 131.2 (C_{IV} , **Bt**), 130.5 (CH, **Bt**), 126.2 (CH, **Bt**), 120.3 (CH, **Bt**), 114.5 (CH, **Bt**), 81.3 (C_{IV} , tBu), 30.9 (CH_2 , **a**), 29.6 (CH_2 , **b**), 28.1 (CH_3 , **c**).

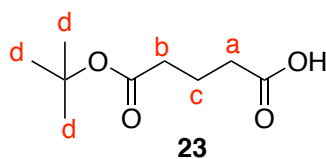


Treatment of **21** (826 mg, 3 mmol) according to the general procedure B lead to *t*Bu-protected **BBAsp 22**. The crude product was purified by flash chromatography with (20:80:0.5 Cyclohexane/EtOAc/AcOH) from 70 % in cyclohexane to 100 % as gradient to afford the pure product in 40 % yield. ¹H-NMR (300 MHz; MeOD): δ 4.60 (dd, *J* = 7.8, 4.6 Hz, 1H, **α**), 3.55 (dd, *J* = 13.9, 4.6 Hz, 1H, **β**), 3.17 (dd, *J* = 13.9, 7.8 Hz, 1H, **β'**), 2.85 (t, *J* = 6.8 Hz, 2H, **a**), 2.55 (t, *J* = 6.8, 2H, **b**), 1.96 (s, 3H, **c**), 1.43 (s, 9H, **d**). ¹³C-NMR (75 MHz; MeOD): δ 198.5 (C_{IV}, CO-S), 173.2 (C_{IV}, CO-NH), 173.0 (C_{IV}, CO-OH), 172.9 (C_{IV}, CO-O*t*Bu), 82.0 (C_{IV}, *t*Bu), 53.3 (CH**α**), 39.4 (CH₂, **a**), 31.2 (CH₂, **b**), 31.0 (CH₂**β**), 28.3 (CH₃, **d**), 22.4 (CH₃, **c**).

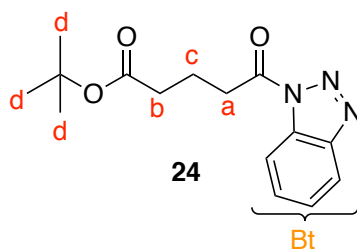


The compound **22** (380 mg, 1.19 mmol) was dissolved in DCM (5 mL) and treated with TFA (5 mL). The mixture was stirred at room temperature for 2 hours. The solvent was removed under reduced pressure. The residue was dissolved in water and freeze-dried to afford the desired product **BBAsp** as a yellow solid in quantitative yield. ¹H-NMR (300 MHz; MeOD): δ 4.61 (dd, *J* = 7.8, 4.6 Hz, 1H, **α**), 3.55 (dd, *J* = 13.9, 4.6 Hz, 1H, **β**), 3.17 (dd, *J* = 13.9, 7.8 Hz, 1H, **β'**), 2.88 (t, *J* = 6.6 Hz, 2H, **a**), 2.61 (t, *J* = 6.6 Hz, 2H, **b**), 1.96 (s, 3H, **c**). ¹³C-NMR (75 MHz; MeOD): δ 198.7 (C_{IV}, CO-S), 175.5 (C_{IV}, CO-NH), 173.4 (C_{IV}, CO-OH), 173.0 (C_{IV}, CO-OH), 53.3 (CH, **α**), 39.3 (CH₂, **β**), 31.0 (CH₂, **a**), 29.8 (CH₂, **b**), 22.4 (CH₃, **c**).
 RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100% MeCN in 10 minutes): *t*_r = 3.8 min, 99 % purity.
 HRMS *m/z* calculated for C₉H₁₃NNaO₆S [MNa⁺] Calc. *m/z* 286.0356 Obs. *m/z* 286.0357

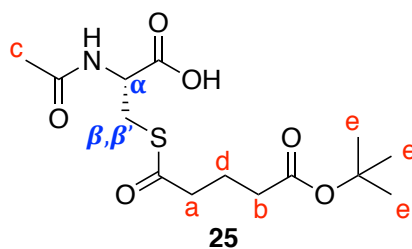
BBGlu



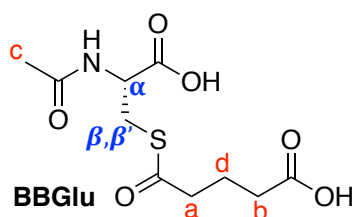
Glutaric anhydride (3.42 g, 30 mmol) and ZnCl_2 (96 mg, 2 mol%) were suspended in *tert*-butanol (17 mL, 180 mmol, 6 eq.). The mixture was heated at 60 °C for 10 days under an argon atmosphere. The mixture was left to cool back to room temperature and 1M NaOH (15 mL) was added to the mixture and stirred for 30 minutes. This aqueous phase was extracted four times with diethylether (15 mL). The combined organic extracts were washed with water (3 x 20 mL), dried over MgSO_4 and filtered before the solvent was removed under reduced pressure. Traces of *tert*-butanol were removed by co-evaporation with cyclohexane. Compound **23** was obtained as a pale yellow oil in 45 % yield. $^1\text{H-NMR}$ (400 MHz; CDCl_3): δ 2.36 (m, 2H, **a**), 2.25 (m, 2H, **b**), 1.86 (quint., $J = 7.3$ Hz, 2H, **c**), 1.40 (s, 9H, **d**). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 178.7 (C_{IV} , CO-OH), 172.5 (C_{IV} , CO-OtBu), 80.5, (C_{IV} , tBu), 34.4 (CH_2 , **b**), 33.0 (CH_2 , **a**), 28.0 (CH_3 , **d**), 20.0 (CH_2 , **c**).



Compound **23** (1.88 g, 10 mmol) was treated according to general procedure A to give **24** as an off-white solid in 86% yield. $^1\text{H-NMR}$ (300 MHz; CDCl_3): δ 8.24 (d, $J = 8.3$ Hz, 1H, **Bt**), 8.08 (d, $J = 8.3$ Hz, 1H, **Bt**), 7.61 (ddd, $J = 8.2, 7.2, 1.0$ Hz, 1H, **Bt**), 7.47 (ddd, $J = 8.2, 7.2, 1.0$ Hz, 1H, **Bt**), 3.47 (t, $J = 7.3$ Hz, 2H, **a**), 2.42 (t, $J = 7.3$ Hz, 2H, **b**), 2.25-2.12 (m, 2H, **c**), 1.42 (s, 9H, **d**). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 172.1 (C_{IV} , CO-N), 172.0 (C_{IV} , CO-OtBu), 146.2 (C_{IV} , **Bt**), 131.1 (C_{IV} , **Bt**), 130.4 (CH, **Bt**), 126.2 (CH, **Bt**), 120.2 (CH, **Bt**), 114.4 (CH, **Bt**), 80.7 (C_{IV} , tBu), 34.7 (CH_2 , **a**), 34.5 (CH_2 , **b**), 28.2 (CH_3 , **d**), 19.7 (CH_2 , **c**).



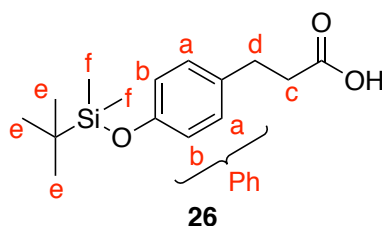
Treatment of **24** (868 mg, 3 mmol) according to the general procedure B lead to *t*Bu-protected **BBGlu 25**. The crude product was purified by flash chromatography with (20:80:0.5 Cyclohexane/EtOAc/AcOH) from 70 % in cyclohexane to 100 % as gradient to afford the pure product in 57 % yield. ¹H-NMR (300 MHz; MeOD): δ 4.61 (dd, *J* = 7.9, 4.7 Hz, 1H, α), 3.52 (dd, *J* = 13.9, 4.7 Hz, 1H, β), 3.17 (dd, *J* = 13.9, 7.9 Hz, 1H, β'), 2.64 (t, *J* = 7.3 Hz, 2H, a), 2.28 (t, *J* = 7.3 Hz, 2H, b), 1.98 (s, 3H, c), 1.88 (quintet, *J* = 7.3 Hz, 2H, d), 1.44 (s, 9H, e). ¹³C NMR (75 MHz; MeOD): δ 199.1 (C_{IV}, CO-S), 173.8 (C_{IV}, CO-NH), 173.1 (C_{IV}, CO-OH), 172.9 (C_{IV}, CO-O*t*Bu), 81.5 (C_{IV}, *t*Bu), 53.1 (CH, α), 43.5 (CH₂, a), 34.9 (CH₂, β), 31.0 (CH₂, b), 28.4 (CH₃, e), 22.4 (CH₃, c), 21.9 (CH₂, d).



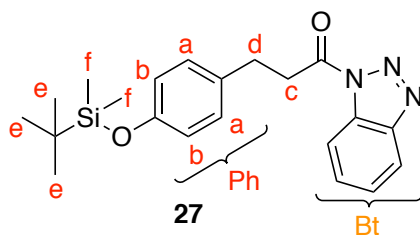
Compound **25** (570 mg, 1,71 mmol) was dissolved in DCM (10 mL) and treated with TFA (10 mL). The mixture was stirred at room temperature under argon atmosphere for 2 hours, until TLC monitoring showed complete deprotection of the *t*-Bu ester (*n*-BuOH/AcOH/H₂O (4:1:5)). The solvent was removed under reduced pressure to afford the desired product **BBGlu** a yellow oil in quantitative yield. ¹H-NMR (300 MHz; MeOD): δ 4.60 (dd, *J* = 8.0, 4.7 Hz, 1H, α), 3.52 (dd, *J* = 13.9, 4.7 Hz, 1H, β), 3.16 (dd, *J* = 13.9, 8.0 Hz, 1H, β'), 2.66 (t, *J* = 7.3 Hz, 2H, a), 2.39 (t, *J* = 7.3 Hz, 2H, b), 1.97 (s, 3H, c), 1.93 (t, *J* = 7.2 Hz, 2H, d). ¹³C-NMR (75 MHz; MeOD): δ 199.4 (C_{IV}, CO-S), 173.8 (C_{IV}, CO-NH), 171.8 (C_{IV}, CO-OH), 170.4 (C_{IV}, CO-OH), 53.3 (CH, α), 43.6 (CH₂, a), 33.6 (CH₂, b), 31.1 (CH₂, β), 22.3 (CH₃, c), 21.8 (CH₂, d).

RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100% MeCN in 10 minutes): *t*_r = 4.2 min, 99 % purity. HRMS *m/z* calculated for C₁₀H₁₄NO₆S [M-H] Calc. *m/z* 276.0547 Obs. *m/z* 276.0549

BBTyr

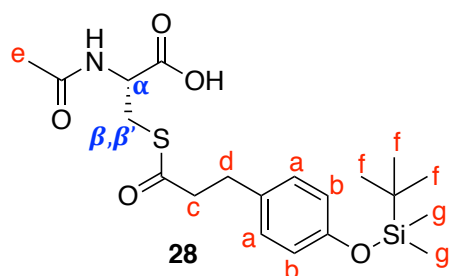


3-(4-hydroxyphenyl)propanoic acid (2.50 g, 15 mmol) was dissolved in 20 mL of anhydrous DMF and *tert*-butyldimethylsilyl chloride (4.75 g, 31.5 mmol, 2.1 eq.) and imidazole (2.14 g, 31.5 mmol, 2.1 eq.) were added. The mixture was stirred overnight under argon atmosphere. TLC (Cyclohexane/EtOAc 5:5) staining solution 4) showed complete transformation of the starting material. The reaction mixture was diluted with water (20 mL) and this aqueous mixture was extracted with cyclohexane (3 x 40 mL). The organic extracts were combined and the solvent was removed under reduced pressure. The residue was re-dissolved in 20 mL THF and 20 mL saturated aqueous NaHCO₃. The mixture was stirred at room temperature for 2 hours then extracted with EtOAc (3 x 40 mL). The aqueous phase was acidified to pH 2-3 with 1M aqueous HCl then re-extracted with EtOAc (3 x 40 mL). The organic extracts were combined and washed with brine (60 mL) then dried over MgSO₄ and filtered before the solvent was removed under reduced pressure. *Tert*-butyldimethylsilanol was removed by co-evaporation with toluene. The pure desired product **26** was obtained as a white solid in 83 % yield. ¹H-NMR (300 MHz; CDCl₃): δ 7.05 (d, *J* = 8.4 Hz, 2H, **a**), 6.76 (d, *J* = 8.4 Hz, 2H, **b**), 2.88 (t, *J* = 7.8 Hz, 2H, **c**), 2.63 (t, *J* = 7.8 Hz, 2H, **d**), 0.98 (s, 9H, **e**), 0.19 (s, 6H, **f**). ¹³C-NMR (75 MHz; CDCl₃): δ 179.4 (C_{IV}, CO-OH), 154.2 (C_{IV}, C-OSi, **Ph**), 133.1 (C_{IV}, C-CH₂, **Ph**), 129.3 (CH, **Ph**), 120.2 (CH, **Ph**), 36.3 (CH₂, **c**), 30.1 (CH₂, **d**), 25.8 (CH₃, **e**), 18.3 (C_{IV}, *t*Bu), -4.3 (CH₃, **f**).

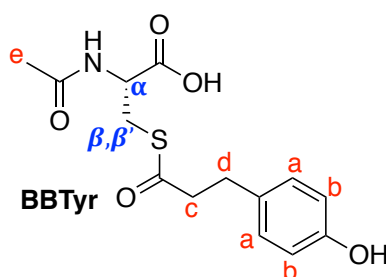


Compound **26** (2.80 g, 10 mmol) was treated according to general procedure A to give **27** as an off-white solid in 90% yield. ¹H-NMR (300 MHz; CDCl₃): δ 8.29 (dt, *J* = 8.3, 0.9 Hz, 1H, **Bt**), 8.12 (dt, *J* = 8.3, 0.9 Hz, 1H, **Bt**), 7.65 (ddd, *J* = 8.3, 7.2, 0.9 Hz, 1H, **Bt**), 7.50 (ddd, *J* = 8.3, 7.2, 0.9 Hz, 1H, **Bt**), 7.18-7.13 (m, 2H, **a**), 6.80-6.74 (m, 2H, **b**), 3.73 (t, *J* = 7.7 Hz, 2H, **c**), 3.17 (t, *J* = 7.7 Hz, 2H, **d**), 0.97 (s, 9H, **e**), 0.17 (s, 6H, **f**). ¹³C-NMR (75 MHz; CDCl₃): δ 171.9 (C_{IV}, CO-N), 154.4 (C_{IV}, C-OSi, **Ph**), 146.3 (C_{IV}, **Bt**), 132.6 (C_{IV}, C-CH₂, **Ph**), 131.1 (C_{IV}, **Bt**), 130.5

(CH, **Bt**), 129.5 (CH, **Ph**), 126.3 (CH, **Bt**), 120.3 (CH, **Ph** and **Bt**), 114.6 (CH, **Bt**), 37.6 (CH₂, **c**), 29.7 (CH₂, **d**), 25.8 (CH₃, **e**), 18.3 (C_{IV}, tBu), -4.3 (CH₃, **f**).



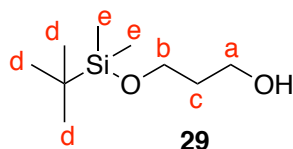
Treatment of **27** (826 mg, 3 mmol) according to the general procedure B lead to the TBS-protected **BBTyr 28**. The crude product was purified by flash chromatography (Cyclohexane/EtOAc/AcOH 20:80:0.5) to afford the pure product as an yellow oil in 53% yield. ¹H-NMR (300 MHz; MeOD): δ 7.09-6.98 (m, 2H, **a**), 6.77-6.66 (m, 2H, **b**), 4.58 (dd, *J* = 7.9, 4.6 Hz, 1H, **α**), 3.52 (dd, *J* = 13.9, 4.6 Hz, 1H, **β**), 3.13 (dd, *J* = 13.9, 7.9 Hz, 1H, **β'**), 2.88-2.81 (m, 4H, **c** and **d**), 1.94-1.93 (s, 3H, **e**), 0.99 (s, 9H, **f**), 0.16 (s, 6H, **g**). ¹³C-NMR (75 MHz; MeOD): δ 199.2 (C_{IV}, CO-S), 173.2 (C_{IV}, CO-NH), 173.1 (C_{IV}, CO-OH), 155.4 (C_{IV}, C-OSi), 134.3 (C_{IV}, C-CH₂), 130.4 (CH, **Ph**), 121.1 (CH, **Ph**), 53.3 (CH, **α**), 46.5 (CH₂, **c**), 31.6 (CH₂, **d**), 31.1 (CH₂, **β**), 26.2 (CH₃, **f**), 22.4 (CH₃, **e**), 19.2 (C_{IV}, tBu) -4.3 (CH₃, **g**).



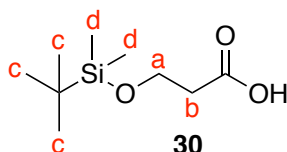
Compound **28** (510 mg, 1.20 mmol) was dissolved in 10 mL of anhydrous THF. 1M TBAF solution in THF (1.44 mL, 1.44 mmol, 1.2 eq.) was added and the mixture was stirred at room temperature for 3 hours under argon atmosphere. TLC monitoring (*n*-BuOH/AcOH/H₂O; 4:1:5, staining with staining solution 3)) showed complete transformation of the starting material. The solvent was removed under reduced pressure and the crude material was directly purified by preparative RP-HPLC. After freeze-drying **BBTyr** was obtained as a white solid in 48% yield. ¹H-NMR (300 MHz; MeOD): δ 7.03-6.98 (m, 2H, **a**), 6.71-6.66 (m, 2H, **b**), 4.57 (dd, *J* = 8.1, 4.6 Hz, 1H, **α**), 3.51 (dd, *J* = 13.9, 4.6 Hz, 1H, **β**), 3.12 (dd, *J* = 13.9, 8.1 Hz, 1H, **β'**), 2.87-2.79 (m, 4H, **c** and **d**), 1.94 (s, 3H, **e**). ¹³C NMR (75 MHz; MeOD): δ 199.3 (C_{IV}, CO-S), 173.3

(C_{IV}, CO-NH), 173.1 (C_{IV}, CO-OH), 156.9 (C_{IV}, C-OH), 132.2 (C_{IV}, C-CH₂), 130.3 (CH, Ph), 116.3 (CH, Ph), 53.3 (CH, α), 46.7 (CH₂, c), 31.6 (CH₂, d), 31.0 (CH₂, β), 22.3 (CH₂, e).
 RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100% MeCN in 10 minutes): t_r= 5.5 min, 99 % purity.
 HRMS m/z calculated for C₁₄H₁₇NO₅SNa [MNa⁺] Calc. m/z 334.0720, Obs. 334.0719

BBSer

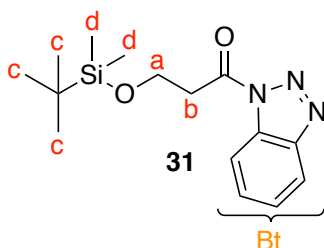


Sodium hydride (60 % suspension in mineral oil, 800 mg, 20 mmol, 1 eq.) was suspended in 35 mL of anhydrous THF. 1,3-propanediol (1.45 mL, 20 mmol) was added dropwise over 10 minutes. The mixture was stirred at room temperature for 45 min then *tert*-butyldimethylsilyl chloride (3.01 g, 20 mmol, 1 eq.) was added in small portions over 10 minutes. The mixture was stirred at room temperature for another 45 minutes. TLC monitoring (Cyclohexane/EtOAc; 6:4; staining solution 1)) showed complete transformation of the starting material. A 5 % aqueous sodium bicarbonate solution (25 mL) was added and this aqueous phase was extracted with diethylether (3 x 25 mL). The combined organic extract was washed with water (2 x 25 mL), brine, dried over MgSO₄ and filtered before the solvent was removed under reduced pressure. The mono TBS-protected diol **29** was obtained as a pale yellow oil in 88 % yield. H-NMR (400 MHz; CDCl₃) : δ 3.79 (t, *J* = 5.7 Hz, 2H, a), 3.75 (t, *J* = 5.7 Hz, 2H, b), 2.76 (br, 1H, OH), 1.74 (quint, *J* = 5.7 Hz, 2H, c), 0.87 (s, 9H, d), 0.05 (s, 6H, e). ¹³C-NMR (75 MHz; CDCl₃) : 62.7 (CH₂, a), 62.1 (CH₂, b), 34.4 (CH₂, c), 26.0 (CH₃, d), 18.3 (C_{IV}, *t*-Bu), -5.4(CH₃, e).



The mono protected alcohol **29** (2.86 g, 15 mmol) was dissolved in a ACN/Water mixture (9:1). TEMPO (469 mg, 3 mmol, 0.2 eq.) and Bis(Acetoxy)IodoBenzene (10.6 g, 33 mmol, 2.2 eq.) were added. The mixture was stirred at room temperature for 2 hours. TLC monitoring (Cyclohexane/EtOAc 8:2) showed complete transformation of the starting material (staining solution 1). The reaction mixture was diluted with 50 mL of DCM. This organic phase was washed with about 100 mL of saturated aqueous Na₂S₂O₃. The organic phase was then dried over MgSO₄ and filtered before the solvent was removed under reduced pressure. The crude material was purified by silica gel column chromatography with Cyclohexane/EtOAc (8:2) as

the mobile phase. The desired carboxylic acid product **30** was obtained as a light orange oil in 80 % yield. $^1\text{H-NMR}$ (300 MHz; CDCl_3): δ 3.90 (t, $J = 6.3$ Hz, 2H, **a**), 2.56 (t, $J = 6.3$ Hz, 2H, **b**), 0.88 (s, 9H, **c**), 0.06 (s, 6H, **d**). $^{13}\text{C-NMR}$ (75 MHz; CDCl_3): δ 177.5 (C_{IV} , CO-OH), 58.9 (CH_2 , **a**), 37.8 (CH_2 , **b**), 25.9 (CH_3 , **c**), 18.4 (C_{IV} , tBu), -5.3 (CH_3 , **d**).



Compound **31** (2.04 g, 10 mmol) was treated according to general procedure A to give **32** as a dark orange oil in 66 % yield. $^1\text{H-NMR}$ (300 MHz; CDCl_3): δ 8.25 (dt, $J = 8.3, 1.0$ Hz, 1H, **Bt**), 8.07 (dt, $J = 8.3, 1.0$ Hz, 1H, **Bt**), 7.60 (ddd, $J = 8.3, 7.2, 0.9$ Hz, 1H, **Bt**), 7.46 (ddd, $J = 8.3, 7.2, 0.9$ Hz, 2H, **Bt**), 4.17 (t, $J = 6.2$ Hz, 2H, **a**), 3.60 (t, $J = 6.2$ Hz, 2H, **b**), 0.79 (s, 9H, **c**), 0.01 (s, 6H, **d**). ^{13}C NMR (75 MHz; CDCl_3): δ 170.7 (C_{IV} , CO-N), 146.2 (C_{IV} , **Bt**), 131.1 (C_{IV} , **Bt**), 130.4 (CH, **Bt**), 126.2 (CH, **Bt**), 120.1 (CH, **Bt**), 114.5 (CH, **Bt**), 58.8 (CH_2 , **a**), 39.1 (CH_2 , **b**), 25.8 (CH_3 , **c**), 18.2 (C_{IV} , tBu), -5.4 (CH_3 , **d**).

II. Peptide Synthesis

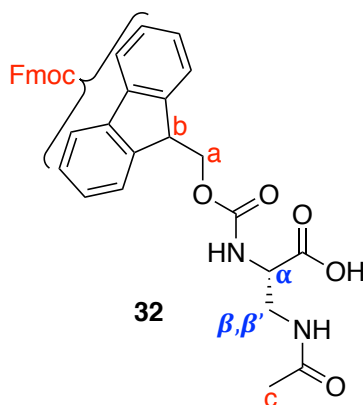
- **General Considerations**

2-Chlorotrytlyl chloride resin (100-200 mesh, loading 1.56 mmol/g), MBHA-Resin (100-200 mesh, loading 0.54 mmol/g) and 2-(1H-9-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from *Iris Biotech GmbH* in peptide synthesis grade. 2-(1H-9-azabenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HATU) was purchased from *Fluorochem* and in peptide synthesis grade. Unless stated otherwise all the amino acids used for peptides synthesis were purchased from *Iris Biotech GmbH* as follow: **Fmoc-Arg(Pbf)-OH**, **Fmoc-Arg(Tos)-OH**, **Fmoc-Cys(Trt)-OH**, **Fmoc-Cys(Mob)-OH**, **Fmoc-Gly-OH**, **Fmoc-Lys(Boc)-OH**, **Fmoc-Lys(2-Cl Z)-OH**, **Fmoc-Pro-OH**, **Fmoc-^DPro-OH**, **Fmoc-Trp(Boc)-OH**, **Fmoc-L-Asp(tBu)-OH** (aa in red were used for all purpose Fmoc SPPS, underlined aa were used for the synthesis of Stapled Helix Peptides). The peptides were synthesized by manual SPPS in polypropylene manual disposable reactors with plunger and frit, pore size 25 μm , with removable PTFE luer stoppers, purchased from

Multisyntech GmbH. Preparative scale purification of peptides was performed by reverse phase HPLC on a Waters systems consisted of a binary pump (Waters 1525) and a dual wavelength UV/Visible Absorbance detector (Waters 2487), piloted by Breeze softwares using the following columns: preparative *Macherey-Nagel Nucleodur* (RP C18, 250 x 16 mm, 5 μ m, 300 Å) and semi-preparative *ACE* (RP C18, 250 x 10 mm, 5 μ m, 300 Å) from *AIT* at a flow rate of 14 mL/min and 6 mL/min respectively. Peptide fractions from purification were analyzed by analytical HPLC on a Dionex system consisted of an analytical automated LC system (Ultimate 3000) equipped with an autosampler, a pump block composed of two ternary gradient pumps and a dual wavelength detector, piloted by Chromeleon software. HPLC analysis were performed on a *Proto 200* (RP C18, 100 x 4.6 mm, 3 μ m, 300 Å) from *Higgins-Analytical Inc.* The pure fractions were gathered according to their purity, partly concentrated under vacuum and freeze-dried using an Alpha 2/4 from Bioblock Scientific to get the expected peptide as a powder. Final peptide purity (> 95 %) was checked on an analytical column using a linear gradient. Peptides characterization was realized using a MALDI-TOF MS (DE-Pro, PerSeptive Biosystems) spectrometer in positive ion reflector mode using the matrix CHCA.

1) Synthesis of unnatural amino acids

Fmoc-L-Dap(Ac)-OH



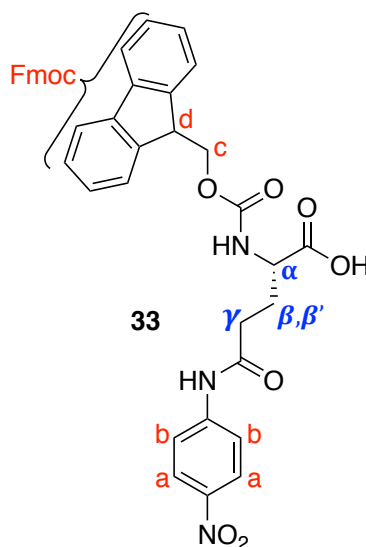
Fmoc-L-Dap(Boc)-OH (5.0 g, 11.72 mmol) was dissolved in a 1:1 DCM/TFA mixture (100 mL). The solution was stirred at room temperature for 2 hours. TLC monitoring (DCM/MeOH/AcOH ; 93:7:0.1 ; staining solution 3)) showed complete deprotection of the Boc group. Volatiles were removed under reduced pressure to afford Fmoc-L-Dap(NH₂.TFA)-OH in quantitative yield which was used in the next step without further treatment.

Fmoc-L-Dap(NH₂.TFA)-OH (5.16 g, 11.7 mmol) was suspended in THF (58.5 mL). Pyridine (2.83 mL, 35.1 mmol, 3 eq.) and acetic anhydride (1.66 mL, 17.6 mmol, 1.5 eq.) in THF (19.5

mL) were added. The mixture was stirred at room temperature for 4 hour. TLC monitoring (*n*-BuOH/AcOH/H₂O; 4:1:5 ; staining solution 3)) showed complete transformation of the free amine. The solvent was removed under reduced pressure. The residue was redissolved in EtOAc (200 mL) and washed successively with 1M aqueous NaHSO₄ (50 mL), H₂O (50 mL) and brine (50 mL). The organic phase was dried over Na₂SO₄ and filtered before the solvent was removed under reduced pressure. The oily residue was lyophilized to afford Fmoc-L-Dap(Ac)-OH as an off-white solid in 89 % yield which was used in SPPS without further purification. ¹H-NMR (300 MHz; CDCl₃): δ 7.89 (br, 2H, NH), 7.70 (d, *J* = 7.5 Hz, 2H, Fmoc), 7.53 (m, 2H, Fmoc), 7.34 (t, *J* = 7.3 Hz, 2H, Fmoc), 7.24 (m, 2H, Fmoc), 4.35 (m, 3H, **a** and **α**), 4.13 (t, *J* = 6.7 Hz, 1H, **b**), 3.64 (br, 2H, **β** and **β'**), 1.93(s, 3H, **c**). ¹³C-NMR (75 MHz; CDCl₃): δ 173.5 (C_{IV}, CO-OH/CO-NH), 172.8 (C_{IV}, CO-OH/CO-NH), 156.8 (C_{IV}, O-CO-NH), 143.8 (C_{IV}, Fmoc), 143.7 (C_{IV}, Fmoc) 141.4 (C_{IV}, Fmoc), 141.3 (C_{IV}, Fmoc), 127.9 (CH, Fmoc), 127.2 (CH, Fmoc), 125.2 (CH, Fmoc), 120.1 (CH, Fmoc), 67.5 (CH₂, **a**), 54.5 (CH, **α**), 47.1 (CH, **b**), 41.6 (CH₂, **β**), 22.7 (CH₃, **c**).

HRMS *m/z* calculated for C₂₀H₂₀N₂NaO₅ [MNa⁺] Calc. 391.1264, Obs. 391.0916

Fmoc-L-Glu(pNA)-OH

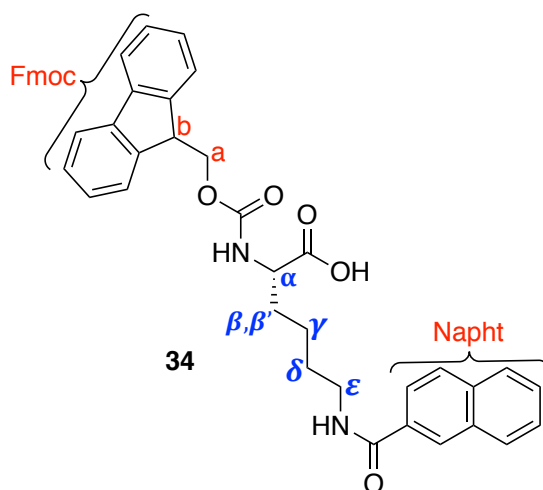


L-Glu(pNA)-OH·H₂O (500 mg, 1.87 mmol) was dissolved in a 1:1 THF/H₂O mixture (20 mL). Fmoc-OSu (693 mg, 2.05 mmol, 1.1 eq.) and K₂CO₃ (283 mg, 2.05 mmol, 1.1 eq.) were added and the mixture was stirred overnight at room temperature. TLC monitoring (*n*-BuOH/AcOH/H₂O; 4:1:5 ; staining solution 3)) showed remaining starting material. Fmoc-OSu (210 mg, 0.56 mmol, 0.3 eq.) and K₂CO₃ (84 mg, 0.56 mmol, 0.3 eq.) were added and the mixture was stirred for two more hours at room temperature. TLC monitoring (*n*-BuOH/AcOH/H₂O; 4:1:5 ; staining solution 3)) showed complete transformation of the starting

material. THF was removed under reduced pressure. The aqueous mixture was acidified to pH 2 with 1M aqueous HCl then extracted with EtOAc. The combined organic layers were dried over MgSO₄ and filtered before the solvent was removed under reduced pressure. The crude solid material was washed with diethyl ether and filtered. The desired product was obtained as a white solid in 96% yield. ¹H-NMR (300 MHz; MeOD): δ 8.11 (d, *J* = 9.2 Hz, 2H, **a**), 7.78-7.74 (m, 4H, **Fmoc**), 7.63 (m, 2H, **Fmoc**), 7.39-7.24 (m, 4H, **b** and **Fmoc**), 4.33 (dd, *J* = 13.1, 7.2 Hz, 1H, **α**), 4.28-4.21 (m, 2H, **c**), 4.16-4.11 (m, 1H, **d**), 2.55 (t, *J* = 7.2 Hz, 2H, **γ**), 2.29 (td, *J* = 13.1, 7.2 Hz, 1H, **β**), 2.05 (td, *J* = 13.1, 7.2 Hz, 1H, **β'**). ¹³C-NMR (75 MHz; MeOD): δ 172.3 (C_{IV}, CO-OH, CO-NH), 157.1 (C_{IV}, O-CO-NH), 146.0 (C_{IV}, **Fmoc**), 145.2 (C_{IV}, **Fmoc**), 145.1 (C_{IV}, C-NO₂), 144.3 (C_{IV}, **Fmoc**), 142.5 (C_{IV}, **Fmoc**), 128.8 (C_{IV}, C-NH-CO), 128.1 (CH, **Fmoc**), 126.2 (CH, **Fmoc**), 125.7 (CH, **Fmoc**), 120.9 (CH, **Fmoc**), 120.9 (CH, **b**), 120.3 (CH, **a**), 68.0 (CH₂, **c**), 55.1 (CH, **α**), 48.3 (CH, **d**), 34.4 (CH₂, **γ**), 28.3 (CH₂, **β**).

HRMS *m/z* calculated for C₂₆H₂₃N₃NaO₇ [MNa⁺] Calc. 512.1428, Obs. 512.1424

Fmoc-L-Lys(2-Naphthoyl)-OH



Fmoc-L-Lys(Boc)-OH (5.0 g, 10.7 mmol) was dissolved in a 1:1 DCM/TFA mixture (100 mL). The solution was stirred at room temperature for 2 hours. Volatiles were removed under reduced pressure to afford Fmoc-L-Lys(NH₂.TFA)-OH in quantitative yield which was used in the next step without further treatment.

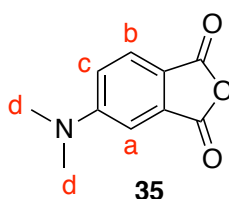
Fmoc-L-Lys(NH₂.TFA)-OH (3.0 g, 6.22 mmol) was dissolved in a 1,4-dioxane/25 % aqueous K₂CO₃ 1:1 mixture (52 mL). 2-Naphthoyl chloride (1.30 g, 6.84 mmol, 1.1 eq.) was dissolved in 1,4-dioxane (30 mL) and this solution was added dropwise to the first one. The mixture was then stirred at room temperature for 2 days. TLC monitoring with DCM/MeOH/AcOH (95:5:0.1) as the mobile phase (staining solution 3) showed complete transformation of the starting material. The reaction mixture was diluted with distilled water (250 mL) then extracted

with Et₂O (3 x 120 mL). The aqueous layer was acidified to pH 2 with 1M aqueous HCl then re-extracted with CHCl₃ (3 x 150 mL). The organic phase was dried over Na₂SO₄ and filtered before the solvent was removed under reduced pressure to afford the expected product in 91 % yield as a white solid, which was used in SPPS without further purification. ¹H-NMR (300 MHz; CDCl₃): δ 8.24 (s, 1H, **Napht**), 7.77-7.64 (m, 6H, **Fmoc** and **Napht**), 7.48-7.36 (m, 5H, **Fmoc** and **Napht**), 7.29 (t, *J* = 7.6 Hz, 2H, **Fmoc**), 7.21-7.18 (m, 2H, **Fmoc**), 4.31 (dd, *J* = 31.7, 5.2 Hz, 2H, **α**), 4.07 (t, *J* = 6.7 Hz, 1H, **ε**), 3.41 (d, *J* = 5.1 Hz, 2H), 2.16 (s, 1H), 1.65 (m, 6H). ¹³C-NMR (75 MHz; CDCl₃): δ 175.3, 168.6, 156.6, 143.9, 143.7, 141.3, 135.8, 134.7, 132.6, 131.9, 131.4, 129.5, 129.0, 128.6, 128.4, 128.3, 127.7, 127.1, 126.7, 125.5, 125.2, 123.6, 120.0, 67.1, 53.7, 47.1, 39.9, 32.0, 28.9, 22.6.

HRMS *m/z* calculated for C₃₂H₃₀N₂NaO₅ [MNa⁺] Calc. 545.2047, Obs. 545.2047

Fmoc-L-Dap(4-DAPA)-OH

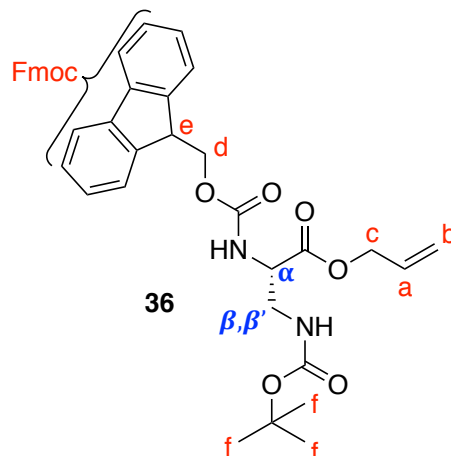
Synthesis of 4-dimethylaminophthalic anhydride



4-aminophthalic acid (500 mg, 2.76 mmol) was suspended in MeOH (150 mL). The mixture was purged with three cycles of evacuation/argon gas saturation. A 37 % aqueous solution of formaldehyde (15 mL) and Pd/C 10 % (200 mg) were added under a gentle stream of argon gas. The mixture was once again purged with three cycles of evacuation/argon gas saturation. A balloon filled with hydrogen gas was connected to the reaction vessel thanks to needle *via* a rubber septum. The mixture was purged with hydrogen and stirred for 3 hours at room temperature under hydrogen atmosphere. The mixture was then purged with argon gas then filtered through a pad of Celite, which was washed with MeOH. The solvent was removed under reduced pressure and dried under vacuum. The solid was suspended in acetic anhydride (50 mL) and stirred at 50 °C overnight in a round-bottom flask equipped with a water-cooled condenser. The solvent was removed under reduced pressure. The residue was redissolved in toluene and dried again under reduced pressure twice. The yellow solid was purified by silica gel column chromatography with DCM as the mobile phase to afford the desired product as a yellow solid in 80 % yield. ¹H-NMR (300 MHz; CDCl₃): δ 7.73 (d, *J* = 8.7 Hz, 1H, **a**), 7.06 (d, *J* = 2.3 Hz, 1H, **b**), 6.94 (dd, *J* = 8.7, 2.4 Hz, 1H, **c**), 3.16 (s, 6H, **d**). ¹³C-NMR (75 MHz; CDCl₃):

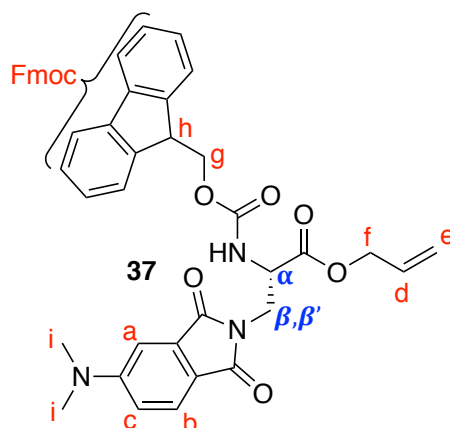
δ 164.4 (C_{IV}, CO-O), 163.1 (C_{IV}, CO-O), 155.4 (C_{IV}, C-N), 134.2 (C_{IV}, C-CO-O), 127.1 (C_{IV}, C-CO-O), 117.3 (CH, **b**), 115.8 (CH, **a**), 106.5 (CH, **c**), 40.7 (CH₃, **d**).

Fmoc-L-Dap(Boc)-OAl



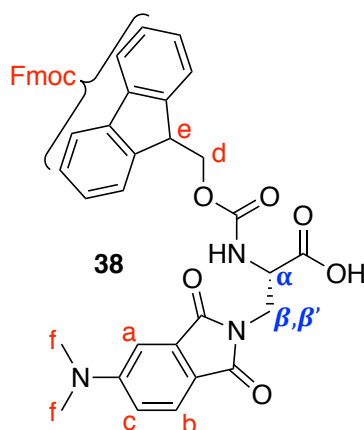
Fmoc-L-Dap(Boc)-OH (1.0 g, 2.34 mmol) was dissolved in MeOH (33 mL). Cesium carbonate (380 mg, 1.17 mmol, 0.5 eq.) was added. The mixture was stirred at room temperature for 30 min. then the solvent was removed under reduced pressure to dryness. The solid was redissolved in dry DMF (67 mL) and allyl bromide (622 μ L, 7.19 mmol, 3.1 eq.) was added. The mixture was stirred at room temperature under argon atmosphere for 1 hour. TLC monitoring (DCM/MeOH; 99:1) showed complete transformation of the starting material. The reaction mixture was poured in 300 mL of a 2 % NaHCO₃ aqueous solution. This water layer was extracted with EtOAc (4 x 100 mL). The combined organic extracts were dried over MgSO₄ then filtered before the solvent was removed under reduced pressure. The residue was purified by silica gel column chromatography (DCM/MeOH; 99:1) to afford the desired product as a white solid in 96 % yield. ¹H-NMR (300 MHz; CDCl₃): δ 7.73 (d, J = 7.4 Hz, 2H, Fmoc), 7.59 (br, 2H, Fmoc), 7.37 (t, J = 7.4 Hz, 2H, Fmoc), 7.27 (t, J = 7.4 Hz, 2H, Fmoc), 6.29 (br, 1H), 5.89 (ddt, J = 17.2, 10.4, 5.6 Hz, 1H, **a**), 5.31 (d, J = 17.2 Hz, 2H, **b**), 5.22 (d, J = 10.4 Hz, 1H, **c**), 4.39 (m, 3H, **d** and α), 4.20 (t, J = 7.0 Hz, 1H, **e**), 3.57 (br, 2H, β and β'), 1.43 (s, 9H, **f**). ¹²C-NMR (75 MHz; CDCl₃): δ 170.2 (C_{IV}, CO-O), 162.6 (C_{IV}, O-CO-NH), 156.4 (C_{IV}, O-CO-NH), 143.7 (C_{IV}, Fmoc), 141.2 (C_{IV}, Fmoc), 131.5 (CH, **a**), 127.6 (CH, Fmoc), 127.0 (CH, Fmoc), 125.1 (CH, Fmoc), 119.9 (CH, Fmoc), 118.8 (CH₂, **b**), 79.7 (C_{IV}, tBu), 67.1 (CH₂, **d**), 66.2 (CH₂, **c**), 55.1 (CH, α), 47.0 (CH₂, β), 42.0 (CH, **e**), 28.2 (CH₃, **f**).

Fmoc-L-Dap(4-DAPA)-OAll



Fmoc-L-Dap(Boc)-OAll (1.05 g, 2.46 mmol) was dissolved in dry DCM (20 mL) and the solution was cooled to 0 °C. TFA (20 mL) was added slowly and the mixture was stirred for 30 minutes at 0°C then 1 hour at room temperature. The solvent was removed under pressure to dryness. The residue was dissolved in dry DMF (30 mL). DIPEA (10 mL, excess) and 4-dimethylaminophthalic anhydride (518 mg, 2.71 mmol, 1.1 eq.) were added. The mixture was stirred overnight at room temperature. A solution of HBTU/HOBt (0.2 M each) in dry DMF (10 mL) was added and the mixture was stirred for 2 more hours at room temperature. 5 mL of the aforementioned solution were added and the mixture was stirred for an additional hour at room temperature. TLC monitoring with Cyclohexane/EtOAc (7:3) as the mobile phase showed complete transformation of the starting material. The reaction mixture was poured in 500 mL of a 5 % NaHCO₃ aqueous solution. This water layer was extracted with EtOAc (3 x 100 mL). The combined organic extracts were dried over MgSO₄ and filtered before the solvent was removed under reduced pressure. The residue was purified by silica gel column chromatography with Cyclohexane/EtOAc (7:3) as the mobile phase to afford the desired product as a yellow solid in 53 % yield. ¹H-NMR (300 MHz; CDCl₃): δ 7.73 (d, J = 7.4 Hz, 2H, a), 7.62-7.55 (m, 3H, Fmoc), 7.37 (t, J = 7.4 Hz, 2H, Fmoc), 7.32-7.27 (m, 2H Fmoc), 7.03 (dd, J = 9.5, 2.3 Hz, 1H, b), 6.79-6.72 (m, 1H, c), 6.00-5.87 (m, 2H, e), 5.38-5.23 (m, 2H, g), 4.76-4.61 (m, 3H, h and α and β), 4.37 (dd, J = 10.0, 6.7 Hz, 1H, β'), 3.05 (s, 6H, i). ¹³C-NMR (75 MHz; CDCl₃): δ 169.8 (C_{IV}, CO-O), 169.2 (C_{IV}, CO-N), 168.6 (C_{IV}, CO-N), 155.9 (C_{IV}, HN-CO-O), 154.5 (C_{IV}, C-N), 143.8 (C_{IV}, Fmoc), 141.2 (C_{IV}, Fmoc), 134.5 (CH, d), 131.5 (C_{IV}, C-CO-O), 127.7 (CH, Fmoc), 127.1 (CH, b), 125.5 (CH, Fmoc), 119.9 (CH, Fmoc), 119.2 (CH₂, e), 117.1 (CH, a), 114.9 (CH, c), 67.5 (CH₂, g), 66.7 (CH₂, f), 53.8 (CH, α), 47.2 (CH₂, β), 40.5 (CH, h), 38.7 (CH₃, i).

Fmoc-L-Dap(4-DAPA)-OH



Fmoc-L-Dap(4-DAPA)-OAll (700 mg, 1.30 mmol) was dissolved in dry DCM (30 mL). Phenylsilane (4.8 mL, 39 mmol, 30 eq.) and Pd(PPh₃)₄ (60 mg, 0.052 mmol, 0.04 eq, ~ 10 %) were added. The mixture was stirred at room temperature under argon atmosphere for 1 hour. TLC monitoring (DCM/MeOH; 95:5) showed complete transformation of the starting material. The reaction mixture was directly loaded onto a silica gel column chromatography and purified (DCM/MeOH; 95:5) to afford the desired product as a yellow solid in 97 % yield. ¹H-NMR (300 MHz; DMSO-d₆): δ 7.85 (d, *J* = 7.4 Hz, 2H, **a**), 7.61 (d, *J* = 7.2 Hz, 1H, **Fmoc**), 7.49 (d, *J* = 7.1 Hz, 2H, **Fmoc** and **b**), 7.38 (td, *J* = 7.4, 3.8 Hz, 2H, **Fmoc**), 7.28 (dt, *J* = 12.7, 6.6 Hz, 2H, **Fmoc**), 6.86 (d, *J* = 1.6 Hz, 1H, **Fmoc**), 6.78 (dd, *J* = 8.3, 2.1 Hz, 1H, **c**), 4.23-4.13 (m, 2H, **d** and **β**), 3.97-3.85 (m, 3H, **e** and **α** and **β'**), 3.73 (dd, *J* = 13.7, 10.4 Hz, 1H, **β'**), 2.93 (s, 6H, **f**). ¹³C-NMR (75 MHz; DMSO): δ 168.4 (C_{IV}, CO-N), 167.9 (C_{IV}, CO-N), 155.7 (C_{IV}, O-CO-NH), 153.9 (C_{IV}, C-N), 144.1 (C_{IV}, **Fmoc**), 143.5 (C_{IV}, **Fmoc**), 140.6 (C_{IV}, **Fmoc**), 140.5 (C_{IV}, **Fmoc**), 134.3 (C_{IV}, C-CO-N), 127.5 (CH, **b**), 127.1 (CH, **Fmoc**), 127.0 (CH, **Fmoc**), 125.4 (CH, **Fmoc**), 125.0 (CH, **Fmoc**), 124.2 (CH, **Fmoc**), 120.0 (CH, **a**), 114.6 (CH, **Fmoc**), 105.0 (CH, **c**), 65.6 (CH₂, **d**), 62.1 (CH₂, **β**), 53.9 (CH, **α**), 46.6 (CH, **e**), 39.9 (CH₃, **f**).

2) Synthesis of β-Hairpin Peptides

• Manual Loading of the first amino acid on 2-Chlorotrityl chloride resin

2-Chlorotrityl chloride resin (1.0 g, 1.6 mmol) was swollen with DCM for 30 min. and a solution of the Fmoc-protected amino acid (0.8 mmol, 0.5 eq.) and DIEA (348 μL, 2 mmol) in DCM (10 mL) is added. The suspension is shaken at room temperature for 3 hours then washed with a mixture of DCM/MeOH/DIEA (17:2:1, 3 times), DCM (3 times), DMF (3 times), DCM (3 times), and MeOH (3 times) then dried under vacuum. The determination of the amino acid loading was achieved by calculating the ratio between the quantity of amino acid used in

millimoles and the mass of dried resin in grams. The loading of the resin was approximately of 0.5 mmol/g.

- **Manual Solid Phase Peptide Synthesis**

The amino acid chain elongation was performed manually starting from 0.25 mmol of resin. The resin was first swollen in DCM for 30 min. and washed with DMF (3 times). For removal of the N-terminal Fmoc groups, the resin was treated with a 20 % solution of piperidine in DMF twice (once 5 min. and once 10 min.) then drained and washed with DMF (3 times).

Standard *N*-Fmoc protected amino acids (1.0 mmol, 4 eq.) and HBTU (360 mg, 0.95 mmol, 3.8 eq.) were dissolved in about 7 mL of DMF. Activation occurred upon the addition of DIPEA (348 μ L, 2 mmol, 8 eq.). The activated amino acid was then transferred to the resin and couplings were performed for 30 minutes with shaking.

Modified amino acids (0.5 mmol, 2 eq.) and HBTU (180 mg, 0.48 mmol, 1.9 eq.) (for Fmoc-L-Dap(4-DAPA)-OH, HATU (185 mg, 0.48 mmol, 1.9 eq.)) were dissolved in about 7 mL of DMF. Activation occurred upon the addition of DIPEA (174 μ L, 1 mmol, 4 eq.). The activated amino acid was then transferred to the resin and couplings were performed for 60 minutes with shaking.

After draining and washing with DMF (3 times), completion of the coupling reaction was assessed by a colorimetric test (Kaiser or Chloranil or TNBS). In the case of a positive test, showing incomplete coupling, the coupling was repeated and the colorimetric test performed once more afterwards. The cycle was repeated as many times as required to obtain the desired sequence assembled on the resin.

The resin was finally washed with DMF (3 times), DCM (3 times) and MeOH (3 times) then dried under vacuum.

- **Cleavage of the Protected Peptide Fragment from the resin**

The linear side chains protected peptide was cleaved from the resin by treatment with a 25 % solution of HFIPA in DCM once for 1 hour and once again for 30 min. The resin was washed with DCM (3 times). The cleavage solutions and rinsing solutions were gathered and the solvents were removed under reduced pressure.

- **Macrolactamization**

The residue obtained from the cleavage step was dissolved in about 300 mL of DMF. HATU (91 mg, 0.24 mmol, 0.95 eq.) and DIPEA (580 μ L, 3.4 mmol, 14 eq.) were added. The mixture

was stirred overnight at room temperature under argon atmosphere. The DMF was removed under reduced pressure to dryness.

- **Deprotection of Peptides**

The residue from the macrolactamization step was treated with about 7 mL of a mixture of TFA/H₂O/TIS/EDT (94:2.5:1:2.5) and stirred at room temperature for 3 hours under argon atmosphere. Volatiles were removed under reduced pressure. The deprotected peptide was precipitated by dropwise addition of the dark oily residue to cooled Et₂O and centrifugation at 8500 rpm for 5 min. The Et₂O was removed and the same operation was repeated twice. The crude product was dissolved in a H₂O/ACN/TFA (90/10/0.1) mixture and lyophilized prior to purification by RP-HPLC. The collected fractions were combined and lyophilized to afford the peptide as a powder.

Hairp0: *cyclo*(Dap(Ac)WDap(Ac)^DPPDap(Ac)ADap(Ac)RG)

Analytical RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): t_r = 4.7 min, 98 % purity

MALDI-MS: m/z calculated for C₅₅H₈₃N₁₉O₁₄ [MH⁺] Calc. 1234.64, Obs. 1234.84

24 % yield

Hairp1: *cyclo*(CWDap(Ac)^DPPDap(Ac)ADap(Ac)RG)

Analytical RP-HPLC (H₂O/MeCN, 0.1 % TFA, 5 to 100 % MeCN in 10 minutes): t_r = 6.9 min, 98 % purity

MALDI-MS: m/z calculated for C₅₀H₇₃N₁₇O₁₃S [MH⁺] Calc. 1152.53, Obs. 1152.24

14 % yield

Hairp1-pNA: *cyclo*(CGlu(pNA)^LPPAMFRG)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 35 to 50 % MeCN in 10 minutes): t_r = 8.2 min, 99 % purity

MALDI-MS: m/z calculated for C₅₅H₈₀N₁₅O₁₃S₂ [MH⁺] Calc. 1222.54, Obs. 1222.42

15 % yield

Hairp1-Naphthoyl-loop: *cyclo*(CRDap(Ac)^DPPDap(Ac)ACK(2-Naphthoyl)G)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 20 to 50 % MeCN in 10 minutes): t_r = 6.5 min,
97 % purity

MALDI-MS: m/z calculated for C₅₆H₈₁N₁₇O₁₄S [MH⁺] Calc. 1248.59, Obs. 1248.70
7 % yield

Hairp1-Naphthoyl-strand: *cyclo*(CK(2-Naphthoyl)Dap(Ac)^DPPDap(Ac)ADap(Ac)RG)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 10 to 60 % MeCN in 10 minutes): t_r = 7.2 min,
96 % purity

MALDI-MS: m/z calculated for C₅₆H₈₁N₁₇O₁₄S [MH⁺] Calc. 1248.59, Obs. 1248.70
4 % yield

Hairp1-DAPA-Ala: *cyclo*(CDap(4-DAPA)Dap(Ac)^DPPDap(Ac)ACRG)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 5 to 100 % MeCN in 10 minutes): t_r = 5.3 min,
99 % purity

MALDI-MS: m/z calculated for C₅₂H₇₆N₁₈O₁₅S [MH⁺] Calc. 1225.55, Obs. 1225.50
6 % yield

Hairp1-DAPA-Lys: *cyclo*(CDap(4-DAPA)Dap(Ac)^DPPDap(Ac)KCRG)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 15 to 30 % MeCN in 10 minutes): t_r = 7.8 min,
99 % purity

MALDI-MS: m/z calculated for C₅₅H₈₃N₁₉O₁₅S [MH⁺] Calc. 1282.60, Obs. 1282.73
8 % yield

Hairp2-Naphthoyl: *cyclo*(CK(2-Naphthoyl)Dap(Ac)^DPPDap(Ac)ACRG)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 10 to 60 % MeCN in 10 minutes): t_r = 7.4 min,
97 % purity

MALDI-MS: m/z calculated for C₅₀H₇₃N₁₇O₁₃S [MH⁺] Calc. 1223.54, Obs. 1223.77
7 % yield

Hairp4-Ala: *cyclo*(CWC^DPPCACRG)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 20 to 50 % MeCN in 10 minutes): t_r = 7.3 min,
99 % purity

MALDI-MS: m/z calculated for C₄₄H₆₄N₁₄O₁₀S₄ [MH⁺] Calc. 1077.38, Obs. 1077.52
13 % yield

Hairp4-Lys: *cyclo*(CWC^DPPCKCRG)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 20 to 50 % MeCN in 10 minutes): t_r = 6.5 min, 98 % purity

MALDI-MS: m/z calculated for C₄₇H₇₁N₁₅O₁₀S₄ [MH⁺] Calc. 1134.44, Obs. 1134.43
19 % yield

Hairp4-DAPA-Lys: *cyclo*(CDap(4-DAPA)C^DPPCKCRG)

Analytical RP-HPLC (H₂O/MeCN, 0.1%TFA, 20 to 50 % MeCN in 10 minutes): t_r = 7.0 min, 99 % purity

MALDI-MS: m/z calculated for C₄₉H₇₄N₁₆O₁₂S₄ [MH⁺] Calc. 1207.46, Obs. 1207.70
8 % yield

3) Synthesis of Stapled Helix Peptides

- **Manual Loading of the first amino acid on MBHA resin**

MBHA resin (3.7 g, 2.0 mmol) was swollen in a glass SPPS vessel with DCM for 1 hour, washed with DMF, neutralised with a 10 % DIEA solution in DMF and washed again with DMF. A solution of Fmoc-L-Asp(tBu)-OH (411 mg, 1.0 mmol, 0.5 eq.), HBTU (379 mg, 1.0 mmol, 0.5 eq.) and DIEA (435 μ L, 2.5 mmol, 2.5 eq.) in DMF (15 mL) is added. The suspension was mixed regularly at room temperature for 2 hours then washed with DMF (3 times). A capping step was performed with a 20 % Ac₂O solution in DCM and an excess of DIEA (1.75 mL, 10 mmol) for 40 min and after washing with DMF (3 times), a Kaiser test showed complete reaction of the free amines from the resin. The resin was washed with MeOH and dried under vacuum. The determination of the amino acid loading was achieved by calculating the ratio between the quantity of amino acid used in millimoles and the mass of dried resin in grams. The loading of the resin was approximately of 0.25 mmol/g.

- **Manual Solid Phase Peptide Synthesis**

- **Synthesis of the common segment** 

The amino acid chain elongation was performed manually starting from 0.75 mmol of resin. The resin was first swollen in DCM for 30 min. and washed with DMF (3 times). For removal of the N-terminal Fmoc groups, the resin was treated with a 20 % solution of piperidine in DMF twice (twice 5 min) then drained and washed with DMF (3 times).

N-Fmoc protected amino acids (Fmoc-L-Dap(Ac)-OH, Fmoc-L-Arg(Tos)-OH, Fmoc-L-Cys(Mob)-OH, Fmoc-L-Lys(2-Cl-Z)-OH or Fmoc-L-Lys(Boc)-OH) (1.0 mmol, 1.33 eq.) and HBTU (360 mg, 0.95 mmol, 1.27 eq.) were dissolved in about 12 mL of DMF. Activation occurred upon the addition of DIPEA (340 μ L, 2 mmol, 2.7 eq.). The activated amino acid was then transferred to the resin and couplings were performed for 2 hours with regular mixing. After draining and washing with DMF (3 times), completion of the coupling reaction was assessed by a colorimetric test (Kaiser or Chloranil or TNBS). In the case of a positive test, showing incomplete coupling, the coupling was repeated and the colorimetric test performed once more afterwards. The cycle was repeated as many times as required to obtain the desired sequence assembled on the resin.

- **Side chain to side chain on resin coupling**

After coupling Fmoc-L-Lys(Boc)-OH and before Fmoc removal, the lysine and the aspartic side chains were deprotected with a TFA/H₂O/TIS (95:2.5:2.5) mixture (four times 2 min and 1 time 10 min). The resin was drained and washed with DCM (3 times), DMF (3 times), neutralized with 10 % DIEA in DMF and washed with DMF (3 times).

A solution of HATU (285 mg, 0.75 mmol, 1.0 eq.) and DIEA (255 μ L, 1.5 mmol, 2 eq.) in DMF (12 mL) was added and coupling was performed for 2 hours with regular mixing. After draining and washing with DMF (3 times), completion of the coupling reaction was assessed by a colorimetric test (Kaiser or Chloranil or TNBS). In the case of a positive test, showing incomplete coupling, the coupling was repeated and the colorimetric test performed once more afterwards.

The *N*-ter Fmoc group was removed with the usual protocol. The resin was washed with DMF (3 times) and MeOH (3 times) and dried under vacuum.

- **Synthesis of longer stapled helix sequences**

The amino acid chain elongation was performed manually starting from 0.25 mmol of resin either from the Fmoc-L-Asp(*t*Bu)-OH loaded resin or from the resin bearing the fragment described above. In the latter case, the resin was split accordingly to reach a 0.25 mmol scale. The resin was first swollen in DCM for 30 min. and washed with DMF (3 times). For removal of the *N*-terminal Fmoc groups, the resin was treated with a 20 % solution of piperidine in DMF twice (once 5 min. and once 10 min.) then drained and washed with DMF (3 times).

N-Fmoc protected amino acids (Fmoc-L-Asp(*O**t*Bu)-OH, Fmoc-L-Trp(Boc)-OH, Fmoc-L-Dap(Ac)-OH, Fmoc-L-Arg(Tos)-OH, Fmoc-L-Cys(Mob)-OH, Fmoc-L-Lys(2-Cl-Z)-OH or Fmoc-L-Lys(Boc)-OH) (1.0 mmol, 4 eq.) and HBTU (360 mg, 0.95 mmol, 3.8 eq.) were

dissolved in about 7 mL of DMF. Activation occurred upon the addition of DIPEA (348 μ L, 2 mmol, 8 eq.). The activated amino acid was then transferred to the resin and couplings were performed for 30 minutes with shaking.

After draining and washing with DMF (3 times), completion of the coupling reaction was assessed by a colorimetric test (Kaiser or Chloranil or TNBS). In the case of a positive test, showing incomplete coupling, the coupling was repeated and the colorimetric test performed once more afterwards. The cycle was repeated as many times as required to obtain the desired sequence assembled on the resin.

The side chain to side chain coupling was performed in the same fashion as described above. After removal of the final Fmoc group, N-terminus acetylation was performed with a 20 % Ac₂O solution in DCM and excess of DIEA for 30 min.

The resin was finally washed with DMF (3 times), DCM (3 times) and MeOH (3 times) then dried under vacuum.

- **Cleavage from the resin and side chain deprotections**

Scavengers were added to the dried resin prior to HF addition accordingly to the following recipe: anisole (1.5 mL/g of resin), Me₂S (0.25 mL/g of resin), *p*-thiocresol (300 mg/g of resin). The peptides were cleaved from the resin and the side chain protecting groups were removed by treatment with HF for 2 hours in an HF ramp. After removal of the HF under vacuum, the deprotected peptide was precipitated by addition of cooled Et₂O. The resin and peptide crude material were filtered through a fritted glass filter under vacuum. The crude product was dissolved in a 10 % AcOH aqueous solution and filtered to remove the resin bead. The aqueous solution was lyophilised prior to purification by RP-HPLC. The collected fractions were combined and lyophilized to afford the peptide as a powder.

Helix1-1: AcW_{cyclo}(2→6)[KCRDap(Ac)D]-NH₂

Analytical RP-HPLC (H₂O/MeCN, 0.1 % TFA, 15 to 30 % MeCN in 10 minutes): t_r = 7.0 min, >99 % purity

MALDI-MS: m/z calculated for C₃₇H₅₆N₁₃O₉S [MH⁺] Calc. 858.40, Obs. 858.35

41 % yield

Helix2-2: AcW_{cyclo}(2→6, 7→11)[KDap(Ac)KCDKCRDap(Ac)D]-NH₂

Analytical RP-HPLC (H₂O/MeCN, 0.1 % TFA, 15 to 30 % MeCN in 10 minutes): t_r = 9.3 min, 97 % purity

MALDI-MS: m/z calculated for C₆₁H₉₆N₂₁O₁₆S₂ [MH⁺] Calc. 1442.67, Obs. 1442.12
6 % yield

Helix2-4: AcW_{cyclo}(2→6, 7→11)[KCKCDKCRCD]-NH₂

Analytical RP-HPLC (H₂O/MeCN, 0.1 % TFA, 20 to 50 % MeCN in 10 minutes): t_r = 5.4 min, 99 % purity

MALDI-MS: m/z calculated for C₅₇H₉₀N₁₉O₉S₄ [MH⁺] Calc. 1392.57, Obs. 1392.12
19 % yield

III. Dynamic Combinatorial Chemistry

- **Preparation of a 100 mM Phosphate Buffer, 10 mM TCEP, 1% TWEEN-20**

In a 100 mL RBF containing a stir bar and 40 mL of water were added successively: 693 mg of Na₂HPO₄·12H₂O, 417 mg of KH₂PO₄ and 143 mg of TCEP. The pH was adjusted to 7, 7.3 or 7.5 with 1M NaOH solution, then 500 μL of TWEEN-20 were added and the volume was adjusted with distilled water at 50 mL before degazing the solution for 15 min. The pH value was checked to be the desired value and re-adjusted accordingly if needed.

For the catalytic experiment, 0.25 mmol (0.85 mg) of imidazole was added before adjusting the pH.

- **Dynamic combinatorial experiments and analysis**

- **Generation of the dynamic combinatorial library**

BB stock solutions of 50 mM and peptides stock solutions at 10 mM concentrations were prepared in one of the buffers described above.

The library was generated by mixing 10 μL of the peptide stock solution and the appropriate amount of BB stock solutions to reach the desired equivalency to the peptide. The volume is adjusted to 100 μL total with buffer to reach a 1 mM concentration of peptide.

The evolution of the mixture was monitored by collecting 20 μL of the mixture, diluted to 0.2 mM of peptide concentration with 80 μL of a quenching mixture (H₂O/MeCN/TFA; 50:50:0.1) and finally analyzed by analytical RP-HPLC analysis (cf Experimental Section, I. Building Blocks Synthesis, General Considerations, About products characterization).

The peaks of the DCC products were manually collected from the analytical RP-HPLC elutions in Eppendorf tubes and analysis by MALDI-MS was performed to identify the product. These fractions were also lyophilized before linearizing them by enzymatic digestion.

- **Peptide linearization by trypsin digestion**

Fractions isolated from DCC experiment after analytical RP-HPLC and lyophilization were redissolved with 5 μ L ACN and 5 μ L H₂O. The walls of the Eppendorf tubes were carefully washed with this mixture and were stored at -4°C . Trypsin digestion was performed by diluting 1 μ L in a mixture of H₂O/ACN of variable proportion (from 0 % ACN up to 50 % ACN) and trypsin solution (0.1 $\mu\text{g}/\mu\text{L}$; 0.5 μL or 1 μL) to a final volume of 10 μL . The digestion mixtures were incubated at 37°C for the desired amount of time before MALDI-MS analysis.

IV. Structural Investigations

- **NMR peptide sample preparation**

The buffer was prepared by dissolving KH₂PO₄ (16.3 mg), Na₂HPO₄·12 H₂O (28.7 mg) and TCEP (25 mg) in 8 mL of mQ water. pH was adjusted to 7 with a 1 M NaOH solution. The volume was adjusted to 10 mL with mQ water to obtain a 20 mM phosphate buffer, 10 mM TCEP at pH 7 (buffer A). Aliquots of 1 mL were prepared in Eppendorf tubes, lyophilized and redissolved in 1 mL of H₂O/D₂O (10 %).

A second buffer was prepared by diluting 1 mL of the above mentioned one in 9 mL of mQ water to obtain a 2 mM phosphate buffer, 1 mM TCEP at pH 7 (buffer B).

NMR sample for **Hairp4-Lys** was prepared by dissolving 4.8 mg of peptide powder in 600 μL of buffer B. The sample was lyophilized and redissolved in 60 μL of a H₂O/D₂O (10%)/DSS 100X-11mM in D₂O (0.1 %) mixture. After dissolution and centrifugation, the supernatant was transferred into a Shigemi tube. Dilution of the sample was performed with the appropriate volume of buffer A from above mentioned aliquots to reach the desired concentration.

• Attribution Table for Hairp4-Lys

Cys 1	δ (ppm)	Multiplicité	J (Hz)		Cys 6	δ (ppm)	Multiplicité	J (Hz)
NH	7,63	d		8,9	NH	7,76	d	9,1
H α	4,76				H α	4,72	dd	15,1 ; 8,4
2H β	3,00 ; 2,84	m			2H β	3,0 ; 2,85	m	
C α	56,9				C α	57,7		
C β	30				C β	30		
Trp 2	δ (ppm)	Multiplicité	J (Hz)		Lys 7	δ (ppm)	Multiplicité	J (Hz)
NH ϵ 1	10,16	s			NH	8,65		
NH	9,02	d		7,1	H α	4,1	m	
H α	4,66	m			2H β	0,99	m	
2H β	3,32 ; 3,08	dd	15,0 ; 2,2 -- 15,1 ; 11,0		2H γ	0,41 ; 0,15	m	
H δ 1	7,3	s			2H δ	1,17 ; 0,48	m	
H ξ 2	7,47	d		8,4	2H ϵ	2,32	m	
H η 2	7,21	t		7,9	C α	55,6		
H ξ 3	6,99	t		7,9	C β	27		
H ϵ 3	7,55	d		8,6	C γ	24,3		
C α	58,5				C δ	32		
C β	29,8				C ϵ	40		
C γ	109,6				Cys 8	δ (ppm)	Multiplicité	J (Hz)
C δ 1	126				NH	8,63		
C ξ 2	114,5				H α	4,51	q	16,4 ; 8,0
C η 2	124,7				2H β	2,84	m	
C ξ 3	122,2				C α	57,9		
C ϵ 3	121,2				C β	29		
C ϵ 2	136,1				Arg 9	δ (ppm)	Multiplicité	J (Hz)
Cys 3	δ (ppm)	Multiplicité	J (Hz)		NH	9,42	br	
NH	9,2	d		9,8	H α	3,93	m	
H α	4,99				2H β	1,93 ; 1,85	m	
2H β	2,86	dd	15,1 ; 7,2		2H γ	1,71 ; 1,58	m	
C β	27,6				2H δ	3,2	m	
Pro 4	δ (ppm)	Multiplicité	J (Hz)		C α	57,5		
H α	4,82				C β	28,2		
2H β	2,33 ; 1,89	m			C γ	27,7		
2H γ	2,13 ; 2,01	m			C δ	43,5		
2H δ	3,90 ; 3,48	m			Gly 10	δ (ppm)	Multiplicité	J (Hz)
C α	61,3				NH	8,61		
C β	30,5				2H CH2	4,12 ; 3,56	m	
C γ	27,9				C α	45,5		
C δ	51							
Pro 5	δ (ppm)	Multiplicité	J (Hz)					
H α	4,65							
2H β	2,26	m						
2H γ	2,09	m						
2H δ	4,05 ; 3,74	m						
C α	63,8							
C β	32,3							
C γ	26,3							
C δ	50,3							

NMR samples for **Hairp0** was prepared by dissolving 4.1 mg of peptide powder in 560 μL of buffer A. The sample was lyophilized and redissolved in 560 μL of a $\text{H}_2\text{O}/\text{D}_2\text{O}$ (10%)/DSS 100X-11mM in D_2O (0.1 %) mixture. After dissolution and centrifugation, the supernatant was transferred into a 5 mm diameter NMR tube.

NMR samples for **Hairp4-Lys-tetraBBGly**^{1,3,6,8} was prepared by dissolving 1.40 mg of peptide powder in 188 μL of buffer A. The sample was lyophilized and redissolved in 188 μL of a $\text{H}_2\text{O}/\text{D}_2\text{O}$ (10%)/DSS 100X-11mM in D_2O (0.1 %) mixture. After dissolution and centrifugation, the supernatant was transferred into a Shigemi tube.

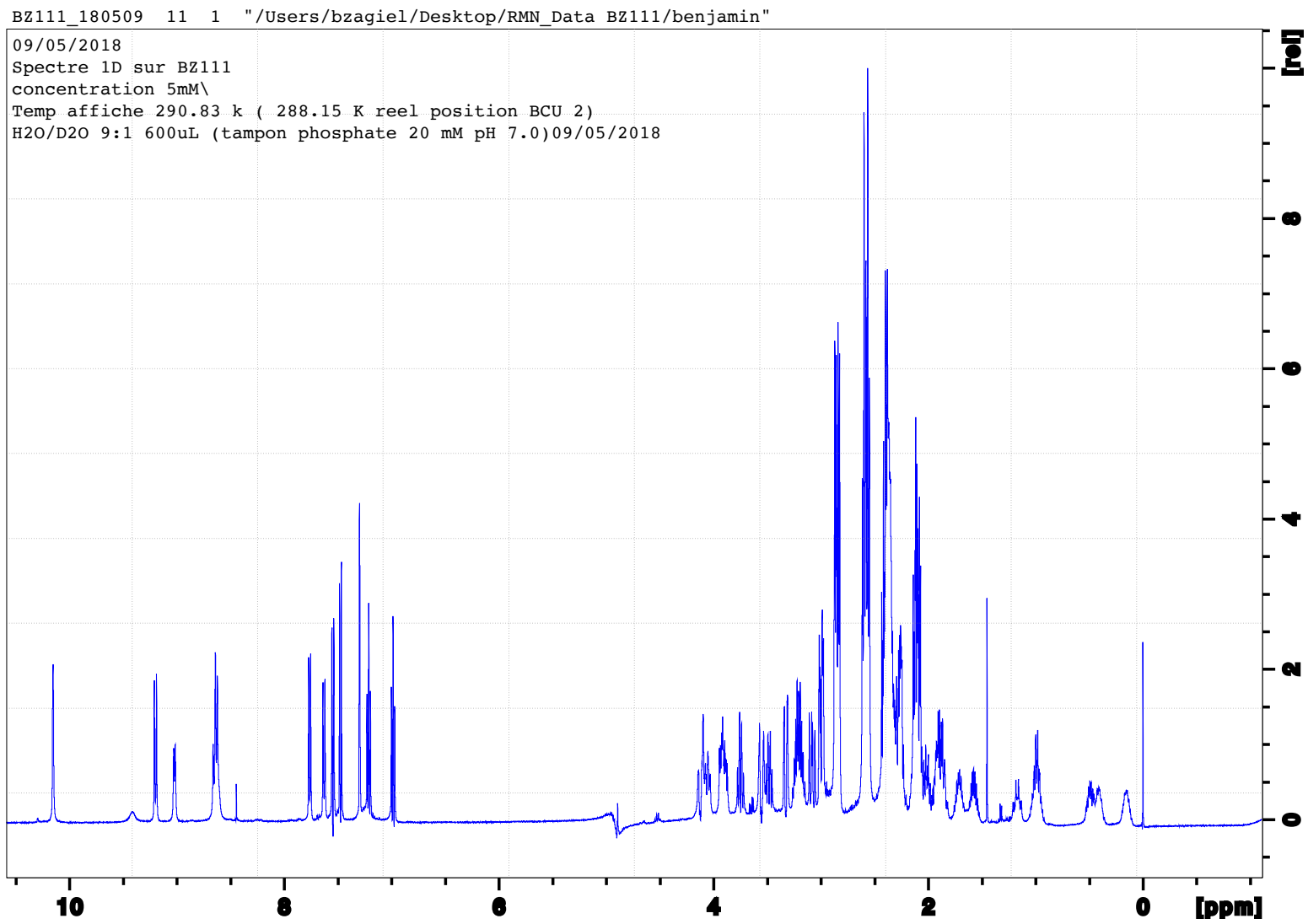
NMR experiments for peptide samples were performed on a Bruker Avance III 500 MHz spectrometer equipped with a TCI cryoprobe.

- **Molecular Dynamics**

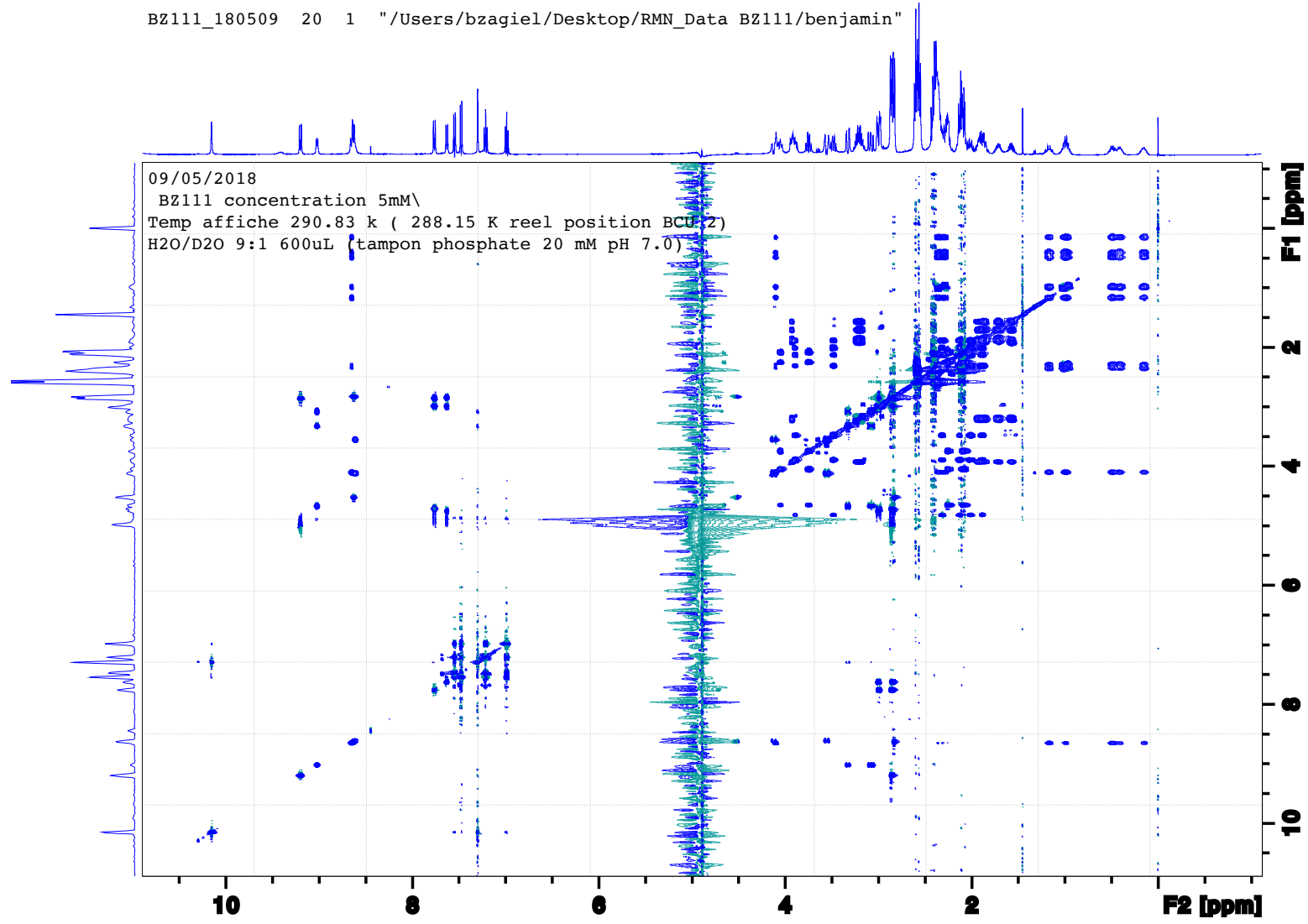
All molecular dynamics simulations were performed using the Amber 18 software package and Amber force fields. The peptide moiety was created using the xLEaP program and parametrized by the Antechamber program suite.

The starting structure has been generated using a simulated annealing protocol consisting of 10000 steps including a heating stage (600 K) followed by a cooling to 0 K. The structure obtained by simulated annealing was subjected to a molecular dynamics simulation in an implicit solvent model. Temperature was equilibrated at 300 K. Molecular dynamics simulations were performed at constant temperature for 10 ns using 1 fs time steps. Coordinates and energies were recorded every 1000 steps, allowing analyses of trajectories. Analysis of hydrogen bonding was conducted by measurement of distances between proton of amides groups and oxygen of carbonyl groups for each snapshots of the trajectory using the cptraj program.

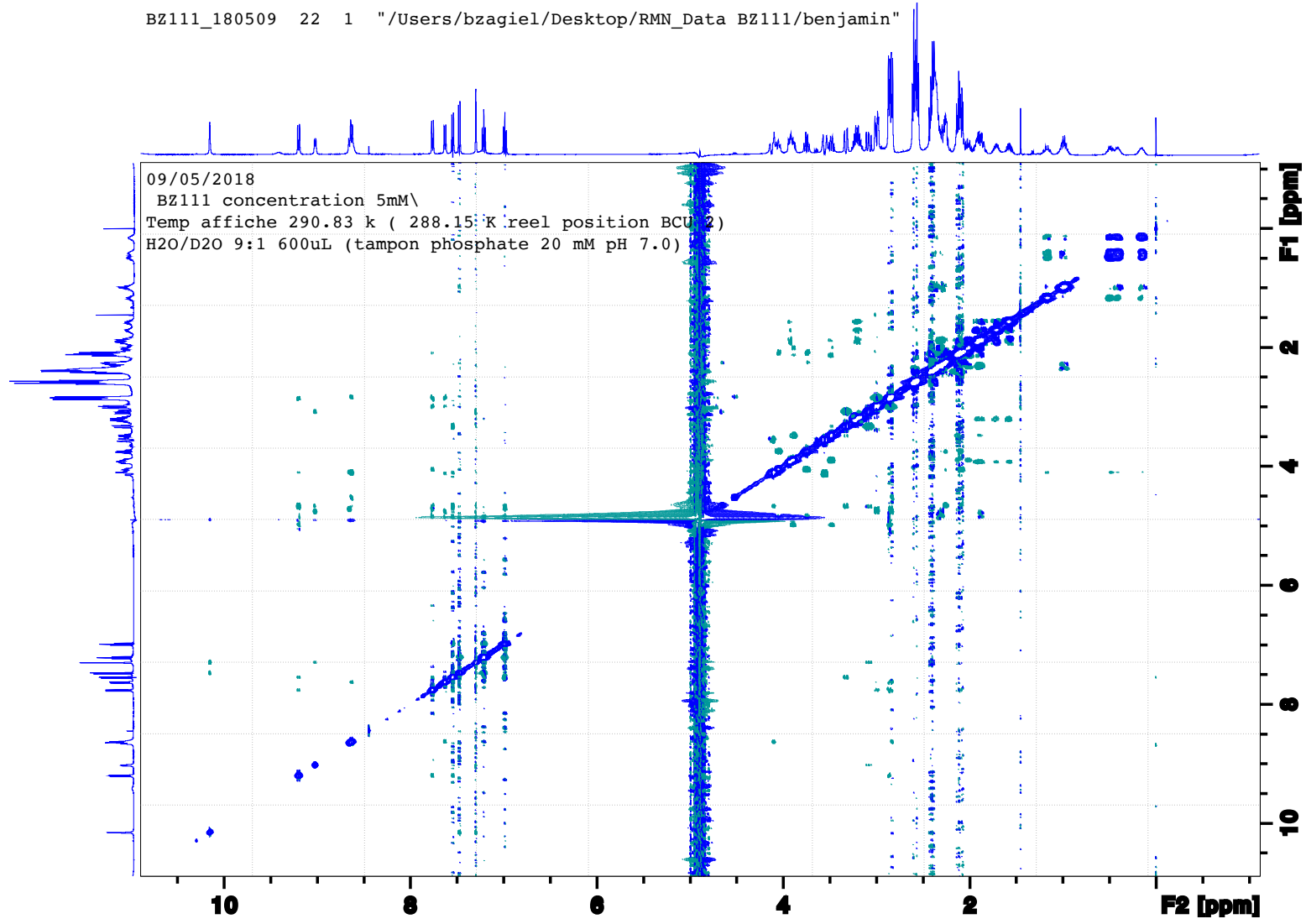
Annexes
Spectre 1H (15 °C) de Hairp4-Lys



Spectre TOCSY de Hairp4 -Lys,

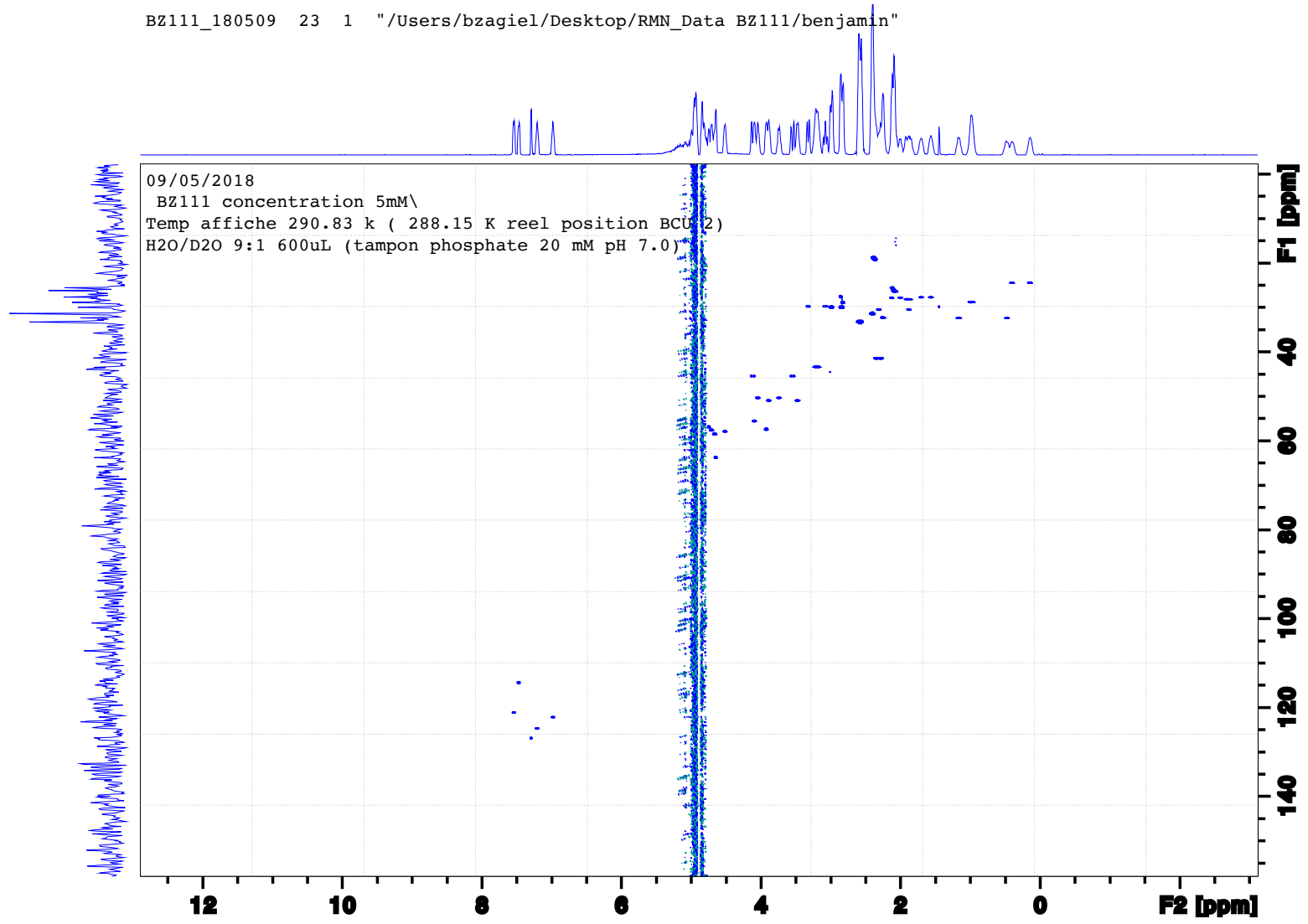


Spectre ROESY de Hairp4-Lys

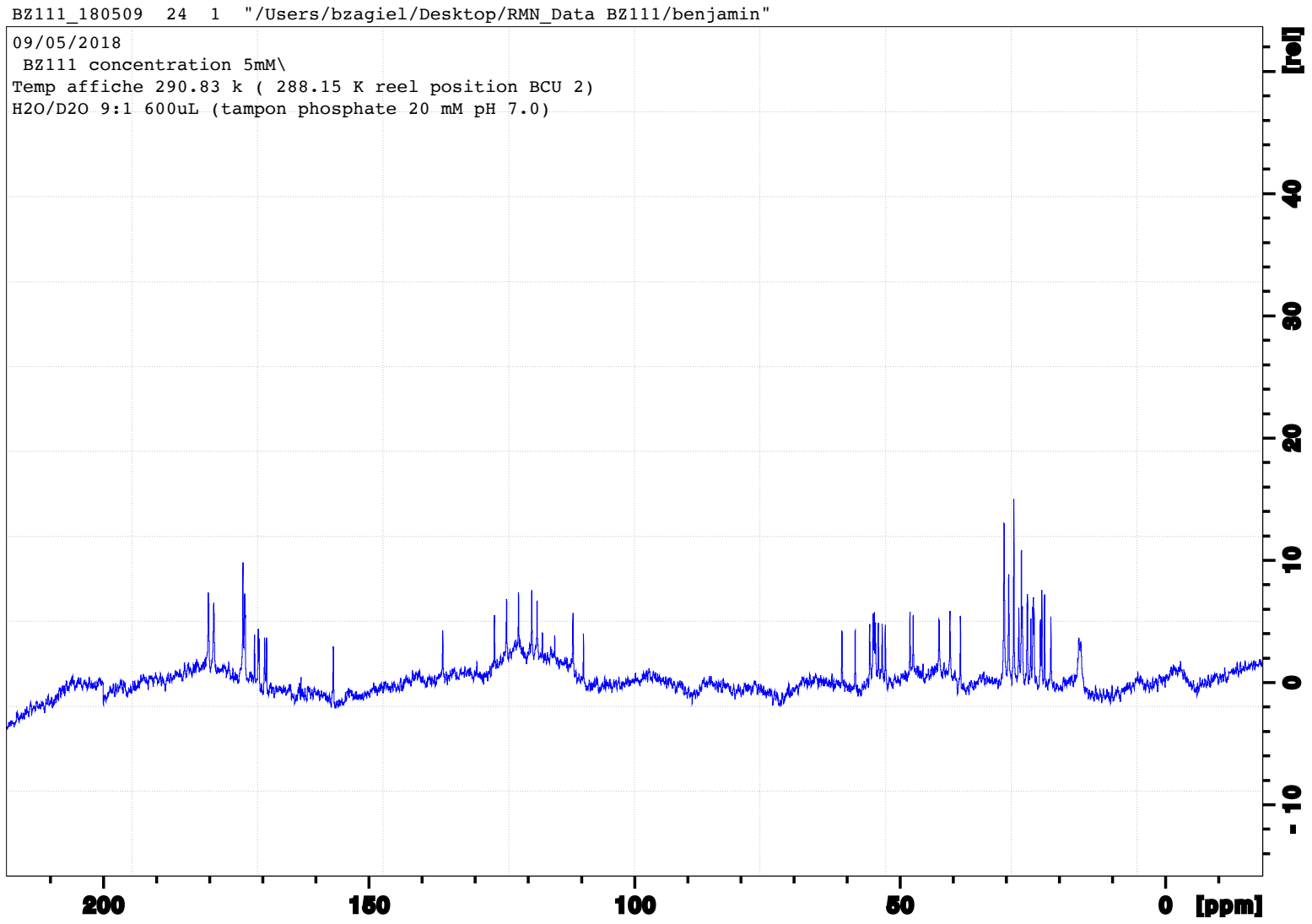


Spectre HSQC de Hairp4-Lys

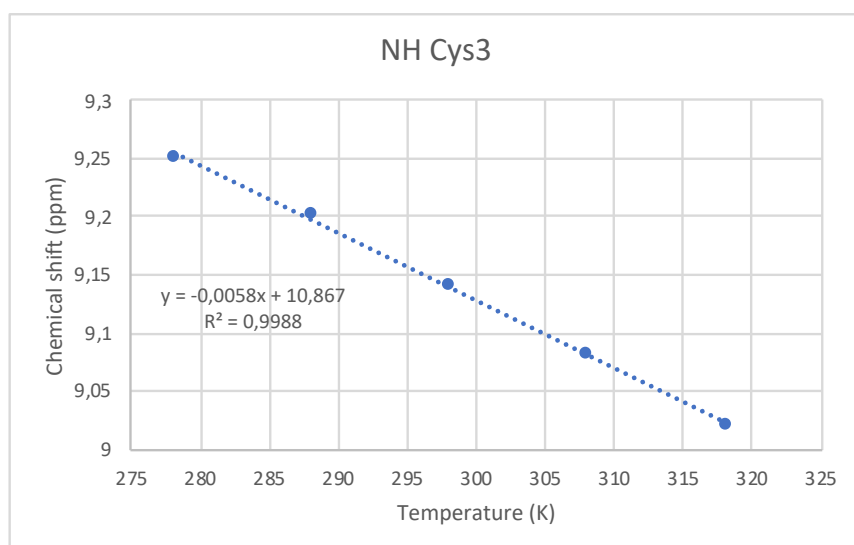
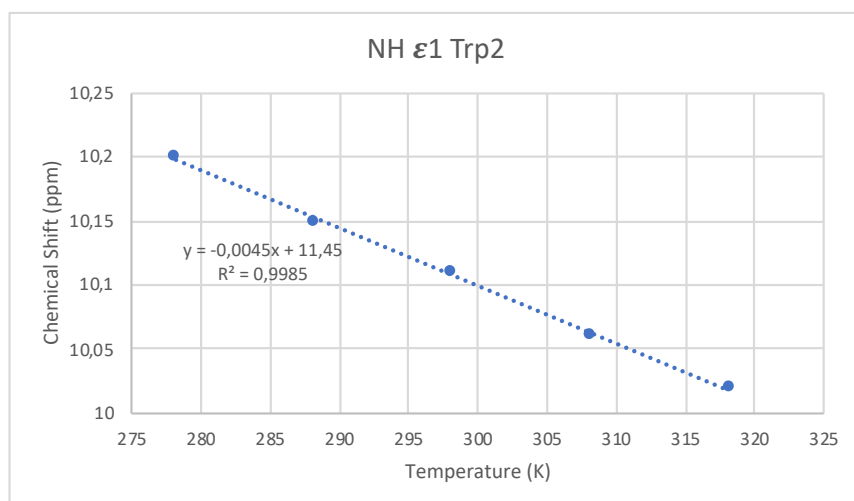
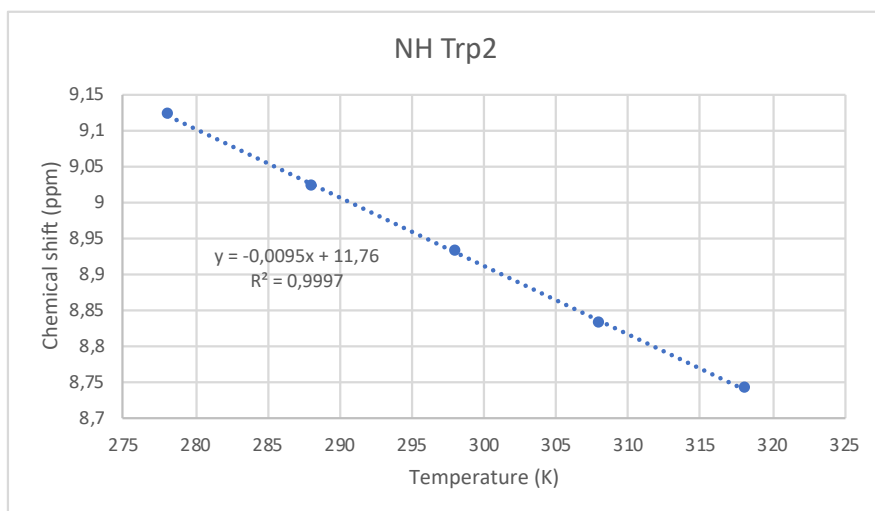
BZ111_180509 23 1 "/Users/bzagiell/Desktop/RMN_Data BZ111/benjamin"

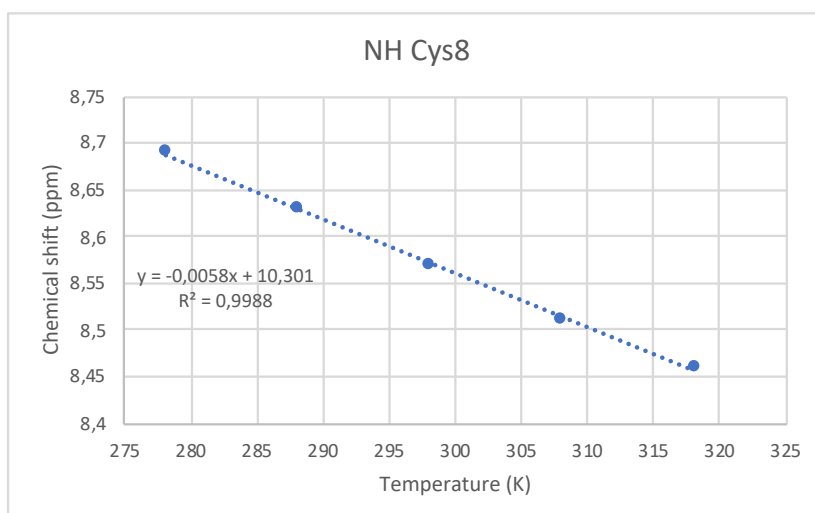
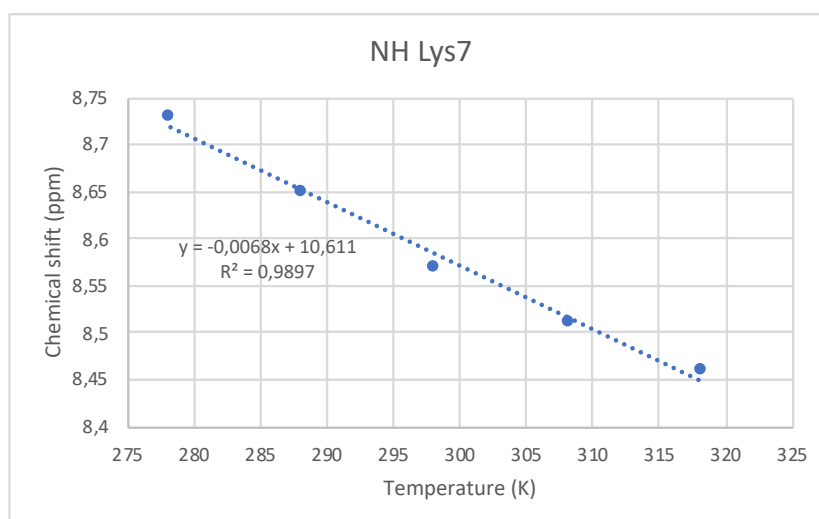
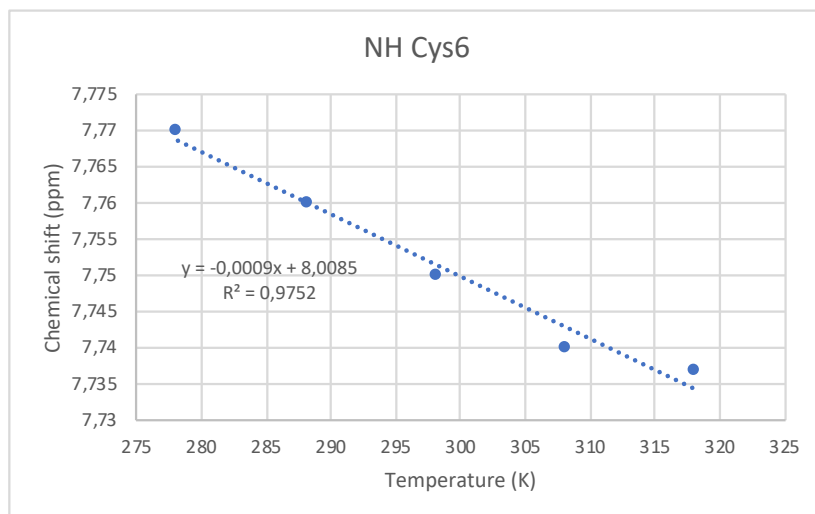


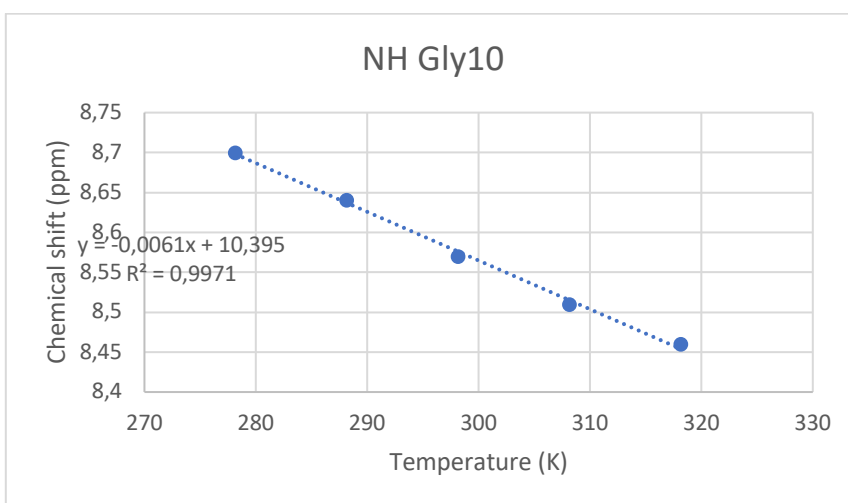
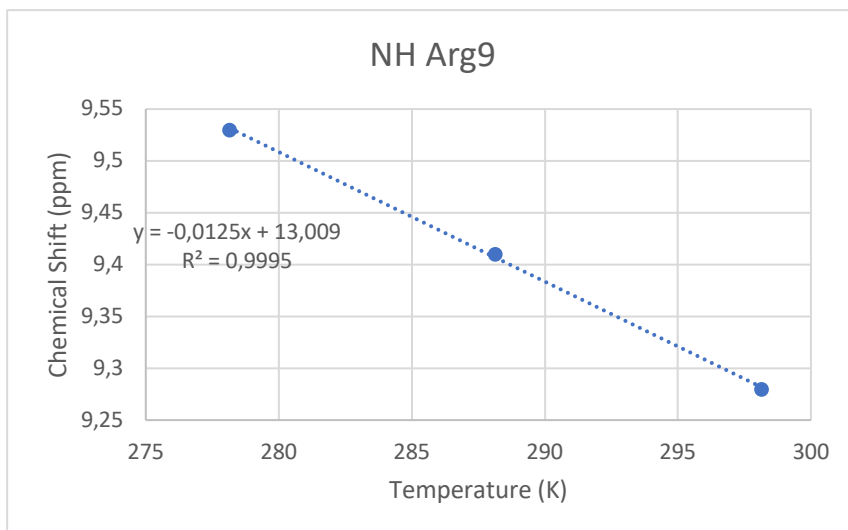
Spectre 13C de Hairp4-Lys



Graphiques de l'évolution des déplacements chimiques des protons amides de Hairp4-Lys en fonction de la température







Tableaux des distances calculées à partir de l'intégration des tâches ROE (Hairp4-Lys) en prenant comme référence de normalisation la tâche de corrélation entre les protons du CH₂ de la glycine qui correspond par convention à une distance de 1,8 Å entre ces protons.

Distances inter-résidus séquentielles le long du squelette peptidique

n°résidu 1	Nom résidu 1	Atome du résidu 1	n°résidu2	Nom résidu 2	Atome du résidu 2	d (Å)
1	Cys	H (NH)	10	Gly	Ha*	1.7
1	Cys	H (NH)	10	Gly	Ha*	3.1
1	Cys	H (NH)	2	Trp	H (NH)	4
2	Trp	H (NH)	1	Cys	Ha	2.5
3	Cys	H (NH)	2	Trp	Ha	3.1
6	Cys	H (NH)	5	Pro	Ha	3.4
7	Lys	H (NH)	6	Cys	Ha	2.1
7	Lys	H (NH)	6	Cys	H (NH)	4
8	Cys	H (NH)	7	Lys	Ha	3
9	Arg	H (NH)	8	Cys	Ha	2.8
9	Arg	Ha	8	Cys	H (NH)	4.1
10	Gly	H (NH)	9	Arg	Ha	2.9
10	Gly	H (NH)	1	Cys	H (NH)	3.1

Distances inter-résidus séquentielles inter-chaînes latérales

n°résidu 1	Nom résidu 1	Atome du résidu 1	n°résidu2	Nom résidu 2	Atome du résidu 2	d (Å)
2	Trp	H (NH)	1	Cys	Hb*	2.7
2	Trp	H (NH)	1	Cys	Hb*	4
3	Cys	H (NH)	2	Trp	He3	2.5
4	Pro	Hd*	3	Cys	Ha	2.2
4	Pro	Hd*	3	Cys	Ha	3.3
5	Pro	Hd*	4	Pro	Hb*	4.6
6	Cys	H (NH)	5	Pro	Hb*	2.6
6	Cys	H (NH)	5	Pro	Hb*	3.9
6	Cys	H (NH)	5	Pro	Hd*	2.6
6	Cys	H (NH)	5	Pro	Hd*	2.6
6	Cys	H (NH)	5	Pro	Hd*	3.9
7	Lys	H (NH)	6	Cys	Hb*	3
8	Cys	H (NH)	7	Lys	Hg*	2.6

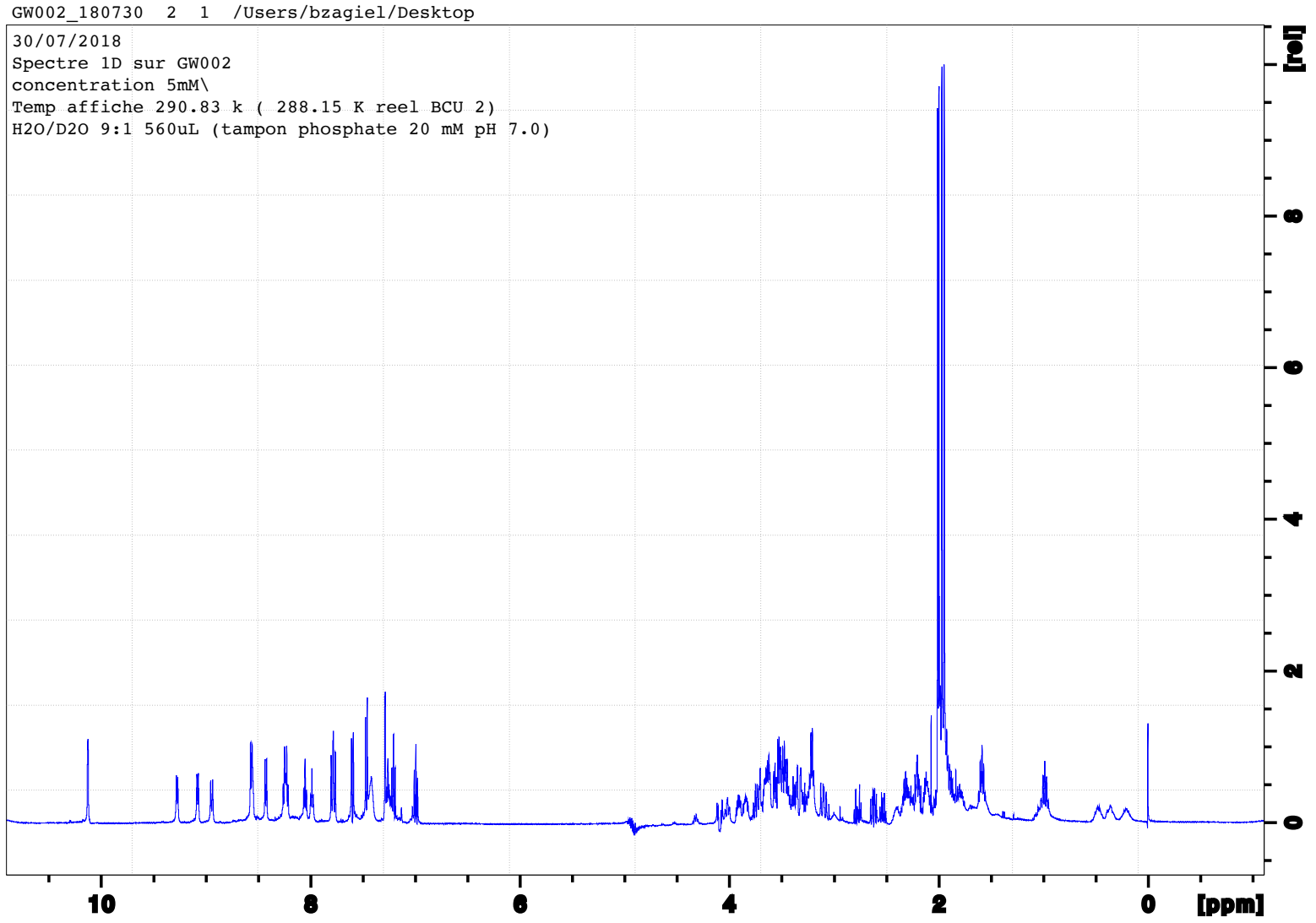
Longues distances entre atomes du squelette

n°résidu 1	Nom résidu 1	Atome du résidu 1	n°résidu2	Nom résidu 2	Atome du résidu 2	d (Å)
3	Cys	H (NH)	6	Cys	H (NH)	3.1
3	Cys	H (NH)	7	Lys	Ha	4
7	Lys	Ha	10	Gly	Ha*	5.3
7	Lys	Ha	2	Trp	Ha	4.6
8	Cys	H (NH)	1	Cys	H (NH)	3.5

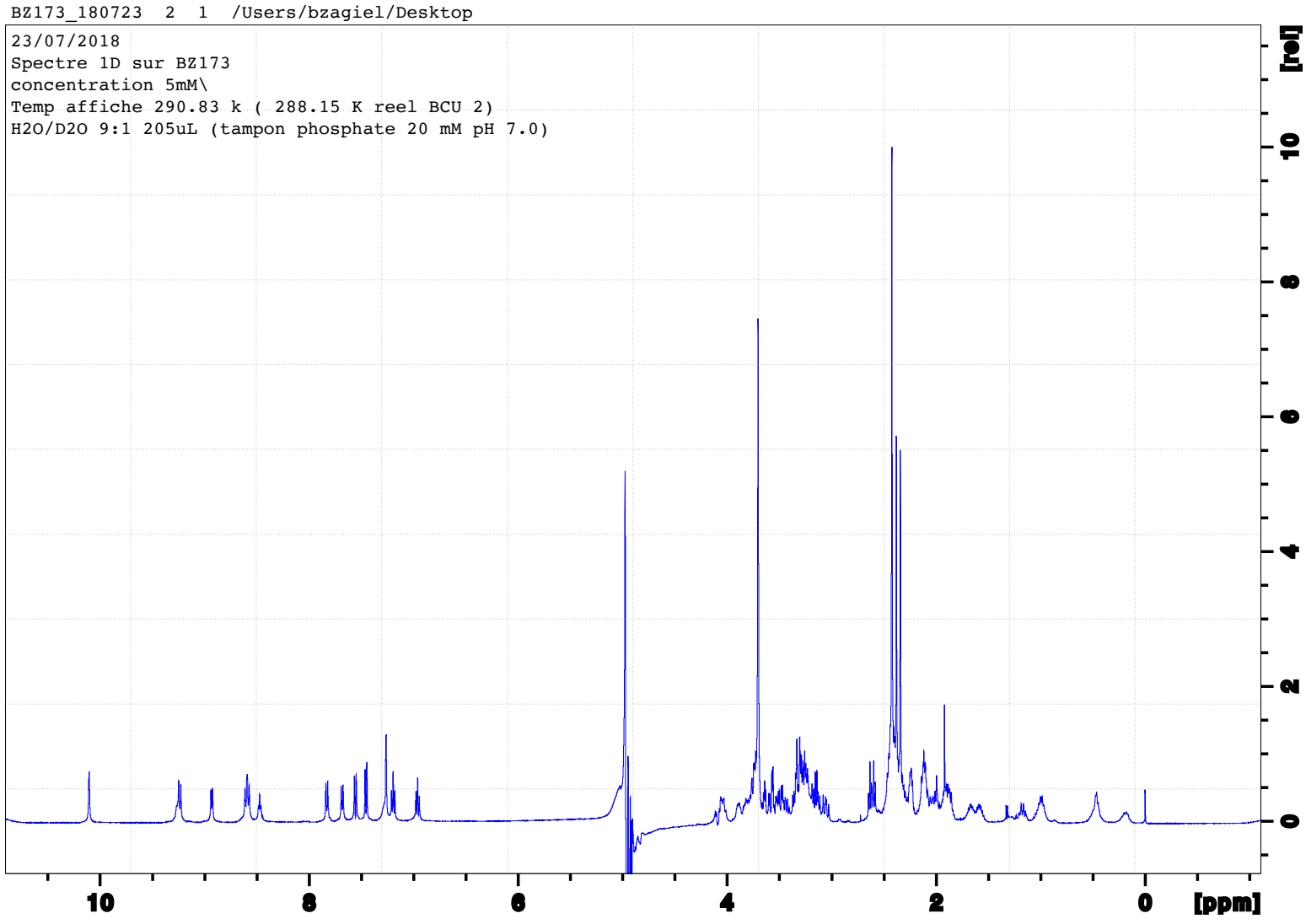
Longues distances entre atomes des chaînes latérales

n°résidu 1	Nom résidu 1	Atome du résidu 1	n°résidu2	Nom résidu 2	Atome du résidu 2	d (Å)
2	Trp	He3	7	Lys	Ha	2.7
3	Cys	H (NH)	2	Trp	Hb*	1.6/3.0
3	Cys	Hb*	6	Cys	H (NH)	3.1
3	Cys	Hb*	6	Cys	H (NH)	4.6
3	Cys	H (NH)	5	Pro	Hd*	4
7	Lys	Hg*	2	Trp	Ha	3.2
7	Lys	Hg*	2	Trp	Ha	4.7
7	Lys	H (NH)	2	Trp	H23	4.4
7	Lys	Hb*	2	Trp	H22	4.3
7	Lys	Hb*	2	Trp	H22	4.9
7	Lys	Hb*	2	Trp	Ha	4
7	Lys	Hb*	2	Trp	Ha	5.5
7	Lys	H (NH)	2	Trp	He3	3.2

Spectre 1H (15°C) de Hairp0



Spectre ^1H (15°C) de Hairp4-Lys-tetraBBGly^{1,3,6,8}



Références bibliographiques

- (1) Joerger, A. C.; Fersht, A. R. The P53 Pathway: Origins, Inactivation in Cancer, and Emerging Therapeutic Approaches. *Annual Review of Biochemistry* **2016**, *85* (1), 375–404. <https://doi.org/10.1146/annurev-biochem-060815-014710>.
- (2) Olivier, M.; Hollstein, M.; Hainaut, P. TP53 Mutations in Human Cancers: Origins, Consequences, and Clinical Use. *Cold Spring Harbor Perspectives in Biology* **2010**, *2* (1), a001008–a001008. <https://doi.org/10.1101/cshperspect.a001008>.
- (3) Gonzalez, M. W.; Kann, M. G. Chapter 4: Protein Interactions and Disease. *PLoS Computational Biology* **2012**, *8* (12), e1002819. <https://doi.org/10.1371/journal.pcbi.1002819>.
- (4) Ideker, T.; Sharan, R. Protein Networks in Disease. *Genome Research* **2008**, *18* (4), 644–652. <https://doi.org/10.1101/gr.071852.107>.
- (5) MacDonald, M. E.; Ambrose, C. M.; Duyao, M. P.; Myers, R. H.; Lin, C.; Srinidhi, L.; Barnes, G.; Taylor, S. A.; James, M.; Groot, N. A Novel Gene Containing a Trinucleotide Repeat That Is Expanded and Unstable on Huntington's Disease Chromosomes. *Cell* **1993**, *72* (6), 971–983.
- (6) Goehler, H.; Lalowski, M.; Stelzl, U.; Waelter, S.; Stroedicke, M.; Worm, U.; Droege, A.; Lindenberg, K. S.; Knoblich, M.; Haenig, C. A Protein Interaction Network Links GIT1, an Enhancer of Huntingtin Aggregation, to Huntington's Disease. *Molecular cell* **2004**, *15* (6), 853–865.
- (7) Duennwald, M. L.; Jagadish, S.; Giorgini, F.; Muchowski, P. J.; Lindquist, S. A Network of Protein Interactions Determines Polyglutamine Toxicity. *Proceedings of the National Academy of Sciences* **2006**, *103* (29), 11051–11056.
- (8) Giorgini, F.; Muchowski, P. J. Connecting the Dots in Huntington's Disease with Protein Interaction Networks. *Genome biology* **2005**, *6* (3), 210.
- (9) Delacourte, A.; Flament, S.; Dibe, E. M.; Hublau, P.; Sablionière, P.; Hemon, B.; Sherrer, V.; Defossez, A. Pathological Proteins Tau 64 and 69 Are Specifically Expressed in the Somatodendritic Domain of the Degenerating Cortical Neurons during Alzheimer's Disease. *Acta neuropathologica* **1990**, *80* (2), 111–117.
- (10) Williams, D. R. Tauopathies: Classification and Clinical Update on Neurodegenerative Diseases Associated with Microtubule-Associated Protein Tau. *Internal Medicine Journal* **2006**, *36* (10), 652–660. <https://doi.org/10.1111/j.1445-5994.2006.01153.x>.
- (11) Wu, F. T. H.; Stefanini, M. O.; Mac Gabhann, F.; Kontos, C. D.; Annex, B. H.; Popel, A. S. Computational Kinetic Model of VEGF Trapping by Soluble VEGF Receptor-1: Effects of Transendothelial and Lymphatic Macromolecular Transport. *Physiological Genomics* **2009**, *38* (1), 29–41. <https://doi.org/10.1152/physiolgenomics.00031.2009>.
- (12) Kuzmanov, U.; Emili, A. Protein-Protein Interaction Networks: Probing Disease Mechanisms Using Model Systems. *Genome medicine* **2013**, *5* (4), 37.
- (13) Jäger, S.; Cimermanic, P.; Gulbahce, N.; Johnson, J. R.; McGovern, K. E.; Clarke, S. C.; Shales, M.; Mercenne, G.; Pache, L.; Li, K.; et al. Global Landscape of HIV–Human Protein Complexes. *Nature* **2012**, *481* (7381), 365–370. <https://doi.org/10.1038/nature10719>.
- (14) Jäger, S.; Gulbahce, N.; Cimermanic, P.; Kane, J.; He, N.; Chou, S.; D'Orso, I.; Fernandes, J.; Jang, G.; Frankel, A. D.; et al. Purification and Characterization of HIV–Human Protein Complexes. *Methods* **2011**, *53* (1), 13–19. <https://doi.org/10.1016/j.ymeth.2010.08.007>.
- (15) Komarova, A. V.; Combredet, C.; Meyniel-Schicklin, L.; Chapelle, M.; Caignard, G.; Camadro, J.-M.; Lotteau, V.; Vidalain, P.-O.; Tangy, F. Proteomic Analysis of Virus-Host Interactions in an Infectious Context Using Recombinant Viruses. *Molecular & Cellular Proteomics* **2011**, *10* (12), M110.007443. <https://doi.org/10.1074/mcp.M110.007443>.
- (16) Chatel-Chaix, L.; Melancon, P.; Racine, M.-E.; Baril, M.; Lamarre, D. Y-Box-Binding Protein 1 Interacts with Hepatitis C Virus NS3/4A and Influences the Equilibrium between Viral RNA Replication and Infectious Particle Production. *Journal of Virology* **2011**, *85* (21), 11022–11037. <https://doi.org/10.1128/JVI.00719-11>.

- (17) Yang, X.; Promnares, K.; Qin, J.; He, M.; Shroder, D. Y.; Kariu, T.; Wang, Y.; Pal, U. Characterization of Multiprotein Complexes of the *Borrelia Burgdorferi* Outer Membrane Vesicles. *Journal of Proteome Research* **2011**, *10* (10), 4556–4566. <https://doi.org/10.1021/pr200395b>.
- (18) Marzoa, J.; Sánchez, S.; Ferreirós, C. M.; Criado, M. T. Identification of *Neisseria Meningitidis* Outer Membrane Vesicle Complexes Using 2-D High Resolution Clear Native/SDS-PAGE. *Journal of Proteome Research* **2010**, *9* (1), 611–619. <https://doi.org/10.1021/pr9006409>.
- (19) Modell, A. E.; Blosser, S. L.; Arora, P. S. Systematic Targeting of Protein–Protein Interactions. *Trends in Pharmacological Sciences* **2016**, *37* (8), 702–713. <https://doi.org/10.1016/j.tips.2016.05.008>.
- (20) Thompson, A. D.; Dugan, A.; Gestwicki, J. E.; Mapp, A. K. Fine-Tuning Multiprotein Complexes Using Small Molecules. *ACS Chemical Biology* **2012**, *7* (8), 1311–1320. <https://doi.org/10.1021/cb300255p>.
- (21) Surade, S.; Blundell, T. L. Structural Biology and Drug Discovery of Difficult Targets: The Limits of Ligandability. *Chemistry & Biology* **2012**, *19* (1), 42–50. <https://doi.org/10.1016/j.chembiol.2011.12.013>.
- (22) Hopkins, A. L.; Groom, C. R. The Druggable Genome. *Nature reviews Drug discovery* **2002**, *1* (9), 727.
- (23) Edfeldt, F. N. B.; Folmer, R. H. A.; Breeze, A. L. Fragment Screening to Predict Druggability (Ligandability) and Lead Discovery Success. *Drug Discovery Today* **2011**, *16* (7–8), 284–287. <https://doi.org/10.1016/j.drudis.2011.02.002>.
- (24) Wells, J. A.; McClendon, C. L. Reaching for High-Hanging Fruit in Drug Discovery at Protein–Protein Interfaces. *Nature* **2007**, *450* (7172), 1001–1009. <https://doi.org/10.1038/nature06526>.
- (25) Jones, S.; Thornton, J. M. Principles of Protein-Protein Interactions. *Proceedings of the National Academy of Sciences* **1996**, *93* (1), 13–20.
- (26) Conte, L. L.; Chothia, C.; Janin, J. The Atomic Structure of Protein-Protein Recognition Sites1. *Journal of molecular biology* **1999**, *285* (5), 2177–2198.
- (27) Cheng, A. C.; Coleman, R. G.; Smyth, K. T.; Cao, Q.; Soulard, P.; Caffrey, D. R.; Salzberg, A. C.; Huang, E. S. Structure-Based Maximal Affinity Model Predicts Small-Molecule Druggability. *Nature Biotechnology* **2007**, *25* (1), 71–75. <https://doi.org/10.1038/nbt1273>.
- (28) Smith, R. D.; Hu, L.; Falkner, J. A.; Benson, M. L.; Nerothin, J. P.; Carlson, H. A. Exploring Protein–Ligand Recognition with Binding MOAD. *Journal of Molecular Graphics and Modelling* **2006**, *24* (6), 414–425. <https://doi.org/10.1016/j.jmgm.2005.08.002>.
- (29) Mayer, T. U. Small Molecule Inhibitor of Mitotic Spindle Bipolarity Identified in a Phenotype-Based Screen. *Science* **1999**, *286* (5441), 971–974. <https://doi.org/10.1126/science.286.5441.971>.
- (30) Kapoor, T. M.; Mayer, T. U.; Coughlin, M. L.; Mitchison, T. J. Probing Spindle Assembly Mechanisms with Monastrol, a Small Molecule Inhibitor of the Mitotic Kinesin, Eg5. *The Journal of cell biology* **2000**, *150* (5), 975–988.
- (31) Lopez-Girona, A.; Mendy, D.; Ito, T.; Miller, K.; Gandhi, A. K.; Kang, J.; Karasawa, S.; Carmel, G.; Jackson, P.; Abbasian, M.; et al. Cereblon Is a Direct Protein Target for Immunomodulatory and Antiproliferative Activities of Lenalidomide and Pomalidomide. *Leukemia* **2012**, *26* (11), 2326–2335. <https://doi.org/10.1038/leu.2012.119>.
- (32) Shortt, J.; Hsu, A. K.; Johnstone, R. W. Thalidomide-Analogue Biology: Immunological, Molecular and Epigenetic Targets in Cancer Therapy. *Oncogene* **2013**, *32* (36), 4191.
- (33) Vassilev, L. T. In Vivo Activation of the P53 Pathway by Small-Molecule Antagonists of MDM2. *Science* **2004**, *303* (5659), 844–848. <https://doi.org/10.1126/science.1092472>.
- (34) Walunas, T. L.; Lenschow, D. J.; Bakker, C. Y.; Linsley, P. S.; Freeman, G. J.; Green,

- J. M.; Thompson, C. B.; Bluestone, J. A. CTLA-4 Can Function as a Negative Regulator of T Cell Activation. *Immunity* **1994**, *1* (5), 405–413.
- (35) Erbe, D. V.; Wang, S.; Xing, Y.; Tobin, J. F. Small Molecule Ligands Define a Binding Site on the Immune Regulatory Protein B7.1. *Journal of Biological Chemistry* **2002**, *277* (9), 7363–7368. <https://doi.org/10.1074/jbc.M110162200>.
- (36) Green, N. J.; Xiang, J.; Chen, J.; Chen, L.; Davies, A. M.; Erbe, D.; Tam, S.; Tobin, J. F. Structure–Activity Studies of a Series of Dipyrazolo[3,4- B :3',4'- d]Pyridin-3-Ones Binding to the Immune Regulatory Protein B7.1. *Bioorganic & Medicinal Chemistry* **2003**, *11* (13), 2991–3013. [https://doi.org/10.1016/S0968-0896\(03\)00183-4](https://doi.org/10.1016/S0968-0896(03)00183-4).
- (37) Huang, Z. The Chemical Biology of Apoptosis: Exploring Protein-Protein Interactions and the Life and Death of Cells with Small Molecules. *Chemistry & biology* **2002**, *9* (10), 1059–1072.
- (38) Degtarev, A.; Lugovskoy, A.; Cardone, M.; Mulley, B.; Wagner, G.; Mitchison, T.; Yuan, J. Identification of Small-Molecule Inhibitors of Interaction between the BH3 Domain and Bcl-x L. *Nature cell biology* **2001**, *3* (2), 173.
- (39) Clackson, T.; Wells, J. A. A Hot Spot of Binding Energy in a Hormone-Receptor-Interface. *Science* **1995**, *267*, 383–386.
- (40) Arkin, M. R.; Randal, M.; DeLano, W. L.; Hyde, J.; Luong, T. N.; Oslob, J. D.; Raphael, D. R.; Taylor, L.; Wang, J.; McDowell, R. S.; et al. Binding of Small Molecules to an Adaptive Protein-Protein Interface. *Proceedings of the National Academy of Sciences* **2003**, *100* (4), 1603–1608. <https://doi.org/10.1073/pnas.252756299>.
- (41) Erlanson, D. A. *Introduction to Fragment-Based Drug Discovery*; Topics in current chemistry; Springer: Heidelberg, 2011.
- (42) Sawyer, N.; Watkins, A. M.; Arora, P. S. Protein Domain Mimics as Modulators of Protein–Protein Interactions. *Accounts of Chemical Research* **2017**, *50* (6), 1313–1322. <https://doi.org/10.1021/acs.accounts.7b00130>.
- (43) Tsomaia, N. Peptide Therapeutics: Targeting the Undruggable Space. *European Journal of Medicinal Chemistry* **2015**, *94*, 459–470. <https://doi.org/10.1016/j.ejmech.2015.01.014>.
- (44) Chang, Y. S.; Graves, B.; Guerlavais, V.; Tovar, C.; Packman, K.; To, K.-H.; Olson, K. A.; Kesavan, K.; Gangurde, P.; Mukherjee, A.; et al. Stapled α -helical Peptide Drug Development: A Potent Dual Inhibitor of MDM2 and MDMX for P53-Dependent Cancer Therapy. *Proceedings of the National Academy of Sciences* **2013**, *110* (36), E3445–E3454. <https://doi.org/10.1073/pnas.1303002110>.
- (45) Johnson, L. M.; Barrick, S.; Hager, M. V.; McFedries, A.; Homan, E. A.; Rabaglia, M. E.; Keller, M. P.; Attie, A. D.; Saghatelian, A.; Bisello, A.; et al. A Potent α/β -Peptide Analogue of GLP-1 with Prolonged Action in Vivo. *Journal of the American Chemical Society* **2014**, *136* (37), 12848–12851. <https://doi.org/10.1021/ja507168t>.
- (46) Hill, T. A.; Shepherd, N. E.; Diness, F.; Fairlie, D. P. Constraining Cyclic Peptides To Mimic Protein Structure Motifs. *Angewandte Chemie International Edition* **2014**, *53* (48), 13020–13041. <https://doi.org/10.1002/anie.201401058>.
- (47) Kritzer, J. A.; Zutshi, R.; Cheah, M.; Ran, F. A.; Webman, R.; Wongjirad, T. M.; Schepartz, A. Miniature Protein Inhibitors of the P53-HDM2 Interaction. *ChemBioChem* **2006**, *7* (1), 29–31. <https://doi.org/10.1002/cbic.200500324>.
- (48) Chin, J. W.; Grotzfeld, R. M.; Fabian, M. A.; Schepartz, A. Methodology for Optimizing Functional Miniature Proteins Based on Avian Pancreatic Polypeptide Using Phage Display. *Bioorganic & Medicinal Chemistry Letters* **2001**, *11* (12), 1501–1505. [https://doi.org/10.1016/S0960-894X\(01\)00139-1](https://doi.org/10.1016/S0960-894X(01)00139-1).
- (49) Vita, C.; Roumestand, C.; Toma, F.; Menez, A. Scorpion Toxins as Natural Scaffolds for Protein Engineering. *Proceedings of the National Academy of Sciences* **1995**, *92* (14), 6404–6408. <https://doi.org/10.1073/pnas.92.14.6404>.
- (50) Lee, J.-H.; Kang, E.; Lee, J.; Kim, J.; Lee, K. H.; Han, J.; Kang, H. Y.; Ahn, S.; Oh, Y.; Shin, D.; et al. Protein Grafting of P53TAD onto a Leucine Zipper Scaffold Generates a Potent

- HDM Dual Inhibitor. *Nature Communications* **2014**, *5* (1). <https://doi.org/10.1038/ncomms4814>.
- (51) Szyk, A.; Lu, W.; Xu, C.; Lubkowski, J. Structure of the Scorpion Toxin BmBKTtx1 Solved from Single Wavelength Anomalous Scattering of Sulfur. *Journal of Structural Biology* **2004**, *145* (3), 289–294. <https://doi.org/10.1016/j.jsb.2003.11.012>.
- (52) O’Shea, E. K.; Klemm, J. D.; KiM, P. S.; Alber, T. X-Ray Structure of the GCN4 Leucine Zipper, a Two-Stranded, Parallel Coiled Coil. *6*.
- (53) Blundell, T. L.; Pitts, J. E.; Tickle, I. J.; Wood, S. P.; Wu, C.-W. X-Ray Analysis (1.4-Å Resolution) of Avian Pancreatic Polypeptide: Small Globular Protein Hormone. *Proceedings of the National Academy of Sciences* **1981**, *78* (7), 4175–4179. <https://doi.org/10.1073/pnas.78.7.4175>.
- (54) Fasan, R.; Dias, R. L. A.; Moehle, K.; Zerbe, O.; Obrecht, D.; Mittl, P. R. E.; Grütter, M. G.; Robinson, J. A. Structure-Activity Studies in a Family of β -Hairpin Protein Epitope Mimetic Inhibitors of the P53-HDM2 Protein-Protein Interaction. *ChemBioChem* **2006**, *7* (3), 515–526. <https://doi.org/10.1002/cbic.200500452>.
- (55) Lehn, J.-M. Dynamic Combinatorial Chemistry and Virtual Combinatorial Libraries. *Chemistry - A European Journal* **1999**, *5* (9), 2455–2463. [https://doi.org/10.1002/\(SICI\)1521-3765\(19990903\)5:9<2455::AID-CHEM2455>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1521-3765(19990903)5:9<2455::AID-CHEM2455>3.0.CO;2-H).
- (56) Lehn, J.-M. From Supramolecular Chemistry towards Constitutional Dynamic Chemistry and Adaptive Chemistry. *Chem. Soc. Rev.* **2007**, *36* (2), 151–160. <https://doi.org/10.1039/B616752G>.
- (57) *Constitutional Dynamic Chemistry*; Barboiu, M., Aastrup, T., Eds.; Topics in current chemistry; Springer: Heidelberg, 2012.
- (58) Sadownik, J. W.; Ulijn, R. V. Dynamic Covalent Chemistry in Aid of Peptide Self-Assembly. *Current Opinion in Biotechnology* **2010**, *21* (4), 401–411. <https://doi.org/10.1016/j.copbio.2010.05.010>.
- (59) Roy, L.; Case, M. A. Protein Core Packing by Dynamic Combinatorial Chemistry. *Journal of the American Chemical Society* **2010**, *132* (26), 8894–8896. <https://doi.org/10.1021/ja1029717>.
- (60) Bilgiçer, B.; Xing, X.; Kumar, K. Programmed Self-Sorting of Coiled Coils with Leucine and Hexafluoroleucine Cores. *Journal of the American Chemical Society* **2001**, *123* (47), 11815–11816. <https://doi.org/10.1021/ja016767o>.
- (61) Woll, M. G.; Gellman, S. H. Backbone Thioester Exchange: A New Approach to Evaluating Higher Order Structural Stability in Polypeptides. *Journal of the American Chemical Society* **2004**, *126* (36), 11172–11174. <https://doi.org/10.1021/ja046891i>.
- (62) Fisher, B. F.; Hong, S. H.; Gellman, S. H. Helix Propensities of Amino Acid Residues via Thioester Exchange. *Journal of the American Chemical Society* **2017**, *139* (38), 13292–13295. <https://doi.org/10.1021/jacs.7b07930>.
- (63) Steinkruger, J. D.; Woolfson, D. N.; Gellman, S. H. Side-Chain Pairing Preferences in the Parallel Coiled-Coil Dimer Motif: Insight on Ion Pairing between Core and Flanking Sites. *Journal of the American Chemical Society* **2010**, *132* (22), 7586–7588. <https://doi.org/10.1021/ja100080q>.
- (64) Steinkruger, J. D.; Bartlett, G. J.; Woolfson, D. N.; Gellman, S. H. Strong Contributions from Vertical Triads to Helix-Partner Preferences in Parallel Coiled Coils. *Journal of the American Chemical Society* **2012**, *134* (38), 15652–15655. <https://doi.org/10.1021/ja3063088>.
- (65) Hadley, E. B.; Gellman, S. H. An Antiparallel α -Helical Coiled-Coil Model System for Rapid Assessment of Side-Chain Recognition at the Hydrophobic Interface. *Journal of the American Chemical Society* **2006**, *128* (51), 16444–16445. <https://doi.org/10.1021/ja067178r>.
- (66) Hadley, E. B.; Testa, O. D.; Woolfson, D. N.; Gellman, S. H. Preferred Side-Chain Constellations at Antiparallel Coiled-Coil Interfaces. *Proceedings of the National Academy of Sciences* **2008**, *105* (2), 530–535. <https://doi.org/10.1073/pnas.0709068105>.
- (67) Hadley, E. B.; Witek, A. M.; Freire, F.; Peoples, A. J.; Gellman, S. H. Thermodynamic

- Analysis of β -Sheet Secondary Structure by Backbone Thioester Exchange. *Angewandte Chemie International Edition* **2007**, *46* (37), 7056–7059. <https://doi.org/10.1002/anie.200702449>.
- (68) Scheike, J. A.; Baldauf, C.; Spengler, J.; Albericio, F.; Pisabarro, M. T.; Kokscha, B. Amide-to-Ester Substitution in Coiled Coils: The Effect of Removing Hydrogen Bonds on Protein Structure. *Angewandte Chemie International Edition* **2007**, *46* (41), 7766–7769. <https://doi.org/10.1002/anie.200702218>.
- (69) Bartus, É.; Hegedüs, Z.; Wéber, E.; Csipak, B.; Szakonyi, G.; Martinek, T. A. De Novo Modular Development of a Foldameric Protein-Protein Interaction Inhibitor for Separate Hot Spots: A Dynamic Covalent Assembly Approach. *ChemistryOpen* **2017**, *6* (2), 236–241. <https://doi.org/10.1002/open.201700012>.
- (70) Krishnan-Ghosh, Y.; Balasubramanian, S. Dynamic Covalent Chemistry on Self-Templating Peptides: Formation of a Disulfide-Linked β -Hairpin Mimic. *Angewandte Chemie International Edition* **2003**, *42* (19), 2171–2173. <https://doi.org/10.1002/anie.200250551>.
- (71) Carnall, J. M. A.; Waudby, C. A.; Belenguer, A. M.; Stuart, M. C. A.; Peyralans, J. J.-P.; Otto, S. Mechano-sensitive Self-Replication Driven by Self-Organization. *Science* **2010**, *327* (5972), 1502–1506. <https://doi.org/10.1126/science.1182767>.
- (72) Lam, R. T. S. Amplification of Acetylcholine-Binding Catenanes from Dynamic Combinatorial Libraries. *Science* **2005**, *308* (5722), 667–669. <https://doi.org/10.1126/science.1109999>.
- (73) Rauschenberg, M.; Bomke, S.; Karst, U.; Ravoo, B. J. Dynamic Peptides as Biomimetic Carbohydrate Receptors. *Angewandte Chemie International Edition* **2010**, *49* (40), 7340–7345. <https://doi.org/10.1002/anie.201002847>.
- (74) Gareiss, P. C.; Sobczak, K.; McNaughton, B. R.; Palde, P. B.; Thornton, C. A.; Miller, B. L. Dynamic Combinatorial Selection of Molecules Capable of Inhibiting the (CUG) Repeat RNA–MBNL1 Interaction In Vitro: Discovery of Lead Compounds Targeting Myotonic Dystrophy (DM1). *Journal of the American Chemical Society* **2008**, *130* (48), 16254–16261. <https://doi.org/10.1021/ja804398y>.
- (75) Toledano, S.; Williams, R. J.; Jayawarna, V.; Ulijn, R. V. Enzyme-Triggered Self-Assembly of Peptide Hydrogels via Reversed Hydrolysis. *Journal of the American Chemical Society* **2006**, *128* (4), 1070–1071. <https://doi.org/10.1021/ja056549l>.
- (76) Ruff, Y.; Garavini, V.; Giuseppone, N. Reversible Native Chemical Ligation: A Facile Access to Dynamic Covalent Peptides. *Journal of the American Chemical Society* **2014**, *136* (17), 6333–6339. <https://doi.org/10.1021/ja4129845>.
- (77) Mahon, C. S.; Jackson, A. W.; Murray, B. S.; Fulton, D. A. Templating a Polymer-Scaffolded Dynamic Combinatorial Library. *Chemical Communications* **2011**, *47* (25), 7209. <https://doi.org/10.1039/c1cc11998b>.
- (78) Mahon, C. S.; Fulton, D. A. Templatation-Induced Re-Equilibration in Polymer-Scaffolded Dynamic Combinatorial Libraries Leads to Enhancements in Binding Affinities. *Chemical Science* **2013**, *4* (9), 3661. <https://doi.org/10.1039/c3sc51413g>.
- (79) Ura, Y.; Beierle, J. M.; Leman, L. J.; Orgel, L. E.; Ghadiri, M. R. Self-Assembling Sequence-Adaptive Peptide Nucleic Acids. *Science* **2009**, *325* (5936), 73–77. <https://doi.org/10.1126/science.1174577>.
- (80) Heemstra, J. M.; Liu, D. R. Templated Synthesis of Peptide Nucleic Acids via Sequence-Selective Base-Filling Reactions. *Journal of the American Chemical Society* **2009**, *131* (32), 11347–11349. <https://doi.org/10.1021/ja904712t>.
- (81) Bartolami, E.; Bessin, Y.; Gervais, V.; Dumy, P.; Ulrich, S. Dynamic Expression of DNA Complexation with Self-Assembled Biomolecular Clusters. *Angewandte Chemie International Edition* **2015**, *54* (35), 10183–10187. <https://doi.org/10.1002/anie.201504047>.
- (82) Huc, I.; Lehn, J.-M. Virtual Combinatorial Libraries: Dynamic Generation of Molecular and Supramolecular Diversity by Self-Assembly. *Proceedings of the National Academy of Sciences* **1997**, *94* (6), 2106–2110. <https://doi.org/10.1073/pnas.94.6.2106>.

- (83) Bunyapaiboonsri et Al. - 2001 - Dynamic Deconvolution of a Pre-Equilibrated Dynamic Library. *Journal of Organic Chemistry* **2001**, *66* (12), 4321–4324. <https://doi.org/10.1021/jo00012a032>.
- (84) Kulchat, S.; Chaur, M. N.; Lehn, J.-M. Kinetic Selectivity and Thermodynamic Features of Competitive Imine Formation in Dynamic Covalent Chemistry. *Chemistry - A European Journal* **2017**, *23* (46), 11108–11118. <https://doi.org/10.1002/chem.201702088>.
- (85) Cousins, G. R. L.; Poulsen, S.-A.; Sanders, J. K. M. Dynamic Combinatorial Libraries of Pseudo-Peptide Hydrazone Macrocycles. *Chemical Communications* **1999**, No. 16, 1575–1576. <https://doi.org/10.1039/a904091i>.
- (86) Frei, P.; Pang, L.; Silbermann, M.; Eriş, D.; Mühlethaler, T.; Schwardt, O.; Ernst, B. Target-Directed Dynamic Combinatorial Chemistry: A Study on Potentials and Pitfalls as Exemplified on a Bacterial Target. *Chemistry - A European Journal* **2017**, *23* (48), 11570–11577. <https://doi.org/10.1002/chem.201701601>.
- (87) Fu, J.; Fu, H.; Dieu, M.; Halloum, I.; Kremer, L.; Xia, Y.; Pan, W.; Vincent, S. P. Identification of Inhibitors Targeting Mycobacterium Tuberculosis Cell Wall Biosynthesis via Dynamic Combinatorial Chemistry. *Chemical Communications* **2017**, *53* (77), 10632–10635. <https://doi.org/10.1039/C7CC05251K>.
- (88) Hewitt, S. H.; Wilson, A. J. Generation of Dynamic Combinatorial Libraries Using Hydrazone-Functionalized Surface Mimetics: Generation of Dynamic Combinatorial Libraries Using Hydrazone-Functionalized Surface Mimetics. *European Journal of Organic Chemistry* **2018**, *2018* (16), 1872–1879. <https://doi.org/10.1002/ejoc.201800022>.
- (89) Atcher, J.; Solà, J.; Alfonso, I. Pseudopeptidic Compounds for the Generation of Dynamic Combinatorial Libraries of Chemically Diverse Macrocycles in Aqueous Media. *Organic & Biomolecular Chemistry* **2017**, *15* (1), 213–219. <https://doi.org/10.1039/C6OB02441F>.
- (90) Drożdż, W.; Kołodziejski, M.; Markiewicz, G.; Jenczak, A.; Stefankiewicz, A. Generation of a Multicomponent Library of Disulfide Donor-Acceptor Architectures Using Dynamic Combinatorial Chemistry. *International Journal of Molecular Sciences* **2015**, *16* (7), 16300–16312. <https://doi.org/10.3390/ijms160716300>.
- (91) Solà, J.; Lafuente, M.; Atcher, J.; Alfonso, I. Constitutional Self-Selection from Dynamic Combinatorial Libraries in Aqueous Solution through Supramolecular Interactions. *Chem. Commun.* **2014**, *50* (35), 4564–4566. <https://doi.org/10.1039/C4CC00245H>.
- (92) Rasmussen, B.; Sørensen, A.; Gotfredsen, H.; Pittelkow, M. Dynamic Combinatorial Chemistry with Diselenides and Disulfides in Water. *Chem. Commun.* **2014**, *50* (28), 3716–3718. <https://doi.org/10.1039/C4CC00523F>.
- (93) Saiz, C.; Castillo, V.; Fontán, P.; Bonilla, M.; Salinas, G.; Rodríguez-Haralambides, A.; Mahler, S. G. Discovering Echinococcus Granulosus Thioredoxin Glutathione Reductase Inhibitors through Site-Specific Dynamic Combinatorial Chemistry. *Molecular Diversity* **2014**, *18* (1), 1–12. <https://doi.org/10.1007/s11030-013-9485-3>.
- (94) Bugaut, A.; Jantos, K.; Wietor, J.-L.; Rodriguez, R.; Sanders, J. K. M.; Balasubramanian, S. Exploring the Differential Recognition of DNA G-Quadruplex Targets by Small Molecules Using Dynamic Combinatorial Chemistry. *Angewandte Chemie International Edition* **2008**, *47* (14), 2677–2680. <https://doi.org/10.1002/anie.200705589>.
- (95) Mansfeld, F. M.; Au-Yeung, H.; Sanders, J. K.; Otto, S. Dynamic Combinatorial Chemistry at the Phospholipid Bilayer Interface. *Journal of Systems Chemistry* **2010**, *1* (1), 12. <https://doi.org/10.1186/1759-2208-1-12>.
- (96) Ghosh, S.; Ingerman, L. A.; Frye, A. G.; Lee, S. J.; Gagné, M. R.; Waters, M. L. Dynamic Cyclic Thiopeptide Libraries From Thiol-Thioester Exchange. *Organic Letters* **2010**, *12* (8), 1860–1863. <https://doi.org/10.1021/ol1004752>.
- (97) Larsson, R.; Pei, Z.; Ramström, O. Catalytic Self-Screening of Cholinesterase Substrates from a Dynamic Combinatorial Thioester Library. *Angewandte Chemie International Edition* **2004**, *43* (28), 3716–3718. <https://doi.org/10.1002/anie.200454165>.
- (98) Katritzky, A. R.; Zhang, Y.; Singh, S. K. Efficient Conversion of Carboxylic Acids into

- N-Acylbenzotriazoles. *Synthesis* **2003**, 2003 (18), 2795–2798.
- (99) Katritzky, A. R.; Tala, S. R.; Abo-Dya, N. E.; Gyanda, K.; El-Gendy, B. E.-D. M.; Abdel-Samii, Z. K.; Steel, P. J. Selective Synthesis and Structural Elucidation of *S*-Acyl- and *N*-Acylcysteines. *The Journal of Organic Chemistry* **2009**, 74 (18), 7165–7167. <https://doi.org/10.1021/jo900853s>.
- (100) Lee, S.; Rosazza, J. P. N. First Total Synthesis of Mycothiol and Mycothiol Disulfide. *Organic Letters* **2004**, 6 (3), 365–368. <https://doi.org/10.1021/ol0362008>.
- (101) Martinez, E. J.; Talley, J. J.; Boehm, T. L. No-Releasing Guanidine-Coxib Anti-Cancer Agents. **2015**.
- (102) Brooks, D. A.; Rito, C. J.; Shuker, A. J.; Dominianni, S. J.; Warshawsky, A. M.; Gossett, L. S.; Matthews, D. P.; Hay, D. A.; Ardecky, R. J.; Michellys, P.-Y. Modulators of Peroxisome Proliferator Activated Receptors. **2002**.
- (103) Pakhomov, A. A.; Kononevich, Y. N.; Stukalova, M. V.; Svidchenko, E. A.; Surin, N. M.; Cherkaev, G. V.; Shchegolikhina, O. I.; Martynov, V. I.; Muzafarov, A. M. Synthesis and Photophysical Properties of a New BODIPY-Based Siloxane Dye. *Tetrahedron Letters* **2016**, 57 (9), 979–982. <https://doi.org/10.1016/j.tetlet.2016.01.059>.
- (104) Mohapatra, D. K.; Umamaheshwar, G.; Rao, M. M.; Umadevi, D.; Yadav, J. S. Concise Total Synthesis of Botryolide B. *RSC Advances* **2014**, 4 (16), 8335–8340.
- (105) Brooks, D. A.; Rito, C. J.; Shuker, A. J.; Dominianni, S. J.; Warshawsky, A. M.; Gossett, L. S.; Matthews, D. P.; Hay, D. A.; Ardecky, R. J.; Michellys, P.-Y. Modulators of Peroxisome Proliferator Activated Receptors. WO 02/100813 A2, July 9, 2002.
- (106) Donner, C. D. Ketyl Radical Cyclization of β -Disubstituted Acrylates: Formal Syntheses of (+)-Secosyrin 1 and Longianone and the Total Synthesis of (+)-4-*Epi*-Secosyrin 1. *Organic Letters* **2013**, 15 (6), 1258–1261. <https://doi.org/10.1021/ol4001894>.
- (107) Lin, L.; Romano, C.; Mazet, C. Palladium-Catalyzed Long-Range Deconjugative Isomerization of Highly Substituted α,β -Unsaturated Carbonyl Compounds. *Journal of the American Chemical Society* **2016**, 138 (32), 10344–10350. <https://doi.org/10.1021/jacs.6b06390>.
- (108) Pauling, L.; Corey, R. B.; Branson, H. R. The Structure of Proteins: Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *Proceedings of the National Academy of Sciences* **1951**, 37 (4), 205–211.
- (109) Dunitz, J. D. Pauling's Left-Handed α -Helix. *Angewandte Chemie International Edition* **2001**, 40 (22), 4167–4173. [https://doi.org/10.1002/1521-3773\(20011119\)40:22<4167::AID-ANIE4167>3.0.CO;2-Q](https://doi.org/10.1002/1521-3773(20011119)40:22<4167::AID-ANIE4167>3.0.CO;2-Q).
- (110) Dantas de Araujo, A.; Perry, S. R.; Fairlie, D. P. Chemically Diverse Helix-Constrained Peptides Using Selenocysteine Crosslinking. *Organic Letters* **2018**, 20 (5), 1453–1456. <https://doi.org/10.1021/acs.orglett.8b00233>.
- (111) Fairlie, D. P.; Dantas de Araujo, A. Stapling Peptides Using Cysteine Crosslinking: Stapling Peptides. *Biopolymers* **2016**, 106 (6), 843–852. <https://doi.org/10.1002/bip.22877>.
- (112) Baxter, D.; Perry, S. R.; Hill, T. A.; Kok, W. M.; Zaccari, N. R.; Brady, R. L.; Fairlie, D. P.; Mason, J. M. Downsizing Proto-Oncogene CFos to Short Helix-Constrained Peptides That Bind Jun. *ACS Chemical Biology* **2017**, 12 (8), 2051–2061. <https://doi.org/10.1021/acscchembio.7b00303>.
- (113) de Araujo, A. D.; Lim, J.; Good, A. C.; Skerlj, R. T.; Fairlie, D. P. Electrophilic Helical Peptides That Bond Covalently, Irreversibly, and Selectively in a Protein–Protein Interaction Site. *ACS Medicinal Chemistry Letters* **2017**, 8 (1), 22–26. <https://doi.org/10.1021/acsmchemlett.6b00395>.
- (114) Shepherd, N. E.; Harrison, R. S.; Ruiz-Gomez, G.; Abbenante, G.; Mason, J. M.; Fairlie, D. P. Downsizing the BAD BH3 Peptide to Small Constrained α -Helices with Improved Ligand Efficiency. *Organic & Biomolecular Chemistry* **2016**, 14 (46), 10939–10945. <https://doi.org/10.1039/C6OB02185A>.
- (115) Perry, S. R.; Hill, T. A.; de Araujo, A. D.; Hoang, H. N.; Fairlie, D. P. Contiguous

- Hydrophobic and Charged Surface Patches in Short Helix-Constrained Peptides Drive Cell Permeability. *Organic & Biomolecular Chemistry* **2018**, *16* (3), 367–371. <https://doi.org/10.1039/C7OB02952G>.
- (116) Hoang, H. N.; Driver, R. W.; Beyer, R. L.; Hill, T. A.; D. de Araujo, A.; Plisson, F.; Harrison, R. S.; Goedecke, L.; Shepherd, N. E.; Fairlie, D. P. Helix Nucleation by the Smallest Known α -Helix in Water. *Angewandte Chemie International Edition* **2016**, *55* (29), 8275–8279. <https://doi.org/10.1002/anie.201602079>.
- (117) Harrison, R. S.; Shepherd, N. E.; Hoang, H. N.; Ruiz-Gomez, G.; Hill, T. A.; Driver, R. W.; Desai, V. S.; Young, P. R.; Abbenante, G.; Fairlie, D. P. Downsizing Human, Bacterial, and Viral Proteins to Short Water-Stable Alpha Helices That Maintain Biological Potency. *Proceedings of the National Academy of Sciences* **2010**, *107* (26), 11686–11691. <https://doi.org/10.1073/pnas.1002498107>.
- (118) Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P. Left- and Right-Handed Alpha-Helical Turns in Homo- and Hetero-Chiral Helical Scaffolds. *Journal of the American Chemical Society* **2009**, *131* (43), 15877–15886. <https://doi.org/10.1021/ja9065283>.
- (119) Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P. Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *Journal of the American Chemical Society* **2005**, *127* (9), 2974–2983. <https://doi.org/10.1021/ja0456003>.
- (120) Shepherd, N. E.; Abbenante, G.; Fairlie, D. P. Consecutive Cyclic Pentapeptide Modules Form Short α -Helices That Are Very Stable to Water and Denaturants. *Angewandte Chemie International Edition* **2004**, *43* (20), 2687–2690. <https://doi.org/10.1002/anie.200352659>.
- (121) Harrison, R. S.; Shepherd, N. E.; Hoang, H. N.; Ruiz-Gomez, G.; Hill, T. A.; Driver, R. W.; Desai, V. S.; Young, P. R.; Abbenante, G.; Fairlie, D. P. Downsizing Human, Bacterial, and Viral Proteins to Short Water-Stable Alpha Helices That Maintain Biological Potency. *Proceedings of the National Academy of Sciences* **2010**, *107* (26), 11686–11691. <https://doi.org/10.1073/pnas.1002498107>.
- (122) Harrison, R. S.; Ruiz-Gómez, G.; Hill, T. A.; Chow, S. Y.; Shepherd, N. E.; Lohman, R.-J.; Abbenante, G.; Hoang, H. N.; Fairlie, D. P. Novel Helix-Constrained Nociceptin Derivatives Are Potent Agonists and Antagonists of ERK Phosphorylation and Thermal Analgesia in Mice. *Journal of Medicinal Chemistry* **2010**, *53* (23), 8400–8408. <https://doi.org/10.1021/jm101139f>.
- (123) Rao, T.; Ruiz-Gómez, G.; Hill, T. A.; Hoang, H. N.; Fairlie, D. P.; Mason, J. M. Truncated and Helix-Constrained Peptides with High Affinity and Specificity for the CFos Coiled-Coil of AP-1. *PLoS ONE* **2013**, *8* (3), e59415. <https://doi.org/10.1371/journal.pone.0059415>.
- (124) Shepherd, N. E.; Harrison, R. S.; Ruiz-Gomez, G.; Abbenante, G.; Mason, J. M.; Fairlie, D. P. Downsizing the BAD BH3 Peptide to Small Constrained α -Helices with Improved Ligand Efficiency. *Org. Biomol. Chem.* **2016**, *14* (46), 10939–10945. <https://doi.org/10.1039/C6OB02185A>.
- (125) Khoo, K. K.; Wilson, M. J.; Smith, B. J.; Zhang, M.-M.; Gulyas, J.; Yoshikami, D.; Rivier, J. E.; Bulaj, G.; Norton, R. S. Lactam-Stabilized Helical Analogues of the Analgesic μ -Conotoxin KIIIA. *Journal of Medicinal Chemistry* **2011**, *54* (21), 7558–7566. <https://doi.org/10.1021/jm200839a>.
- (126) Philippe, G.; Huang, Y.-H.; Cheneval, O.; Lawrence, N.; Zhang, Z.; Fairlie, D. P.; Craik, D. J.; de Araujo, A. D.; Henriques, S. T. Development of Cell-Penetrating Peptide-Based Drug Leads to Inhibit MDMX:P53 and MDM2:P53 Interactions: Cell-Penetrating Stapled Peptide Inhibitors of MDMX/MDM2:P53 Interactions. *Biopolymers* **2016**, *106* (6), 853–863. <https://doi.org/10.1002/bip.22893>.
- (127) Robinson, J. A. β -Hairpin Peptidomimetics: Design, Structures and Biological Activities. *Accounts of Chemical Research* **2008**, *41* (10), 1278–1288. <https://doi.org/10.1021/ar700259k>.

- (128) Robinson, J. A. The Design, Synthesis and Conformation of Some New β -Hairpin Mimetics: Novel Reagents for Drug and Vaccine Discovery. *New York* **2000**, No. 4, 13.
- (129) IVENKATACHALAPATHI, A. X-Ray Crystal Structure of Pivaloyl-D-Pro-L-Pro-L-Ala-N-Methylamide; Observation of a Consecutive F3 -Turn Conformation. 2.
- (130) Bean, J. W.; Kopple, K. D.; Peishoff, C. E. Conformational Analysis of Cyclic Hexapeptides Containing the D-Pro-L-Pro Sequence to Fix β -Turn Positions. *Journal of the American Chemical Society* **1992**, *114* (13), 5328–5334. <https://doi.org/10.1021/ja00039a050>.
- (131) Chalmers, D. K.; Marshall, G. R. Pro-D-NMe-Amino Acid and D-Pro-NMe-Amino Acid: Simple, Efficient Reverse-Turn Constraints. *Journal of the American Chemical Society* **1995**, *117* (22), 5927–5937. <https://doi.org/10.1021/ja00127a004>.
- (132) Sarver, A.; Scheffler, N. K.; Shetlar, M. D.; Gibson, B. W. Analysis of Peptides and Proteins Containing Nitrotyrosine by Matrix-Assisted Laser Desorption/Ionization Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2001**, *12* (4), 439–448. [https://doi.org/10.1016/S1044-0305\(01\)00213-6](https://doi.org/10.1016/S1044-0305(01)00213-6).
- (133) Sainlos, M.; Imperiali, B. Synthesis of Anhydride Precursors of the Environment-Sensitive Fluorophores 4-DMAP and 6-DMN. *Nature protocols* **2007**, *2* (12), 3219.
- (134) Sainlos, M.; Imperiali, B. Tools for Investigating Peptide–Protein Interactions: Peptide Incorporation of Environment-Sensitive Fluorophores through SPPS-Based ‘building Block’ approach. *Nature protocols* **2007**, *2* (12), 3210.
- (135) Favre, M.; Moehle, K.; Jiang, L.; Pfeiffer, B.; Robinson, J. A. Structural Mimicry of Canonical Conformations in Antibody Hypervariable Loops Using Cyclic Peptides Containing a Heterochiral Diproline Template. *Journal of the American Chemical Society* **1999**, *121* (12), 2679–2685. <https://doi.org/10.1021/ja984016p>.
- (136) Poole, L. B. The Basics of Thiols and Cysteines in Redox Biology and Chemistry. *Free Radical Biology and Medicine* **2015**, *80*, 148–157. <https://doi.org/10.1016/j.freeradbiomed.2014.11.013>.
- (137) Bracher, P. J.; Snyder, P. W.; Bohall, B. R.; Whitesides, G. M. The Relative Rates of Thiol–Thioester Exchange and Hydrolysis for Alkyl and Aryl Thioalkanoates in Water. *Origins of Life and Evolution of Biospheres* **2011**, *41* (5), 399–412. <https://doi.org/10.1007/s11084-011-9243-4>.
- (138) Sakamoto, K.; Tsuda, S.; Mochizuki, M.; Nohara, Y.; Nishio, H.; Yoshiya, T. Imidazole-Aided Native Chemical Ligation: Imidazole as a One-Pot Desulfurization-Amenable Non-Thiol-Type Alternative to 4-Mercaptophenylacetic Acid. *Chemistry - A European Journal* **2016**, *22* (50), 17940–17944. <https://doi.org/10.1002/chem.201604320>.
- (139) Bechara, C.; Sagan, S. Cell-Penetrating Peptides: 20 Years Later, Where Do We Stand? *FEBS Letters* **2013**, *587* (12), 1693–1702. <https://doi.org/10.1016/j.febslet.2013.04.031>.
- (140) Ludlow, R. F.; Otto, S. Two-Vial, LC–MS Identification of Ephedrine Receptors from a Solution-Phase Dynamic Combinatorial Library of over 9000 Components. *Journal of the American Chemical Society* **2008**, *130* (37), 12218–12219. <https://doi.org/10.1021/ja803317k>.
- (141) Liang, X.; Girard, A.; Biron, E. Practical Ring-Opening Strategy for the Sequence Determination of Cyclic Peptides from One-Bead-One-Compound Libraries. *ACS Combinatorial Science* **2013**, *15* (10), 535–540. <https://doi.org/10.1021/co4000979>.
- (142) Lee, K. J.; Lim, H.-S. Facile Method To Sequence Cyclic Peptides/Peptoids via One-Pot Ring-Opening/Cleavage Reaction. *Organic Letters* **2014**, *16* (21), 5710–5713. <https://doi.org/10.1021/ol502788e>.
- (143) Liang, X.; Vézina-Dawod, S.; Bédard, F.; Porte, K.; Biron, E. One-Pot Photochemical Ring-Opening/Cleavage Approach for the Synthesis and Decoding of Cyclic Peptide Libraries. *Organic Letters* **2016**, *18* (5), 1174–1177. <https://doi.org/10.1021/acs.orglett.6b00296>.
- (144) Gurevich-Messina, J. M.; Giudicessi, S. L.; Martínez-Ceron, M. C.; Acosta, G.; Erra-Balsells, R.; Cascone, O.; Albericio, F.; Camperi, S. A. A Simple Protocol for Combinatorial Cyclic Depsipeptide Libraries Sequencing by Matrix-Assisted Laser Desorption/Ionisation

- Mass Spectrometry: COMBINATORIAL CYCLIC LIBRARIES SEQUENCING BY MASS SPECTROMETRY. *Journal of Peptide Science* **2015**, *21* (1), 40–45. <https://doi.org/10.1002/psc.2716>.
- (145) Chu, I. K.; Siu, C.-K.; Lau, J. K.-C.; Tang, W. K.; Mu, X.; Lai, C. K.; Guo, X.; Wang, X.; Li, N.; Xia, Y.; et al. Proposed Nomenclature for Peptide Ion Fragmentation. *International Journal of Mass Spectrometry* **2015**, *390*, 24–27. <https://doi.org/10.1016/j.ijms.2015.07.021>.
- (146) Demeure, K.; Gabelica, V.; De Pauw, E. A. New Advances in the Understanding of the In-Source Decay Fragmentation of Peptides in MALDI-TOF-MS. *Journal of the American Society for Mass Spectrometry* **2010**. <https://doi.org/10.1016/j.jasms.2010.07.009>.
- (147) Coon, J. J.; Syka, J. E. P.; Shabanowitz, J.; Hunt, D. F. Tandem Mass Spectrometry for Peptide and Protein Sequence Analysis. *BioTechniques* **2005**, *38* (4), 519–523. <https://doi.org/10.2144/05384TE01>.
- (148) Asakawa, D.; Smargiasso, N.; Quinton, L.; De Pauw, E. Influences of Proline and Cysteine Residues on Fragment Yield in Matrix-Assisted Laser Desorption/Ionization In-Source Decay Mass Spectrometry. *Journal of The American Society for Mass Spectrometry* **2014**, *25* (6), 1040–1048. <https://doi.org/10.1007/s13361-014-0868-1>.
- (149) Sipost, T.; Merkelt, J. R. An Effect of Calcium Ions on the Activity, Heat Stability, and Structure of Trypsin". 10.
- (150) Endo, M.; Nakayama, K.; Kaida, Y.; Majima, T. Design and Synthesis of Photochemically Controllable Caspase-3. *Angewandte Chemie International Edition* **2004**, *43* (42), 5643–5645. <https://doi.org/10.1002/anie.200460889>.
- (151) Peters, F. B.; Brock, A.; Wang, J.; Schultz, P. G. Photocleavage of the Polypeptide Backbone by 2-Nitrophenylalanine. *Chemistry & Biology* **2009**, *16* (2), 148–152. <https://doi.org/10.1016/j.chembiol.2009.01.013>.
- (152) Salomon-Ferrer, R.; Case, D. A.; Walker, R. C. An Overview of the Amber Biomolecular Simulation Package: Amber Biomolecular Simulation Package. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3* (2), 198–210. <https://doi.org/10.1002/wcms.1121>.
- (153) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *Journal of Computational Chemistry* **2005**, *26* (16), 1668–1688. <https://doi.org/10.1002/jcc.20290>.
- (154) Markley, J. L.; Bax, A.; Arata, Y.; Hilbers, C. W.; Kaptein, R.; Sykes, B. D.; Wright, P. E.; Wuthrich, K. Recommendations for the Presentation of NMR Structures of Proteins and Nucleic Acids. IUPAC-IUBMB-IUPAB Inter-Union Task Group on the Standardization of Data Bases of Protein and Nucleic Acid Structures Determined by NMR Spectroscopy. *European Journal of Biochemistry* **1998**, *256* (1), 1–15. <https://doi.org/10.1046/j.1432-1327.1998.2560001.x>.
- (155) Cierpicki, T.; Otlewski, J. Amide Proton Temperature Coefficients as Hydrogen Bond Indicators in Proteins. *Journal of Biomolecular NMR* **2001**, *21* (3), 249–261. <https://doi.org/10.1023/A:1012911329730>.

Chimie combinatoire dynamique pour la conception de mimes de protéines :

Les interactions protéine-protéine (PPI) constituent des cibles thérapeutiques de premier intérêt mais la complexité des interfaces impliquées dans ces interactions rend le développement d'inhibiteurs de PPI difficile.

Les peptides présentent un grand potentiel pour cibler les PPI mais sans structure 3D définie, ils ne présentent généralement pas d'affinité. Les peptides contraints dans une conformation contrôlée ou *scaffolds* constituent une solution efficace mais leur conception requiert une compréhension significative du complexe protéique à l'échelle de l'épitope.

Ce travail a pour but de développer une approche de chimie combinatoire dynamique (DCC) pour la fonctionnalisation de *scaffolds* peptidiques. Cette approche permettrait de générer des bibliothèques dynamiques par des réactions réversibles équilibrées entre des blocs moléculaires. Ces systèmes dynamiques étant sensibles aux stimuli, il est alors possible d'introduire une cible d'intérêt pour les cribler, le tout en une seule étape. Ce manuscrit présente la conception et la synthèse de blocs moléculaires et de *scaffolds* pour cette stratégie de DCC. La mise au point de cette méthodologie de DCC est ensuite présentée dans une approche pas à pas jusqu'à une tentative d'application à une cible d'intérêt : l'héparine. Le développement d'une méthode d'analyse par spectrométrie de masse des composés générés par DCC est ensuite abordée. Enfin, l'analyse structurale des plateformes et l'influence de leur fonctionnalisation sur leur structuration est abordée.

Mots-clés : mimes de protéines, greffage d'épitope, chimie combinatoire dynamique (DCC), *scaffolds* peptidiques, peptides cycliques, thioesters

Dynamic combinatorial chemistry for the design of protein mimics :

Protein protein interactions (PPI) are therapeutic targets of prime interest but the complexity of their interaction interfaces makes difficult the development of PPIs inhibitors.

Peptides display great potential for targeting PPIs but without a defined 3D structure, they usually show no affinity. Constrained peptides with a controlled conformation or scaffolds constitute an effective solution but their design requires a significant comprehension of the protein complex at the epitope scale.

This work aims at developing a dynamic combinatorial chemistry (DCC) approach to functionalize peptide scaffolds. This approach allows to generate dynamic libraries thanks to equilibrated reversible reactions between molecular building blocks. These dynamic systems being stimuli-responsive, it is therefore possible to introduce a relevant target to screen them in one step. This thesis presents the design and the synthesis of the building blocks and scaffolds for this DCC strategy. The development of the DCC methodology is then presented in a step by step approach up to an application attempt with heparin as the target. Next, the mass spectrometry method for the analysis of the DCC generated compounds and its setting up is addressed. Finally, the analysis of the peptide scaffolds structure and the influence of their functionalization on it is broached.

Keywords : protein mimics, epitope grafting, dynamic combinatorial chemistry (DCC), peptide scaffolds, cyclic peptides, thioesters