



**HAL**  
open science

# Conception de peptides de novo s'autoassemblant en hélices $\beta$ gauche

Maxime Naudé

► **To cite this version:**

Maxime Naudé. Conception de peptides de novo s'autoassemblant en hélices  $\beta$  gauche. Autre. Université de Strasbourg, 2022. Français. NNT : 2022STRAF065 . tel-04189472

**HAL Id: tel-04189472**

**<https://theses.hal.science/tel-04189472v1>**

Submitted on 28 Aug 2023

**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.

**ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES** Institut  
de chimie-UMR711

**THÈSE** présentée par :

**Maxime NAUDÉ**

Soutenue le : 28 octobre 2022

pour obtenir le grade de : **Docteur de l'Université de Strasbourg**

Discipline/ Spécialité : Chimie

**Conception de peptides *de novo*  
s'autoassemblant en hélice  $\beta$  gauche**

**THÈSE dirigée par :**

**M. FALLER Peter**

Professeur des universités, Université de Strasbourg

**THÈSE coencadrée par :**

**M. LEBRUN Vincent**

Chargé de recherche CNRS, Université de Strasbourg

**RAPPORTEURS :**

**M. FORGE Vincent**

Chercheur CEA, CEA-Grenoble

**M. SENEQUE Olivier**

Directeur de recherche CNRS, CEA-Grenoble

**AUTRES MEMBRES DU JURY :**

**Mme PATERNOSTRE Maïté**

Directrice de recherche CNRS, I2BC Université Paris-Saclay

**M. GHATTAS Wadih**

Chargé de recherche, LCBB – ICMMO Université Paris-Saclay

**Mme MÉNARD-MOYON Cécilia**

Directrice de recherche, IBMC Université de Strasbourg

*To all the wonderful people who  
have contributed to this project.*





# Table des matières

<b>Abréviations.....</b>	<b>9</b>
<b>Remerciements/Aknowledgments.....</b>	<b>11</b>
<b>1 Introduction.....</b>	<b>15</b>
1.1 Généralités sur les métalloenzymes artificielles.....	16
1.2 Autoassemblages minimalistes fonctionnels comme métalloenzymes artificielles .....	18
1.3 Des briques peptidiques pour mimer les métalloenzymes multisites .....	21
1.4 Les solénoïdes $\beta$ .....	24
1.5 Propriétés d'autoassemblage des solénoïdes $\beta$ .....	25
1.6 Le cas des hélices $\beta$ gauche.....	27
1.7 Les hélices $\beta$ gauche de type I.....	29
1.8 Objectifs de thèse .....	32
1.9 Références bibliographiques .....	34
<b>2 Analyse bio-informatique des structures en hélice <math>\beta</math> gauche naturelles .....</b>	<b>39</b>
2.1 Introduction .....	40
2.1.1 Limites des études précédentes .....	40
2.2 Construction de l'outil.....	40
2.2.1 Recherche de structures PDB.....	41
2.3 Restriction de l'analyse au pur repliement hélice $\beta$ gauche de type I .....	42
2.3.1 Le cœur hydrophobe.....	44
2.3.2 Les tours .....	45
2.3.2.1 Les tours de nucléation (N) .....	46
2.3.2.1.1 Identification d'un motif de 3 résidus comme motif de nucléation.....	46
2.3.2.1.2 La flexibilité des coudes $\beta$ .....	47
2.3.2.1.3 La position $B_2^\circ$ est la plus variable en fonction des faces .....	48
2.3.2.2 Les tours de terminaison (C) .....	48
2.3.2.3 Les tours canoniques (R).....	50
2.3.2.4 Motifs de capping (« bouchon ») en C-ter de l'hélice.....	52
2.4 Conclusion.....	53
2.5 Perspective : Application du nouvel outil à la conception de peptides solubles de $L\beta H$ .....	55
2.6 Références bibliographiques .....	59
<b>3 Conception de peptides de novo s'autoassemblant en hélice <math>\beta</math> gauche .....</b>	<b>61</b>
3.1 Introduction .....	62



3.2	Conception rationnelle de la brique élémentaire .....	62
3.2.1	Choix de la taille de séquence (étape 1) .....	63
3.2.2	Coupure, formylation et amidation de la brique (étape 2).....	64
3.2.3	Le cœur hydrophobe (étape 3).....	65
3.2.4	Rôle des prolines (étape 4) .....	66
3.2.5	Importance des glycines et lysines (étape 5) .....	66
3.2.6	Complémentarité de charges entre les tours (étape 6) .....	67
3.2.7	Formation d'échelles d'asparagines (étape 7) .....	67
3.2.8	Histidines au centre du trimère (étape 8).....	68
3.2.9	Les positions B <sub>2</sub> non attribuées à l'étape 6 (étape 9) .....	69
3.2.10	Brique élémentaire finale.....	69
3.2.11	35m2 : Mutation du 35m1 pour défavoriser l'agrégation latérale (étape 10).....	70
3.2.12	Modélisation .....	70
3.3	Production des peptides .....	71
3.3.1	Synthèse des briques de base par SPPS .....	71
3.3.2	Purification par HPLC semi-préparative .....	72
3.3.3	Caractérisations .....	73
3.3.3.1	Suivi cinétique et mécanisme d'assemblage avec la thioflavine-T .....	73
3.3.3.2	Évolution de la fluorescence de la Thioflavine-T en présence des peptides 35m1 et 35m2 ...	75
3.3.3.2.1	Fibrillation en fonction du temps.....	75
3.3.3.2.2	Fibrillation en fonction du pH .....	77
3.4	Observation de la morphologie par microscopie électronique .....	78
3.4.1	Observations TEM et optimisation des conditions.....	79
3.4.2	Optimisation des conditions de fibrillation en vue d'une analyse en Cryo-TEM.....	82
3.4.3	Suivi de la fibrillation par TEM .....	83
3.4.4	Analyses préliminaires en Cryo-TEM .....	85
3.4.4.1	Préparation des grilles .....	85
3.4.4.2	Résultats de l'analyse en Cryo-TEM du peptide 35m2 .....	87
3.4.5	Informations structurales par dichroïsme circulaire .....	88
3.4.5.1	Comparaison en CD et SR-CD des peptides fibrillés .....	91
3.4.5.2	Différences entre les peptides en solution et fibrillés .....	92
3.4.5.3	Suivi du repliement de 35m2 au cours du temps et en fonction de la température .....	93
3.4.5.4	Alignement des fibres visible par dichroïsme linéaire.....	95
3.4.5.5	Problèmes rencontrés lors des analyses SR-CD et SR-LC .....	96



3.4.6	Analyse des fibres par spectroscopie infrarouge .....	97
3.5	Discussion et conclusion .....	98
3.6	Références bibliographiques .....	101
<b>4</b>	<b>Séquence peptidique minimale afin de former une hélice <math>\beta</math> gauche.....</b>	<b>105</b>
4.1	Introduction .....	106
4.2	Contexte.....	106
4.3	Conception des séquences .....	107
4.4	Propriétés des peptides .....	109
4.4.1	Synthèses .....	109
4.4.2	Solubilité .....	109
4.5	Vers une structure en L $\beta$ H ?.....	109
4.5.1	Dichroïsme circulaire .....	110
4.5.2	Spectroscopie Infra-rouge .....	111
4.5.3	Diffraction des rayons X .....	111
4.5.3.1	Cristallogénèse en goutte assise .....	111
4.5.3.1.1	L'ensemencement : apport de nucléi pour optimiser la cristallisation .....	115
4.5.3.1.2	Influence de l'agarose.....	116
4.5.3.1.3	Applications à d'autres peptides.....	117
4.5.3.2	Diffraction des rayons X sur plaques .....	118
4.5.3.3	Cristallogénèse en micro-batch : à la recherche de cristaux de meilleure qualité .....	120
4.5.3.4	Cristallogénèse en tube.....	122
4.5.3.4.1	Production de cristaux en grandes quantités.....	122
4.5.3.4.2	Préparation de cristaux optimisés pour la DRX .....	123
4.5.3.5	Résultats de diffraction.....	124
4.5.3.5.1	Structures obtenues.....	125
4.5.3.5.2	f-VGPLY-Am .....	125
4.5.3.5.3	f-VGPNT-Am.....	126
4.5.4	RMN du solide .....	126
4.6	Conclusion.....	127
4.7	Références bibliographiques .....	129
<b>5</b>	<b>Conclusion .....</b>	<b>131</b>
<b>6</b>	<b>Perspectives .....</b>	<b>137</b>
6.1	Perspectives à court terme : suite de la caractérisation .....	138
6.2	Vers la conception rationnelle d'ArMs .....	139



<b>7</b>	<b>Partie expérimentale.....</b>	<b>143</b>
7.1	Matériel.....	144
7.1.1	Analyse bio-informatique des structures en hélice $\beta$ gauche naturelles .....	147
7.1.2	Logiciels et bases de données.....	147
7.1.3	Démarche d'analyse .....	147
7.1.3.1	Visualisation.....	147
7.1.3.2	Report et tri dans une feuille de calcul .....	148
7.1.4	Conception de peptides de novo s'autoassemblant en hélice $\beta$ gauche.....	148
7.1.5	Méthodes .....	148
7.1.6	Synthèse Peptidique en Phase Solide (SPPS), stratégie Fmoc/tBu : <sup>[2,3]</sup> .....	149
7.1.7	Purification .....	152
7.1.7.1	Protocole.....	152
7.1.8	Suivi cinétique à l'aide de la thioflavine-T.....	153
7.1.8.1	Préparation de la solution de ThT.....	153
7.1.8.2	Préparation des tampons.....	153
7.1.8.3	Préparation des plaques pour la mesure de la fluorescence.....	154
7.1.8.3.1	Paramètres de mesures.....	154
7.1.8.3.2	Préparation des plaques .....	154
7.1.8.3.3	Collecte et traitement des données .....	154
7.1.9	Microscopie électronique .....	155
7.1.9.1	Incubation des solutions pour fibrillation.....	155
7.1.9.2	Préparation des grilles de microscopie électronique à transmission (TEM) et Cryo-TEM... 155	
7.1.9.2.1	Grilles de TEM .....	155
7.1.9.2.2	Grille de Cryo-TEM .....	156
7.1.10	Dichroïsme circulaire et linéaire.....	157
7.1.10.1	Préparation des cuves/cellules.....	157
7.1.10.2	Acquisition et traitement des spectres .....	157
7.1.11	Spectroscopie infrarouge .....	159
7.1.11.1	Incubation et centrifugation des peptides .....	159
7.1.11.2	Acquisition des spectres infrarouges .....	159
7.2	Motif minimal permettant la formation d'une hélice $\beta$ gauche .....	160
7.2.1	Cristallisation.....	161
7.2.1.1	Cristallisation en goutte assise.....	161
7.2.1.1.1	Protocole.....	161





7.2.1.2	Cristallisation en microbatch.....	162
7.2.1.2.1	Protocole.....	163
7.2.1.3	Cristallisation en Tube.....	163
7.2.1.3.1	Pour obtenir des cristaux en quantité.....	163
7.2.1.3.2	Pour obtenir des cristaux de qualité.....	164
7.2.2	Diffraction des Rayons X.....	164
7.2.2.1	Méthodes.....	164
7.2.2.2	Collecte et traitement des données.....	165
7.2.3	RMN du solide.....	165
7.2.3.1	Préparation de l'expérience RMN du solide.....	165
7.2.3.2	Méthodes d'analyse en RMN du solide.....	166
7.3	Références bibliographiques.....	166
<b>Annexes</b>	.....	<b>167</b>





# Abréviations

Abs : absorbance

Ac<sub>2</sub>O : anhydride acétique

AcU : Acétate d'uranyle

AFM : microscopie à force atomique

ArMs : métalloenzymes artificielles

A $\beta$  : amyloïde  $\beta$

Cryo-TEM : microscopie électronique à basse température

C-ter/C<sub>ter</sub> : C-terminus

DCM : dichlorométhane

DIC : N,N'-Diisopropylcarbodiimide

DIEA : N,N-Diisopropylethylamine

DMF : diméthyle formamide

DRN : diffusion des neutrons

DSC : calorimétrie différentielle à balayage

Et<sub>2</sub>O : éther diéthylique

FA : acide formique

Fmoc : fluorénylméthoxycarbone

HCTU : O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate

HEPES : 2-[4-(2-hydroxyethyl)piperazin-1-yl]éthanesulfonique acide

HPLC : Chromatographie liquide à haute performance

ITC : titration calorimétrique isotherme

LC-MS : Chromatographie Liquide couplée à la Spectrométrie Masse

L $\beta$ H : hélice  $\beta$  gauche

MD : dynamique moléculaire

MeCN : Acétonitrile

micro-ED : micro-diffraction électronique

NCL : chimie de ligation native

NMP : N-Méthyl-2-pyrrolidone

N-ter/N<sub>ter</sub> : N-terminus

Pbf : 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl

PDB : Protein Data Bank

pI : point isolélectrique



pip-DBF : pipéridine-dibenzofulvène

RMN : résonance magnétique nucléaire

Rpm : rotations par minute

R $\beta$ H : hélice  $\beta$  droite

SPPS : Synthèse Peptidique en Phase Solide

TEM : microscopie électronique à transmission

TFA : acide trifluoroacétique

ThT: thioflavine-T

TIS : Triisopropyle silane

TNBS : 2,4,6 - Trinitrobenzenesulfonique Acide

UHPLC : Chromatographie liquide à ultra-haute performance

WAXS : diffusion des rayons X aux grands angles

# Remerciements/Aknowledgments

Il me semble important de commencer les remerciements par dire un grand merci à tous les gens qui m'ont accompagné durant mes années d'études et je tiens à leur exprimer toute ma gratitude.

C'est ainsi que je termine mes 9 ans d'études post-bac (oui je suis vieux...) où j'ai traversé bien des épreuves mais j'ai survécu !!!!!!! Même s'il y a quelques séquelles (d'après les dires de certaines personnes). Je n'aurais pas pu surmonter cela sans mes amis, mes collègues, souvent devenus amis, mes professeurs, mes collaborateurs, les échanges informels, les discussions autour d'une (ou souvent plusieurs) bière, les apéros et j'en passe (et non, je ne suis pas alcoolique, c'est promis).

Mais c'est lors de ce projet de thèse que je me suis le plus éclaté quand bien même ce n'était pas facile tous les jours (hein Vincent). J'ai rencontré une vraie famille dans ce laboratoire, où la cohésion est si importante que nous avons même décidé d'attraper la COVID-19 tous ensemble, solidaires jusqu'au bout ! Mais même lorsque nous avons dû être confinés, et arrêter les entraînements, nous étions toujours épaulés, et soutenus, pour cela, un grand merci à toute l'équipe du laboratoire BCB. Les groupes meeting, les réunions, les apéros, les sorties, et toute l'aide apportée par chacun a été très importante pour moi.

À tous les membres de mon jury, Dr. Olivier Sénèque, Dr. Vincent Forge, Dr. Maïté Paternostre, Dr. Wadih Ghattas, Dr. Cécilia Ménard-Moyon un grand merci. Vos expertises me sont précieuses, et je suis très reconnaissant du temps que vous m'accordez à lire et examiner mon projet de thèse. Je tiens à remercier Dr. O. Sénèque et la Dr. M. Paternostre pour les discussions informelles et les conseils apportés sur mon projet. Un grand merci à la Dr. C. Ménard-Moyon pour ta bienveillance et tes conseils avisés.

Un énorme merci au Pr. Peter Fallor, qui a pris bien des surnoms au fil de la thèse, comme « papa Selian » ou « Ôh grand chef ». Tu as su m'accueillir chaleureusement dans ce laboratoire où j'ai appris énormément, notamment sur les fromages suisses. Mais, avec ta connaissance qui semble sans limites, tu as toujours su m'orienter, et me conseiller lorsque les temps étaient difficiles, et cela même si ce projet se trouve en dehors de ta zone de connaissances habituelle. Tu es une personne remarquable, bienveillante, toujours disponible, stimulante et nous faisant sentir comme tes égaux. Que ce soit scientifiquement et humainement, tu restes pour moi (mais aussi pour les autres) un modèle, même si tu chahutes un peu sur le champagne...

Un grand merci au Dr. Vincent Lebrun pour m'avoir choisi parmi tant d'autres afin de mener à bien ce projet. Je sais que la pression était forte de ton côté, notamment de la familia, mais tu n'as pas cédé et j'ai été l' élu ! Mais ce que tu ne savais pas, c'étaient les conséquences de ce choix, j'espère que maintenant tu sais que « un grand pouvoir implique de grandes responsabilités ».<sup>[1]</sup> Je parie qu'à maintes reprises tu l'as regretté, que ce soit à cause de ma gâchette facile (comme lors de moments pas très avantageux pour toi), mes jeux de mots (qui ne nous voilons pas la face, sont les meilleurs) ou pour mon engagement à te réveiller à n'importe quelle heure de la journée lorsque tu ne t'y attends pas. Plus sérieusement, je suis très reconnaissant d'avoir été ton premier thésard, on dit que la première crêpe est ratée mais dans ce cas c'est « partiellement » faux (j'espère). Tu es un

peu têtu, mais on s'y fait, et j'ai énormément appris à tes côtés sur tous les sujets abordés lors de ce projet très innovant. Tu as toujours su te rendre disponible lorsque le besoin se faisait ressentir. Les discussions et les conseils que tu m'as prodigués me seront d'une aide précieuse dans l'avenir et ta vision d'ensemble sur le projet a été remarquable et pertinente. Et non le respect n'est pas mort dès la première semaine, j'exprime mon respect de manière différente, c'est tout. Et tu es pardonné pour tous les « a 你好 » et toutes les autres choses que tu as dites ou faites. Je me souviendrai toute ma vie des voyages faits ensemble, notamment au synchrotron. Encore un grand merci à toi.

Le célèbre MCF5 Laurent Raibaut n'est plus à présenter. Je tiens à te remercier pour tous tes enseignements, tes conseils et les discussions sur les matériels, les synthèses au laboratoire. Grâce à toi, Timmy, j'ai pu bricoler au laboratoire et connaître les secrets des machines utilisées. C'est pour ces raisons que je ne t'en veux pas te t'avoir financé les repas du Crous, ou nourrit lors de l'heure du goûter à 16H00, et si je peux aider un pauvre maître de conférences... Et je peux le dire, je suis un des seuls à avoir pu pénétrer dans ta modeste tente, alors merci. Tu m'as marqué par ton sens du capitalisme et de la logique, je me suis même amélioré en blagues en devenant ton fidèle padawan. Mais un des moments les plus marquants a été le retour de confinement, tu es revenu avec tant d'enthousiasme ! cela faisait plaisir à voir. Bientôt la DTCLR003...

Sylvie, tu t'es élevée au rang de CR-HC et pour cela bravo. Angélique, tu t'es élevée au rang de CR-HC et pour cela bravo (**n'oublie pas l'apéro quand même...**). Merci aussi pour tous les conseils et la bonne humeur que tu as apportés lors de ces trois années de thèse. Et évidemment, BONSOIIIIIIIRRRRRRR.

Un immense merci à toi Thibaut, nous avons commencé ensemble et nous finissons ensemble ! Je me souviens de mon premier jour où je suis arrivé au laboratoire, nous nous sommes tout de suite entendus, et nous sommes devenus très bons amis, et je peux même le dire, tu es mon meilleur ami. Je ne peux être plus heureux d'avoir croisé ta route. Tu m'as beaucoup appris que ce soit culinairement, sur les vins, les champignons et tellement d'autres choses. Évidemment tu resteras aussi mon ami de Pokémon, où, je l'espère, nous deviendrons de grands maîtres chinpoko extramon ! Et merci pour tous les repas, les apéros, les sorties... Dans tous les cas, je n'aurais pas pu avoir un co-bureau aussi bien que toi. Merci pour tout, mon pote de Galler. Et je tiens aussi à remercier toute ta famille pour leur hospitalité et pour m'avoir accueilli comme l'un des vôtres.

Ensuite, Capucine, tu es devenue une très bonne amie aussi, avec qui j'ai pu parler, cuisiner, boire. Tu as su me sortir de ma zone de confort et je t'en remercie, j'ai découvert énormément de choses, comme les différents vins, la spéléo, des restaurants et j'en passe ! Tu m'as aussi présenté ta famille dont l'hospitalité est remarquable, et pour cela, merci à toi et à eux. Merci à tous les deux pour ces repas, soirées... et merci à Osmium de m'avoir presque accepté haha.

Enrico, Ma Ciao, mon bello, tu es devenu un très bon ami au fil du temps et tu as su être toujours présent lorsque j'avais besoin de toi, que ce soit scientifiquement ou non. Évidemment, je suis très heureux que tu fasses des efforts pour faire des blagues, et surtout, mon plus grand succès lors de cette thèse a été de te montrer H. Et promis, les blagues sur les ananas, j'essaie d'arrêter. Je te remercie pour tous les moments passés ensemble et je te souhaite beaucoup de réussites, même si tu es déjà un très bon chercheur.

Pauline, ma meilleure amie, je te remercie d'avoir été là depuis notre rencontre en master. Je me rappelle nos apéros, notre coloc en Estonie et tous les bons moments passés ensemble. Merci d'avoir été là pour moi, même à distance. Je suis super heureux pour toi que tu puisses t'épanouir dans un autre pays, toi qui avais peur de quitter la France. Et merci à tes parents pour leur hospitalité, avec Thibaut, lorsque nous sommes venus et pour les découvertes.

Un grand merci à Ryan, pour ses enseignements sur la modération, pour ces apéros et repas passés ensemble, et pour m'avoir fait participer à cette sortie luge. Tu es et resteras un très bon ami sur qui je peux compter. Et bien sûr, bon retour parmi nous.

Je tiens aussi à remercier tous mes collègues, Michael (pour tes raccourcis et tes préparations maisons, même si desfois...) et Merwan (pour tes histoires et ta détente en toutes circonstances), puis Marie-Julie, Katharina et Jordan.

I would like to say a big thanks to Lucie for being here, supporting and being my friend, Alice for your advice and good mood.

Un énorme merci à mes stagiaires, Louis et Solène, qui ont effectué un travail remarquable, qui ont apporté une bonne ambiance et qui m'ont supporté durant leurs stages.

Je tiens à remercier particulièrement toutes les personnes avec qui j'ai eu la chance de collaborer. Un grand merci à C. Sauter pour nous avoir donné la fièvre du synchrotron, grâce à qui nous sommes allé plusieurs fois. Évidemment, merci à Kévin (qui s'est arraché les cheveux sur la résolution de nos structures) et Petr pour leur implication. Merci à P. Bertani, pour la ssNMR, W. Shepard et D. Jeangerard pour la DRX sur la ligne PX-2A et F. Jamme et F. Wien pour la SR-CD/LD sur la ligne DISCO du synchrotron SOLEIL. Je tiens à remercier V. Demais et M. Schmutz pour les analyses de TEM. Un grand merci à S. Bressanelli et L. Pieri pour leur aide dans les expériences de TEM et Cryo-TEM. Merci à P. Hellwig et Y. El Khoury pour les mesures en ATR-IR et à M. Heinrich pour la CD.

Je tiens aussi à remercier le GDR Médyna pour son soutien dans la réalisation des expériences de TEM.

Merci à Georges, Antoine, et tous les autres du magasin pour leur serviabilité et leur sympathie.

Merci au soutien que ma famille m'a apporté durant toutes ces années d'études (et à Solène ma petite sœur préférée).

Pour finir, un grand merci à vous tous, pour votre soutien tout au long de mes années d'études.

**Merci à tous !**

[1] O. Ben, *Spider-Man*, 2002.





# 1 Introduction

## 1.1 Généralités sur les métalloenzymes artificielles

L'éventail d'objets/structures disponibles dans la Nature est immense, et chacun joue un rôle. C'est notamment le cas des enzymes qui ont pour rôle de catalyser des réactions chimiques essentielles à la survie, au développement et à la reproduction de l'organisme.<sup>[1]</sup> Les enzymes sont des protéines ayant une activité catalytique, qu'elle soit centrée ou non sur un ion métallique. Lorsque c'est le cas, nous parlons de métalloenzymes.<sup>[2]</sup> Les performances qu'offrent les enzymes en termes de chemo- régio- et énantio-sélectivité ainsi que leurs vitesses de réaction sont souvent bien meilleures que les activités des catalyseurs conventionnels.<sup>[3]</sup> Ceci fait que, dans l'industrie, les catalyseurs « classiques » peuvent être moins compétitifs lors d'étapes complexes (par exemple la fonctionnalisation « tardive » de petites molécules), malgré leur faible coût. Les enzymes souffrent cependant d'inconvénients non négligeables : elles demandent des conditions bien définies pour leur activité, mais aussi, leur coût de production est élevé. Les conditions de pH, température, mais aussi la stabilité, le prix et les quantités de catalyseurs d'enzymes disponibles représentent un frein à leur utilisation. Une comparaison entre les propriétés des catalyseurs organiques et enzymatiques est montrée en Tableau 1.<sup>[4]</sup>

Tableau 1 : Comparaison des propriétés des catalyseurs homogènes et enzymatiques (adapté de la ref <sup>[4]</sup>).

	<b>Catalyseur homogène</b>	<b>Catalyseur enzymatique</b>
<b>Énantiomères</b>	Chaque énantiomère accessible	Un énantiomère
<b>Répertoire de réactions</b>	<b>Grand</b>	Petit
<b>Nombre cycles catalytique</b>	Petit	<b>Grand</b>
<b>Compatibilité avec le solvant</b>	Organique > Aqueux	<b>Aqueux &gt; Organique</b>
<b>Optimisation</b>	<b>Chimique</b>	génétique
<b>Seconde sphère de coordination</b>	Pas bien définie	Bien définie
<b>Tolérance vis-à-vis du groupe fonctionnel</b>	Petite	<b>Grande</b>
<b>Éventail de substrats</b>	<b>Grand</b>	Limité
<b>Substrat typique</b>	Apolaire	Flexible

Aujourd'hui, où la question de l'écologie est plus que jamais un enjeu mondial, la conception de catalyseurs à partir de ressources renouvelables, et qui combinerait les avantages des deux types de catalyseurs (Tableau 1), devient de plus en plus importante pour remplacer ceux utilisés industriellement.<sup>[5]</sup> Pour cela, une transition

doit avoir lieu pour passer de catalyseurs issus de la chimie du pétrole (organiques) et comportant des métaux rares à des enzymes accessibles naturellement, à base de métaux disponibles biologiquement comme le fer, zinc, cuivre par exemple.<sup>[6]</sup> Afin de combiner les avantages des deux types de catalyseurs énumérés précédemment (en gras dans le Tableau 1), les chercheurs ont eu pour volonté de créer des métalloenzymes artificielles (ArMs).<sup>[7]</sup> Le but général étant de trouver la bonne correspondance clé - serrure, comme l'a explicité E. Fischer en 1894 et repris depuis pour expliquer la complémentarité entre la conformation d'une enzyme et celle de son substrat.<sup>[8,9]</sup> Les métalloenzymes artificielles sont des protéines modifiées chimiquement dans le but d'incorporer un groupe catalytique, basé sur un ion métallique. Elles sont donc non-naturelles. La première modification d'une protéine a eu lieu en 1956 et a été réalisée par adsorption et réduction d'un complexe palladium-soie fibroïne.<sup>[10]</sup> La conception de ces ArMs a été réalisée dans un premier temps afin d'identifier les sites responsables de la catalyse. Ce n'est que quelques années plus tard que des molécules synthétiques ont été greffées sur des squelettes protéiques déjà existants dans le but d'y apporter une activité catalytique plus

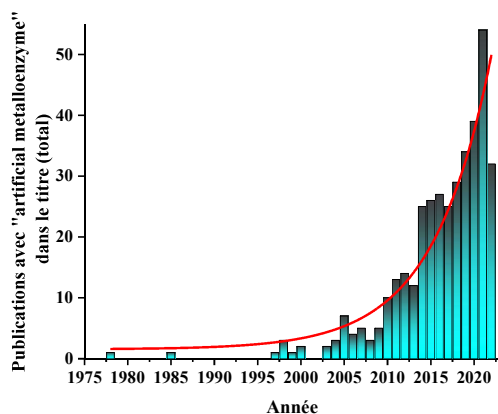


Figure 1 : Nombre de publications contenant le mot « artificial metalloenzyme » dans le titre en fonction des années sur *pubmed.gov*.

importante ou plus spécifique.<sup>[11,12]</sup> Mais la mutation de résidus ou la fonctionnalisation de ceux-ci peut changer drastiquement les propriétés de l'enzyme, comme une perte d'activité pouvant être due à une modification de son repliement, et par conséquent de sa poche catalytique.<sup>[13,14]</sup> Au fil des années et depuis la fin des années 1990, le nombre de publications contenant le mot « artificial metalloenzyme » dans son titre n'a fait qu'augmenter (Figure 1). La croissance exponentielle de publications liées aux ArMs démontre bien l'intérêt grandissant de ce champ de recherche et la volonté de développer de nouveaux catalyseurs.

Ces nouveaux catalyseurs se veulent donc plus efficaces et sélectifs que les catalyseurs homogènes ou enzymatiques. Grâce à cela, des réactions non accessibles naturellement, souvent réalisées en industries par des catalyseurs à base de métaux rares, ont pu être réalisées. C'est notamment le cas des réactions de polymérisations radicalaires contrôlées<sup>[15]</sup>, des réactions de métathèses<sup>[16]</sup> ou encore des réactions de Diels-Alder (même si récemment a été découverte une enzyme naturelle capable de faire cela).<sup>[17,18]</sup> Ces dernières étant des réactions extrêmement utilisées en industrie, il y a un intérêt pour remplacer les catalyseurs actuels. De plus en plus de réactions sont accessibles et optimisées avec des ArMs pour des applications diverses et variées en industrie ou médecine.<sup>[19]</sup>

Leur conception tend aussi à respecter de plus en plus les différents principes de la chimie verte afin d'avoir des catalyseurs qui soient les plus durables possibles.<sup>[20-22]</sup> Mais l'incorporation de cofacteurs/ligands synthétiques aux protéines existantes requiert des étapes de synthèses organiques supplémentaires, et parfois, des métaux rares. Plusieurs stratégies sont employées aujourd'hui afin de concevoir des métalloenzymes artificielles<sup>[23]</sup> que ce soient par les développements d'anticorps catalytiques<sup>[24]</sup>, d'enzymes *de novo* par conception informatique<sup>[25]</sup>, de métalloenzymes par évolution dirigée<sup>[26-28]</sup>, de DNazymes<sup>[29]</sup> ou encore, et la plus répandue de nos jours, par l'incorporation d'un cofacteur (artificiel) métallique.<sup>[30]</sup> Pour cette dernière, il existe quatre différentes stratégies, celles-ci étant schématisées en Figure 2. La première consiste à ancrer

covalentement un cofacteur métallique (complexe) sur un résidu nucléophile de la protéine (a).<sup>[31]</sup> La seconde stratégie est l'ancrage par interactions supramoléculaires (b).<sup>[32]</sup> La troisième stratégie se fait par incorporation du cofacteur par création d'une liaison dative entre le métal et un/des résidu(s) coordinant de la protéine, c'est-à-dire qu'on forme un complexe ternaire (c).<sup>[33]</sup> La dernière se résume à substituer un métal présent naturellement dans l'enzyme par un autre (d).<sup>[34]</sup>

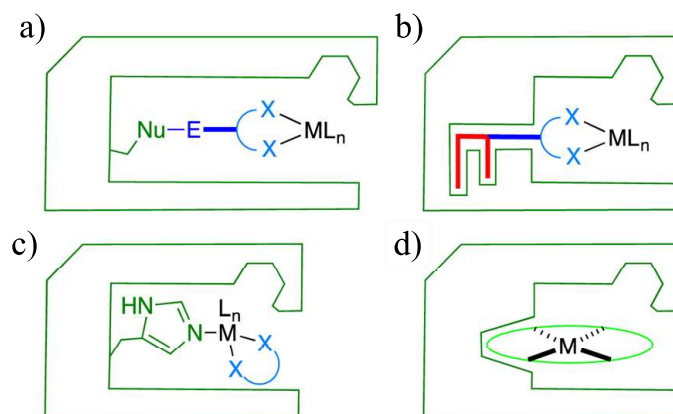


Figure 2 : Différents types d'ancrages d'un cofacteur métallique à l'intérieur d'un squelette protéique, que ce soit a) covalent, b) supramoléculaire, c) datif ou d) par substitution d'un métal. En vert sont représentés les protéines et cofacteurs naturels, l'ancrage en rouge, les ligands et espaceurs en bleu, et en noir le métal. (repris de la ref<sup>[7]</sup>)

Néanmoins, toutes ces méthodes sont basées sur l'utilisation de protéines, bien repliées, entières, ce qui veut dire que les propriétés de l'enzyme artificielle résultante sont limitées par celles de la protéine hôte (domaine de stabilité, solubilité, etc.), et ce qui pose la problématique de leur production. En effet, produire des protéines n'est pas toujours facile, ni même abordable, et les quantités obtenues peuvent être faibles.<sup>[35]</sup> Aussi, une des voies explorées dans la conception d'enzymes artificielles est la conception dite « minimaliste », reposant sur des séquences courtes, faciles à produire.

## 1.2 Autoassemblages minimalistes fonctionnels comme métalloenzymes artificielles

Inspiré par la nature, le but ici est de former des catalyseurs mimant les sites catalytiques naturels à partir de peptides de séquences courtes.<sup>[36,37]</sup> Ces peptides peuvent avoir des tailles très variables et avoir des repliements différents.<sup>[38]</sup> Dans ce domaine, l'équipe d'I. Korendovych a beaucoup œuvré. Il a notamment montré qu'on pouvait diminuer la flexibilité des peptides courts (principal défaut des peptides courts lorsqu'on veut les utiliser pour des complexes catalytiques) en les « verrouillant » dans un assemblage de type amyloïde.<sup>[39]</sup> Ces assemblages sont fibrillaires et sont constitués d'un empilement de feuillets  $\beta$  perpendiculaires à l'axe de la fibre, formant des feuillets  $\beta$  croisés de longueur infinie.<sup>[40]</sup> En effet, en suivant une approche minimaliste de conception de métallopeptides catalytiques, ils ont conçu des séquences courtes (i.e. 7 acides

aminés) capables de s'autoassembler en fibres, de coordonner des métaux, permettant ainsi la catalyse de réactions d'hydrolyse et/ou d'oxydation.<sup>[38,41]</sup> Un des peptides les plus connus, dont la séquence est directement inspirée de l'anhydrase carbonique humaine, reproduit le site catalytique à trois histidines de l'enzyme simplement par superposition de deux feuillets  $\beta$  parallèles (Figure 3).<sup>[42]</sup> Cette conception minimaliste augmente considérablement le ratio entre le site métallique et la taille du support (peptide court vs enzyme entière), et cela tout en conservant un assemblage stable. À noter qu'une des histidines pontre deux ions  $Zn^{II}$ .<sup>[42]</sup> La séquence choisie, Ac-IHIHIQI-amide, est optimisée pour former des feuillets  $\beta$  grâce à une séquence riche en résidus apolaires (hydrophobes) alternativement avec des résidus polaires (hydrophiles).<sup>[43]</sup> Les résidus hydrophobes seront alors sur une face alors que les hydrophiles sur l'autre, ce qui pousse le système à former des interactions entre les feuillets  $\beta$  pour masquer les résidus apolaires, et *in fine*, amènent à bien exposer au solvant les sites métalliques (Figure 3 c et d). Cet exemple illustre bien la conception « d'enzymes artificielles » avec un motif minimal, par identification des sites d'intérêts et reproduction sur des peptides de taille réduite. Ceci facilite grandement la synthèse, la purification et la tolérance vis-à-vis des mutations, permettant de personnaliser les activités que ceux-ci auront.<sup>[44]</sup>

Un autre exemple, plus récent, est celui de l'équipe de D. Das. Ils ont synthétisé un nanotube de peptides amyloïdes (7 acides aminés (A.As)) fonctionnalisé avec un imidazole en  $N_{ter}$ . Celui-ci se retrouvant en surface, il peut coordonner l'hémine et interagir avec la sarcosine oxydase. De cette manière, une réaction en cascade a pu être réalisée.<sup>[45]</sup>

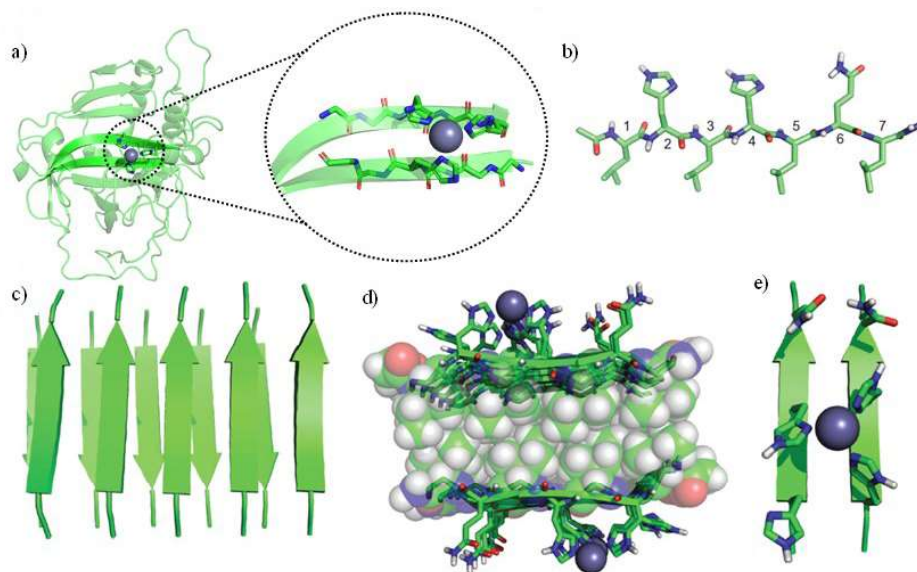


Figure 3 : Étapes de conception de l'heptapeptide Ac-IHIHIQI-amide s'inspirant de l'enzyme naturelle. a) Structure cristallographique de l'anhydrase carbonique humaine avec le site catalytique formé de deux His d'un feuillet et d'une His d'un autre feuillet coordonnant un cation de Zinc. b) Modèle du peptide conçu montrant la position des résidus dans la séquence. c) Modèle fibrillaire global du peptide, d) mise en évidence du cœur hydrophobe et e) première sphère de coordination du zinc avec le peptide. (repris de la ref<sup>[42]</sup>).

Le point commun entre ces peptides est qu'ils adoptent une structure fibrillaire de type amyloïde. C'est pour ces questions de stabilités que des peptides ont été conçus à partir de fragments de l'amyloïde  $\beta$  ( $A\beta$ ) et permettent de former une structure fibrillaire stable par autoassemblage.<sup>[46-48]</sup> Cette stabilité est conférée par la formation d'un réseau de liaisons hydrogène inter-chaines principales stables dues à l'enchaînement de feuillets  $\beta$  parallèles le long de la fibre. Les feuillets  $\beta$  peuvent adopter des conformations parallèles et anti-parallèles, avec une certaine torsion, afin de s'adapter à leurs environnements.<sup>[49]</sup> De manière générale, dans l'optique de concevoir des métalloenzymes artificielles plus respectueuses de l'environnement, la conception de fibres peptidiques représente un choix prometteur. Industriellement, la séparation des produits et leur recyclage peuvent être relativement faciles (i.e. par filtration). Les amyloïdes catalytiques représentent encore un jeune champ de recherche en pleine expansion mais ayant déjà montré des résultats surpassant parfois ceux des enzymes naturelles.<sup>[42,45]</sup>

De plus en plus de matériaux sont conçus à partir de peptides capables de former des feuillets  $\beta$  adoptant des structures de types amyloïdes qui sont connues, produites et caractérisées.<sup>[39,50]</sup> Une multitude d'applications a déjà été développée pour ces fibres dans le but de concevoir des matériaux fonctionnels.<sup>[51,52]</sup> La stabilité des structures amyloïdes permettent la mutation de leurs séquences sans pour autant perturber leur assemblage.<sup>[53]</sup> Ces séquences protéiques peuvent notamment être ajoutées à d'autres protéines pour leur donner des propriétés d'assemblage. Et en particulier, les protéines prion, étant basées sur une structure amyloïde, sont formées de feuillets  $\beta$  parallèles par enroulement de la chaîne protéique, constituant un cœur hydrophobe. Cette définition correspond à la définition plus générale d'un solénoïde  $\beta$ . Plusieurs exemples de construction d'assemblages à partir de ce type de repliement a été réalisé. C'est notamment le cas de fils moléculaires chimériques conçus à partir de protéines prion, de taille nanométrique, pouvant servir de fils conducteurs d'électrons pour des applications catalytiques ou photovoltaïques par exemple.<sup>[54,55]</sup>

Un premier exemple se base sur la séquence canonique de la protéine prion, HET-s nommée PFD(HET-s) pour « prion forming domain ». C'est une séquence connue pour former des autoassemblages fibrillaires stables et canoniques. À partir de ce PFD(HET-s), correspondant à deux pas de solénoïde, il est alors possible de concevoir des fils ayant une fonction définie, par fonctionnalisation des séquences. Ceci a été réalisé par l'équipe de V. Forge grâce à la fonctionnalisation du  $N_{\text{ter}}$  de ce PFD(HET-s) avec une protéine rubredoxine (Rb) contenant un site actif à  $Fe^{2+/3+}$  coordonné par quatre cystéines. Les Rb pouvant alors s'aligner le long de la fibre formée, avec une distance entre site de coordination d'environ 10 Å permettant le transfert d'électrons. Pour cela, ils ont exprimé la protéine chimérique Rb-HET et ils ont montré qu'il était possible de former des fibres avec, apportant une propriété de conduction d'électrons le long de la fibre formée (Figure 4a). Ces travaux montrent qu'il est possible de concevoir des hydrogels RedOx à base d'amyloïdes (AbRxHs).

Toujours à base du motif PFD(HET-s), cette équipe a montré qu'en changeant la nature du site RedOx, il est alors possible de changer le potentiel RedOx, cela pouvant permettre de balayer une large gamme de potentiels, pour une large gamme de réactions chimiques.<sup>[55]</sup> Pour cela, un mutant de PFD(HET-s) contenant une seule cystéine est conçu et exprimé. Cette cystéine seule permet la fonctionnalisation d'une seule molécule de méthyle viologène, utilisé comme relais d'électrons. Comme précédemment, des fibres ont été formées par

autoassemblage des unités PFD(HET-s) modifiées de manière homogène et prédictible (Figure 4b). Ces fibres permettent de conduire le courant par saut d'électrons entre les différents sites de résidence de ceux-ci. Il a aussi été démontré qu'en fonction du pH, des hydrogels ont pu être conçus et que ceux-ci peuvent être utilisés comme transporteurs d'électrons entre une hydrogénase et une électrode.

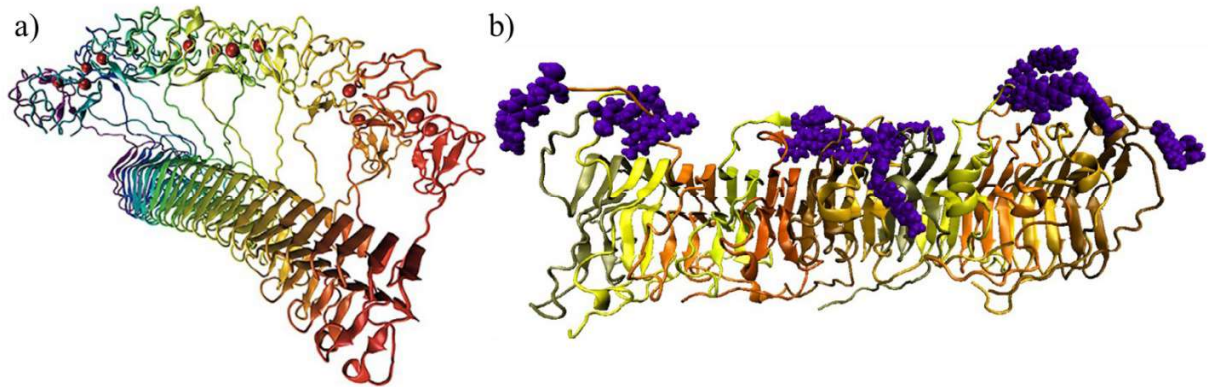


Figure 4 : Modèle de fibres formées à partir du motif s'autoassemblant d'une protéine prion HET[218-289] a) chimérique Rb-HET[218-289] b) fonctionnalisée avec du méthyle viologène (en violet). (Figures reprises des ref<sup>[54,55]</sup>).

Ces deux exemples montrent bien qu'il est possible, à partir de motifs répétitifs, de concevoir des autoassemblages fonctionnels, pouvant être contrôlés en fonction du pH. Les propriétés RedOx de ces assemblages fonctionnalisés par des protéines ou molécules pouvant être changées. Il serait alors possible d'assembler des briques différentes, avec des sites fonctionnels différents afin de pouvoir apporter des propriétés catalytiques par exemple.

### 1.3 Des briques peptidiques pour mimer les métalloenzymes multisites

Les propriétés des fibres amyloïdes sont très avantageuses pour former de nouveaux catalyseurs hétérogènes plus verts et en accord avec les enjeux écologiques actuels. Comme nous l'avons vu dans la partie précédente, l'introduction d'un site catalytique mimant un site enzymatique est relativement accessible. Cependant, lorsque nous voulons combiner plusieurs fonctions sur une même fibre (i.e. catalyser plusieurs réactions) ou mimer une métalloenzyme multisite, cela devient plus complexe. Une métalloenzyme multisite est une enzyme contenant plusieurs sites métalliques, qui peuvent travailler de concert (i.e. plusieurs sites permettent aux électrons d'arriver au site catalytique). Il en existe une multitude et peuvent contenir différents métaux espacés de plusieurs dizaines d'angströms.<sup>[56,57]</sup> Ainsi, avec un seul peptide très court, il est difficile de former, de manière sélective et contrôlée, deux (ou plusieurs) sites métalliques, et même d'en moduler la distance. Il faut donc un système plus versatile que ces peptides très courts. Une des manières de faire serait de pouvoir concevoir des briques (rappelant les Lego), dont on contrôle l'assemblage entre les différentes briques,

et que l'on peut fonctionnaliser de diverses manières sans affecter leurs propriétés d'assemblage (Figure 5). Chaque brique ne pouvant interagir qu'avec une autre bien définie, il serait alors possible de concevoir des assemblages avec des fonctions dans un ordre précis. L'assemblage devrait donc permettre de rapprocher, et de maintenir proches, les différents sites dans l'espace. Ceci permettrait la coopération entre les différents sites.<sup>[58]</sup> De plus, si l'on peut aussi faire varier leur taille, alors cela permet de faire varier la distance entre les sites (après formation des assemblages). On peut voir un exemple schématique explicité en Figure 5.

Il est alors nécessaire d'identifier, pour concevoir ces briques, un repliement protéique qui réponde à ce cahier des charges.

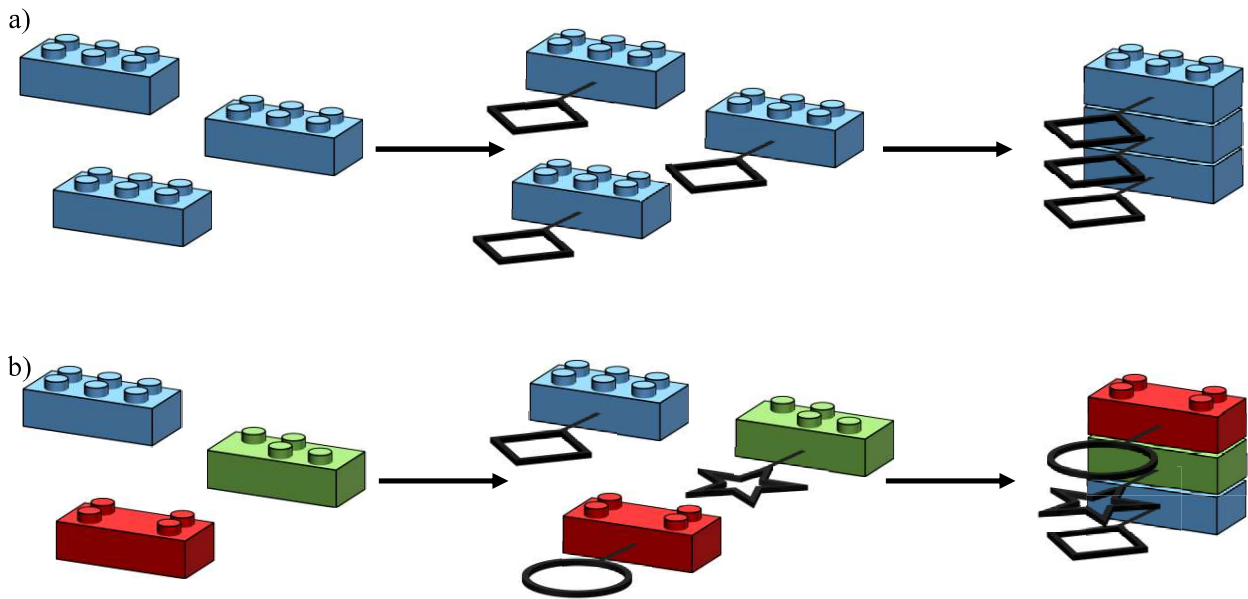


Figure 5 : a) conception de fibres peptidiques formant une métalloenzyme artificielle, b) conception de fibres peptidiques formant une métalloenzyme artificielle multisite, la brique verte ne pouvant s'empiler que sur la bleue, la rouge sur la verte et, la bleue sur la rouge permettant un contrôle de l'assemblage précis.

Dans cette optique, les nanotubes formés de peptides cycliques composés d'une alternance de résidus de configuration L et D, ou pseudopeptides cycliques, pourraient être une bonne option.<sup>[59,60]</sup> Cependant, dans ces systèmes, il est difficile de contrôler le nombre de briques empilées. Contrairement à des systèmes de chaînes naturelles (tout L) et linéaires, ils ne se prêtent pas à la production par expression recombinante. De plus, ces systèmes étant aussi très symétriques, ils ont une forte tendance à s'agréger latéralement, bloquant les interfaces entre les possibles substrats et sites catalytiques.<sup>[61]</sup> L'incorporation de boucles comme retrouvées dans les protéines n'est pas possible dans le cas de peptides cycliques L,D.<sup>[62]</sup> Malgré cela, ces peptides sont étudiés pour d'autres applications, comme la conception de fils conducteurs, des capteurs ou encore transporteurs d'ions.<sup>[61,63-65]</sup> Des systèmes catalytiques à partir de ces peptides cycliques commencent également à être développés pour être utilisés en tant que métalloenzymes artificielles, malgré les limitations énoncées.<sup>[66]</sup>



Pour ces raisons, nous pensons que baser ce système de briques sur des chaînes protéiques linéaires capables d'adopter un repliement stable par formation d'une structure de type prion/amyloïde est une stratégie plus judicieuse et plus généralisable en ce qui concerne le développement de métalloenzymes artificielles, notamment multi-sites.

Nous nous sommes tournés vers les solénoïdes protéiques qui correspondent à des types de structures répétitives retrouvées fréquemment dans les protéines.<sup>[67]</sup> Il en existe de différentes tailles, et de structures très variables, ce qui permet d'amener un large éventail de possibilités pour concevoir les briques de base. Elles peuvent être composées d'hélices  $\alpha$ , de feuillets  $\beta$  ou d'un mélange des deux avec des stabilités structurales différentes en fonctions des structures secondaires formées (Figure 6).<sup>[67]</sup>

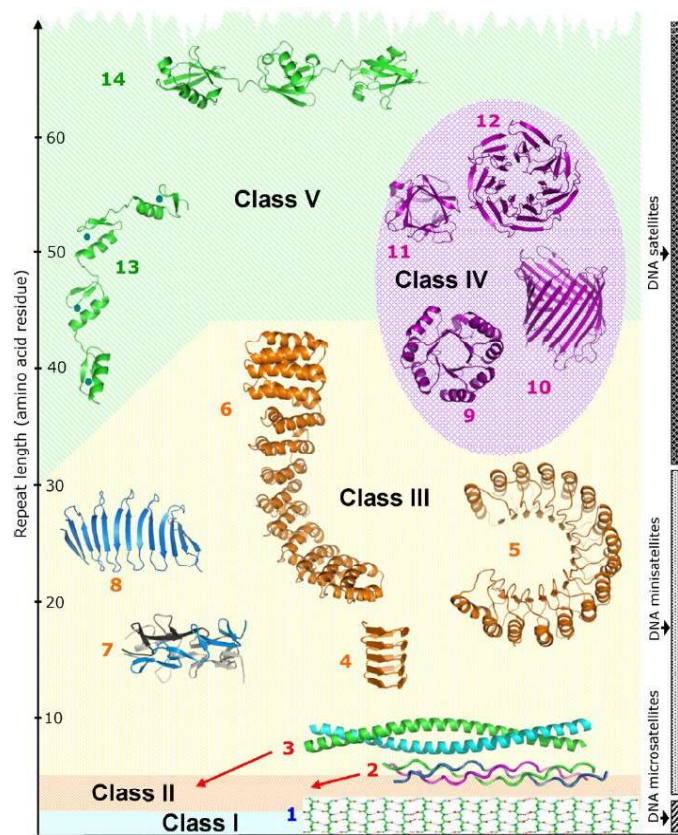


Figure 6 : Classification structurelle des protéines à motif répétitif en fonction de la longueur de ce motif. Les exemples montrés par classe sont : I, 1) la structure  $\beta$  cristalline d'une polyalanine, II, 2) une triple hélice de collagène et 3) une bobine enroulée d'hélices  $\alpha$ , III, 4) un solénoïde  $\beta$ , 5) un solénoïde  $\alpha/\beta$  de protéine LRR (répétition riche en leucines), 6) un solénoïde  $\alpha$ , 7) un trimère de spirales  $\beta$  et 8) une structure antiparallèle  $\beta$  plane. IV, 9) un TIM-barrel, 10) un tonneau  $\beta$  transmembranaire, 11) une structure en trèfle  $\beta$ , 12) une hélice marine  $\beta$ . V, 13) un domaine de doigt de zinc, et 14) une chaîne poly-ubiquitine.

Figure reprise de la ref<sup>[67]</sup>.

Des banques de données permettent aujourd'hui de classer les protéines répétitives en fonction de leur structure grâce à l'identification de celle-ci par différents algorithmes. La première ayant été construite en 1999,

PROSITE<sup>[68]</sup>, d'autres ont suivi et sont énumérées en référence <sup>[69]</sup>. C'est la méthodologie d'identification de motifs répétitifs qu'a utilisée l'équipe de P.L Davies pour concevoir une brique de base à partir d'un motif répétitif naturel en solénoïde  $\beta$ , mais pouvant s'appliquer à toutes les structures répétitives.<sup>[70]</sup> Ils ont décrit une méthode de conception consistant à identifier les résidus les plus fréquemment retrouvés dans un motif, ceux qui sont critiques au repliement, les autres pouvant être mutés afin d'apporter une fonction car ils ne sont pas critiques pour la structure. Puis, par modélisation, il a été suggéré d'en faire une boucle, de les empiler et de les connecter pour ensuite évaluer leur stabilité par dynamiques moléculaires. Une fois le modèle stable, le peptide peut être synthétisé et utilisé comme brique de base pour la conception de fibres.<sup>[70]</sup> La conception de Lego stables peut être réalisée facilement par cette méthode, par l'étude de structures répétitives (Figure 7).

Cette méthode représente une méthode générale de conception issue d'un exemple publié quelques années auparavant. En s'inspirant d'une protéine à rouleau- $\beta$ , un type de solénoïde  $\beta$ , ils ont réussi à concevoir un peptide capable de s'autoassembler en fibres en présence de lanthane.<sup>[71]</sup> Dans cet exemple, la taille des fibres formées a été contrôlée par l'ajout d'un motif de terminaison, permettant de stopper la continuité de l'assemblage.<sup>[72]</sup>

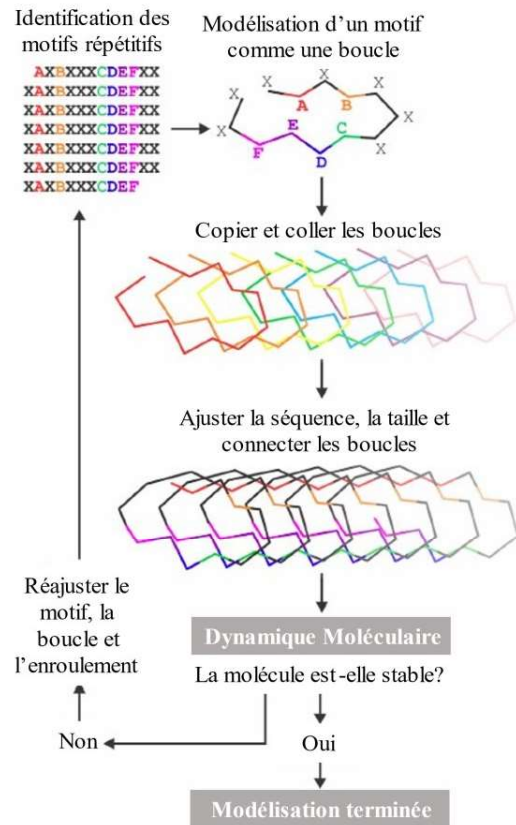


Figure 7 : Méthode générale de conception de protéines à motifs répétitifs avec A-F étant des résidus participants aux motifs et X ceux qui ne se répètent pas. Figure reprise de la ref<sup>[70]</sup>.

## 1.4 Les solénoïdes $\beta$

Un solénoïde  $\beta$  est un repliement de la chaîne polypeptidique s'enroulant autour d'un axe formant des feuillettes  $\beta$  parallèles (perpendiculaires à l'axe) connectés par des coudes dits  $\beta$ .<sup>[73]</sup> Le nombre de feuillettes présents par pas d'hélice diffère en fonction des structures et on parlera de rouleaux  $\beta$  parallèles lorsqu'il y en a deux et d'hélices  $\beta$  parallèles lorsqu'il y en a trois par exemple (Figure 8).<sup>[74,75]</sup> Les toutes premières structures de ce type ont été découvertes en 1993, la pectate lyase C d'*Erwinia chrysanthemi*<sup>[75]</sup> et la protéase alcaline de *Pseudomonas aeruginosa*.<sup>[74]</sup> Ces structures adoptent des formes différentes (Figure 8)<sup>[76-82]</sup> et elles sont retrouvées dans la nature ayant différentes fonctions, comme protéines antigel par exemple, dans certains poissons ou plantes afin d'empêcher la formation de cristaux de glace lors du gel, pour que les cellules n'éclatent pas.<sup>[83]</sup> Ceci est dû au fait que les solénoïdes  $\beta$  de ces protéines exposent une surface qui permet

d'organiser les molécules d'eau, les empêchant de cristalliser.<sup>[84]</sup> Ils peuvent aussi être retrouvés dans des enzymes comme des lyases<sup>[85]</sup>, des transférases<sup>[86]</sup> mais aussi dans les protéines de type prion.<sup>[87,88]</sup> Ils reposent sur un réseau de liaisons hydrogène conféré par l'empilement des feuillets  $\beta$  parallèles dans leurs structures qui représente la contribution principale à leur stabilité. La seconde contribution à leur stabilité est celle du cœur hydrophobe, autour duquel s'enroule la chaîne principale. Ce cœur est formé de chaînes latérales apolaires dirigées vers l'intérieur du solénoïde. Il contribue à la stabilité générale de la structure 3D et permet de la renforcer.<sup>[89]</sup> D'autres interactions entre les différentes chaînes latérales peuvent avoir lieu, permettant de renforcer encore plus le repliement, c'est notamment le cas des chaînes latérales des résidus asparagines ou glutamines qui forment des échelles de liaisons hydrogène lorsqu'elles sont empilées.<sup>[90,91]</sup> L'empilement de chaînes latérales aromatiques telles que celle des phénylalanines, tyrosines ou encore histidines sont retrouvées dans ces structures. Tout cela confère une régularité structurale permettant au squelette d'avoir des propriétés mécaniques très intéressantes, pouvant résister à des contraintes diverses.<sup>[92,93]</sup>

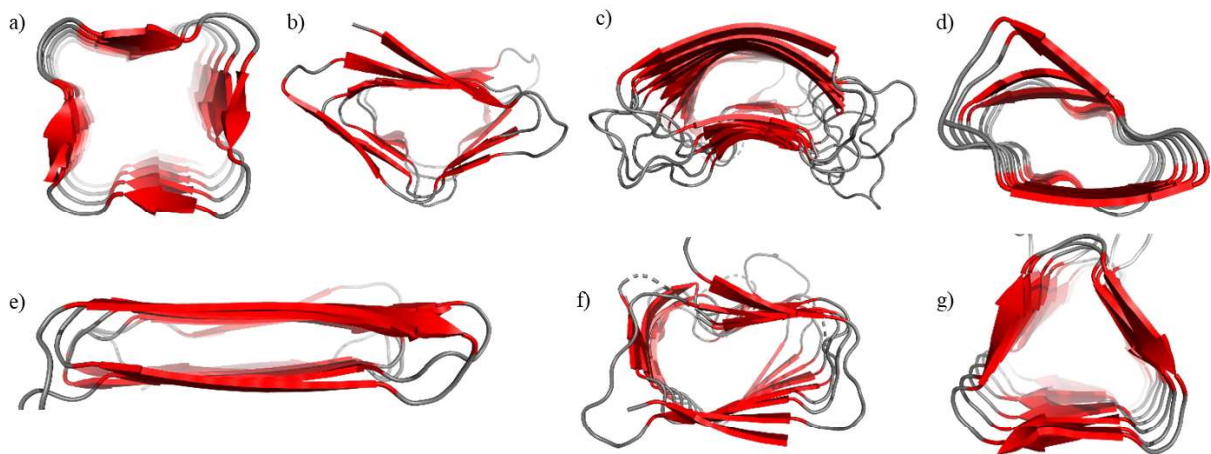


Figure 8 : Exemples de solénoïdes  $\beta$  naturels, PDB : a) 4YEI, b) 4W8Q, c) 3SUC, d) 1KAP, e) 6XNR, f) 1O8G et g) 1LXA.  
[76-82]

## 1.5 Propriétés d'autoassemblage des solénoïdes $\beta$

Aux vues de leurs structures et compositions, les solénoïdes  $\beta$  représentent une classe de protéines très intéressante pour la conception de biomatériaux à base de protéines (i.e. nanofils, fibres). Nous pensons que leurs motifs répétitifs, leur conférant une régularité et une stabilité remarquable, se prêtent à la conception de briques de base pour concevoir des métalloenzymes artificielles par autoassemblage de ceux-ci.<sup>[70]</sup> Les exemples naturels de protéines à solénoïde  $\beta$  étant composés de motifs de terminaisons afin de garder un assemblage fini et le plus souvent soluble (i.e. évite l'agrégation de ces protéines dans la cellule), il est nécessaire de les enlever lorsque nous souhaitons concevoir des autoassemblages. Dans un deuxième temps, ces motifs de terminaison pourront être utilisés pour contrôler l'assemblage.

Plusieurs études ont été réalisées sur l'ingénierie de motifs solénoïdes  $\beta$  pour leur faculté à s'autoassembler. Les premières ont été réalisées *in silico* (i.e. modélisation, par F. Cohen et collaborateurs).<sup>[94]</sup> Les auteurs ont montré qu'il était possible d'empiler des motifs en solénoïdes  $\beta$ , par autoassemblage, formant ainsi des fibres. Pour cela, ils ont émis l'hypothèse qu'une protéine prion se replie en solénoïde  $\beta$  et forme un trimère, celle-ci composée de trois hélices  $\beta$  parallèles puis, par dynamique moléculaire (MD), ils ont observé un empilement de ces types de motifs, hélice sur hélice.

Nous pouvons aussi penser que l'autoassemblage de motifs canoniques seuls correspondant à la séquence formant uniquement le solénoïde suffirait à la fibre d'être stable. Ce sera la colonne vertébrale de notre système de briques, sur lequel nous viendrons ajouter diverses structures fonctionnelles.

L'équipe de D. L. Cox a montré que l'assemblage pouvait se faire avec des types de solénoïdes  $\beta$  différents, et que l'assemblage restait stable dans le temps.<sup>[95]</sup> Beaucoup de modélisation sur les protéines se repliant en solénoïdes  $\beta$  a été réalisée pour en savoir plus sur la stabilité des édifices qui seront formés par autoassemblage.<sup>[93]</sup> Il en est ressorti que les solénoïdes  $\beta$  formaient des assemblages stables pouvant permettre la conception de nouveaux matériaux, pour diverses applications.<sup>[96]</sup> D'après les mesures réalisées *in silico*, la structure la plus stable correspond à celle d'une hélice  $\beta$  gauche. Cependant, il manque encore beaucoup de travaux expérimentaux permettant déjà de confirmer les études *in silico* et de mettre en avant les propriétés très intéressantes de ces solénoïdes estimées par dynamique moléculaire. Cette équipe a réalisé des travaux expérimentaux sur une structure étudiée par MD afin de concevoir des biomatériaux à partir de solénoïdes  $\beta$ .<sup>[96]</sup>

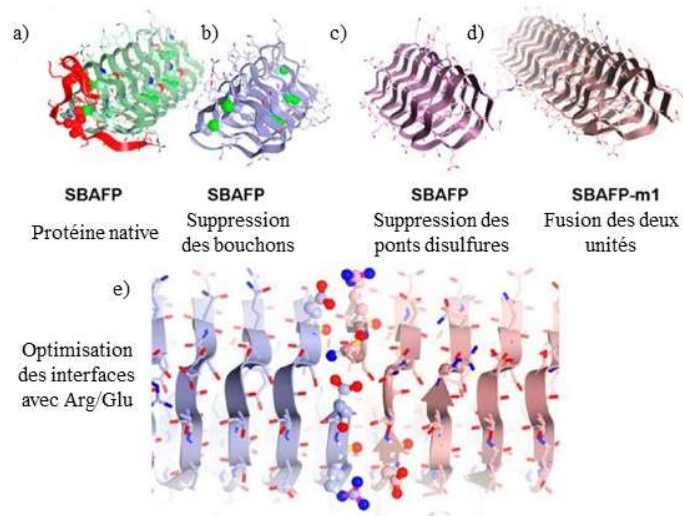


Figure 9 : Étapes de conception de biomatériaux à partir d'une protéine ayant un repliement en hélice  $\beta$  gauche de type II. a) Protéine native, b) Suppression des motifs de « bouchon », c) Suppression des ponts disulfures, d) Fusion de deux unités formant le monomère SPAFP-m1 et e) Optimisation des surfaces de contact par incorporation de couples arginines/acides glutamiques aux extrémités. Figure reprise de la ref<sup>[96]</sup>.

Afin de réaliser cela, ils ont choisi la *Spruce Budworm Antifreeze Protein* (PDB : 1M8N), une protéine se repliant en six tours d'hélice  $\beta$  gauche de type II. Ils ont commencé par enlever les motifs de bouchon présents naturellement. Ils ont ensuite enlevé les cystéines formant des ponts disulfures pouvant poser des problèmes

lors de l'expression et les ont remplacées par des sérines pouvant faire des liaisons hydrogène avec un azote de la chaîne principale. En second lieu, ils ont fusionné deux unités permettant d'obtenir un objet plus maniable expérimentalement, et avec une meilleure préorganisation. La dernière étape de la conception a été l'optimisation de l'interface entre deux monomères afin de former des ponts salins entre eux, notamment grâce à l'incorporation de paires d'arginines/acides glutamiques. Les différentes étapes de conception sont résumées en Figure 9. Cet exemple est le premier en ce qui concerne la conception de biomatériaux à partir d'un type particulier de solénoïde  $\beta$ , l'hélice  $\beta$  gauche de type II. Ils ont produit ces monomères par expression recombinante, et ont montré que ces monomères s'autoassemblent pour former des fibres. Les caractérisations associées sont compatibles avec le repliement attendu. Les études par spectroscopie de dichroïsme circulaire (CD) montrent une structure secondaire en feuillet  $\beta$  avec un signal négatif autour de 218 nm et positif à 200 nm (Figure 10 a). La microscopie à force atomique (AFM) a, elle, prouvé la présence de fibres formées par autoassemblage (Figure 10 b). Ils ont aussi montré qu'il était possible de suivre la formation de fibres formées par fluorescence de la thioflavine-T (ThT), un marqueur utilisé permettant de suivre la formation de fibres amyloïdes.

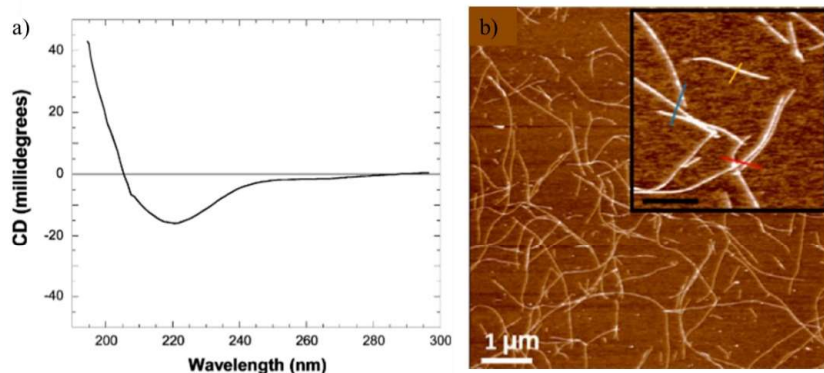


Figure 10 : a) Spectre de dichroïsme circulaire de l'autoassemblage formé par le SPAFP-m1 après incubation et b) Microscopie à force atomique des fibres formées. Figure reprise de la ref<sup>[96]</sup>.

Ce travail représente donc une base solide démontrant la faisabilité de la formation de fibres à partir de solénoïdes  $\beta$ , et plus particulièrement d'hélices  $\beta$  gauche. Une démarche similaire a été entreprise avec un autre repliement, toujours en solénoïde  $\beta$ , et a montré une très forte stabilité de ce type d'assemblage.<sup>[97]</sup>

## 1.6 Le cas des hélices $\beta$ gauche

Les hélices  $\beta$  ( $\beta$ Hs) représentent un type particulier de repliement en solénoïde  $\beta$ , possédant 3 brins  $\beta$  par pas d'hélice, c'est-à-dire qu'elles sont de section triangulaire. Il y existe deux conformations différentes, une droite<sup>[75]</sup> ( $R\beta$ H) et une gauche<sup>[98]</sup> ( $L\beta$ H). On retrouve deux types de  $L\beta$ Hs différents, le type I et II.<sup>[99,100]</sup> Les hélices  $\beta$  droites ( $R\beta$ Hs) (i.e. enroulement de la chaîne vers la droite du  $N_{\text{ter}}$  vers le  $C_{\text{ter}}$ ), sont celles qui sont les moins régulières avec des feuillets non réguliers connectés par des boucles et/ou coudes de type variable et une

section variable le long de la structure (Figure 8 b et e). La bactofiline est une protéine qui possède cette structure, qui s'autoassemble dans la cellule pour former des fibres dont on pense qu'elles jouent le rôle de cytosquelette chez certaines bactéries.<sup>[101]</sup> À l'inverse, les hélices  $\beta$  gauche (i.e. enroulement de la chaîne vers la gauche du  $N_{ter}$  vers le  $C_{ter}$ ), sont beaucoup plus régulières, avec des feuillets  $\beta$  ainsi que des coudes  $\beta$  de taille unique, et comportant peu de boucles (Figure 8 f). Ces différences entre les deux conformations sont à l'origine de la différence de stabilité entre les deux structures, la  $L\beta H$  étant plus régulière et rigide que la  $R\beta H$ .<sup>[102]</sup> La conception et l'empilement de brique à base d'hélices  $\beta$  gauche nous apparaît donc plus évident, plus prédictible, qu'avec la conformation droite. Les assemblages ayant des repliements en hélices  $\beta$  gauches auront une stabilité importante, se reposant sur un cœur hydrophobe compact, ainsi qu'un réseau de liaisons hydrogène répétitif.<sup>[91]</sup>

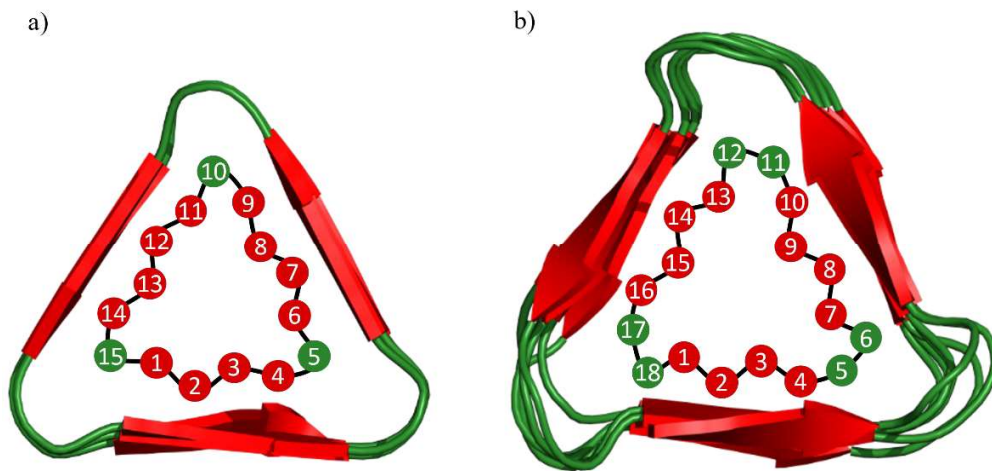


Figure 11 : Comparaison des sections des deux types d'hélices  $\beta$  gauche, a) de type II (PDB : 1M8N) et b) de type I (PDB : 1HM9). Les feuillets  $\beta$  sont représentés par les flèches rouges et les coudes  $\beta$  sont montrés en vert.

Ensuite, les différences entre les deux types de  $L\beta H$ s seront discutées ici. Pour commencer, la  $L\beta H$  de type II est une hélice composée de 15 résidus par pas d'hélice. Ses trois faces sont constituées de brins  $\beta$  de quatre résidus de long connectés par un résidu constituant le coude  $\beta$  (Figure 11 a). La  $L\beta H$  type I est constituée de 18 A.As par pas d'hélice, répartie dans trois feuillets  $\beta$  de quatre acides aminés connectés par des coudes de 2 résidus (Figure 11 b). Comme nous pouvons le voir, le diamètre est différent entre les deux types, ce qui est dû au nombre de résidus par tour. Aussi, il a été reporté que la rigidité de la  $L\beta H$  de type II est supérieure à celle du type I.<sup>[100]</sup> Cette rigidité est due aux ponts disulfures présents afin de renforcer le coude dans les hélices de type II naturelles, mais aussi aux feuillets  $\beta$  qui sont plats. La  $L\beta H$  de type I, est beaucoup plus flexible, comme nous pouvons le voir sur la Figure 11 b, avec les brins  $\beta$  des feuillets ayant des angles différents entre eux et les coudes plus libres et flexibles. Une autre différence est la stabilité de ces types de repliements, le type I étant plus stable que le II.<sup>[100]</sup> Le type II est plus rigide mais ne tolère pas les fonctionnalisations, mutations ou incorporation de boucles (i.e. qui permettraient d'accueillir un site métallique), ce qui n'est pas le cas pour le type I. Le type II serait alors une bonne option pour concevoir des nanofils linéaires et rigides, alors que le type I serait plus adapté aux applications nécessitant une diversité structurale plus importante, comme la

fonctionnalisation avec des boucles ou des domaines catalytiques par exemple. Enfin, une autre propriété intéressante est la capacité des hélices  $\beta$  gauche de type I à former, naturellement, des structures trimériques. Ceci peut s'avérer intéressant aussi lorsque l'on veut contrôler l'état d'oligomérisation d'une protéine par ingénierie. On retrouve naturellement dans certaines structures des métaux à l'interface entre deux hélices, coordonnés par deux résidus d'une hélice, et un d'une autre, comme c'est le cas pour l'anhydrase carbonique  $\gamma$ .<sup>[86]</sup> En revanche, les L $\beta$ Hs de type II sont monomériques et sont souvent utilisées par la Nature comme protéines antigels grâce à leur régularité permettant d'avoir des résidus répartis régulièrement sur une surface plane.<sup>[103]</sup>

Dans le but de concevoir des ArMs par autoassemblage, il nous paraît donc judicieux de se baser sur le repliement en hélice  $\beta$  gauche de type I, permettant une large palette de fonctionnalisation et des mutations, sans pour autant casser l'assemblage.

Dans la suite de ce manuscrit, les termes « hélice  $\beta$  gauche » et « L $\beta$ H » désigneront une hélice  $\beta$  parallèle gauche de type I.

## 1.7 Les hélices $\beta$ gauche de type I

Les hélices  $\beta$  gauche de type I sont des structures retrouvées principalement dans les bactéries.<sup>[104]</sup> Elles sont le plus souvent retrouvées dans des transférases.<sup>[100]</sup>

L'hélice  $\beta$  gauche de type I, comme montré en Figure 11 est un repliement plus gros que le type II, lui permettant une flexibilité plus importante, qui a pour conséquence de garder l'assemblage solidaire lorsque celui-ci est mis sous contrainte sans causer sa rupture et lui permet aussi d'accommoder des boucles de structures très diverses.<sup>[92,100,105]</sup> Le volume du cœur hydrophobe de ce repliement est relativement important, et est composé de deux résidus par feuillets, donc 6 résidus par tour d'hélice, principalement composé de valines, leucines ou isoleucines aux positions orientées vers l'intérieur, à savoir  $B_1^i$  et  $B_3^i$ . Un schéma montrant les positions des résidus ainsi que leur orientation est montré en Figure 12. Les positions sont nommées  $B_x^{i,o}$  avec  $x = 1,2,3,4$ , « B » faisant référence au feuillet  $\beta$ , « x » correspondant au numéro du résidu dans le feuillet et « i » pour un résidu orienté vers l'intérieur (in) et « o » vers l'extérieur (out). Ceux étant orientés vers l'extérieur sont généralement plus hydrophiles que ceux orientés vers l'intérieur.<sup>[100,106]</sup> La cohésion du repliement est assurée, en majorité, par les feuillets  $\beta$ , avec une moyenne de 13 liaisons hydrogène dues aux interactions inter-feuillets, mais aussi par le cœur hydrophobe.<sup>[100]</sup> La distance entre les feuillets  $\beta$  est la même que celle retrouvée classiquement dans les structures comportant des feuillets

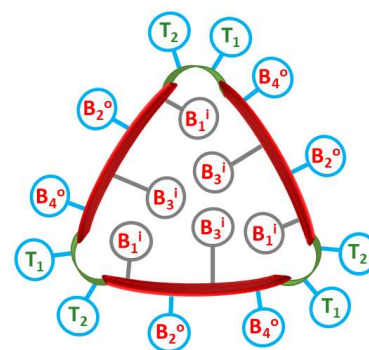


Figure 12 : Représentation schématique d'une L $\beta$ H de type I illustrant l'orientation des différents résidus, participants aux feuillets  $\beta$  en rouge, aux coudes  $\beta$  en vert. En bleu sont montrées les positions orientées vers l'extérieur de l'hélice et en gris vert l'intérieur (cœur hydrophobe).<sup>[100,106]</sup>

que ce soit parallèles ou anti-parallèles, ce qui fait que la structure a une croissance axiale d'environ 4.8 Å par pas d'hélice.

Depuis sa découverte en 1995, peu d'études ont été réalisées sur ce type particulier de repliement. C'est en 2008 que l'équipe de Cohen publie une analyse des séquences et des structures de huit protéines différentes se repliant en hélice  $\beta$  gauche. Pour se faire, les pas d'hélice, de 18 résidus (= 3 hexades), ont été moyennés sur une hexade (Figure 13).<sup>[100,106]</sup>

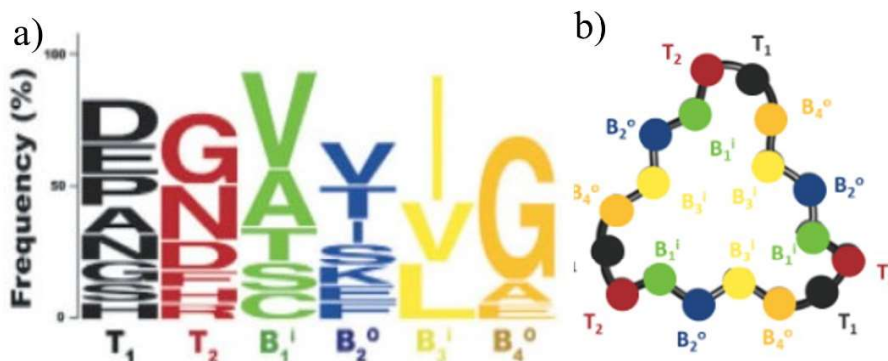


Figure 13 : a) Fréquence d'apparition des résidus en fonction de leurs positions dans la L $\beta$ H ; b) représentation schématique d'un pas d'hélice avec les couleurs correspondantes. Figure adaptée des ref<sup>[100,106]</sup>

Ils ont déterminé la fréquence d'apparition des résidus pour chaque position de l'hélice (Figure 13 a). Cela a permis de mettre en évidence plusieurs choses comme la présence importante de résidus hydrophobes orientés à l'intérieur de l'hélice, mais aussi la prépondérance de la glycine en position B<sub>4</sub>, probablement présente ici dans le but d'apporter de la flexibilité et ne pas encombrer le coude  $\beta$ . En ce qui concerne les autres positions, pointant vers l'extérieur de l'hélice, la distribution des acides aminés est plus diversifiée. La même équipe a aussi publié une étude de compatibilité des différents résidus pour chaque position, en mutant certains résidus de l'hélice  $\beta$  gauche de la protéine E. coli UDP-N-acétylglucosamine acyltransférase (LpxA) (Tableau 2).

Tableau 2 : Tolérance des résidus du domaine L $\beta$ H de la protéine LpxA à partir d'essais de repliement *in vivo*. Résidus chargés (en vert), encombrants (en gris), proline en dehors des positions T<sub>1</sub> et T<sub>2</sub> (en rouge). Tableau adapté de la ref<sup>[106]</sup>.

type	position (LpxA)	résidu connu	autorisé	non autorisé
T <sub>1</sub>	G52, R58	P, D, G, A, E, N, S, H, K, T, Y, M, F, R, Q, I, C	L, V, W	
T <sub>2</sub>	H53, D59	N, G, D, P, F, R, E, K, Q, H, S, Y, T, L, A, M, W, C	I, V	
B <sub>1</sub> <sup>i</sup>	T54, N60	V, A, T, S, C, I, N, L, F	G, Q, M	R, K, W, Y, H, E, D, P
B <sub>2</sub> <sup>o</sup>	K55, E61	V, T, I, S, E, Y, K, F, H, R, M, N, W, L, Q, A, G, C	D	P
B <sub>3</sub> <sup>i</sup>	I56, I62	I, V, L, A, F, T, M, C	G, N, Q, S	R, K, W, Y, H, E, D, P
B <sub>4</sub> <sup>o</sup>	G57, N51	G, E, A, D, H, V, N, S, I, K, R, M, T, L, Y, Q, C, F, W		P



Cette étude fait notamment apparaître le caractère déstabilisateur de la proline dans les positions autres que  $T_1$  et  $T_2$  (coude). Elle montre également l'incompatibilité d'avoir des résidus hydrophobes trop importants (i.e. W, Y) ainsi que des résidus chargés (i.e. H, K, D, E) aux positions  $B_1^i$  et  $B_3^i$  qui empêchent le repliement de l'hélice  $\beta$  gauche. Ces études ont été les premières mettant en évidence la relation séquence-structure pour ce type de repliement, en termes de compatibilité et de fréquences des résidus le composant.<sup>[106]</sup>

Une autre caractéristique structurale intéressante de ce type de  $L\beta H$  est son arrangement en trimère d'hélices, où les trois hélices interagissent latéralement. Ces interactions peuvent se faire au niveau du repliement canonique (surtout), des motifs de terminaison (« bouchons ») ou des boucles présentes le long de l'hélice

(Figure 14). Généralement, les motifs de terminaison en  $N_{ter}$  du motif canonique sont les plus gros, ils préviennent l'agrégation latérale avec d'autres hélices par leur taille conséquente. Les motifs de terminaison peuvent aussi être responsables de la cohésion par interaction de ceux-ci avec les derniers tours de motifs canoniques d'une autre hélice, servant par la même occasion à masquer le cœur hydrophobe. L'assemblage est alors d'autant plus stabilisé qu'il y a des interactions entre les hélices, comme à l'intérieur du trimère d'hélices avec des échelles d'asparagines ou encore d'empilements  $\pi$ .

Prenant comme base une hélice  $\beta$  gauche de type I, des dynamiques moléculaires ont été effectuées afin de savoir s'il était possible de former des autoassemblages stables avec ce repliement. C'est ce que l'équipe de R. Nussinov a réalisé.<sup>[91,107]</sup> Ils ont modélisé et étudié la stabilité de l'assemblage d'un motif correspondant à deux pas d'hélice  $\beta$  (type I) (Figure 15). Avec toutes ces analyses, il en a été conclu que, pour ce type de structure, les glycines et prolines étaient importantes dans les coudes pour maintenir la stabilité de l'assemblage. L'incorporation de résidus non naturels dans les coudes de cette structure permettrait aussi d'augmenter la stabilité de l'assemblage par rigidification de celui-ci.<sup>[107,108]</sup>

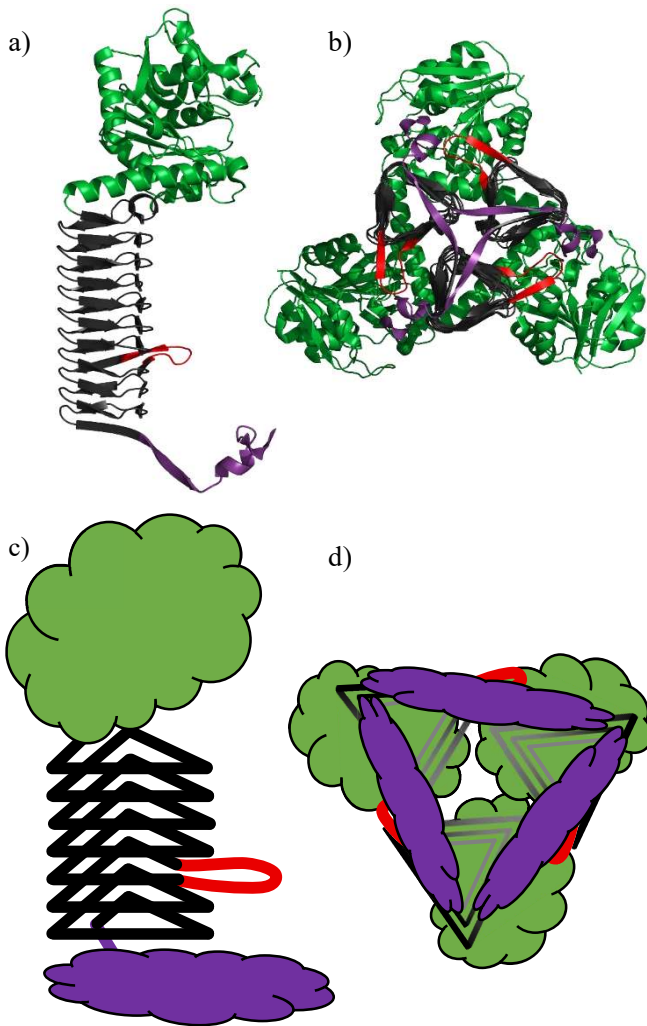


Figure 14 : Exemple d'une structure naturelle d'hélice  $\beta$  gauche de type I (PDB : 1HM9) (haut), et sa représentation schématique (bas). Vu de profil d'un monomère (a, c) et vu du « dessus » du trimère (b, d). Les motifs de terminaison (ou « bouchons ») en  $N_{ter}$  sont montrés en verts et ceux en  $C_{ter}$  en violet, les boucles présentes le long du repliement canonique de l'hélice (noir).

Plus tard, les travaux de D. L. Cox sur des solénoïdes similaires se repliant en hélices  $\beta$  (type II) ont permis de confirmer expérimentalement les dynamiques moléculaires réalisées précédemment sur l'autoassemblage de ces structures. Même lorsque le motif était plus grand (unité de base avec un plus grand nombre de pas d'hélice), des fibres ont été obtenues pour des motifs se repliant en hélices  $\beta$  parallèles gauche.<sup>[96]</sup>

Ces travaux se rapprochent beaucoup de ce qui a été fait par P. L. Davies avec des structures en solénoïdes  $\beta$  mais de type différent. Tous montrent que l'autoassemblage de solénoïdes  $\beta$  est favorable et stable, grâce aux diverses interactions inhérentes à ce repliement. Ceci montre la possibilité d'obtenir des assemblages intéressants par conception *de novo* de séquences se repliant en L $\beta$ Hs. La conception *de novo* se résume à concevoir des séquences basées sur la compréhension de leur repliement. Cela nécessite donc de faire le lien entre la séquence d'une protéine et la structure de celle-ci, i.e. identifier les résidus responsables de tel ou tel repliement. Une conception rationnelle, ou *de novo*, est alors basée sur une dépendance entre la structure, la séquence et la fonction.<sup>[17,109–112]</sup> Par la suite, une fois les résidus essentiels au repliement identifiés, des fonctions (i.e. catalytiques ou biologiques) peuvent être apportées. De plus, de nos jours, la conception de protéines est facilitée par les progrès récents dans le domaine des modélisations moléculaires avec Rosetta et plus récemment AlphaFold.<sup>[113,114]</sup>

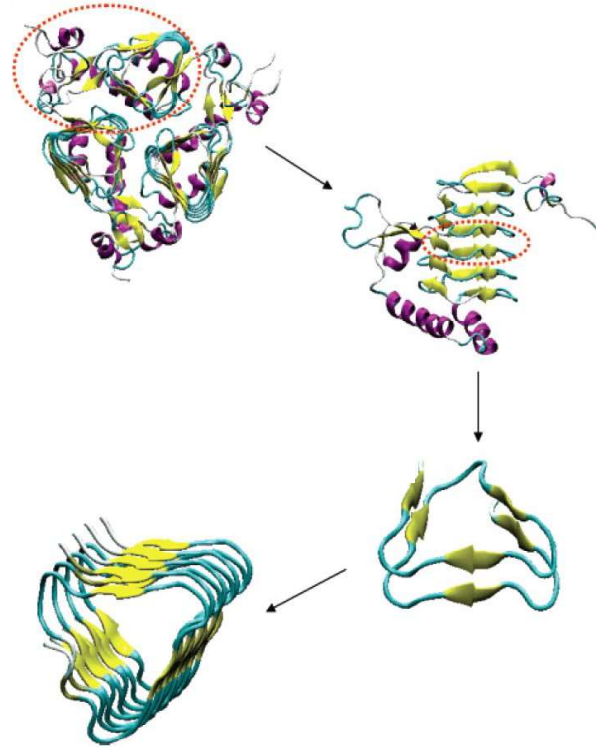


Figure 15 : Illustration schématique de conception de fibres d'hélices  $\beta$  gauche par autoassemblage d'une unité de base. a) structure cristallographique trimérique de la galactoside acétyltransférase (PDB : 1KRR), b) structure monomérique montrant le motif répétitif choisi, c) motif répétitif en solénoïde  $\beta$  composé de deux pas et c) empilement de ces motifs. Figure reprise de la ref<sup>[107]</sup>.

## 1.8 Objectifs de thèse

En se basant sur ces différents travaux, notre but général est de construire une ou des briques de base, pouvant s'assembler en hélice  $\beta$  gauche de type I de manière prédictible et stable, celles-ci pouvant être fonctionnalisées avec divers domaines et/ou mutées à certaines positions. Cela correspond donc aux premiers pas de l'idée proposée schématiquement dans la Figure 5 avec la brique choisie constituée d'un peptide se repliant en hélice  $\beta$  gauche. Cela afin de développer une méthodologie nouvelle et versatile de conception de biomatériaux par autoassemblage et plus particulièrement de métalloenzymes artificielles multisites. Les travaux précédents démontrent une stabilité importante/suffisante pour les assemblages de briques conçues à partir de plusieurs solénoïdes  $\beta$ . De plus, ils montrent que l'hélice  $\beta$  gauche de type I est un repliement stable,

permettant d'accommoder une variété structurale importante (boucles) à certains coudes. C'est alors pour cela que nous avons choisi ce repliement afin de concevoir la brique de base.

Ces travaux de thèse ont pour but d'établir les premières étapes vers le développement de cette méthodologie de conception d'enzymes artificielles multisites. Entre autres, il a fallu concevoir des séquences relativement courtes se repliant en hélice  $\beta$  gauche. Pour cela, mon travail s'est divisé en plusieurs sous-projets.

Le premier projet a visé à établir un outil plus fiable pour la conception de séquences favorables au repliement L $\beta$ H. Les études précédentes se limitant à l'objet monomérique, nous proposons une analyse plus détaillée en prenant en compte que les L $\beta$ Hs existent en trimère dans les protéines naturelles. C'est pour cela que nous avons étudié des structures naturelles de L $\beta$ Hs afin de recenser les différents acides aminés et leurs positions dans l'hélice tout en montrant le rôle dans de la séquence dans la formation trimérique des assemblages et des premiers et derniers tours de chaque hélice. De plus, une étude sur plusieurs dizaines de structures cristallographiques a permis d'obtenir des informations sur les résidus et différents motifs structuraux de ce repliement. Cette analyse a alors permis la construction d'une table rendant compte plus fidèlement des différentes faces et parties des hélices, notamment les premiers et derniers tours ayant pour rôle de nucléer et de terminer le repliement. Elle pourra alors être utilisée comme outil de conception 1) de briques de base de petite taille, pour la formation de fibres trimériques d'hélices  $\beta$  gauche 2) de séquences peptidiques (voire de protéines entières) pour la formation d'assemblages trimériques de L $\beta$ Hs solubles par l'utilisation de motifs de terminaison.

Le second projet a constitué la partie la plus conséquente de mon projet. Il a s'agit ici de concevoir des briques peptidiques de taille minimale pouvant servir de plateforme pour diverses applications. Nous avons pu, à partir de l'analyse précédente, concevoir, synthétiser et caractériser des peptides susceptibles de se replier et s'autoassembler en trimères d'hélices  $\beta$  gauche. Afin de caractériser la structure adoptée par ces peptides, nous avons utilisé une approche multitechnique permettant de combiner et confirmer les résultats de chacune des techniques utilisées. Le but étant de dresser une « carte d'identité » du repliement hélice  $\beta$  gauche, ce repliement n'étant jamais retrouvé seul naturellement du fait des motifs de terminaison et boucles. Nous nous sommes appuyés sur différentes techniques comme les spectroscopies de fluorescence, dichroïsme circulaire et infrarouge ainsi que des techniques de microscopie comme la microscopie électronique à transmission et la microscopie électronique à transmission à basse température.

Enfin, le dernier projet de ma thèse est différent des deux autres. Pour celui-ci, l'objectif principal a été de concevoir des peptides, de séquence minimale afin d'encoder le repliement hélice  $\beta$  gauche. Pour réaliser cela, nous avons conçu des séquences correspondant à un tiers de pas d'hélice, ce que nous avons supposé être le motif minimal. L'hélice  $\beta$  serait donc reconstituée ici par assemblage dans les 3 directions de l'espace. Aussi, et à l'inverse des autres parties, l'assemblage ne pourra pas être limité à la formation d'un trimère. C'est pour cette raison que nous nous attendions à ce que l'assemblage de ces peptides conduise à la formation de cristaux. Par conséquent, ces objets ont été caractérisés par diffraction des rayons X. Cette partie décrit donc

l'optimisation des conditions de cristallisation des peptides menant à la résolution de la structure de certains d'entre eux.

## 1.9 Références bibliographiques

- [1] S. Edition, *Enzymes: Biochemistry, Biotechnology and Clinical Chemistry*, Woodhead Publishing, **2011**.
- [2] G. M. Cooper, *The Eukaryotic Cell Cycle*, **2019**.
- [3] R. Wolfenden, M. J. Snider, *Acc. Chem. Res.* **2001**, *34*, 938–945.
- [4] C. M. Thomas, T. R. Ward, *Chem. Soc. Rev.* **2005**, *34*, 337–346.
- [5] S. Wu, R. Snajdrova, J. C. Moore, K. Baldenius, U. T. Bornscheuer, *Angew. Chemie - Int. Ed.* **2021**, *60*, 88–119.
- [6] J. J. G. M. Maia, M. Isabel, M. Luisa B., Eds., *Enzymes for Solving Humankind's Problems*, Springer Cham, **2021**.
- [7] F. Schwizer, Y. Okamoto, T. Heinisch, Y. Gu, M. M. Pellizzoni, V. Lebrun, R. Reuter, V. Köhler, J. C. Lewis, T. R. Ward, *Chem. Rev.* **2018**, *118*, 142–231.
- [8] E. Fischer, *Berichte der Dtsch. Chem. Gesellschaft* **1894**, *27*, 2985–2993.
- [9] Z. Prokop, A. Gora, J. Brezovsky, R. Chaloupkova, V. Stepankova, J. Damborsky, *Protein Eng. Handb.* **2012**, *3*, 421–464.
- [10] S. AKABORI, S. SAKURAI, Y. IZUMI, Y. FUJII, *Nature* **1956**, *178*, 323–324.
- [11] L. Polgar, M. L. Bender, P. E. Peterson, L. Polgar, M. L. Bender, *J. Am. Chem. Soc.* **1966**, *88*, 3153–3154.
- [12] M. E. Wilson, G. M. Whitesides, R. L. Baughn, Ó. Adalsteinsson, G. M. Whitesides, *J. Am. Chem. Soc.* **1978**, *100*, 306–307.
- [13] K. E. Neet, D. E. Koshland, *Proc. Natl. Acad. Sci. U. S. A.* **1966**, *56*, 1606–1611.
- [14] E. T. Kaiser, D. S. Lawrence, *Science (80-. )*. **1984**, *226*, 505–511.
- [15] K. J. Rodriguez, B. Gajewska, J. Pollard, M. M. Pellizzoni, C. Fodor, N. Bruns, *ACS Macro Lett.* **2018**, *7*, 1111–1119.
- [16] M. Jeschek, R. Reuter, T. Heinisch, C. Trindler, J. Klehr, S. Panke, T. R. Ward, *Nature* **2016**, *537*, 661–665.
- [17] W. Ghattas, J. P. Mahy, M. Réglie, A. J. Simaan, *ChemBioChem* **2021**, *22*, 443–459.
- [18] L. Gao, C. Su, X. Du, R. Wang, S. Chen, Y. Zhou, C. Liu, X. Liu, R. Tian, L. Zhang, et al., *Nat. Chem.* **2020**, *12*, 620–628.
- [19] W. Ghattas, J. P. Mahy, M. Réglie, A. J. Simaan, *ChemBioChem* **2021**, *22*, 443–459.
- [20] P. T. Anastas, J. C. Warner, *Green Chemistry: Theory and Practice*, Oxford University Press, **1998**.
- [21] P. T. Anastas, M. M. Kirchhoff, *Acc. Chem. Res.* **2002**, *35*, 686–694.
- [22] P. Anastas, N. Eghbali, *Chem. Soc. Rev.* **2010**, *39*, 301–312.
- [23] T. Heinisch, T. R. Ward, *Curr. Opin. Chem. Biol.* **2010**, *14*, 184–199.
- [24] W. P. Jencks, *Catalysis in Chemistry and Enzymology*, Dover, **1987**.
- [25] H. Kries, R. Blomberg, D. Hilvert, *Curr. Opin. Chem. Biol.* **2013**, *17*, 221–228.
- [26] F. H. Arnold, *Acc. Chem. Res.* **1998**, *31*, 125–131.
- [27] E. M. Brustad, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2011**, *15*, 201–210.
- [28] R. K. Zhang, D. K. Romney, S. B. J. Kan, F. H. Arnold, *Artif. Met. Met. Catal.* **2018**, 137–170.

- [29] G. Steger, H. Rosenbach, I. Span, J. M. Walker, *DNAzymes: Methods and Protocols*, Springer US, **2022**.
- [30] F. Rosati, G. Roelfes, *ChemCatChem* **2010**, *2*, 916–927.
- [31] G. T. Hermanson, *Bioconjugate Techniques*, Elsevier Science, **2013**.
- [32] L. Pocquet, N. Vologdin, G. F. Mangiatordi, I. Ciofini, O. Nicolotti, S. Thorimbert, M. Salmain, *Eur. J. Inorg. Chem.* **2017**, *2017*, 3622–3634.
- [33] J. Podtetenieff, A. Taglieber, E. Bill, E. J. Reijerse, M. T. Reetz, *Angew. Chemie* **2010**, *122*, 5277–5281.
- [34] A. Fernández-Gacio, A. Codina, J. Fastrez, O. Riant, P. Soumillion, *ChemBioChem* **2006**, *7*, 1013–1016.
- [35] M. Kafri, E. Metzl-Raz, G. Jona, N. Barkai, *Cell Rep.* **2016**, *14*, 22–31.
- [36] W. F. DeGrado, Z. R. Wasserman, J. D. Lear, E. Vincent, W. F. DeGrado, Z. R. Wasserman, J. D. Lear, *Science (80-. )*. **1989**, *243*, 622–628.
- [37] M. S. Ekiz, G. Cinar, M. A. Khalily, M. O. Guler, *Nanotechnology* **2016**, *27*, 402002.
- [38] L. R. Marshall, O. Zozulia, Z. Lengyel-Zhand, I. V. Korendovych, *ACS Catal.* **2019**, *9*, 9265–9275.
- [39] L. R. Marshall, I. V. Korendovych, *Curr. Opin. Chem. Biol.* **2021**, *64*, 145–153.
- [40] J. Greenwald, R. Riek, *Structure* **2010**, *18*, 1244–1260.
- [41] O. Zozulia, M. A. Dolan, I. V. Korendovych, *Chem. Soc. Rev.* **2018**, *47*, 3621–3639.
- [42] C. M. Rufo, Y. S. Moroz, O. V. Moroz, J. Stöhr, T. A. Smith, X. Hu, W. F. Degrado, I. V. Korendovych, *Nat. Chem.* **2014**, *6*, 303–309.
- [43] M. W. West, W. Wang, J. Patterson, J. D. Mancias, J. R. Beasley, M. H. Hecht, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 11211–11216.
- [44] S. Pal, S. Goswami, D. Das, *Chem. Commun.* **2021**, *57*, 7597–7609.
- [45] A. Chatterjee, C. Mahato, D. Das, *Angew. Chemie - Int. Ed.* **2021**, *60*, 202–207.
- [46] J. D. Pham, R. K. Spencer, K. H. Chen, J. S. Nowick, *J. Am. Chem. Soc.* **2014**, *136*, 12682–12690.
- [47] N. L. Truex, Y. Wang, J. S. Nowick, *J. Am. Chem. Soc.* **2016**, *138*, 13882–13890.
- [48] N. L. Truex, J. S. Nowick, *J. Am. Chem. Soc.* **2016**, *138*, 13891–13900.
- [49] P.-N. Cheng, J. D. Pham, J. S. Nowick, *J. Am. Chem. Soc.* **2013**, *135*, 5477–5492.
- [50] R. Gallardo, N. A. Ranson, S. E. Radford, *Curr. Opin. Struct. Biol.* **2020**, *60*, 7–16.
- [51] T. P. J. Knowles, R. Mezzenga, *Adv. Mater.* **2016**, *28*, 6546–6561.
- [52] C. Valéry, F. Artzner, M. Paternostre, *Soft Matter* **2011**, *7*, 9583–9594.
- [53] D. Pramanik, S. G. Dey, *J. Am. Chem. Soc.* **2011**, *133*, 81–87.
- [54] L. Altamura, C. Horvath, S. Rengaraj, A. Rongier, K. Elouarzaki, C. Gondran, A. L. B. Maçon, C. Vendrely, V. Bouchiat, M. Fontecave, et al., *Nat. Chem.* **2017**, *9*, 157–163.
- [55] N. Duraffourg, M. Leprince, S. Crouzy, O. Hamelin, Y. Usson, L. Signor, C. Cavazza, V. Forge, L. Albertin, *Angew. Chemie - Int. Ed.* **2021**, *60*, 14488–14497.
- [56] J. R. Winkler, B. G. Malmström, H. B. Gray, *Biophys. Chem.* **1995**, *54*, 199–209.
- [57] H. Michel, J. Behr, A. Harrenga, A. Kannt, *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 329–356.
- [58] J. M. Palomo, *Curr. Opin. Green Sustain. Chem.* **2021**, *29*, 100452.
- [59] N. Khazanovich, J. R. Granja, D. E. McRee, R. A. Milligan, M. R. Ghadiri, *J. Am. Chem. Soc.* **1994**, *116*, 6011–6012.

- [60] N. Sakai, J. Mareda, S. Matile, *Acc. Chem. Res.* **2008**, *41*, 1354–1365.
- [61] J. Montenegro, M. R. Ghadiri, J. R. Granja, *Acc. Chem. Res.* **2013**, *46*, 2955–2965.
- [62] W. S. Horne, N. Ashkenasy, M. R. Ghadiri, *Chem. - A Eur. J.* **2005**, *11*, 1137–1144.
- [63] L. Adler-Abramovich, E. Gazit, *Chem. Soc. Rev.* **2014**, *43*, 6881–6893.
- [64] A. Wilson, G. Gasparini, S. Matile, *Chem. Soc. Rev.* **2014**, *43*, 1948–1962.
- [65] A. Montero, P. Gastaminza, M. Law, G. Cheng, F. V. V. Chisari, M. R. R. Ghadiri, *Chem. Biol.* **2011**, *18*, 1453–1462.
- [66] Y. Chen, Y. Yang, A. A. Orr, P. Makam, B. Redko, E. Haimov, Y. Wang, L. J. W. Shimon, S. Rencus-Lazar, M. Ju, et al., *Angew. Chemie Int. Ed.* **2021**, 17164–17170.
- [67] B. Kobe, A. V. Kajava, *Trend Biochem. Sci.* **2000**, *25*, 509–515.
- [68] L. Falquet, M. Pagni, P. Bucher, N. Hulo, C. J. A. Sigrist, K. Hofmann, A. Bairoch, *Nucleic Acids Res.* **1999**, *30*, 235–238.
- [69] A. V. Kajava, *J. Struct. Biol.* **2012**, *179*, 279–288.
- [70] K. Basu, R. L. Campbell, S. Guo, T. Sun, P. L. Davies, *Protein Sci.* **2016**, *25*, 946–958.
- [71] A. J. Scotter, M. Guo, M. M. Tomczak, M. E. Daley, R. L. Campbell, R. J. Oko, D. A. Bateman, A. Chakrabarty, B. D. Sykes, P. L. Davies, *BMC Struct. Biol.* **2007**, *7*, 63.
- [72] L. A. Graham, W. Qin, S. C. Loughheed, P. L. Davies, V. K. Walker, *J. Mol. Evol.* **2007**, *64*, 387–398.
- [73] A. V. Kajava, A. C. Steven, in *Adv. Protein Chem.*, Elsevier, **2006**, pp. 55–96.
- [74] U. Baumann, S. Wu, K. M. Flaherty, D. B. McKay, *EMBO J.* **1993**, *12*, 3357–3364.
- [75] M. D. Yoder, S. E. Lietzke, F. Jurnak, *Structure* **1993**, *1*, 241–251.
- [76] J. T. MacDonald, B. V. Kabasakal, D. Godding, S. Kraatz, L. Henderson, J. Barber, P. S. Freemont, J. W. Murray, *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 10346–10351.
- [77] W. R. P. Novak, B. Bhattacharyya, D. P. Grilley, T. M. Weaver, *Acta Crystallogr. Sect. F* **2017**, *73*, 138–145.
- [78] Y. Xiang, P. G. Leiman, L. Li, S. Grimes, D. L. Anderson, M. G. Rossmann, *Mol. Cell* **2009**, *34*, 375–386.
- [79] A. J. Middleton, C. B. Marshall, F. Faucher, M. Bar-Dolev, I. Braslavsky, R. L. Campbell, V. K. Walker, P. L. Davies, *J. Mol. Biol.* **2012**, *416*, 713–724.
- [80] Q. Ye, R. Eves, R. L. Campbell, P. L. Davies, *Biochem. J.* **2020**, *477*, 3271–3286.
- [81] S. R. Herron, R. D. Scavetta, M. Garrett, M. Legner, F. Jurnak, *J. Biol. Chem.* **2003**, *278*, 12271–12277.
- [82] J. Badger, B. Chie-Leon, C. Logan, V. Sridhar, B. Sankaran, P. H. Zwart, V. Nienaber, *Acta Crystallogr. Sect. F* **2011**, *67*, 749–752.
- [83] P. L. Davies, *Trends Biochem. Sci.* **2014**, *39*, 548–555.
- [84] M. Banach, L. Konieczny, I. Roterman, *Biochimie* **2018**, *144*, 74–84.
- [85] M. Banach, I. Roterman, *Solenoid - an Amyloid under Control*, Elsevier Ltd, **2019**.
- [86] J. G. Ferry, *Biochim. Biophys. Acta - Proteins Proteomics* **2010**, *1804*, 374–381.
- [87] K. M. Pan, M. Baldwin, J. Nguyen, M. Gasset, A. Serban, D. Groth, I. Mehlhorn, Z. Huang, R. J. Fletterick, F. E. Cohen, et al., *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 10962–10966.
- [88] F. E. Cohen, S. B. Prusiner, *Annu. Rev. Biochem.* **1998**, *67*, 793–819.
- [89] T. N. Petersen, S. Kauppinen, S. Larsen, *Structure* **1997**, *5*, 533–544.
- [90] P. Iengar, N. V. Joshi, P. Balaram, *Structure* **2006**, *14*, 529–542.

- [91] N. Haspel, D. Zanuy, C. Alemán, H. Wolfson, R. Nussinov, *Structure* **2006**, *14*, 1137–1148.
- [92] A. S. Parker, K. M. Ravikumar, D. L. Cox, *arXiv:1706.05451 [q-bio]* **2017**.
- [93] L. P. Heinz, K. M. Ravikumar, D. L. Cox, *Nano Lett.* **2015**, *15*, 3035–3040.
- [94] C. Govaerts, H. Wille, S. B. Prusiner, F. E. Cohen, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 8342–8347.
- [95] A. Parker, K. Ravikumar, D. Cox, *Soft Matter* **2017**, *13*, 6218–6226.
- [96] M. D. R. R. Peralta, A. Karsai, A. Ngo, C. Sierra, K. T. Fong, N. R. Hayre, N. Mirzaee, K. M. Ravikumar, A. J. Kluber, X. Chen, et al., *ACS Nano* **2015**, *9*, 449–463.
- [97] Z. Peng, M. D. R. Peralta, M. D. Toney, *Biochemistry* **2017**, *56*, 6041–6050.
- [98] C. R. H. Raetz, S. L. Roderick, *Science (80-. )*. **1995**, *270*, 997–1000.
- [99] J. Jenkins, R. Pickersgill, *Prog. Biophys. Mol. Biol.* **2001**, *77*, 111–175.
- [100] J. H. Choi, C. Govaerts, B. C. H. H. May, F. E. Cohen, *Proteins Struct. Funct. Bioinforma.* **2008**, *73*, 150–160.
- [101] X. Deng, A. Gonzalez Llamazares, J. M. Wagstaff, V. L. Hale, G. Cannone, S. H. McLaughlin, D. Kureisaite-Ciziene, J. Löwe, *Nat. Microbiol.* **2019**, *4*, 2357–2368.
- [102] J. P. M. Langedijk, G. Fuentes, R. Boshuizen, A. M. J. J. Bonvin, *J. Mol. Biol.* **2006**, *360*, 907–920.
- [103] E. K. Leinala, P. L. Davies, Z. Jia, *Structure* **2002**, *10*, 619–627.
- [104] J. Jenkins, R. Pickersgill, *Mol. Biol.* **2001**, 65.
- [105] L. P. Heinz, K. M. Ravikumar, D. L. Cox, *Nano Lett.* **n.d.**, 8.
- [106] J. H. Choi, B. C. H. H. May, C. Govaerts, F. E. Cohen, *Structure* **2009**, *17*, 1014–1023.
- [107] J. Zheng, D. Zanuy, N. Haspel, C.-J. J. Tsai, C. Alemán, R. Nussinov, *Biochemistry* **2007**, *46*, 1205–1218.
- [108] G. Ballano, D. Zanuy, A. I. Jiménez, C. Cativiela, R. Nussinov, C. Alemán, *J. Phys. Chem. B* **2008**, *112*, 13101–13115.
- [109] Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall, *Nature* **2009**, *460*, 855–862.
- [110] R. A. Sheldon, J. M. Woodley, *Chem. Rev.* **2018**, *118*, 801–838.
- [111] I. V. Korendovych, W. F. DeGrado, *Q. Rev. Biophys.* **2020**, *53*, e3.
- [112] H. Shen, J. A. Fallas, E. Lynch, W. Sheffler, B. Parry, N. Jannetty, J. Decarreau, M. Wagenbach, J. J. Vicente, J. Chen, et al., *Science (80-. )*. **2018**, *362*, 705–709.
- [113] C. A. Rohl, C. E. M. Strauss, K. M. S. Misura, D. Baker, *Methods Enzymol.* **2004**, *383*, 66–93.
- [114] A. W. Senior, R. Evans, J. Jumper, J. Kirkpatrick, L. Sifre, T. Green, C. Qin, A. Židek, A. W. R. Nelson, A. Bridgland, et al., *Nature* **2020**, *577*, 706–710.





## **2 Analyse bio-informatique des structures en hélice $\beta$ gauche naturelles**

## 2.1 Introduction

### 2.1.1 Limites des études précédentes

Comme présenté dans l'Introduction, l'équipe de F. Cohen a publié deux articles étudiant la relation séquence-structure pour les hélices  $\beta$  gauche de type I (L $\beta$ Hs).<sup>[1,2]</sup> Ce fut un pas important dans la compréhension des L $\beta$ Hs mais ils sont limitées car plusieurs aspects de ce repliement ont été ignorés :

-Le premier est que ces hélices (de type I) sont des domaines de trimérisation, ce qui veut dire que chaque face de l'hélice n'est pas exposée au même environnement. La différenciation des faces a été mentionnée auparavant par d'autres que l'équipe de Cohen, mais sans réaliser de séquence consensus sur ce type de structure.<sup>[3]</sup>

-Le deuxième est que tous les pas d'hélice ne sont pas équivalents. En effet, les tours situés aux extrémités sont bien différents de ceux retrouvés dans le corps de l'hélice. Ils ont des rôles de nucléation et des rôles de terminaison, avec des motifs particuliers. Ces motifs ne sont pas retrouvés dans les pas du milieu de l'hélice.

Aussi, dans les travaux de Cohen, ne pas avoir fait ces distinctions a amené à gommer ces différences, les noyant dans une « hexade moyenne » (Figure 13).

Dans l'optique de concevoir des peptides se repliant en hélice  $\beta$  gauche trimérique, de taille infinie (i.e. nanofils) ou finie (i.e. structure soluble) nous avons besoin de connaître le rôle des A.As présents dans l'hélice et leurs fréquences, à la lumière du caractère trimérique de ces objets ainsi que des différents rôles joués par les différents tours de l'hélice. Cela n'étant pas décrit fidèlement dans la littérature, nous nous sommes mis à construire un outil de conception d'hélices  $\beta$  gauche de type I, plus précis et plus fidèle que les études déjà réalisées.

## 2.2 Construction de l'outil

La construction de ce nouvel outil pour la conception peptides/protéines se repliant en hélices  $\beta$  gauche est illustrée en Figure 16. Elle s'est faite similairement à l'analyse qui a été réalisée par Cohen et collaborateurs, mais avec différenciation des faces (Figure 16 c) et parties (Figure 16 d) de l'hélice. Il en résulte un tableau de fréquences d'apparition des résidus suivant les positions (Figure 16 e), pour les 9 (3 faces x 3 parties) hexades distinguées.<sup>[1]</sup> Dans l'optique de développer un outil se basant sur la méthodologie de Davies (identification de motifs structuraux répétitifs, à partir de séquences consensus, pour former des assemblages) l'étude permettra de différencier les parties canoniques et celles qui le sont moins.<sup>[4]</sup>

La première étape a été la recherche de toutes les structures d'hélices  $\beta$  gauche de Type I présentes sur la Protein Data Bank (PDB) (Figure 16 a). Nous avons ensuite identifié les parties canoniques des protéines correspondantes à la L $\beta$ H (Figure 16 b). Les différentes hexades ont ensuite été identifiées (I, II, III), de même pour les différents tours (N, R, C) (Figure 16 c et d). Les séquences ont ensuite été reportées dans un tableau, séparées suivant les hexades et tours et analysées pour donner un tableau 3x3 (Figure 16 e).

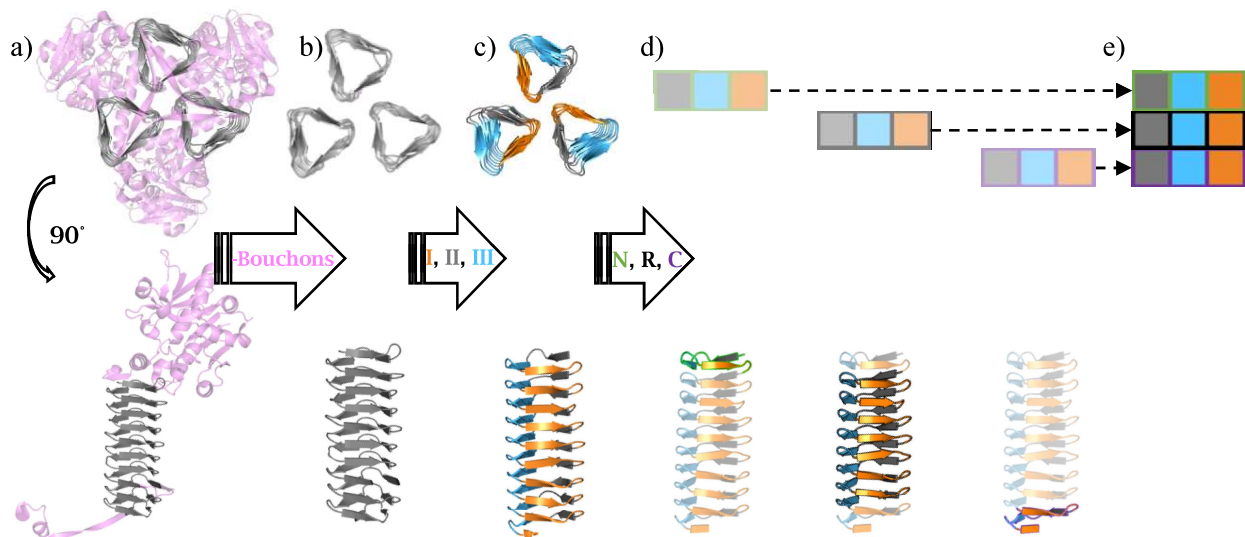


Figure 16 : Schéma synthétique de la méthodologie appliquée pour construire l'outil. a) Protéine naturelle composée d'un trimère d'hélices (gris) et de motifs de terminaison et boucles (rose) ; b) Trimère d'hélices seul ; c) Discrimination des faces selon l'arrangement trimérique ; d) différenciation des parties des hélices et représentations de construction de la matrice 3x3.

## 2.2.1 Recherche de structures PDB

La recherche de structures PDB correspondant aux hélices  $\beta$  gauche n'est pas triviale. Partant de structures ayant un repliement en L $\beta$ H, une recherche par similarité de structure sur la PDB a été réalisée. Cela nous a permis d'identifier une quarantaine de structures avec un repliement en L $\beta$ H (Annexe 1). La recherche par mot clé s'est avérée très utile car la majorité de ces structures sont celles d'UDP N-Acétyletransférases. La recherche de structure a ensuite été grandement aidée par la publication d'une banque de données regroupant tous les motifs structuraux répétitifs retrouvés dans la PDB.<sup>[5]</sup> Les tables de données par structures peuvent être téléchargées sous format .xlsx permettant le traitement par la suite. Cependant, la PDB et la banque de données regroupent tout par code PDB et il a donc fallu trier toutes ces structures par code Uniprot afin de s'assurer que les protéines sont différentes. Cela a permis de passer de 344 structures à 86, ce qui a très grandement simplifié l'étude. Toutes les structures n'étant pas des hélices  $\beta$  gauche trimériques ont été mises de côté. Parmi les 86 structures référencées, 19 d'entre elles ne sont pas soit trimériques, soit de section triangulaire. Il y a beaucoup de solénoïdes  $\beta$  mais aussi des structures non organisées en solénoïdes  $\beta$ . Après avoir trié et identifié les structures d'intérêt, nous avons commencé à mettre en place une méthodologie de travail afin de pouvoir construire le nouvel outil se présentant sous forme de matrice 3 x 3. La recherche de structure a été stoppée le

15/08/2021, ce qui veut dire que toutes les protéines reportées ont été mises en ligne avant cette date. Il est à noter qu'à la date du 30/06/2022 seulement 2 autres structures différentes de celles précédentes ont été déposées sur la PDB correspondant à un repliement en hélice  $\beta$  gauche de type I. Ce travail a pris énormément de temps puisqu'il a été réalisé sur 78 structures différentes, ce qui correspond à un peu plus de 10000 résidus, triés et reportés « à la main ».

## 2.3 Restriction de l'analyse au pur repliement hélice $\beta$ gauche de type I

Les structures disponibles dans la Protein Data Bank sont celles de protéines cristallisées contenant des hélices  $\beta$  gauche mais pas seulement. Il ne s'agit pas d'hélices pures, elles sont en réalité composées de motifs de terminaison (bouchons) en  $N_{\text{ter}}$  et en  $C_{\text{ter}}$  présents naturellement afin de stopper la croissance de l'hélice, et parfois d'autres domaines encore (i.e. domaines globulaires). Mais ce n'est pas tout, il y a aussi des boucles, dans le corps l'hélice, qui ne sont pas canoniques. Toutes ces parties de la protéine n'ont donc pas été prises en compte lors de l'étude, cette dernière se focalisant seulement sur le repliement hélice  $\beta$  gauche de Type I (Figure 16 et Figure 14). Nous avons de même observé des parties de l'hélice qui comprennent des ajouts de résidus (en moyenne 2) par rapport au tour de référence de 18 A.As et ne formant pas de boucles, idem pour les suppressions d'acides aminés. Lors de cette étape, il est important d'appliquer toujours la même méthodologie concernant ce qui est canonique ou pas. Le postulat de départ est que tous les résidus hors des positions canoniques de l'hélice ne font pas partie du repliement canonique. En d'autres termes, les acides aminés ne formant pas de feuilletts ou coudes  $\beta$  canoniques n'ont pas été pris en compte, ce qui permet d'éliminer facilement les chaînes peptidiques de début et de fin de boucle par exemple. L'étape suivante a alors été l'identification des différentes faces ou hexades en fonction de leur position dans le trimère. Une hexade est définie ici comme étant assimilée à une face car chacune est composée de six résidus. Le trimère est symétrique, et est un assemblage de trois hélices identiques qui sont-elles même asymétriques, c'est-à-dire que leurs faces ne sont pas équivalentes.

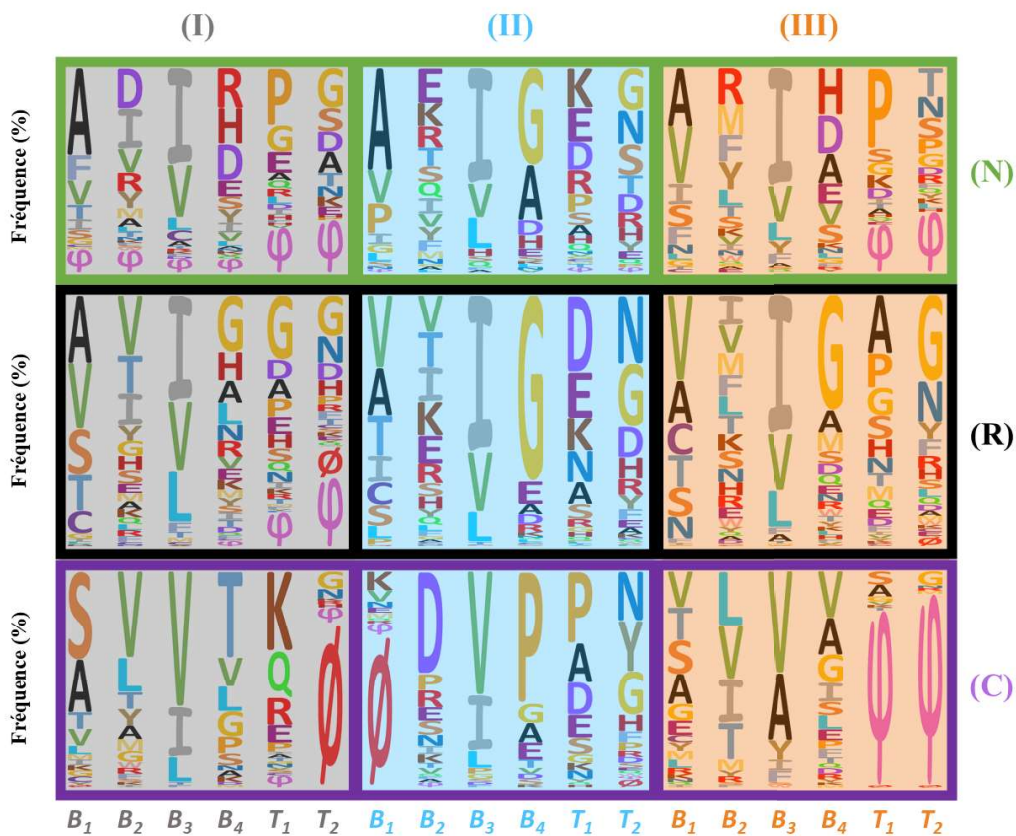
Ces faces (ainsi que leur hexade correspondante :  $B_1, B_2, B_3, B_4, T_1, T_2$  sans leurs exposants) ont été colorées et nommées (I), (II) et (III) (Figure 16 c). L'hexade (II) fait référence à la face la plus exposée au solvant, donc, où les résidus les plus hydrophiles seront retrouvés. La (III) correspond à l'hexade dont les positions  $T_1$  et  $T_2$  forment le coude étant au centre du trimère, cette face pouvant interagir avec les résidus de la (I). De ce fait, l'hexade (I) est la moins exposée au solvant, avec des A.As faisant front à la (III) d'une autre hélice.

Enfin, chacune des hélices étudiées a été décomposée en 3 parties distinctes (N), (R) et (C) comme montrées en Figure 16 d. Cette séparation permet de mettre en valeur les différents motifs structuraux associés à chaque partie. Ce qui nous a donné accès à l'identification de motifs (enchaînements d'acides aminés) favorisant (i) l'initiation de l'hélice, pour le tour (N) (premier tour de l'hélice après le  $N_{\text{ter}}$ ), (ii) la stabilisation pour les tours

(R) et (iii) stoppant l'hélice pour le tour (C) (dernier tour avant le C<sub>ter</sub>). L'identification de ces parties est importante concernant la conception du tableau comme nouvel outil pouvant servir à la conception de structures se repliant en hélices  $\beta$  gauche. En effet, grâce à celles-ci, la taille des objets conçus pourrait être contrôlée par simple ajout d'un tour (N) et (C) pouvant mener à la formation d'objets solubles de taille finie ou non. Les protéines naturelles retrouvées dans la PDB ont un nombre de tours limité grâce à ces parties (de 4 à 11 au maximum avec une moyenne de 7 tours, 6 étant le nombre de tours le plus fréquent).

Pour chaque structure identifiée, le fichier .pdb a été téléchargé et visualisé avec Pymol.<sup>[6]</sup> La visualisation de ces protéines est primordiale dans le but de pouvoir identifier précisément les hélices et surtout les positions des résidus, tout en omettant les boucles et motifs globulaires des protéines naturelles. Nous avons créé un tableur dans lequel les acides aminés sont répertoriés suivant leur position et la partie de l'hélice auxquelles ils appartiennent (i.e. matrice 3x3 d'hexades). Il en résulte que pour chaque position (3 types de tours x 3 côtés x 6 résidus = 54 positions), le pourcentage de présence de chaque acide aminé est déterminé (Tableau 3). Un tableau récapitulatif des structures étudiées est montré en Annexe 1.

Tableau 3 : Matrice 3x3 représentant l'analyse effectuée sur les différentes protéines où est représentée la fréquence de présence des résidus en fonction des tours, des faces et des positions. Les lettres correspondent aux résidus, Ø fait référence à des positions manquantes alors que  $\phi$  se rapporte à des résidus participants aux boucles.



Quant aux formes trimériques des hélices  $\beta$  gauche, les faces ne sont pas équivalentes, notamment pour les résidus étant orientés vers l'extérieur de chaque hélice. Cela peut se comprendre par le fait que l'exposition au solvant y est différente, ce qui est également le cas pour d'autres chaînes peptidiques comme une hélice différente, un bouchon ou encore d'autres molécules. Cependant, les chaînes latérales étant orientées vers l'intérieur de l'hélice, les résidus participant au cœur hydrophobe restent immuables et sont comparables à l'étude précédente.<sup>[1]</sup>

### 2.3.1 Le cœur hydrophobe

Étant commun à toutes les hexades, le cœur hydrophobe est la seconde interaction la plus forte maintenant la cohésion de l'hélice et lui est essentiel.<sup>[1,3,7,8]</sup> Il est composé de résidus présents aux positions  $B_1^i$  et  $B_3^i$ , qui sont tournées vers l'intérieur de l'hélice (Figure 17). De plus, l'hélice étant extrêmement régulière en termes de structure, les résidus composant cet intérieur ne peuvent varier que très peu. Il s'agit ici donc des résidus les plus conservés dans ce type de structure, surtout dans les tours canoniques (R). Évidemment, il existe des différences entre les deux positions, qui sont principalement dues à un encombrement stérique distinct. En position  $B_1^i$ , les résidus les plus fréquemment retrouvés sont des alanines, valines, sérines et thréonines, l'encombrement stérique limitant la taille des résidus. Les sérines et thréonines participent à la cohésion par formation d'une liaison hydrogène avec la chaîne principale d'un autre résidu situé en face ( $B_4^o$ ) (Figure 18 a). Nous pouvons aussi noter la présence non négligeable d'une cystéine à cette position, pouvant faire cette même liaison hydrogène. Il est cependant possible de retrouver à cette position des résidus plus gros tels que des leucines, isoleucines ou encore phénylalanine, mais cela est rare, notamment pour la phénylalanine qui n'est retrouvée que dans la première partie de l'hélice (N). Comparé aux autres tours, le premier est le moins contraint (puisqu'exposé au solvant), et des résidus plus gros y sont retrouvés afin d'apporter une contribution non négligeable au cœur hydrophobe.

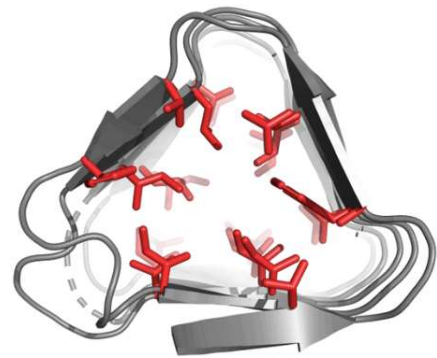


Figure 17: Exemple d'une hélice  $\beta$  gauche mettant en évidence les résidus du cœur hydrophobe en rouge.

Une chose surprenante que l'on peut y remarquer est la possibilité de rencontrer une proline en  $B_1^i$  qui participe au cœur hydrophobe de l'hélice. Cela se produit lorsqu'il y a une boucle juste avant (Figure 18 b). Un peu plus courante est la présence d'asparagines à cette position (tours R) consolidant l'assemblage par formation d'échelles de liaisons H et par interactions avec le résidu situé en face en  $B_4^o$  (Figure 18 c). Toutes ces interactions  $B_1^i$ - $B_4^o$  consolident grandement le coude  $\beta$ , et *ipso facto* le repliement entier.



Figure 18 : Résidus aux positions  $B_1^i$  illustrant a) la contribution des thréonines, sérines et cystéines à la formation de liaisons hydrogène avec la chaîne principale du résidu en  $B_4^o$  en face (PDB : 3IXC) ; b) la proline participant au cœur hydrophobe juste après une boucle (en rose) (PDB : 1XHD) ; c) une échelle d'asparagine interne (position  $B_1$ ) interagissant comme les résidus en a) (PDB : 4EQY).

En ce qui concerne la position  $B_3^i$ , les résidus les plus fréquemment retrouvés sont l'isoleucine, la valine et la leucine. L'isoleucine est présente en majorité dans les tours (N) et (R) et non dans le tour (C), ce qui peut s'expliquer par le fait que le dernier tour a souvent une section plus petite que les autres tours (à cause d'As manquants dans les coudes), ne permettant pas d'accommoder d'isoleucines. Dans ce cas-ci, la valine est préférée. Néanmoins, la fréquence des isoleucines reportée dans le Tableau 4 ne reflète pas réellement ce qu'il se passe dans le cœur hydrophobe. En effet, il est possible de trouver des ensembles de trois isoleucines dans le même tour, mais cela reste relativement rare (6.7% pour (N), 16.3% pour (R) et 0% pour (C)). La raison pourrait être la stabilité qui ne serait pas optimale lorsque 3 isoleucines se font face dans le même tour. C'est pour cela qu'en général, le motif retrouvé est de 1-2 Ile pour 2-1 Val (ou leucine). Il n'est aussi pas retrouvé, ou très peu, d'empilement d'isoleucines pour une même position entre différents tours. 22 % des tours (C) sont composés de trois Val ; signe que ce dernier tour est composé d'un cœur hydrophobe moins important que les autres.

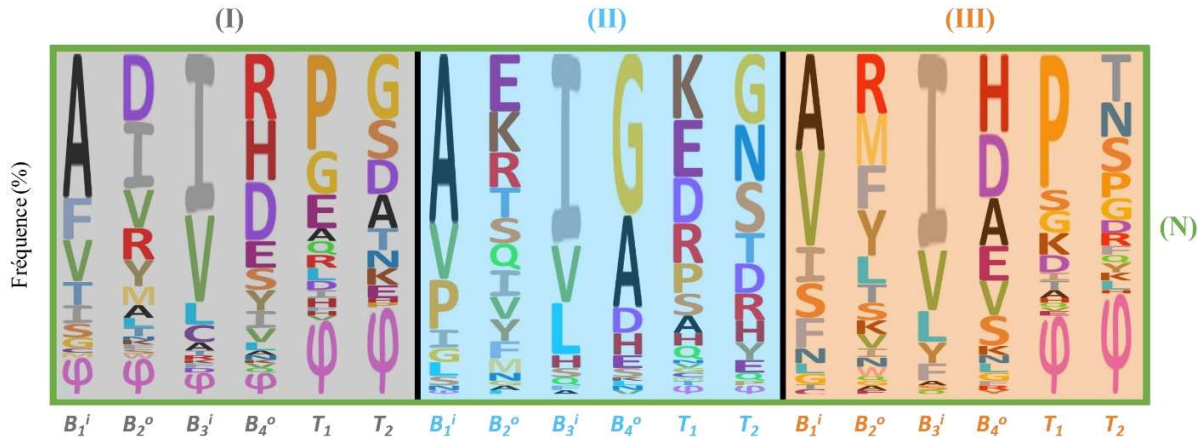
Alors que les résidus orientés vers l'intérieur de l'hélice sont très conservés, ceux orientés vers l'extérieur sont très différents, et varient selon les positions, les faces mais aussi les types de tours. Nous traiterons ces positions selon leurs appartenances aux différentes parties de l'hélice car cela permet de mettre en lumière les différents motifs identifiés en fonction des parties, les faces pouvant être directement comparées entre-elles.

## 2.3.2 Les tours

Comme dit précédemment, l'identification des différentes parties des hélices nous permettra de mieux comprendre les interactions en jeu lors de la formation de  $L\beta$ Hs et s'en servir pour la conception rationnelle de protéines adoptant ce repliement. Nous pourrions notamment déterminer des motifs, favorisant ce repliement, mais aussi ceux qui le défavorisent.

### 2.3.2.1 Les tours de nucléation (N)

Tableau 4 : Fréquence d'observation des acides aminés suivant les positions retrouvées sur le tour de nucléation. Les caractères  $\phi$  indiquent que les résidus à cette position adoptent une conformation non-canonique.



Le premier tour, en N-ter de l'hélice, fait le lien entre le bouchon en N-ter de cette structure et le reste de l'hélice. Il est alors très important que ce tour soit bien défini. Ensuite, sur les structures, nous remarquons que ce tour admet une courbure (perpendiculaire aux feuillet) plus importante au niveau de ses coudes  $\beta$  que les autres parties (visible Figure 16), qui sont relativement plates et parallèles aux feuillet  $\beta$ . De plus, la fréquence des prolines dans ces tours est plus importante que dans les autres tours, aux positions T<sub>1</sub>.<sup>[1]</sup> Enfin, des motifs spécifiques renforçant la cohésion des coudes  $\beta$  par interactions avec des résidus de la chaîne principale du tour suivant sont observés (voir ci-dessus). Pour toutes ces considérations, ce premier tour est très important, et nous l'appellerons « tour de nucléation » par la suite.

#### 2.3.2.1.1 Identification d'un motif de 3 résidus comme motif de nucléation

Il a été reporté de multiples fois dans la littérature que la proline est une inductrice de coude car elle induit un angle dans la chaîne principale, du fait de sa structure cyclique. En position T<sub>1</sub>, la conformation *cis* de cette dernière permet de former un coude  $\beta$  qui a pour avantage de placer les positions B<sub>4</sub><sup>o</sup> et B<sub>1</sub><sup>i</sup> en face. Cela permettant la formation d'une liaison hydrogène entre chaîne principale et chaîne latérale des acides aminés qui y sont respectivement présents, stabilisant la formation du coude. En plus de former des liaisons hydrogène inter-coudes, celles-ci s'orientent de façon à former une interaction avec le coude du tour suivant, stabilisant ainsi la formation de l'hélice. Ces motifs ont pour séquence générique X-P-Z avec X = D ou H, et Z = S, T, D. Ils permettent d'induire la formation de coudes  $\beta$  tout en formant des interactions avec le tour en n+1 comme nous pouvons le voir en Figure 19.



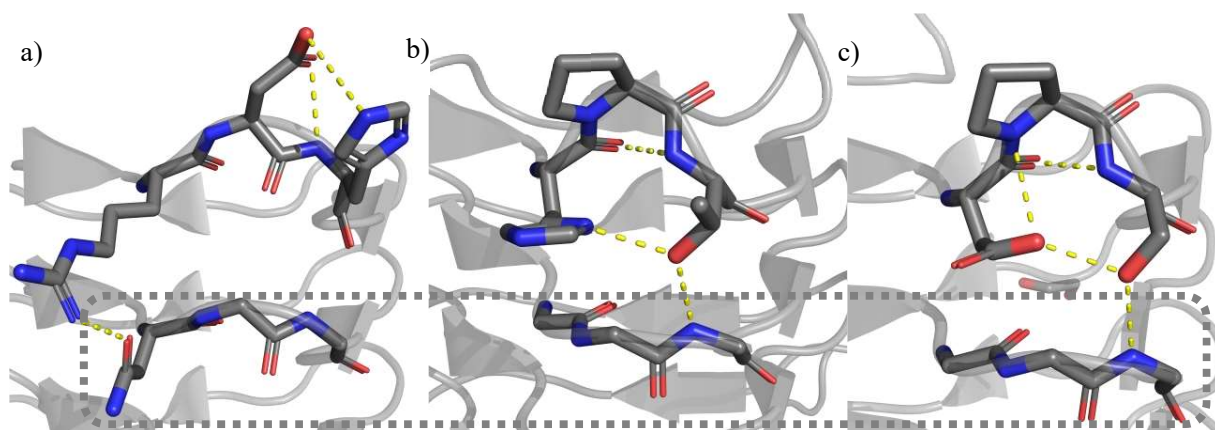


Figure 19: Exemples de motifs retrouvés en tour (N) renforçant la formation de l'hélice avec a) RDH (PDB : 3R8Y), b) HPT (PDB : 3HSQ) et c) DPS (PDB : 4E6U). En pointillés sont encadrés les résidus faisant partie du tour n+1.

Ces motifs sont surtout présents dans les faces (I) et (III) comme nous pouvons le voir sur le Tableau 4. En ce qui concerne la face (II) qui est la plus exposée à l'eau, ces motifs ne sont retrouvés qu'en très faibles proportions. Une explication pourrait être que ces résidus sont ceux qui interagissent avec l'eau afin de garder l'assemblage soluble. C'est d'ailleurs l'hexade où la lysine est la plus présente en  $T_1$ .

Nous n'avons pas observé de tours (N) avec trois prolines en  $T_1$ , mais aussi que très peu de tours où les prolines sont présentes au nombre de deux (quatre structures, 5%). Nous pouvons cependant noter la présence en très large majorité d'au moins une proline dans ce premier tour dans 56 structures, soit 78 % des tours (N) étudiés. Il reste néanmoins 22 structures où aucune proline n'est présente dans ce tour (N) à la position  $T_1$ , essentiellement dû à la présence de chaînes non-canoniques, commençant très souvent en  $T_1$  (dans 47 % des cas pour le tour (N) ou pour les autres).

### 2.3.2.1.2 La flexibilité des coudes $\beta$

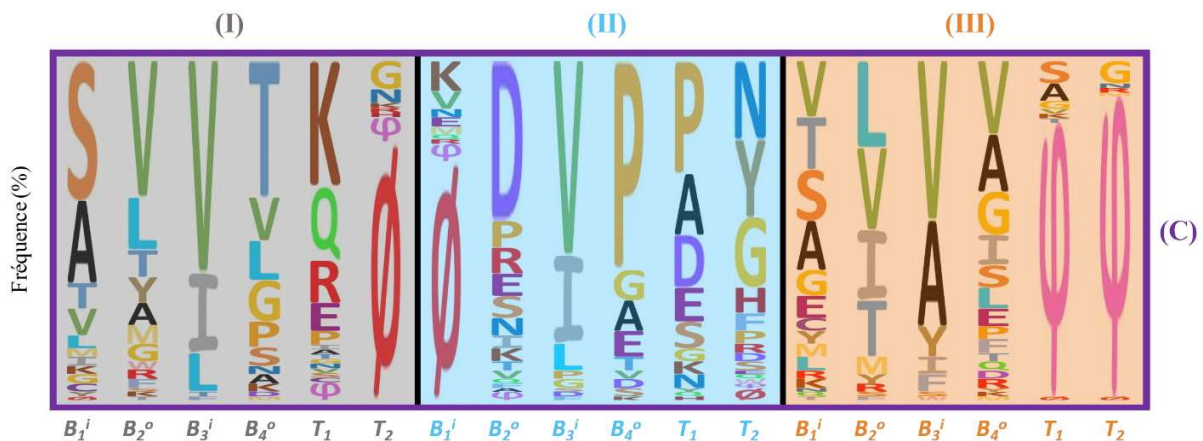
Un peu à l'opposé de la contrainte qu'apporte la proline en position  $T_1$ , des glycines sont, elles, responsables de la flexibilité de la chaîne, à cause d'une absence de chaîne latérale. Elle est plus particulièrement présente aux positions  $B_4^\circ$ ,  $T_1$ ,  $T_2$  qui sont les positions les moins contraintes à adopter un repliement en feuillets  $\beta$  canoniques et ne contribuant pas au cœur hydrophobe. Il est donc logique de la retrouver dans ces positions, ce qui est particulièrement vrai pour les tours canoniques (R). Mais comme nous pouvons le voir dans le Tableau 4, leur présence est bien moins importante. Cela peut être expliqué par le fait que le premier tour (N) doit avoir un degré de liberté limité, c'est-à-dire que, dans ce tour, ce qui est important c'est la stabilité. Elle doit être relativement importante afin de pouvoir amorcer le repliement. C'est pour ces raisons qu'à la place de glycines, nous retrouvons des résidus capables de faire des interactions chaînes principales-chaînes latérales. Il s'agit ici de ce que nous appelons des motifs de nucléations.

### 2.3.2.1.3 La position $B_2^o$ est la plus variable en fonction des faces

En revanche, la position  $B_2^o$  admet beaucoup plus de liberté et accommode préférentiellement des résidus assez imposants. Mais encore une fois, et logiquement, la face (II) est celle qui porte le plus d'acides aminés chargés et polaires à cette position. Pour les deux autres faces, nous observons des résidus hydrophiles et hydrophobes en fonction de l'environnement de ceux-ci. En effet, dans les motifs formant le bouchon, les boucles sont présentes et sont souvent imposantes (domaines globulaires), ce qui les amènent à interagir avec les résidus tournés vers l'extérieur de l'hélice et plus particulièrement  $B_2^o$ . Nous avons choisi, dans notre analyse, de ne pas prendre en compte les parties autres que celle de l'hélice  $\beta$  gauche, et donc les interactions dans lesquelles elles sont impliquées ne sont pas détaillées ici.

### 2.3.2.2 Les tours de terminaison (C)

Tableau 5 : Fréquence d'observation des acides aminés suivant les positions identifiées dans le tour de terminaison (C). Les suppressions de résidus sont ici montrées en  $\emptyset$  et les  $\phi$  indiquent que les résidus à cette position adoptent une conformation non-canonique.



Les tours de terminaison (C) ont pour rôle d'achever le repliement, c'est pour cela que nous allons y retrouver des motifs structuraux déstabilisants. Cela sert alors à prévenir les interactions avec d'autres tours  $n+1$  hypothétiques. Ce qui veut dire que dans ces dernières séquences, les motifs qui y sont présents sont bien différents des autres parties. L'enjeu est d'empêcher la continuation de l'hélice vers le  $C_{\text{ter}}$ , sans pour autant déstabiliser le repliement en amont (en  $N_{\text{ter}}$  de ce tour-là).

D'abord, comme dit plus tôt, le cœur hydrophobe y est différent, majoritairement composé de valines en position  $B_3^i$ , mais nous y retrouvons aussi des résidus comme la tyrosine et la phénylalanine, ce qui a été décrit comme interdit dans l'étude précédente (Tableau 2). Nous observons la présence non négligeable de sérines, et thréonines en position  $B_1^i$  des faces (I) et (III) (Figure 20 a). Et encore une fois, l'hexade (II) est différente puisque dans plus de 70 % des structures analysées, cette position  $B_1^i$  n'était pas présente, déstabilisant le

repliement. De plus, lorsqu'un résidu occupe cette position, plusieurs acides aminés y sont retrouvés, dont la lysine, ce qui n'est pas observé dans les autres tours et n'est normalement pas compatible avec la continuation de l'hélice puisque cela l'orient vers le cœur hydrophobe. Cette suppression va de pair avec la position T<sub>2</sub> adjacente (i.e. hexade (I)) qui, elle aussi, n'est pas présente dans la majorité des cas. Nous avons attribué cela à un motif de terminaison permettant de « casser » le réseau de liaisons hydrogène de l'hélice (Figure 20 b).

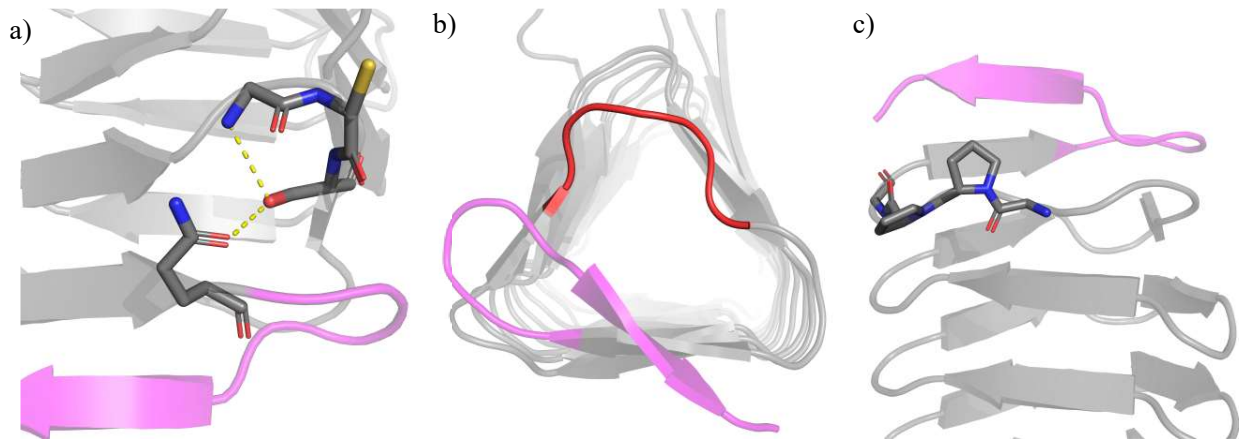


Figure 20: Motifs caractéristiques des tours de terminaison (C) avec a) l'interaction de la sérine en B<sub>1</sub><sup>o</sup>, b) la suppression de deux résidus cassant le réseau de liaisons H (rouge), c) la présence d'une proline en B<sub>4</sub><sup>o</sup> perpendiculaire aux feuillets β et T<sub>1</sub> induisant le coude. En rose est montrée l'épingle β comme motif de « bouchon » (PDB : 1LXA).

Pour rappel, tous les résidus ne contribuant pas à l'hélice que ce soit par interactions de type feuillets β, cœur hydrophobe, hydrogènes ou électrostatiques ne sont pas pris en compte dans l'analyse. Ce qui veut dire ici que les positions T<sub>1</sub> et T<sub>2</sub> de la face (III) ne contribuent pas à l'hélice puisque n'occupent pas des positions canoniques. Comme vu ci-dessus, très souvent, l'hélice se termine par un motif en épingle β. Ces épingle commencent à la position T<sub>1</sub> de cette même face, ce qui explique le taux de présence très faible d'acides aminés aux positions T<sub>1</sub> et T<sub>2</sub>.

Un autre motif de terminaison très fréquent est la présence de proline en position B<sub>4</sub><sup>o</sup>, retrouvée dans plus de 60 % des structures étudiées. L'étude précédente montrant l'incompatibilité de retrouver une proline à cette position confirme le rôle déstabilisateur de la proline.<sup>[8]</sup> Cette proline, très souvent suivie d'une autre (motif Pro-Pro), se place perpendiculairement aux feuillets β, ce qui prévient l'empilement d'autres feuillets (Figure 20 c). Cela permettant de terminer la croissance de l'hélice et d'avoir un assemblage fini.

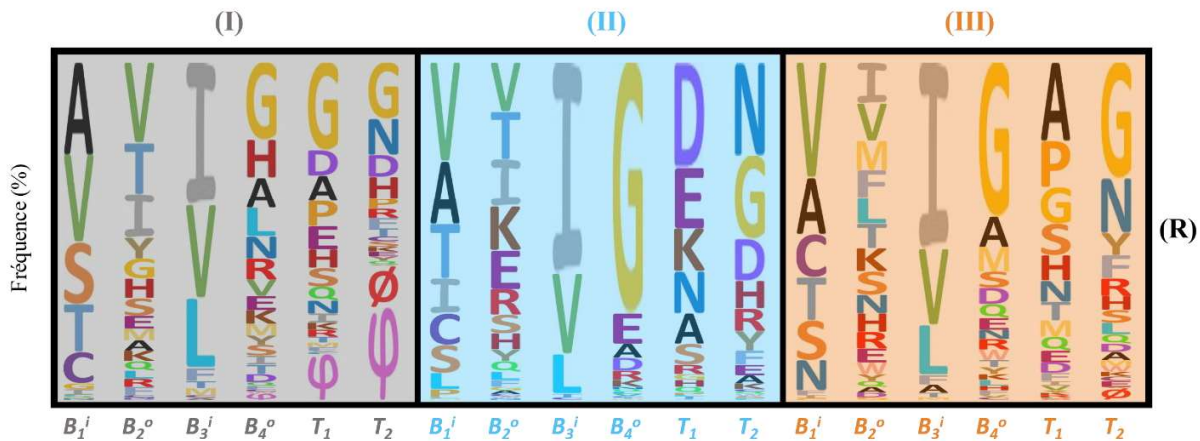
Il paraît aussi nécessaire de mentionner le fait que les acides aspartiques présents en B<sub>2</sub><sup>o</sup> de la face (II) font une liaison hydrogène avec les chaînes principales situées dans le tour n-1. Alors que les résidus hydrophobes comme la leucine, la valine, ou l'isoleucine retrouvées à ces positions interagissent très souvent avec le bouchon C-terminal de l'hélice qui peut être en épingle β mais aussi plus désordonné, ou encore en hélice α.

L'importance de ce tour final dans l'assemblage de l'hélice β gauche de type I est donc démontrée ici par ces motifs spécifiques, il sert à interdire le prolongement de l'hélice en C<sub>ter</sub> de la séquence, ainsi que la possible

agrégation par interaction avec une autre hélice. Ce tour-ci, au même titre que le tour de nucléation est donc primordial dans le but de garder un repliement fini, et soluble.

### 2.3.2.3 Les tours canoniques (R)

Tableau 6 : Fréquence d'observation des acides aminés suivant les positions retrouvées sur les tours canoniques. Ø et φ faisant respectivement référence aux positions supprimées et n'étant pas canoniques (boucles).



Les tours dits « canoniques » représentent la majorité des résidus comptabilisés dans cette étude. En effet, ceux-ci sont les plus représentés avec plus de 7300 résidus analysés (72.6%) dans ces tours (et seulement un peu moins de 1400 (13.7%) pour chaque tour (N) et (C)). Ceci explique pourquoi le Tableau 6 ressemble beaucoup à celui de Cohen *et coll.* (Figure 13 a de l'Introduction) qui représente une hexade moyenne issue la protéine UDP N-acetylglucosamine acyltransferase (PDB : 1LXA). Cependant, à la différence de cette précédente étude, la discrimination entre les faces permet de mettre en valeur certaines différences dans leur composition.

Ces tours R sont les plus réguliers, au regard de la conformation, car ils sont situés au centre de l'édifice, et sont responsables de la stabilité générale de l'hélice. Les résidus les plus conservés sont, comme dans les autres parties, ceux du cœur hydrophobe. Il est aussi frappant de voir la fréquence à laquelle la glycine est présente en position B<sub>4</sub><sup>o</sup>, en particulier dans l'hexade (II). La logique voudrait que nous ayons une grande diversité de résidus présents à cette position (puisque'elle pointe vers le solvant), toutefois ce n'est pas le cas. Une explication possible serait le besoin de flexibilité ou plutôt, le besoin de ne pas faire d'interaction, tirant le feuillet vers l'extérieur, ce qui aurait pour conséquence de déstabiliser l'assemblage. La grande variabilité de résidus dans cette partie, pour les résidus extérieurs, permet de dire que les séquences sont très variables et qu'il n'y a pas vraiment d'acides aminés préférentiels, notamment pour l'hexade (I) et plus généralement pour les positions B<sub>2</sub><sup>o</sup>. La nature des résidus dans ces tours est déterminée par ce qu'il y a en face (i.e. coude d'une autre hélice en vis-à-vis), ce qui fait que pour les faces (I) et (III) on observe plus de résidus hydrophobes afin de favoriser les interactions avec les chaînes protéiques non loin.

La présence de boucles ( $\phi$ ), non canoniques, est aussi importante dans ces tours et commence quasiment toujours à la position T<sub>1</sub> de la face (I), ces coudes pointant directement vers le solvant. Elles ne peuvent pas être présentes sur l'hexade (III) car les positions T<sub>1</sub> et T<sub>2</sub> pointent directement vers l'intérieur du trimère, ne permettant pas d'accommoder des boucles. Elles ne sont pas non plus présentes sur l'hexade (II) car elles font face, très souvent, à d'autres chaînes protéiques provenant d'un bouchon par exemple (Figure 20 a et c).

Dans ces tours, la proline est aussi présente mais dans de plus faibles proportions que les tours (N) et (C), en T<sub>1</sub>, ce qui confirme l'importance des prolines dans les premiers tours, et moins dans les tours canoniques. Nous pouvons aussi remarquer qu'il y a seulement 1 % de prolines sur la face (II) alors que les autres faces en contiennent 8 et 14 % ((I) et (III) respectivement).

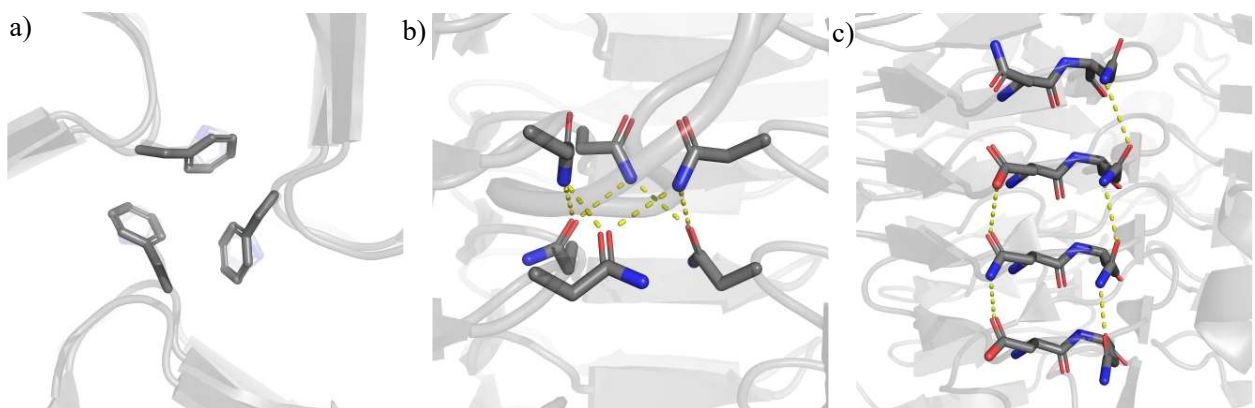


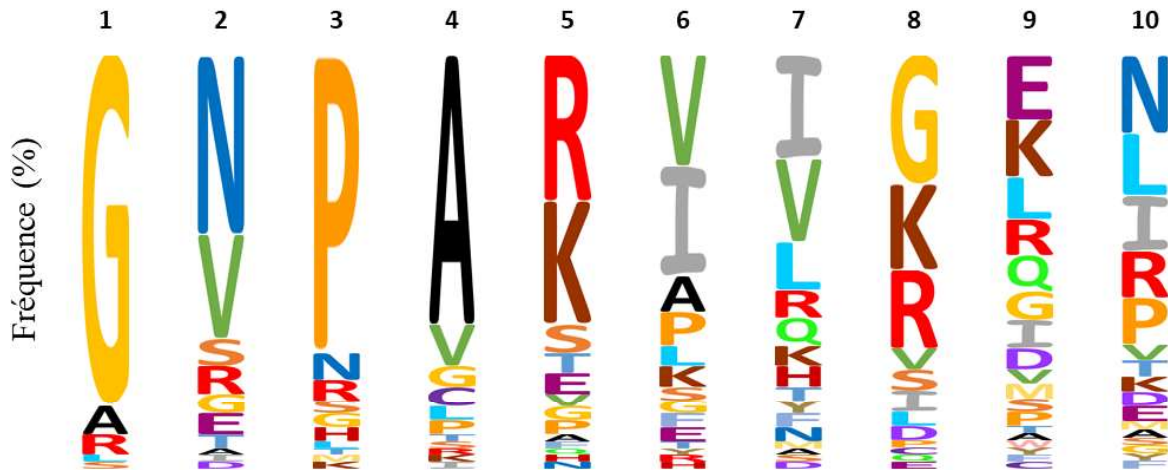
Figure 21 : Exemples d'interactions stabilisantes typiquement retrouvées entre les tours (R). a) de trois phénylalanines en position T<sub>2</sub> (PDB : 1LXA) et b) de trois asparagines en position T<sub>1</sub> se faisant face au centre du trimère réalisant des échelles avec d'autres asparagines (PDB : 4EQY), c) échelles d'asparagines en position T<sub>2</sub> et alternances avec des acides aspartiques vers l'extérieur du trimère (PDB : 1OCX).

Les résidus sont plus chargés sur cette hexade (II) aux positions T<sub>1</sub> et T<sub>2</sub> afin de favoriser les interactions avec l'eau (et donc la solubilité de la protéine) mais aussi avec les autres chaînes latérales, ou principales (Asp, Glu, Lys, Asn). Il est cependant possible de retrouver des phénylalanines ou histidines en T<sub>2</sub> permettant la formation d'interactions  $\pi$ , notamment au centre du trimère (Figure 21 a).

Des interactions permettant de stabiliser l'assemblage sont également retrouvées ici en T<sub>1</sub> et T<sub>2</sub>. Cela est notamment le cas avec la présence d'acides aspartiques et glutamiques en T<sub>1</sub> effectuant des liaisons hydrogène avec la chaîne principale du tour n-1. Toutefois nous retrouvons également des asparagines en position T<sub>1</sub> et T<sub>2</sub> formant des échelles d'asparagines (Figure 21 c), comme décrit précédemment par Perutz.<sup>[9]</sup> Les motifs de stabilisation les plus retrouvés dans les protéines étudiées sont représentés dans le Tableau 6, notamment avec la séquence GDN retrouvée fréquemment.

### 2.3.2.4 Motifs de capping (« bouchon ») en C-ter de l'hélice

Tableau 7 : Fréquence des résidus retrouvés dans le motif épingle  $\beta$  masquant l'extrémité C-terminale de l'hélice. Les résidus sont notés de 1 à 10 dans le sens d'observation après l'hélice  $\beta$ .



Naturellement, les tours canoniques ont des motifs de terminaison, des bouchons, en  $N_{\text{ter}}$  et  $C_{\text{ter}}$  permettant de masquer l'extrémité de l'hélice pour éviter sa continuation, en plus du tour de terminaison (C) (Figure 14). Nous avons remarqué, lors de l'étude de ces hélices, qu'un grand nombre de ces structures se terminaient par un bouchon après la face (III). Nous avons observé que 92 % des structures étudiées ont cette face se prolongeant par un motif de terminaison. Lors de cette étude des structures en hélices  $\beta$  gauche de type I, nous avons remarqué une part non négligeable de motifs formant le bouchon se terminant par une épingle  $\beta$  avec 77 % des hélices se terminant par ce motif, signe que ce motif est un paramètre important à considérer pour la conception d'hélice de taille finie (Figure 20 a). Nous avons alors accompli la même chose précédemment avec l'hélice, et le résultat est montré en Tableau 7. Dans cet histogramme nous visualisons particulièrement bien l'augmentation de la diversité de la séquence au fur et à mesure qu'on s'éloigne de la séquence de la L $\beta$ H canonique. Ce qui traduit une conservation de la séquence lorsque nous sommes au début du motif, débutant quasi systématiquement (85%) à la place de la position  $T_1$  de l'hexade (III), correspondant à la glycine au début du motif (Tableau 7). Les interactions retrouvées ici permettent de masquer le cœur hydrophobe d'une part, et d'autre part de casser la régularité et l'enchaînement des feuillettes  $\beta$  parallèles : le « bouchon » se mettant souvent « en travers » de l'hélice.

À partir de ce petit tableau regroupant les séquences de 10 acides aminés des épingles  $\beta$  issus des différentes structures, nous pouvons identifier des séquences permettant d'ajouter un motif court lors de la conception de la séquence d'une L $\beta$ H, afin de stopper l'hélice.

## 2.4 Conclusion

Les études précédentes ont permis de comprendre le repliement des protéines naturelles en hélice  $\beta$  gauche. La force motrice de son repliement est principalement la formation de feuillets  $\beta$  parallèles ainsi que la formation d'un cœur hydrophobe. D'autres contributions ont été décrites, comme les échelles d'asparagines. La proline a un rôle important dans les tours aux extrémités de chaque hélice.

Cependant, dans ces études, ni la structure trimérique de l'hélice  $\beta$  gauche de type I, ni les différentes parties la composant n'ont été prises en compte. Une étude plus poussée, permettant de mettre en valeur les résidus ou motifs spécifiques à chaque face et tour, au nombre de 3 chacun est décrite ici. L'analyse alors réalisée est présentée sous forme d'un tableau de 9 hexades qui pourra servir d'outil à la conception de peptides et de protéines se repliant en L $\beta$ Hs (comme dans le Chapitre 2 de cette thèse par exemple).

De la recherche de structure en hélices  $\beta$  gauche à l'analyse, tout a été mis en place pour rendre compte de la manière la plus fidèle possible de ce qu'on peut retrouver dans la Nature. Nous retrouvons d'abord la composition du cœur hydrophobe, élément essentiel de ce repliement, tel que décrit dans la littérature. Elle est complétée par l'observation d'interactions particulières, impliquant la chaîne latérale de l'acide aminé en position  $B_1^i$ . Ensuite, nous avons mis en lumière de nouveaux éléments participant à la stabilité, à l'initiation ou à la terminaison de ce repliement. Ces derniers illustrent l'importance des tours aux extrémités N-ter et C-ter pour la formation de ce repliement et sa terminaison.

En effet, dans le tour N-terminal de l'hélice (N), nous avons identifié des motifs X-P-Z qu'on ne retrouve que là, stabilisant la formation des coudes ainsi que leur interaction avec le tour  $n+1$ . Aussi, nous voyons en ce tour N-terminal un rôle très important de nucléation du repliement en L $\beta$ H.

Dans le tour C-terminal de l'hélice (C) en revanche, ce sont des motifs qui déstabilisent (voir empêchent complètement) l'interaction avec un éventuel tour  $n+1$  qui sont observés. C'est le cas de la présence de prolines en  $B_4^o$  (notamment dans le motif Pro-Pro), retrouvées nulle part ailleurs, ainsi que la suppression de résidus en  $T_1$  et  $T_2$  par exemple. Le cœur hydrophobe, ainsi que la section, de ce tour-ci est aussi moins important que les autres tours. Nous découvrons donc en ce tour C-terminal un rôle très important de terminaison du repliement.

Les tours au centre de l'hélice, dits canoniques, sont eux les plus réguliers et similaires à ce qu'avait déjà décrit Cohen et ses collaborateurs.<sup>[27,31]</sup> Néanmoins, en discriminant les trois faces de l'hélice, nous avons pu mettre en évidence des motifs telle que la séquence G-D-N ( $B_4-T_2$ ) fréquemment retrouvée dans les protéines étudiées. D'autres interactions stabilisant la forme trimérique ont été observées et identifiées comme la formation d'échelles asparagines ou l'empilement de phénylalanines avec des résidus présents en positions  $T_1$  ou  $T_2$ .

Malgré toutes les considérations prises en compte lors de cette étude, il est important de rappeler ici que nous ne nous sommes intéressés qu'aux résidus contribuant seulement à l'hélice  $\beta$ . Toutes les interactions autres, que ce soit avec ligands, métaux, ou chaînes protéiques de domaines globulaires n'ont pas été prises en compte.

Malgré tout, nous avons identifié un motif formant un bouchon en C-ter, en épingle  $\beta$ , fréquent pour ce type de structure.

Tableau 8 : Tableau regroupant les différentes statistiques observées lors de l'analyse des structures en L $\beta$ Hs de type I. (Nb=nombre). Les couleurs font référence aux faces et parties définies précédemment.

Statistiques par pas d'hélice			I			II			III			Tours
Nombre d'A.As (+ manquants & boucles)	N <sub>ter</sub> (N)	Nb	461			463			453			1377
		par	Tour	Face	Total	Tour	Face	Total	Tour	Face	Total	Total
		%	33.5	14.3	4.6	33.6	14.0	4.6	32.9	12.8	4.5	13.7
	Canonique (R)	Nb	2310			2388			2619			7317
		par	Tour	Face	Total	Tour	Face	Total	Tour	Face	Total	Total
		%	31.6	71.6	22.9	32.6	71.9	23.7	35.8	74.0	26.0	72.6
	C <sub>ter</sub> (C)	Nb	455			468			465			1388
		par	Tour	Face	Total	Tour	Face	Total	Tour	Face	Total	Total
		%	32.8	14.1	4.5	33.7	14.1	4.6	33.5	13.2	4.6	13.7
	Résidus au total			3226 (32%)			3319 (32.9%)			3537 (35.1%)		
% Structures contenant des métaux			27.0 % (=21/78); 38.1 % (8) Zn; 33.3 % (7) Mg 28.6 % (6) Ca & Na, Ni									
Nombre de tours complets			Moyen = 7 ; Max = 11; Min = 4									551
Structures débutant par face			20 (25.6 %)			19 (24.4%)			39 (50.0 %)			78
Bouchon C <sub>ter</sub>	Autres		18 (23.1 %) chaînes et hélice									
	Épingles $\beta$		60 (76.9 %) sur la face(III)									
Structures se terminant (C <sub>ter</sub> ) par face			5 (6.4%)			1 (1.3%)			72 (92.3%)			78

Un tableau récapitulatif des différentes observations statistiques des L $\beta$ Hs est montré dans le Tableau 8. Il regroupe le nombre d'A.As par tour, par face, et total, la présence de métaux ou non, le nombre de tours maximum, minimum et moyen ainsi que les début et fin d'hélices avec la présence de bouchons. Cela permet de rendre compte de l'importance des acides aminés de la partie canonique (R), les plus représentés. Nous pouvons aussi observer le nombre de structures se terminant par la face (III), et le nombre d'épingles  $\beta$  comme bouchon en C<sub>ter</sub> de cette face est majoritaire.

Nous avons de plus réalisé un diagramme pour rendre mieux compte de la répartition des tours par structure, nous permettant d'affirmer que le minimum de tours est de quatre pour ce type de structure (naturelle) avec une moyenne de sept tours par structure, le nombre le plus retrouvé étant de six (Figure 22).

L'analyse de toutes les séquences des protéines naturelles des L $\beta$ Hs de type I a permis de construire un tableau présentant les résidus les plus fréquents pour chaque position, ainsi des motifs

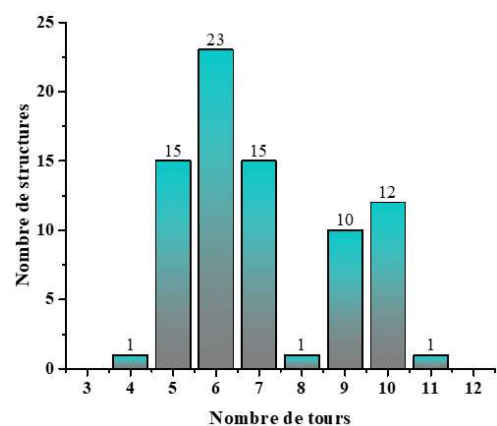


Figure 22 : Distribution du nombre de tours d'hélice en fonction des structures étudiées.



stabilisant ce type de repliement. Celui-ci sera très utile pour concevoir des séquences de *novo* de peptides et de protéines se repliant en hélices  $\beta$  gauche de type I.

Dans le second chapitre de ma thèse, nous verrons comment nous nous sommes aidés de cette analyse dans le but de concevoir des peptides s'autoassemblant et se repliant en L $\beta$ Hs.

## 2.5 Perspective : Application du nouvel outil à la conception de peptides solubles de L $\beta$ H

À partir du nouvel outil, nous avons conçu la séquence d'un peptide, visant à ce qu'il se replie en L $\beta$ H soluble. Les caractérisations n'étant pas finies, la conception et les résultats préliminaires obtenus sont ici présentés en tant que perspective d'application de l'outil développé dans ce chapitre.

Le peptide conçu correspond à trois tours de L $\beta$ H, un tour (N), un tour (R) et un tour (C). Des bouchons ont été ajoutés aux extrémités afin de garder un assemblage fini et le plus soluble possible. Dans l'optique d'en obtenir la structure, nous essaierons d'obtenir des cristaux de ce peptide. Le but étant d'obtenir une preuve de concept, démontrant la pertinence de ce tableau comme nouvel outil de conception d'hélices  $\beta$  gauche de type I. La séquence conçue fait 59 acides aminés de long et est nommée 59M (Figure 23).

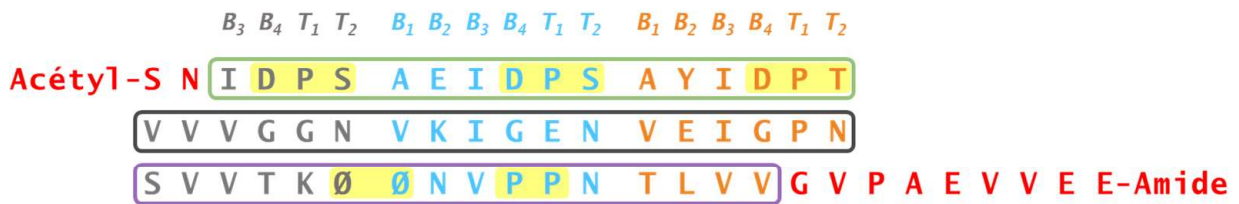


Figure 23 : Séquence du peptide 59M conçu en fonction des hexades (I), (II) et (III) et des tours (N), (R) et (C). Les motifs de bouchon sont représentés en rouge. Surlignés en jaune, nous retrouvons les motifs caractéristiques pour chaque tour identifié précédemment.

Pour concevoir cette séquence, nous avons directement lu le Tableau 3 puis nous avons choisi les résidus les plus favorables pour chaque position. Nous avons introduit les motifs de nucléations XPZ identifiés précédemment pour la conception du tour (N), ceux-ci étant surlignés en jaune. Pour le tour (R), nous avons lu le tableau en choisissant aux positions  $B_2^\circ$  des résidus permettant une complémentarité de charge avec les tours en  $n+1$ , de même pour les asparagines en  $T_2$ , et enfin une proline a été intégrée dans le tour (R) pour aider un peu plus au repliement. Le tour (C) est conçu avec les résidus les plus présents pour chaque position, les suppressions de résidus étant aussi prises en compte, ainsi que le motif Pro-Pro, surligné en jaune (Figure 23). Le bouchon commence, comme dans beaucoup de structures, par la position  $T_1$  de ce dernier tour, ce qui enlève deux positions à l'hélice. Le bouchon choisi est un motif d'épingle  $\beta$  construit à partir de la séquence consensus des motifs étudiés (Tableau 7). Il est important ici de préciser que la conception de cette séquence s'est fait avec une volonté d'avoir un repliement globalement chargé 7- à pH=7.4 et sera donc probablement déstructuré

à ce pH. Ce qui veut dire qu'à un pH relativement acide (pH environ 3.8), le peptide sera neutre (pI = 3.8) minimisant ainsi les répulsions de charges. Il y a des chances que le peptide se structure dans un domaine proche de ce pH. Le bouchon en N<sub>ter</sub> composé de Ser-Asn-Ile- est un motif formant un pli en fin de chaîne (séquence venant de ce qu'on retrouve chez l'anhydrase carbonique  $\gamma$  (PDB : 3OTM)). La sérine, interagissant avec la chaîne principale de l'isoleucine et l'asparagine avec celle d'un coude  $\beta$  situé en n+1, renforce l'assemblage tout en déstabilisant l'approche d'un autre feuillet (Figure 24 a, milieu). L'isoleucine elle, est le premier résidu participant au cœur hydrophobe. La Figure 24 montre comment un modèle pour le peptide conçu à partir de l'outil précédent, avec des interactions et motifs caractéristiques pour chaque partie d'hélice (i.e. D-P-S-, Ø- Ø- ou encore P-P-).

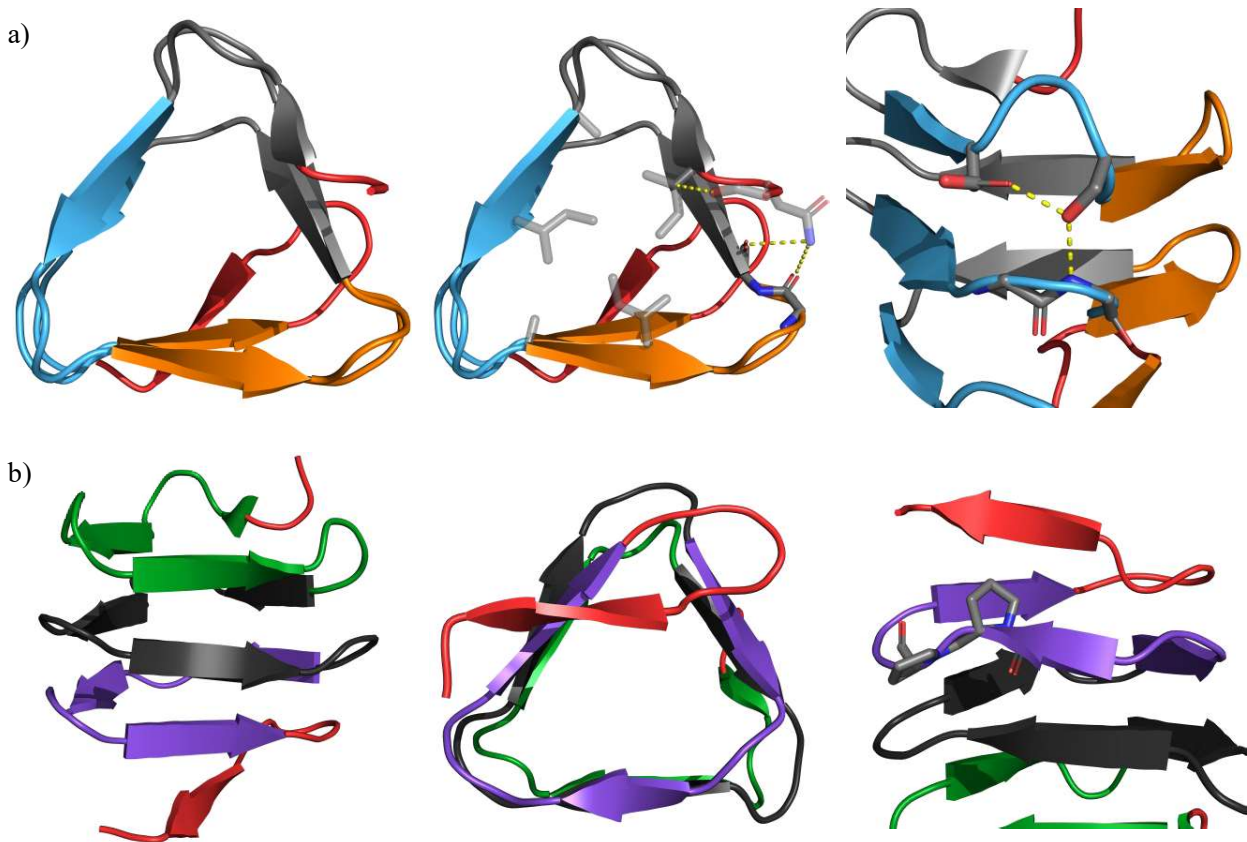


Figure 24 : : Modélisation AlphaFold du peptide 59M montrant a) les différentes faces (I), (II), (III) et interactions pour le tour de nucléation et b) les différents tours (N), (R), (C) témoignant des interactions dans ce dernier tour. Les bouchons sont montrés en rouge.

La synthèse de ce peptide a été réalisée, comme pour les peptides du Chapitre 2, par SPPS classique en stratégie Fmoc/tBu avec un robot de synthèse. La purification a été réalisée par HPLC semi-préparative à un pH=9 (formate d'ammonium), avec une colonne C4 suivie d'une analyse HPLC et LCMS afin de confirmer que le peptide voulu est obtenu et avec une pureté suffisante. La pureté maximum obtenue n'a été que de 90 % (estimée par HPLC analytique), ce qui est acceptable vu la longueur du peptide. Les fractions les plus pures ont été lyophilisées.

Pour le moment, seules des analyses préliminaires ont été réalisées afin de voir si la structure secondaire est compatible avec le repliement attendu. Des essais de cristallisation ont été faits, mais pour l'instant, aucun cristal n'a pu être obtenu. Nous avons analysé ce peptide par spectroscopie de dichroïsme circulaire afin d'identifier sa structure secondaire. Le détail lié aux mesures sur synchrotron est décrit dans le chapitre 2 et en partie expérimentale. Comme nous pouvons le voir sur la Figure 25, après dissolution dans l'eau (pH environ 7.5-8) le peptide est très peu structuré à 25 °C (pelote statistique) bien qu'un épaulement à 220 nm indique la présence de brins  $\beta$ . Ce qui est le résultat attendu au vu de la séquence conçue. Nous avons ajusté le pH avec un tampon phosphate de potassium à pH=3, et ajouté du sel (sulfate d'ammonium), pour voir si le peptide se replierait. Le signal de brins  $\beta$  (bande négative à 220 nm) augmente, mais le peptide reste majoritairement désordonné. Il semblerait alors que le pH et la salinité ne suffisent pas à replier le peptide. Cela nous a donc amenés à penser que le pH est certes important mais aussi que la température peut jouer un rôle. Nous avons alors réalisé une cinétique de repliement du peptide 59M à pH=7.4 en présence de 500 mM de sulfate d'ammonium (pour augmenter la force ionique, qui pourrait aider au repliement de l'hélice).

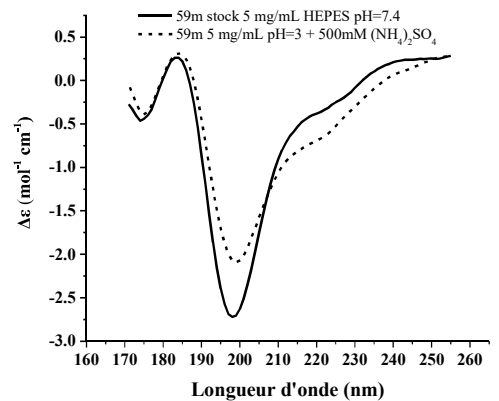


Figure 25 : Spectres SR-CD du peptide 59M en solution à pH 7.4 et 3 avec du sel (sulfate d'ammonium).

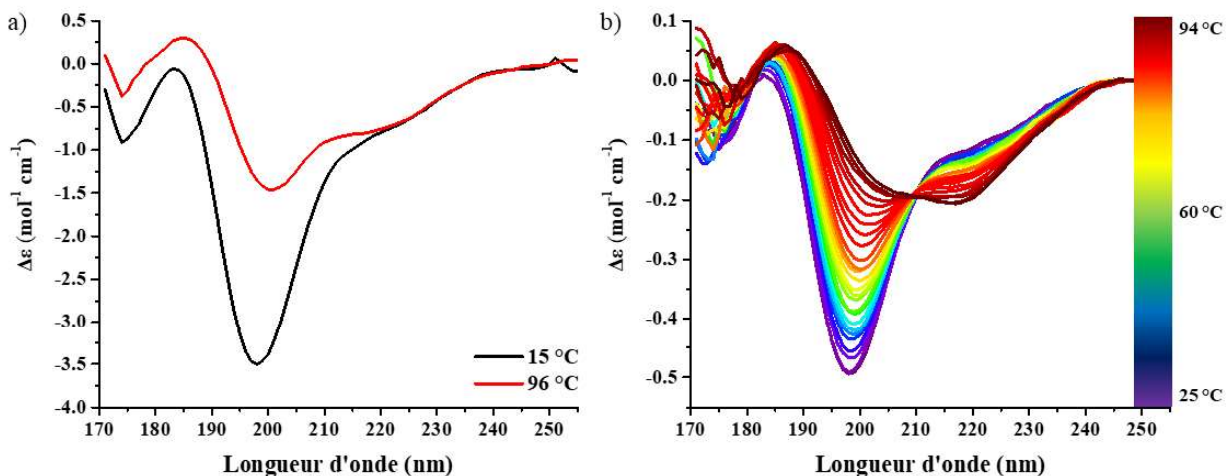


Figure 26: a) Spectres SR-CD du peptide 59M à 5 mg/mL dans l'HEPES pH=7.4 à 50 mM à 15 et 96 °C et b) suivi cinétique SR-CD du peptide 59M à 5 mg/mL dans le MES pH=6 à 50 mM en fonction de la température. Les spectres a) et b) ont été réalisés en présence d'une concentration de sulfate d'ammonium de 500 mM pour aider au repliement.

Une montée en température allant de 15 à 96 °C a été réalisée afin de pouvoir observer si, à ce pH (7.4), le repliement de l'hélice avait lieu. Comme le montre la Figure 26 a), ce n'est visiblement pas le cas, même si la bande négative à 198 nm caractéristique des pelotes statistiques a bien diminuée, la bande correspondante aux feuillets  $\beta$  n'étant pas devenue plus intense. La bande positive à 186 nm est déplacée à 189 nm et augmentée en fin d'expérience ce qui pourrait présager d'une réorganisation, mais minime. Nous pouvons alors en conclure

qu'à un pH=7.4, même en présence de sel et avec une température élevée, le repliement n'a pas lieu. Pour vérifier la complémentarité entre le pH et la température, nous avons réalisé une cinétique de repliement à pH=6 en fonction de la température. Nous observons que lorsque le pH est abaissé à 6, un changement de structure secondaire a lieu lorsque la température est élevée (Figure 26 b). Ce qui permet de mettre en évidence la complémentarité entre le pH et la température dans le repliement en l'hélice  $\beta$  gauche de ce peptide. Même si nous n'observons pas de conversion complète en feuillets  $\beta$ , nous observons très clairement, contrairement au pH=7.4, la formation de ceux-ci avec la bande à 218 nm devenant plus intense, le début d'inversion de la bande à 198 nm et le déplacement de la bande passant de 183 à 187 nm. Cela indique alors le repliement, au moins partiel, du peptide lorsque la température est augmentée, avec un pH de 6 en présence de sel, ce qui va dans le sens de la conception réalisée.

Afin d'être sûr de la relation pH/température, nous avons réalisé des spectres SR-CD à différentes températures, 25, 60, 90 °C puis redescendu à 15 °C à des pH de 3 et 4. Pour réaliser ces mesures, nous avons fixé la concentration de peptide comme précédemment à 5 mg/mL dans un tampon phosphate de potassium pH=3 ou acétate pH=4 à 50 mM, puis nous avons ajouté 500 mM de  $(\text{NH}_4)_2\text{SO}_4$ . Nous avons par la suite mesuré les spectres à différentes températures, un temps d'équilibration de 15 minutes pour chaque température a été observé afin de permettre au système de s'équilibrer. Les résultats obtenus sont montrés en Figure 27. Les différences d'intensités entre les spectres noirs aux pH 3 et 4 pourraient indiquer la présence d'une structure secondaire moins importante à pH= 4 comparée à celle à pH= 3. Nous remarquons également que pour ces deux pH, nous obtenons, à la fin, une structure secondaire en feuillets  $\beta$  (parallèle). La seule différence entre les deux pH est le repliement partiel pour le pH=4 à 60 °C (Figure 27 a) alors que pour le pH 3 celui-ci est figé et ne bouge plus beaucoup (Figure 27 b). Cependant, nous remarquons, dans les deux cas un déplacement de la bande à  $\sim 218$  nm (vers 215 nm) lorsque la température passe de 90°C à 15 °C, mettant en évidence le rôle non-négligeable de la température au signal CD.

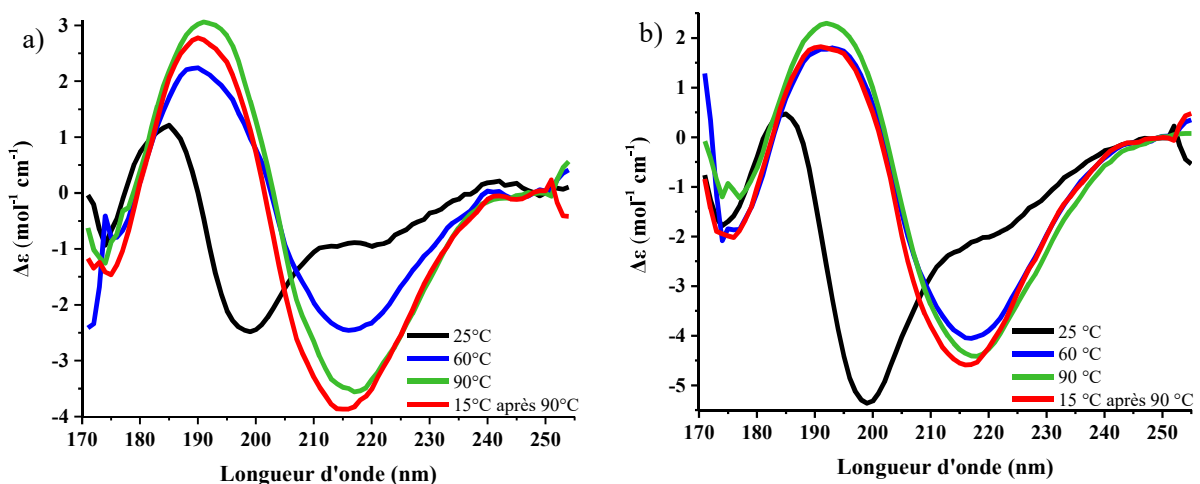


Figure 27: Spectres SR-CD du peptide 59M à 5 mg/mL avec 500 mM de  $(\text{NH}_4)_2\text{SO}_4$  avec une concentration de tampon de 50 mM pour un pH de a) 4 et b) 3 avec des températures différentes.

Nous remarquons en outre que le repliement, une fois formé, reste stable lorsque la température est abaissée après la chauffe, signe que celui-ci est robuste et inchangé après sa formation. Et encore une fois, cela est compatible avec la formation d'hélices  $\beta$  gauche à partir de la séquence conçue pour le peptide 59M. Néanmoins, ici nous n'avons pas vérifié que le peptide était resté bien soluble et n'a pas précipité ou formé de fibres.

À travers cette étude préliminaire en SR-CD d'un peptide conçu à partir du nouvel outil de conception de séquences pouvant se replier en hélices  $\beta$  gauche, nous avons montré que ce peptide a besoin d'un pH bas et d'une température élevée pour se replier. Le repliement formé n'est pas réversible puisque le peptide reste replié même lorsque la température est redescendue. Les spectres nous montrent la formation de feuilletts  $\beta$  probablement de conformation majoritairement parallèle pour des pH inférieurs ou égaux à 4, ce qui correspond au pKa des acides glutamiques et aspartiques présents dans la séquence. Par ailleurs, il faudra absolument s'assurer de la solubilité de ce peptide dans toutes ces conditions. Les étapes suivantes seront de vérifier l'état d'oligomérisation du peptide et d'en obtenir la structure par diffraction des rayons X si possible.

## 2.6 Références bibliographiques

- [1] J. H. Choi, C. Govaerts, B. C. H. H. May, F. E. Cohen, *Proteins Struct. Funct. Bioinforma.* **2008**, *73*, 150–160.
- [2] J. H. Choi, B. C. H. H. May, C. Govaerts, F. E. Cohen, *Structure* **2009**, *17*, 1014–1023.
- [3] P. Iengar, N. V. Joshi, P. Balaram, *Structure* **2006**, *14*, 529–542.
- [4] K. Basu, R. L. Campbell, S. Guo, T. Sun, P. L. Davies, *Protein Sci.* **2016**, *25*, 946–958.
- [5] B. Chakrabarty, N. Parekh, *Protein Sci.* **2021**, 1–14.
- [6] Schrödinger, LLC, *The {PyMOL} Molecular Graphics System, Version~1.8*, **2015**.
- [7] C. Govaerts, H. Wille, S. B. Prusiner, F. E. Cohen, *Proc. Natl. Acad. Sci.* **2004**, *101*, 8342–8347.
- [8] J. H. Choi, B. C. H. H. May, C. Govaerts, F. E. Cohen, *Structure* **2009**, *17*, 1014–1023.
- [9] M. F. Perutz, J. T. Finch, J. Berriman, A. Lesk, *Proc. Natl. Acad. Sci.* **2002**, *99*, 5591–5595.



# **3 Conception de peptides de novo s'autoassemblant en hélice $\beta$ gauche**

## 3.1 Introduction

Ce chapitre deux décrit la conception de briques de base pour la création de fibres, ou de fils moléculaires, par autoassemblage de peptides se repliant en L $\beta$ Hs. Pour faire cela, le nouvel outil développé dans le chapitre un, prenant en compte les différentes parties d'hélice ainsi que la non-équivalence de ses faces, servira de base. La première raison est justement la conception de trimères, ce qui veut dire que la séquence de la brique élémentaire de l'hélice ne doit pas être symétrique, car les séquences, correspondant aux faces, ne sont pas équivalentes. Le besoin de concevoir une face favorable à l'exposition au solvant (II), une propice à l'exposition d'un coude d'une autre hélice (I) et une pouvant s'accommoder de solvant mais aussi d'une face d'hélice (III) est alors justifié lorsque l'on veut former un trimère d'hélices. De ce fait, l'agrégation latérale d'autres unités sera moins favorable et la fibre aura plus de probabilités de former une structure trimérique.

La stratégie utilisée s'inspire de celle décrite par l'équipe de P. L. Davies qui se base sur l'identification de motifs répétitifs dans les solénoïdes  $\beta$ . L'empilement de ces motifs permettra de former un solénoïde  $\beta$  à partir d'autoassemblages de briques de base.<sup>[1]</sup> Également inspiré par les travaux de l'équipe de D.L. Cox, qui a formé des fibres de solénoïdes  $\beta$ , plus précisément des L $\beta$ Hs de type II, à partir de protéines exprimées et modifiées dans le but de favoriser la polymérisation des motifs (voir la section 1.5 de l'Introduction).<sup>[2]</sup> Mais ces derniers travaux reposent sur une stratégie Top-Down, où, à partir d'un objet de taille importante que l'on « réduit/nettoie » on obtient une brique de taille réduite (Figure 9). À l'inverse, nous visons à concevoir des objets de taille importante à partir d'une brique minimale dont la séquence est conçue rationnellement (approche Bottom-Up).

Pour cette partie, les six positions correspondantes à celles des résidus dans l'hélice seront nommées sans leurs indices et exposants faisant référence à l'orientation vers l'intérieur ou l'extérieur, cela pour des raisons de simplicité.

## 3.2 Conception rationnelle de la brique élémentaire

Nous cherchons ici à synthétiser, chimiquement, des briques de bases pouvant par la suite s'autoassembler en fibres d'hélices  $\beta$  gauche de type I à partir des travaux du chapitre I. La conception d'une séquence, comme brique élémentaire se repliant en hélice  $\beta$  gauche, doit se faire en favorisant au maximum le repliement de cette brique. C'est pour cette raison que la séquence minimale a été choisie comme correspondant à 2 tours d'hélice, ce qui permet une préorganisation par interactions intramoléculaires, notamment chaîne-principale-chaîne principale (entre feuilletts  $\beta$ ).<sup>[3-7]</sup> Nous présentons, dans le Tableau 9, les différentes étapes menant à la conception de deux peptides. Ils seront nommés par la suite 35m1 et 35m2, leurs séquences correspondant à 35 acides aminés (35 mer).







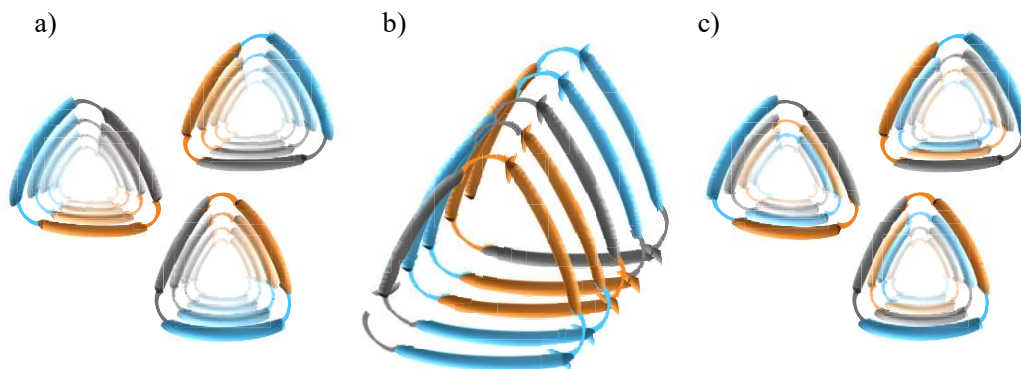


Figure 28 : Schéma de a) structure trimérique attendue composée d'autoassemblages de briques de base avec un nombre entier de tours ; b) monomère d'hélice et c) trimère d'hélices conçues avec des nombres de tours non entiers.

Il est alors primordial d'avoir un nombre de tours entier afin de pouvoir mimer les structures trimériques des L $\beta$ Hs naturelles. La brique de base pourra correspondre alors à un tour, 18 A.As ou un multiple. Or, comme l'a été montré dans la littérature par des simulations, la conception d'une unité faite d'un seul tour n'est pas stable. Mais lorsque ce nombre est augmenté à deux, l'assemblage reste inchangé.<sup>[9]</sup> Au-delà, l'assemblage reste stable et ne bouge pas non plus, mais plus la séquence est longue, plus elle est difficile à synthétiser. Il nous a alors paru opportun de concevoir une brique de base correspondant à 2 tours, comme brique minimale. Un peptide de 2 tours de L $\beta$ H correspond donc à  $2 \times 18 = 36$  acides aminés. Ces deux tours permettent une préorganisation du système par interactions intra-moléculaires : formation du cœur hydrophobe, interactions des feuilletts  $\beta$  parallèles entre les tours  $n$  et  $n+1$  ou encore ponts salins entre acides aminés de charges opposées des tours  $n$  et  $n+1$ .

### 3.2.2 Coupure, formylation et amidation de la brique (étape 2)

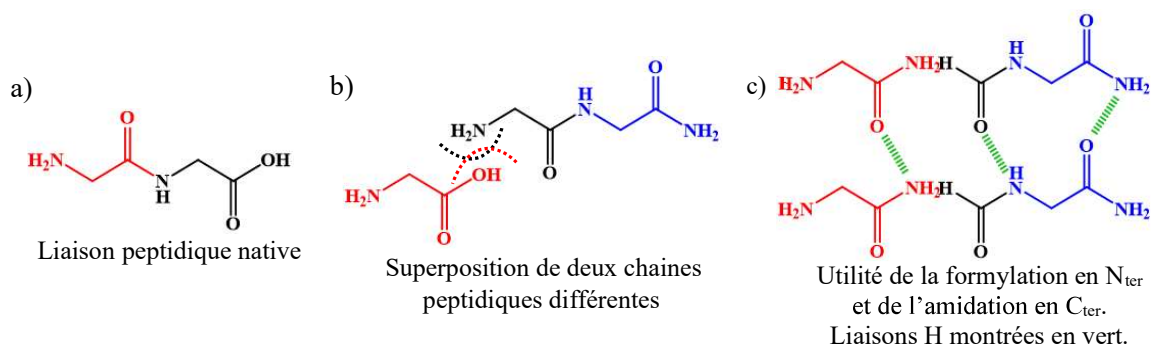


Figure 29 : Conception mettant en évidence l'importance de la suppression d'un résidu avec les modifications associées pour que l'assemblage des briques puisse conduire au même réseau de liaisons hydrogène que dans les L $\beta$ Hs.

Maintenant que la taille de séquence est fixée, il est important de choisir l'endroit où la « coupure » aura lieu, en d'autres termes, à quelle position la brique de base sera stoppée. Dans notre cas, la coupure correspond à un sacrifice d'acide aminé afin d'éviter une superposition d'atomes lors de l'assemblage pouvant

déstabiliser ce dernier (Figure 29 b). Ici, le but est de former des hélices, par autoassemblage et donc par interaction de plusieurs unités peptidiques (briques) entre elles, celles-ci s'enroulant autour d'un même axe. Comme il s'agit de plusieurs unités s'additionnant bout à bout, les N-ter et C-ter se feront face, mais comparé à une séquence linéaire, il y a trop d'atomes au point de contact des deux chaînes (Figure 29 a vs b). Le résultat est que, pour s'assembler en L $\beta$ H canonique, deux atomes, l'oxygène de l'acide carboxylique en C-terminal et l'azote de l'amine en N-terminal, doivent occuper la même position dans l'espace (Figure 29 a et b). Ceci est impossible. Ainsi, dans l'optique de reproduire le plus fidèlement possible ces structures, nous avons décidé de sacrifier une position, c'est-à-dire de concevoir des séquences de 35 résidus au lieu de 36. Celles-ci ont été amidées en C<sub>ter</sub> et formylées en N<sub>ter</sub>. L'idée derrière ces modifications est de préserver le réseau de liaisons hydrogène, lors de l'autoassemblage, malgré un résidu en moins (Figure 29 c).<sup>[10-12]</sup> Il en résulte alors une séquence générique de type Formyle(f)-peptide-Amide(Am).

Pour le résidu à sacrifier, nous avons le choix entre 18 positions (18 résidus d'un tour d'hélice). Comme nous ne pouvons pas toucher au cœur hydrophobe, les positions B<sub>1</sub> et B<sub>3</sub> sont verrouillées. La question s'est posée d'effectuer la coupure en plein milieu du feuillet  $\beta$ , en position B<sub>2</sub> ou B<sub>4</sub>, qui aurait pu se reformer avec la force du cœur hydrophobe et l'effet template de l'assemblage. Nous pouvons aussi supprimer une position dans le coude, T<sub>1</sub> ou T<sub>2</sub> mais le problème identifié ici serait la perte de structuration dans cette partie, ainsi que la perte d'interactions stabilisatrices comme les échelles d'asparagine. La position qui nous a semblé la moins dommageable pour la structuration est donc la position B<sub>4</sub> contenant en majorité des glycines et ne contribuant que pour une liaison hydrogène dans l'assemblage, sans chaîne latérale.

### 3.2.3 Le cœur hydrophobe (étape 3)



Figure 30 : modélisation (Rosetta) du peptide 35m1 montrant le cœur hydrophobe conçu en rouge.<sup>[25]</sup>

Élément central des structures ciblées, ce cœur hydrophobe représente la seconde interaction la plus importante en ce qui concerne la cohésion de l'assemblage.<sup>[13,14]</sup> Il est donc impératif de conserver les résidus constituant le cœur aux positions propres B<sub>1</sub> et B<sub>3</sub> (Figure 30). Comme décrit précédemment dans le chapitre 1, la position B<sub>1</sub> est la plus contrainte (en termes d'encombrement) du cœur hydrophobe. Lorsqu'une thréonine ou une serine est à cette position, elle peut former une liaison hydrogène avec la chaîne principale du résidu en B<sub>4</sub>. Néanmoins, les séquences riches en Ser/Thr peuvent s'avérer difficiles à synthétiser par SPPS. C'est pour cela que nous avons choisi de mettre des valines à toutes les positions B1.

Les positions B<sub>3</sub> étant moins encombrées que les précédentes, la liberté des résidus à choisir est plus importante. Il faut cependant prendre en compte le fait que des résidus trop encombrants ne sont pas compatibles avec la structure et risquent même de défavoriser le repliement. C'est pour cela que nous nous sommes contraints à des acides aminés tels que la valine ou l'isoleucine. Comme expliqué dans le chapitre I, un tour avec une

séquence I, I, I en B<sub>3</sub> ne serait pas favorable, alors nous avons choisi d'effectuer une alternance entre valines et isoleucines. Nous avons alors construit une séquence interne, pour les positions B<sub>3</sub> égale à V, V, I, ce qui permet de ne pas observer un cœur trop encombré et est très similaire à ce qui est retrouvé dans les structures naturelles.

### 3.2.4 Rôle des prolines (étape 4)

À la différence des structures naturelles, ayant des tours canoniques compris entre des tours de nucléation/terminaison, la brique de base que nous visons à concevoir sera constituée seulement de tours canoniques. Cela car nous voulons synthétiser des fibres à partir de ces briques (assemblage virtuellement infini). Nous pensons qu'introduire des prolines dans la séquence de 35 acides aminés favorisera le repliement, en plus des autres contributions apportées, notamment au niveau des coudes β. La proline induisant un angle permettant de courber la chaîne principale, le coude β sera favorisé par l'introduction de ce résidu.<sup>[15,16]</sup> Cependant, les protéines naturelles n'ont pas de proline à chaque position T<sub>1</sub>. Nous avons choisi d'introduire seulement deux prolines sur 35 résidus, une par pas d'hélice, dans le T<sub>1</sub> du coude orienté vers l'intérieur du trimère (III) comme montré en Figure 31. La formation du cœur hydrophobe et des feuillets β aidant à replier les autres coudes n'ayant pas de proline, laisse de la place pour d'autres résidus à ces positions T<sub>1</sub>.

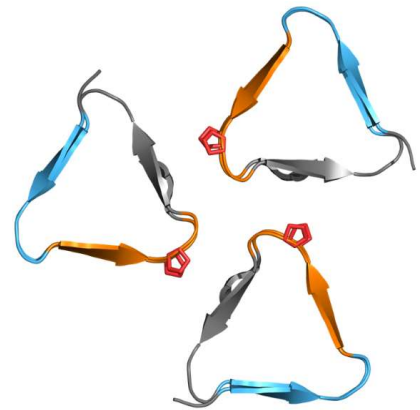


Figure 31 : Insertion des prolines dans les coudes β de l'intérieur du trimère par modélisation et reconstitution de 3 hélices.

### 3.2.5 Importance des glycines et lysines (étape 5)

Tout comme les prolines aux positions T<sub>1</sub>, les glycines présentes en position B<sub>4</sub> sont importantes quant au repliement de l'hélice β gauche. En effet, le rôle des glycines dans l'assemblage, ici, est d'apporter un peu de flexibilité, mais aussi car la contrainte stérique à cette position est importante comme il l'a été décrit précédemment.<sup>[13]</sup> Nous avons alors gardé la majorité de ces positions avec des glycines, sauf une, avec une lysine, qui pointe vers le solvant. Cette lysine a également été apportée afin d'avoir une chaîne longue et prévenir d'autres hélices de venir s'agglutiner aux autres, le but étant de garder une forme trimérique et non agrégée latéralement (Figure 32). De plus, cette lysine apportera une charge (+) supplémentaire au système.

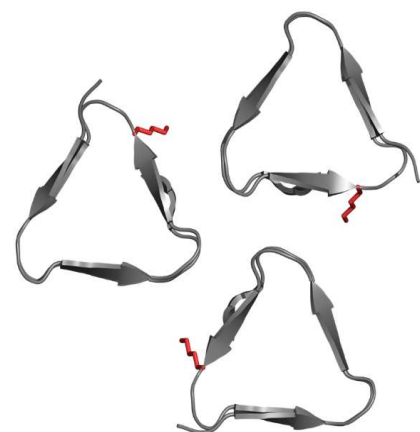


Figure 32 : Modèle montrant le placement et l'orientation des lysines (rouge) pour le trimère d'hélices.

### 3.2.6 Complémentarité de charges entre les tours (étape 6)

Afin de former un assemblage stable et de favoriser son repliement en L $\beta$ H, nous avons incorporé dans la séquence, des résidus chargés, de manière à ce que le repliement en L $\beta$ H rapproche dans l'espace les charges complémentaires.<sup>[17]</sup> Cela est retrouvé dans les structures naturelles, mais assez peu. Nous avons choisi d'alterner des résidus de charges positives et négatives, aux mêmes positions, mais à un tour n+1. Des paires complémentaires aux positions T<sub>1</sub> (II), et B<sub>2</sub> (III), ont été sélectionnées, toutes deux faisant face au solvant et n'étant pas encombrées. Ces interactions intra-moléculaires ont aussi pour but de préorganiser le peptide monomérique. Notre choix s'est tourné sur le couple Glu-Arg et Glu-Lys respectivement (Figure 33).

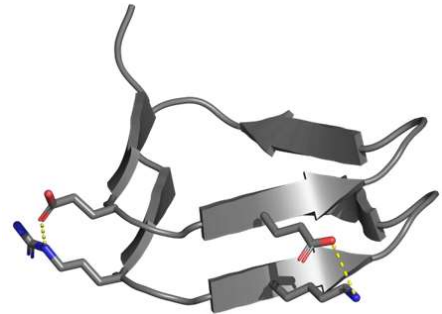


Figure 33 : Illustration de la complémentarité de charges des deux tours, introduites aux positions B<sub>2</sub> et B<sub>4</sub>.

L'acide glutamique, ayant un pKa d'environ 4.15, est préféré à un acide aspartique pour des raisons de facilité de synthèse (ce dernier ayant tendance à former des aspartimides lors de la SPPS).<sup>[18]</sup> La lysine et l'arginine ayant des pKa respectifs d'environ 11 et 12, apportent chacune une charge positive à l'assemblage, et vont interagir avec les glutamates lorsque les conditions de pH se trouvent entre 4 et 11. L'autre raison de l'incorporation de ces résidus est l'apport de charges positives afin d'avoir une molécule qui soit soluble en conditions acides (pH=2), dans le but de faciliter la purification du peptide.

### 3.2.7 Formation d'échelles d'asparagines (étape 7)

L'addition de résidus à chaîne latérale amidée (permettant de faire des liaisons hydrogène) est un facteur stabilisant la cohésion de l'assemblage, que ce soit du monomère replié, ou de la fibre. Le rôle de ces résidus est important (i.e. échelle d'asparagines) lorsque la question de la conception d'objets répétitifs par autoassemblage se pose. Ils forment un réseau de liaisons hydrogène important en gardant l'assemblage solidaire. Ici, l'asparagine a été utilisée afin de former ce réseau. Nous en avons ajouté à deux positions que nous avons identifiées comme importantes en Chapitre 1. La première est la position T<sub>2</sub> des deux coudes  $\beta$  n'étant pas situés à l'intérieur du trimère.

Une autre question importante à se poser est le degré de liberté de la chaîne peptidique en N<sub>ter</sub> et en C<sub>ter</sub>. En effet, les extrémités sont très souvent plus libres et moins contraintes, donc plus flexibles, ce qui peut s'avérer très utile lors de la chélation de métaux ou la reconnaissance de protéines, substrats, etc...<sup>[19,20]</sup> Or dans notre cas, nous souhaitons avoir ces parties terminales les moins libres possibles afin de pouvoir former une hélice  $\beta$  gauche stable. Nous voulons alors verrouiller ces résidus dans un réseau de liaisons hydrogène, c'est pour cela que des asparagines ont été ajoutées en position T<sub>1</sub> (I) qui correspond à la première position du N<sub>ter</sub>. Nous pensons qu'en formant des échelles d'asparagines à ces positions, la liberté des motifs terminaux diminue et l'hélice sera plus stable. En N<sub>ter</sub>, les modélisations Alphafold et Rosetta ne montrent pas la formation d'échelles

avec seulement deux asparagines, cependant, sous forme de fibre, celles-ci nous montrent bien la formation d'échelles (Figure 34). À noter que le motif terminal présent en  $C_{ter}$  est ici stabilisé par le cœur hydrophobe, puisque le dernier résidu est en position  $B_3$ .

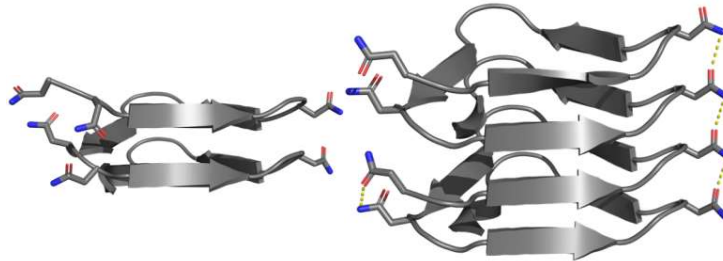


Figure 34 : Modèle d'échelle d'asparagines à gauche sur un monomère (non structuré) et à droite sur un dimère autoassemblé (structuré).

### 3.2.8 Histidines au centre du trimère (étape 8)

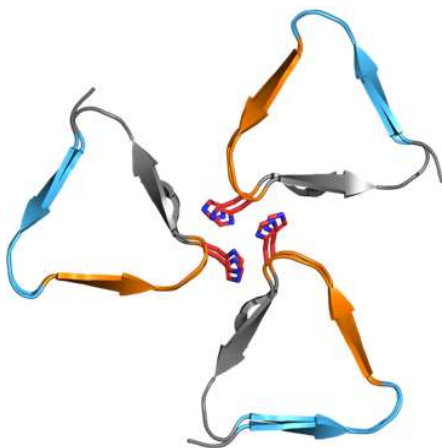


Figure 35 : Visualisation des histidines au centre du trimère (modèle).

Les positions tournées vers l'intérieur du trimère d'hélice sont peu nombreuses ( $T_1$  (III)). Les résidus présents à ces positions doivent alors être bien choisis dans le but de former et favoriser la formation sous forme trimérique des hélices. Favoriser l'assemblage trimérique revient à considérer des résidus avec des chaînes latérales pas trop longues et neutres. Pour cela, les A.As hydrophobes pourraient être un choix judicieux, mais le contrôle de l'assemblage est beaucoup moins favorable même si la formation d'un cœur entre les hélices stabiliserait la forme trimérique. Un autre problème serait aussi la solubilité des monomères avec des résidus hydrophobes supplémentaires. Afin d'éviter cela, un résidu avec une chaîne latérale pouvant être chargée ou neutre en fonction du pH et pas trop longue semble être une bonne option. C'est pour cette raison qu'il nous a paru logique de choisir

l'histidine comme résidu à l'intérieur du trimère (Figure 35). Elle a la capacité de former des empilements de type  $\pi$  et peut permettre de contrôler l'assemblage via un changement de pH.<sup>[21,22]</sup> Lorsque les histidines sont protonées, l'assemblage devrait être défavorisé, au moins sous forme trimérique. Ce qui permet de garder la solution de monomères stables, à un pH acide, sans avoir besoin de prendre trop de précautions. Puis, lorsque le pH est augmenté (au-dessus de 6), les fibres devraient alors se former.

### 3.2.9 Les positions B<sub>2</sub> non attribuées à l'étape 6 (étape 9)

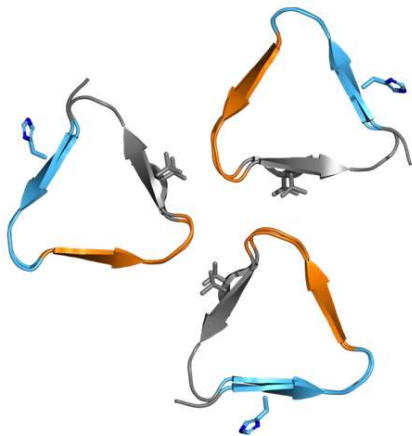


Figure 36 : Modèle trimérique des résidus choisis pour les positions B<sub>2</sub> restantes.

Les résidus présents à ces positions ont été choisis en fonction de la face à laquelle ils appartenaient. Il ne restait à attribuer que quatre positions B<sub>2</sub> sur six, une sur l'hexade (II) et deux sur l'hexade (I). Nous avons choisi de mettre une valine à chacune des positions B<sub>2</sub> de la face (I), qui est retrouvée beaucoup dans les séquences naturelles (Tableau 6), mais aussi car ici, cette position fait directement face à un coin d'une hélice (T<sub>1</sub> de la face (III) soit des prolines). Ce coin possédant une proline, nous avons pensé qu'une valine serait un résidu approprié à la formation trimérique des hélices. En ce qui concerne l'hexade (II), notre choix s'est porté sur l'histidine car elle fait face au solvant (i.e. nécessité d'un acide aminé polaire) et permet d'apporter une charge positive que l'on peut contrôler en fonction du pH (Figure 36).

### 3.2.10 Brique élémentaire finale

La brique élémentaire résultant de toutes les considérations précédentes, et se basant sur l'outil élaboré en chapitre 1, est alors une séquence de 35 acides aminés, notés 35m1 (Figure 37). Nous nous attendons à des peptides qui ont une pré-organisation relativement importante, par formation du cœur hydrophobe et des feuillets β parallèles intra-moléculairement, ainsi que par la complémentarité de charge entre les tours. Aussi, les deux séquences ont été conçues de manière à ce qu'à pH acide, les peptides soient solubles et l'assemblage défavorisé. En revanche, à un pH au-dessus de 6-6.5, la formation d'autoassemblage de peptide et son repliement en hélice β gauche (donc de fibres) devraient être favorisés. Tout cela devrait permettre de former des fibres par autoassemblage de ces briques élémentaires en hélice β gauche trimériques.

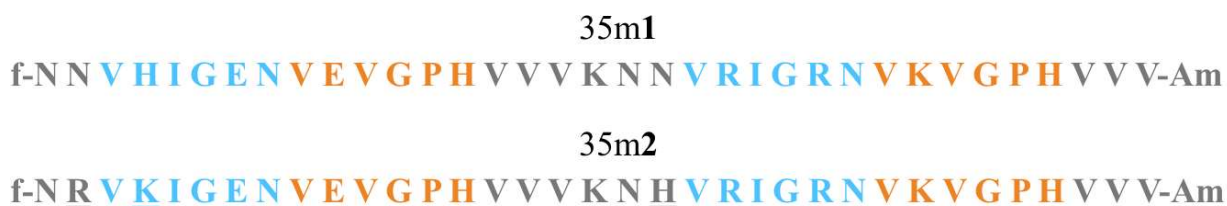


Figure 37 : Séquence des deux peptides conçus 35m1 et 35m2 avec les mutations soulignées.

### 3.2.11 35m2 : Mutation du 35m1 pour défavoriser l'agrégation latérale (étape 10)

Au cours des différentes caractérisations effectuées pour le peptide 35m1, nous nous sommes rendu compte que ce dernier avait une forte tendance à former des faisceaux de fibres relativement courtes agrégées latéralement (voir plus bas, microscopie électronique). Malgré la recherche de conditions défavorisant ce phénomène, nous n'avons pas réussi à l'éviter. Cela indique clairement un assemblage latéral plus important que trimérique. Nous avons alors conçu une séquence peptidique, 35m2, comportant trois mutations, soulignées en étape 10 du Tableau 9 et Figure 37, par rapport au peptide de départ. Ces mutations ont été faites pour défavoriser l'assemblage latéral des hélices au-delà du trimère. Cependant, ceci requiert des sacrifices au niveau des interactions servant à stabiliser l'autoassemblage comme les échelles d'asparagines. Nous avons remplacé deux Asn en position T<sub>1</sub> (I), aux positions 34 et 16, par une lysine et une histidine respectivement. Cela dans le but d'apporter des charges positives supplémentaires pour repousser d'autres éventuelles hélices pouvant venir s'y coller. De même, l'histidine en position 31 a été mutée par une lysine, plus encombrante et ayant un pKa plus élevé, ce qui permettrait de « chasser » les autres hélices. Cela a pour but d'augmenter la charge globale du peptide à pH neutre passant de + 4 pour le 35m1 à + 6 pour le 35m2. Ces différences sont illustrées en Figure 38.

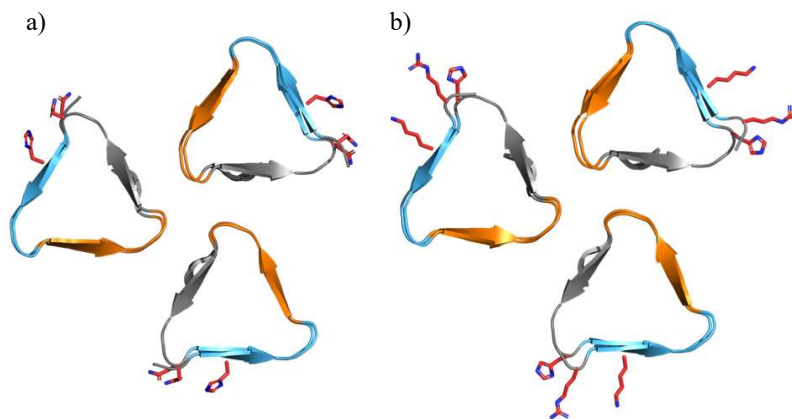


Figure 38 : Modèle trimérique des peptides 35m1 (a) et 35m2 (b) montrant les mutations (rouge) entre les deux.

### 3.2.12 Modélisation

La modélisation est un outil intéressant lors de la conception de nouvelles séquences peptidiques visant à adopter une conformation particulière. Elle est notamment très utilisée afin de comprendre les interactions entre plusieurs protéines, ou ligands, ainsi que la conformation de complexes. Dans notre cas, les logiciels de modélisation tels que Rosetta<sup>[23]</sup> ou plus récemment Alphafold<sup>[24]</sup> permettent une prédiction du repliement. C'est ce que nous avons fait pour les deux peptides précédents afin de vérifier la conception, et la pertinence



du modèle. Nous avons obtenu une prédiction de repliement en hélice  $\beta$  gauche pour chacun des peptides, que ce soit pour un monomère ou pour plusieurs molécules. La précision des prédictions a été très élevée (> 95 % de confiance) traduisant des prédictions fiables. Ces modèles nous indiquent que les peptides sont susceptibles de s'autoassembler en L $\beta$ Hs, mais aucune information n'est donnée quant à l'association des hélices en trimères (Figure 39). Ce qui peut se comprendre par le fait que les interactions principales sont les empilements entre peptides et non les interactions entre monomères au sein du trimère. Les trimères visibles sur les figures ci-dessus (Figure 38 a et b), ont été « trimérisés » manuellement en les alignant aux structures existantes. Il faut cependant noter que certains modèles réalisés ont donné des conformations en hélices  $\beta$  droites, mais avec un score de confiance bien moins élevé (environ 70 % au maximum). Nous pensons que cela est dû aux paramètres de modélisation, qui ont été pris par défaut via les plateformes Robetta<sup>[25]</sup> et colabfold<sup>[26]</sup>, mais aussi car ces algorithmes de prédiction ne sont pas optimisés pour des objets dont le repliement est stabilisé par autoassemblage comme dans notre cas, mais plus pour des objets plus gros et globulaires comme des protéines solubles.

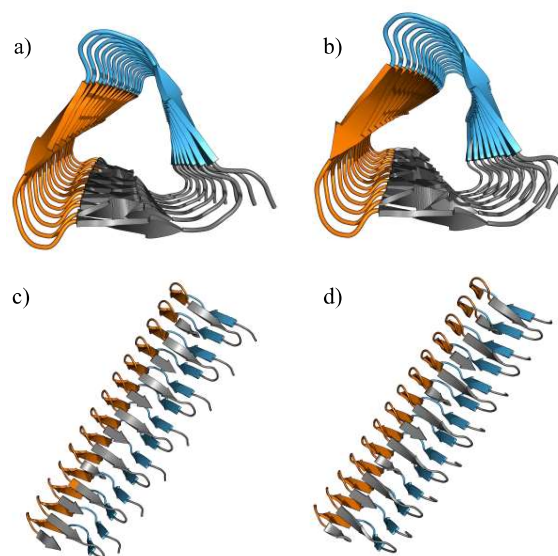


Figure 39 : Modélisation AlphaFold de 7 monomères pour les peptides a et c) 35m1 et b et d) 35m2.

## 3.3 Production des peptides

### 3.3.1 Synthèse des briques de base par SPPS

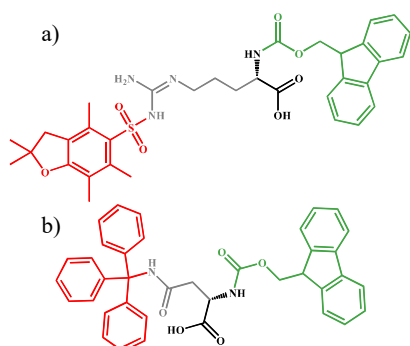


Figure 40 : Fmoc-arginine avec sa chaîne secondaire protégée avec un Pbf (a) et Fmoc-asparagine avec sa chaîne latérale protégée avec un tryptile (b).

Les peptides conçus ont été synthétisés par synthèse peptidique en phase solide (SPPS) suivant la stratégie Fmoc/tBu classique, à l'aide d'un robot de synthèse.<sup>[27]</sup> Les synthèses sont décrites dans la partie matériel et méthodes. Comme décrit dans l'étape 2 du Tableau 9, les séquences conçues sont amidées en C<sub>ter</sub>, et formylées en N<sub>ter</sub>. Dans un premier temps, nous avons tenté de formyler le N<sub>ter</sub> par la N-formylsaccharine, après clivage de la résine, mais nous n'avons pas réussi à obtenir ce composé en quantité suffisante. Ainsi, la formylation des peptides s'est alors réalisée « sur résine », avec de l'acide formique, DIC et DIEA. Lors des synthèses, nous avons remarqué que certains couplages en phase solide étaient très difficiles à faire, et étaient responsables d'une perte de

rendement. C'est notamment le cas du couplage arginine-asparagine (R-N) qui a lieu au 11<sup>e</sup> couplage. Ce couplage cause une perte de rendement d'environ 30% de la charge totale (estimé par HPLC) et lors de l'analyse LC-MS du peptide brut, nous avons remarqué qu'un pic correspondant au peptide acétylé après l'asparagine 10 représentait une part importante du peptide avec une masse de 1088.50 g/mol. Nous pensons que ceci vient d'une gêne stérique entre les groupements protecteurs des chaînes latérales de ces acides aminés. En effet, l'acide aminé à réagir étant protégé sur sa chaîne principale par un groupement Fmoc et en chaîne latérale par divers groupements, l'arginine par un groupement 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) et l'asparagine par un groupement trityle comme illustré en Figure 40. Nous avons pu résoudre cela en préactivant et augmentant le temps de couplage. Pour preuve, comparée aux synthèses précédentes, la mesure des groupements Fmoc après synthèse nous a donné un rendement juste après synthèse qui a été amélioré de 2.8 fois et le ratio entre les deux pics correspondant au peptide et au peptide tronqué a augmenté en faveur du peptide (Figure 41). Les rendements totaux ont été, pour le peptide 35m1 et 35m2 sont de 4 à 5.5 % et de 5 à 6.8 %, respectivement.

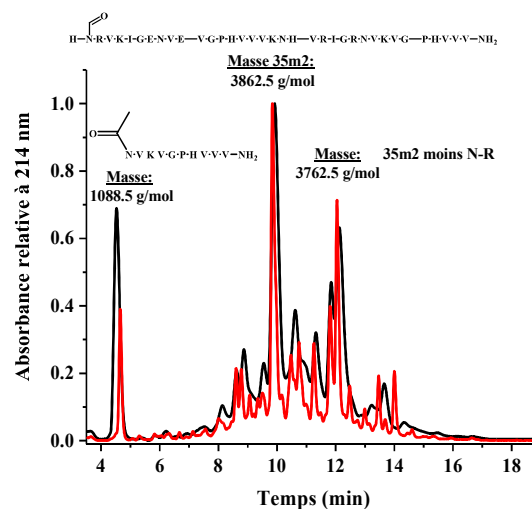


Figure 41 : Chromatogrammes bruts du peptide 35m2 à 214 nm acquis avec une méthode 20-40% de B en 30 min à 1 mL/min à 60°C sur colonne C4. En noir, la première synthèse, en rouge, la synthèse optimisée montrant une amélioration du couplage n° 11 R-N.

### 3.3.2 Purification par HPLC semi-préparative

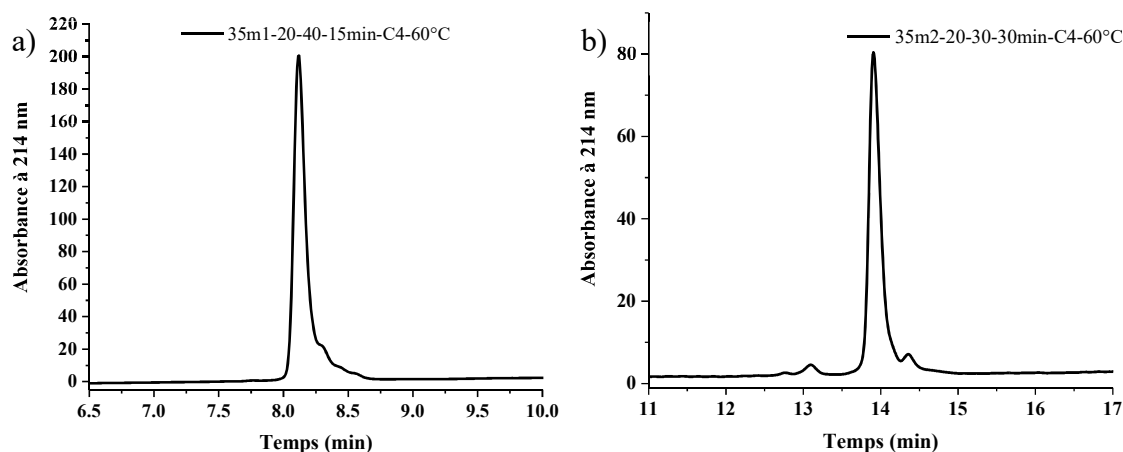


Figure 42 : Chromatogrammes HPLC après purification des peptides étudiés montrant une pureté finale de 90 % pour le 35m1 et 89 % pour le 35m2.

Des puretés estimées par HPLC (par intégration des aires sous courbes) de peptides allant jusque 90 % et 89 % pour les peptides 35m1 et 35m2 ont pu être obtenues (Figure 42). Grâce à ces puretés relativement

élevées, nous avons pu utiliser ces peptides pour de plus amples analyses afin de déterminer leurs structures fibrillaires.

### 3.3.3 Caractérisations

Avec les peptides *de novo*, il n'est pas toujours trivial de confirmer que la structure obtenue est bien celle imaginée/conçue. C'est tout le but de cette partie, et plus généralement de ma thèse, déterminer comment, et par quelles techniques nous pouvions obtenir des informations structurales de ces peptides. Mais peu de techniques permettent d'avoir la structure tridimensionnelle. Nous avons alors adopté une approche combinatoire de plusieurs méthodes de caractérisation dans le but d'avoir la structure de l'autoassemblage. En effet, une fois la structure résolue, le but serait d'avoir des signatures spectroscopiques, et microscopiques spécifiques de ce type d'assemblage, en hélice  $\beta$  gauche de type I. Ceci permettrait à l'avenir de les identifier plus facilement, par des techniques simples, disponibles au laboratoire. Cela revient à dresser une carte d'identité de ce type de structure, à laquelle nous pourrions faire référence par la suite. Les techniques utilisées ci-après sont l'émission de fluorescence de la thioflavine-T (ThT), la microscopie électronique à transmission (TEM et Cryo-TEM), le dichroïsme circulaire (CD) et l'infrarouge à réflectance totale atténuée (ATR-IR). Chacune nous permettant de confirmer les résultats des autres techniques et de vérifier la pertinence des analyses réalisées. Au vu des caractéristiques structurales de l'objet étudié, nous nous attendons à des signatures similaires aux amyloïdes pour la fluorescence de la ThT<sup>[28]</sup>, à des fibres régulières et homogènes de section définie par TEM<sup>[2,7]</sup>, à des informations sur l'arrangement des feuillets  $\beta$  dans la fibre (voir obtenir la structure) via Cryo-TEM<sup>[29]</sup>, à des indications sur la structure secondaire en feuillets  $\beta$  par CD<sup>[30,31]</sup> ainsi qu'en ATR-IR<sup>[32,33]</sup> (notamment l'orientation des feuillets et la présence de coudes  $\beta$ ).

#### 3.3.3.1 Suivi cinétique et mécanisme d'assemblage avec la thioflavine-T

La thioflavine-T est une molécule utilisée pour suivre la formation de fibres amyloïdes ainsi que leur mécanisme d'association.<sup>[28]</sup> En solution, la rotation de la liaison C-C entre les cycles aniline et benzothiazole de cette molécule possible. La molécule est alors peu fluorescente. En revanche, lorsque celle-ci est liée aux fibres amyloïdes, la rotation n'est plus possible, ce qui permet l'émission de fluorescence de la ThT. Le mécanisme proposé le plus probable est l'interaction de la ThT avec des sillons, plutôt hydrophobes, formés par les résidus le long des feuillets  $\beta$ .<sup>[34]</sup> C'est une technique très utilisée et cela peut permettre d'obtenir des informations sur le mécanisme de formation des fibres. Un comportement modèle peut être décomposé en trois parties différentes, une phase dite latente, une de croissance et le plateau final (Figure 43). L'intensité de la fluorescence est alors dépendante de la quantité de fibres et de ThT, en solution. Il est important de noter ici que la concentration de peptide ou protéine doit être supérieure à la concentration de fibrillation critique afin de pouvoir observer la formation de fibres.

La première étape, la phase latente, correspond au temps de formation de nucléi ou d'agrégats solubles non détectables, aussi appelée nucléation primaire. Ici un nucléi est défini comme la formation de structures pré-fibrillaires solubles. C'est à partir de ces nucléi que la fibre pourra croître en objet de taille plus importante. Dans notre cas, l'étape limitante pourrait être le repliement du peptide passant de la forme dépliée à repliée en brique de LβH. Cette étape est poussée ici par la formation du cœur hydrophobe et des feuillets β constituant chaque Lego (au nombre de 6 par peptide). Dans cette étape, les molécules de thioflavine-T ne se lient pas, ou très peu aux feuillets β.

La seconde étape du processus de formation de fibres amyloïdes est la phase de croissance. C'est lors de cette phase que les fibres vont croître par association de molécules monomériques sur les nucléi et les fibres en croissance.

Enfin, la dernière phase correspond à l'équilibre thermodynamique atteint lors de la fibrillation. Elle est représentée par un plateau d'émission de fluorescence de la ThT. Dans cette étape, les fibres sont (théoriquement) stables et ne se dissocient pas. Cependant, les fibres peuvent se dissocier, ce qui a pour conséquence une décroissance de l'émission où le plateau n'est plus stable. Plusieurs facteurs peuvent causer cela, comme la stabilité des fibres, la température, la force ionique ou encore le pH élevé de la solution. Une hypothèse serait que dans certains cas, l'agrégation latérale des fibres en croissance pouvant chasser les molécules de ThT déjà présentes, ce qui provoquerait une chute de l'émission de fluorescence (voir schéma en Figure 44).<sup>[35]</sup>

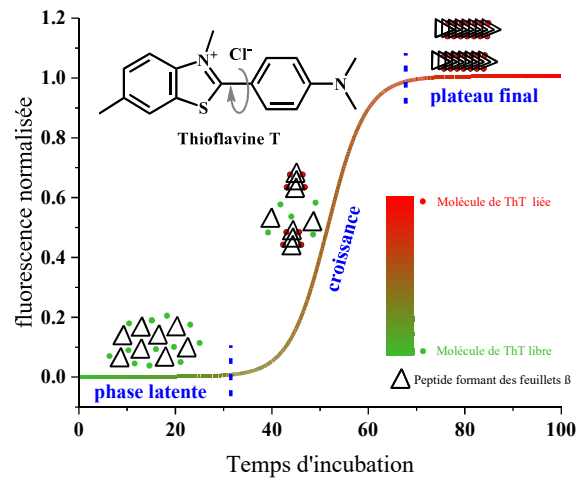


Figure 43 : Courbe d'émission de fluorescence typique de la thioflavine-T pour un comportement de type amyloïde ( $\lambda_{ex}=450$  nm ;  $\lambda_{em}=465$  nm).

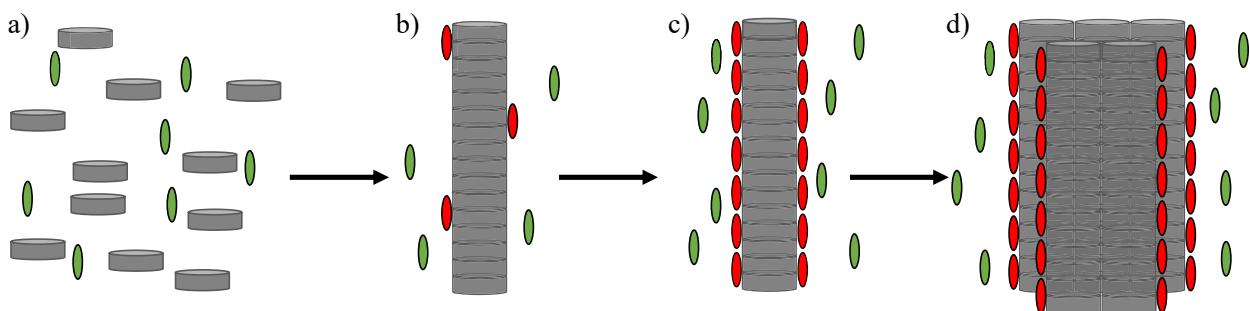


Figure 44 : Représentation schématique de la fibrillation en présence de molécules de ThT. a) étape de nucléation avec les monomères en solution (gris) et les molécules de ThT libres (vert), b) phase de croissance, formation d'une fibre et interaction de molécules de ThT (rouge), c) plateau symbolisé par une émission de fluorescence constante liée aux molécules de ThT liées aux fibres et d) baisse d'émission liée à l'agrégation latérale des fibres entre elles, chassant les molécules de ThT.

Pour les peptides 35m1 et 35m2, nous nous attendons à un comportement de type amyloïde caractérisé par les trois phases présentes ci-dessus. En revanche, nous n'aurons aucun indice en ce qui concerne la structure

tridimensionnelle adoptée par les peptides ni même concernant la formation de trimère d'hélices par cette technique.

### 3.3.3.2 Évolution de la fluorescence de la Thioflavine-T en présence des peptides 35m1 et 35m2

Les solutions stock de peptides de 2 mM, ont été préparées par dissolution de poudres lyophilisées dans l'eau en conditions acides. Les deux peptides ont été utilisés directement à partir des solutions stock à un pH d'environ 3.5. Les solutions de ThT à 2.2 mM, préparées par dissolution dans l'eau et dosage à 412 nm ( $\epsilon=36000 \text{ L.mol}^{-1}.\text{cm}^{-1}$ ) ont un pH d'environ 5,7.<sup>[36]</sup> Les solutions de tampon utilisées ont été préparées à 500 mM stock et leur pH ajusté. Dans les puits, les concentrations finales en peptide ont été de 1 à 100  $\mu\text{M}$  de 200  $\mu\text{M}$  pour la ThT et de 50 mM en tampons. L'addition dans les puits s'est faite par ajout des peptides puis de la ThT, de l'eau et enfin l'ajout des tampons. Cela car les peptides sont conçus pour se replier et s'assembler seulement à un pH supérieur à celui du pKa des histidines et l'ajout du tampon serait alors censé gouverner l'assemblage. Ces solutions ont alors été incubées à 37°C dans une plaque 96 puits scellée et suivi au lecteur de plaque par mesure de l'émission de fluorescence.

#### 3.3.3.2.1 Fibrillation en fonction du temps

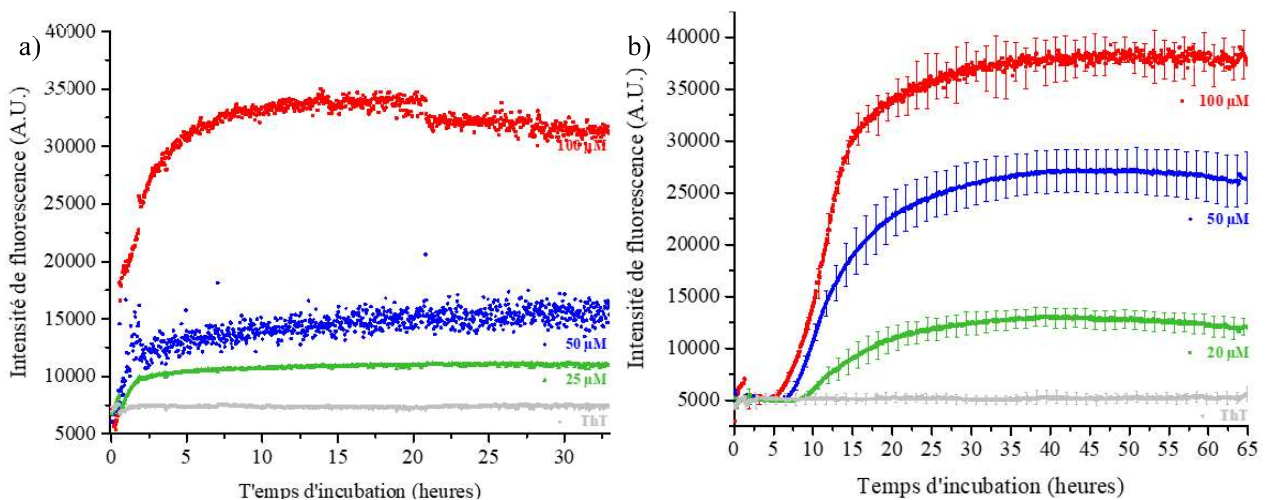


Figure 45 : Suivi de l'émission de fluorescence de la thioflavine-T au cours du temps pour les peptides à pH=8.5 a) 35m1 et b) 35m2 qui sont une moyenne de 3 expériences différentes.  $t=0$  heure correspond au temps d'ajout du tampon aux puits contenant le peptide + la ThT.

La formation de fibres a été suivie par incubation des peptides directement au lecteur de plaque à une température de 37°C à un pH de 8.5 afin d'être sûr de la déprotonation des histidines : la charge au centre du trimère (hypothétique) passe de (+6) à 0, ce qui autorise la formation de fibres sous forme trimérique. La formation de fibres est aussi favorisée par la formation de ponts salins à ce pH. La ThT est excitée à 450 ( $\pm 15$

nm) et la fluorescence émise à 480 ( $\pm 20$  nm) est enregistrée en fonction du temps pour suivre la fibrillation des deux peptides 35m1 et 35m2 (Figure 45). Les deux peptides montrent deux comportements différents. Le peptide 35m1 (a) a un comportement s'apparentant à une solution de peptideensemencée avec des fibres déjà formées : l'aspect de la courbe est concave. Le comportement est très différent pour le 35m2, plus classique, avec une phase latente. Nous pouvons aussi remarquer que l'intensité de fluorescence au plateau final correspondant à l'équilibre thermodynamique est globalement la même, ce qui laisse suggérer que la fibrillation des peptides est similaire. En d'autres termes, la quantité finale de fibres formées est similaire. Le mécanisme de fibrillation pourrait être un mécanisme dominé par une nucléation secondaire multi-étapes, ce qui correspond aux fibres déjà formées catalysant la formation de fibres parallèlement à celles-ci.<sup>[37,38]</sup> Nous pouvons cependant observer une baisse de l'émission de fluorescence qui pourrait être due à un mécanisme de fragmentation des fibres qui seraient trop longues, ou alors tout simplement due à la ThT qui pourrait s'hydroxyler en milieu basique (pH=8.5) après un temps important.<sup>[39]</sup> Pour le peptide 35m2, la limite d'agrégation a été déterminée par fluorescence comme étant comprise entre 1 et 5  $\mu$ M alors que pour le peptide 35m1 sa limite d'agrégation est inférieure à 1  $\mu$ M. Cela est aussi compatible avec l'absence de phase latente pour le peptide 35m1 qui suggère une cinétique d'agrégation bien plus rapide que le 35m2. Dans sa globalité, les essais de fluorescence de la Thioflavine-T nous montrent la formation de fibres de type amyloïdes, en feuillets  $\beta$ .

Afin de mieux comprendre le comportement des peptides et si la formation de fibres était bien contrôlée par le pH, nous avons réalisé plusieurs expériences. Une a été réalisée par l'ajout du tampon en fonction du temps (HEPES pH=8.5). Nous avons choisi d'ajouter le tampon et non le peptide pour plusieurs raisons, la première étant de vérifier que l'augmentation de fluorescence était bien due à la fibrillation du peptide en conditions basiques. Ensuite, cela permet de vérifier la stabilité des peptides sur un temps donné : vérifier qu'en conditions acides, il n'y a pas de formation de nucléi. À noter que ceci ne se vérifiera que dans le cas où la formation de nuclei est l'étape cinétiquement déterminante.

Cette expérience de fibrillation pour le premier peptide (35m1) est montrée en Figure 46. Nous voyons que les comportements du peptide à trois temps d'ajout différents sont très similaires sinon les mêmes. Nous remarquons qu'à chaque ajout de tampon, autrement dit, qu'à chaque augmentation de pH passant d'un pH env. 5 à un pH de 8.5, l'émission augmente de manière significative très rapidement jusqu'à arriver à un plateau cinq heures après l'ajout. Comme précédemment, nous confirmons la cinétique très rapide pour ce peptide avec un  $t_{1/2}$  (qui correspond au temps où la fluorescence est à la moitié de sa valeur max) d'en moyenne 35 min (à 100  $\mu$ M). Nous pensons cependant que le phénomène n'est pas

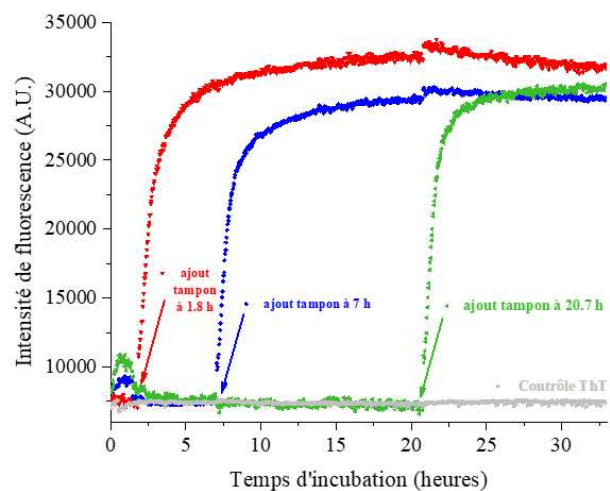


Figure 46 : Cinétique de fibrillation du peptide 35m1 à 100 $\mu$ M suivie en émission de fluorescence de la ThT (200  $\mu$ M), avec des ajouts à temps différents de tampon HEPES à pH=8.5 (50 mM). 37°C, avec agitation orbitaire.

dû à la présence de nucléi en solution au début de l'expérience, car nous voyons que la fluorescence n'augmente pas en fonction du temps si le pH n'est pas augmenté. Nous avons aussi calculé la pente de la phase de croissance pour les différents temps d'ajouts de tampon et nous observons les mêmes valeurs de pentes (Tableau 10). Les valeurs sont du même ordre de grandeur, ce que nous interprétons comme une concentration de nucléi constante au cours du temps. Ceci est montré pour une concentration de 100  $\mu$ M, mais pour d'autres concentrations le comportement est le même.

Tableau 10 : Régression linéaire de la pente pour le peptide 35m1 à 100  $\mu$ M en fonction du temps d'ajout de tampon.

Équation	$y = a + b \cdot x$				
Temps d'ajout	0 h	0.5 h	1.8 h	7 h	20.7 h
Pente (b)	$260 \pm 20$	$290 \pm 20$	$250 \pm 20$	$240 \pm 20$	$270 \pm 20$

Le peptide 35m2 a lui un  $t_{1/2}$  d'environ 11.6 heures pour une même concentration lors de la mesure. Nous pouvons interpréter cela comme une préorganisation plus rapide et une formation de l'assemblage plus rapide du 35m1 comparée au peptide 35m2 qui a plus de répulsion de charges et donc peut-être moins de préorganisation. Cela pourrait être dû aux résidus choisis qui ont été apportés dans la séquence pour repousser les autres hélices, ceci ralentissant la formation des fibres. Nous remarquons aussi que pour le peptide 35m2, le comportement est le même concernant la phase de croissance même si les valeurs de pentes sont différentes, lorsque la phase de latence est exclue.

### 3.3.3.2.2 Fibrillation en fonction du pH

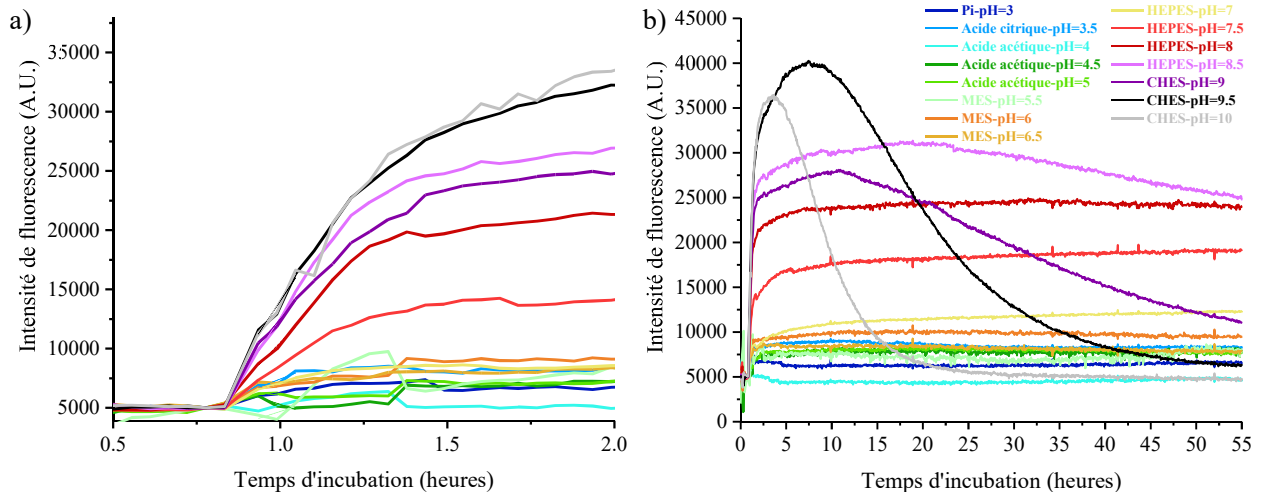


Figure 47 : Suivi cinétique de formation de fibres avec le peptide 35m1 en fonction du pH montrant la partie linéaire juste après ajout (a) et vue globale (b).

Les peptides synthétisés ayant été conçus pour s'autoassembler en fonction du pH, il est important pour nous de regarder le comportement de la fibrillation en fonction du pH par suivi de l'émission de fluorescence

de la ThT. Des solutions stock de tampons à des pH allant de 3 à 10 avec des incréments de 0.5 (tampons et préparations détaillés en partie matériel et méthodes) ont été préparées. Pour chacune des conditions, le puits est préparé par ajout de peptide et ThT dans un milieu non tamponné, donc acide, et la fluorescence est suivie pendant 50 minutes. Au bout de 50 minutes, les tampons, pour fixer le pH à la valeur voulue, ont été rapidement ajoutés, considérant la rapidité de formation des fibres, et la plaque a été remise à incuber. Nous pouvons voir sur la Figure 47 que les comportements, en fonction du pH, sont très différents. En effet, nous observons une augmentation de fluorescence de la ThT de manière significative à partir d'un pH de 7.5. En deçà, la formation de fibres se fait très peu. Nous observons une augmentation de fluorescence, mais très faible, ce qui traduit la formation d'agrégats de petite taille ou de précipités. Nous remarquons également que les valeurs de pente pour la cinétique à des pH supérieurs à 8-8.5 sont du même ordre de grandeur comme montré en Tableau 11. Cela traduit une cinétique de formation des fibres similaire pour des pH supérieurs à 8. Nous observons aussi que plus le pH est élevé, plus l'intensité de la fluorescence est importante, mais moins longue. Ce qui pourrait signifier que les histidines, supposées au centre du trimère, forment des empilements rapides, ce qui amène la formation de fibres. Mais ce peut aussi être dû à une hydroxylation de la ThT sur le carbone induisant la rotation, mais de même à une stabilité des fibres qui serait moins favorable à ces pH.<sup>[40]</sup> Cela pourrait de plus venir du fait que l'assemblage se forme rapidement et que très vite d'autres fibres viennent se coller aux précédentes, ce qui chasserait les molécules de ThT déjà présentes le long des feuillettes (Figure 44).

Tableau 11 : Calcul de pente de la section linéaire en fonction du pH pour le peptide 35m1.

Equation	$y = a + b*x$														
pH	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5	8	8.5	9	9.5	10
Pente (b)	70 ± 25	90 ± 25	80 ± 25	-30 ± 25	-20 ± 25	320 ± 25	60 ± 25	20 ± 25	100 ± 25	290 ± 25	540 ± 25	700 ± 25	510 ± 25	690 ± 25	680 ± 75

### 3.4 Observation de la morphologie par microscopie électronique

La microscopie électronique à transmission (TEM) nous permet d'observer la morphologie des objets formés, dont au moins une des dimensions est sub-micrométrique.<sup>[41,42]</sup> Elle permet de nous renseigner sur la forme, la longueur, la largeur, l'homogénéité, et la possible agrégation des objets. Cette technique peut même, si tant est qu'on ait des échantillons qui s'y prêtent, nous fournir directement la structure via analyse à basse température (Cryo-TEM).<sup>[29,43,44]</sup> D'après les caractéristiques des hélices  $\beta$  gauche de type I, la section d'une hélice seule est d'environ 2.5-2.7 nm et la section du trimère serait d'environ 4.5-5 nm. Nous nous attendons alors à des fibres d'environ 5 nm de diamètre, plutôt cylindriques et de longueur importante (quelques centaines de nm -  $\mu$ m) si la conception a bien été réalisée. Nous nous attendons également à une possible périodicité du système, même si ce n'est pas toujours le cas dans les structures naturelles, cela n'est donc pas primordial.



### 3.4.1 Observations TEM et optimisation des conditions

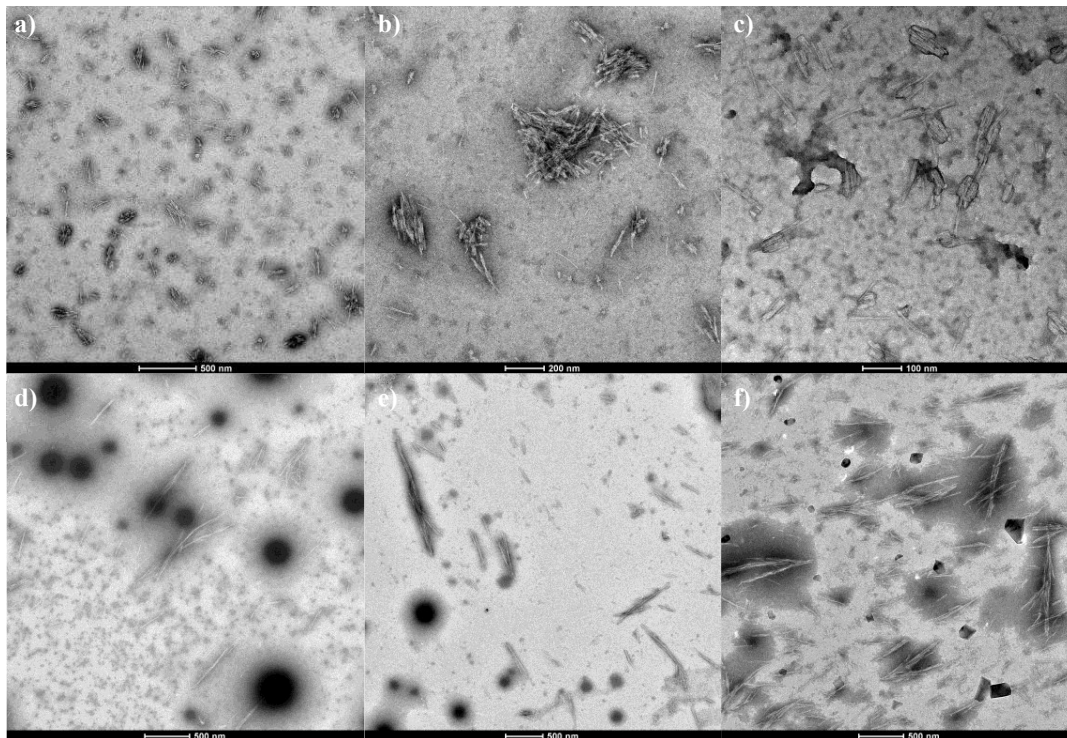


Figure 48 : Images TEM du peptide a,b,c) 35m1 incubé à 500  $\mu$ M dans  $\text{NH}_3/\text{HCOOH}$  à pH=8.5 (50 mM) à 40 °C pendant 2h à 1600 rpm (agent de contraste Acétate d'uranyle à pH=4), d,e,f) 35m2 à 500  $\mu$ M dans Pi à pH=7,4 (50 mM) à 1000 rpm pendant 2h à 37°C (agent de contraste : phosphotungstate de sodium à pH=7).

Comme vu précédemment lors de l'étape de conception de nos peptides, et du suivi cinétique en fluorescence de la ThT, les peptides sont confectionnés pour s'autoassembler et former des fibres à un pH bien supérieur à 6 (pKa des histidines). Nous nous sommes alors placés à pH=7.4 ou 8.5 afin de déprotoner les histidines présentes dans les séquences. Un Thermomixer C a été utilisé afin de pouvoir réaliser des incubations à des température et agitation contrôlées dans des tubes. Ici les deux peptides seront traités ensemble, dans un premier temps, puis séparément car le peptide 35m2 a été le plus investigué puisqu'il a été optimisé afin de défavoriser l'agrégation latérale par rapport au 35m1. Le dépôt des échantillons peut aussi jouer un rôle important sur la forme des objets à observer.<sup>[45]</sup> L'obtention des premières images des deux peptides a été réalisée en collaboration avec le Dr. Schmutz (Figure 48).

À partir de ces premières expériences, nous avons pu remarquer la forte tendance du peptide 35m1 à former des faisceaux compacts de fibres contrairement au peptide 35m2 qui a plus tendance à former des fibres beaucoup moins agrégées que le premier. La coloration négative nous permet d'observer comment les fibres se collent entre-elles et met en évidence les réseaux formés. D'après ces premières analyses, nous avons pu voir que la longueur des fibres était différente entre les deux peptides, le premier formant des fibres plutôt courtes et en faisceaux alors que le second forme des fibres beaucoup plus longues et moins collées. Ceci va

dans le sens de la conception de nos peptides, notamment pour le peptide 35m2. Néanmoins, ici, les conditions d'incubation des peptides sont différentes et cela peut bien sûr jouer un rôle important dans les morphologies observées. C'est pour cette raison que nous avons réalisé une analyse où les deux peptides ont été incubés de la même manière et préparés de la même façon. Nous avons incubé 20  $\mu\text{L}$  de solution de chaque peptide à 500  $\mu\text{M}$  dans un tampon HEPES à  $\text{pH}=8.5$  à 50 mM, dans un tube de 100  $\mu\text{L}$  fermé et scellé avec du parafilm à 40°C avec une agitation de 1400 rpm pendant 60 heures. Juste après incubation, les suspensions de fibres (d'aspect trouble) sont diluées et déposées sur les grilles de TEM puis contrastés avec de l'acétate d'uranyle (AcU). Ensuite, la grille est mise à sécher pendant quelques minutes et mise dans le microscope électronique. Ces expériences ont été réalisées à la plateforme de microscopie électronique de l'INCI en collaboration avec la Dr. Demais. Les résultats sont montrés en Figure 49. Tout d'abord, nous pouvons remarquer que le contraste est positif, coloration des fibres en noires sur fond blanc, ce qui n'est pas dû à un agent de contraste, où on devait voir les fibres en blanc sur fond noir. Ce qui montre que nous n'avons pas encore réussi à optimiser cette étape, ou que les colorants ne se lient pas bien à nos objets. Mais, même sans cela, nos fibres étaient très bien visibles mais les détails comme des faisceaux ne sont plus observables contrairement aux images de la Figure 48.

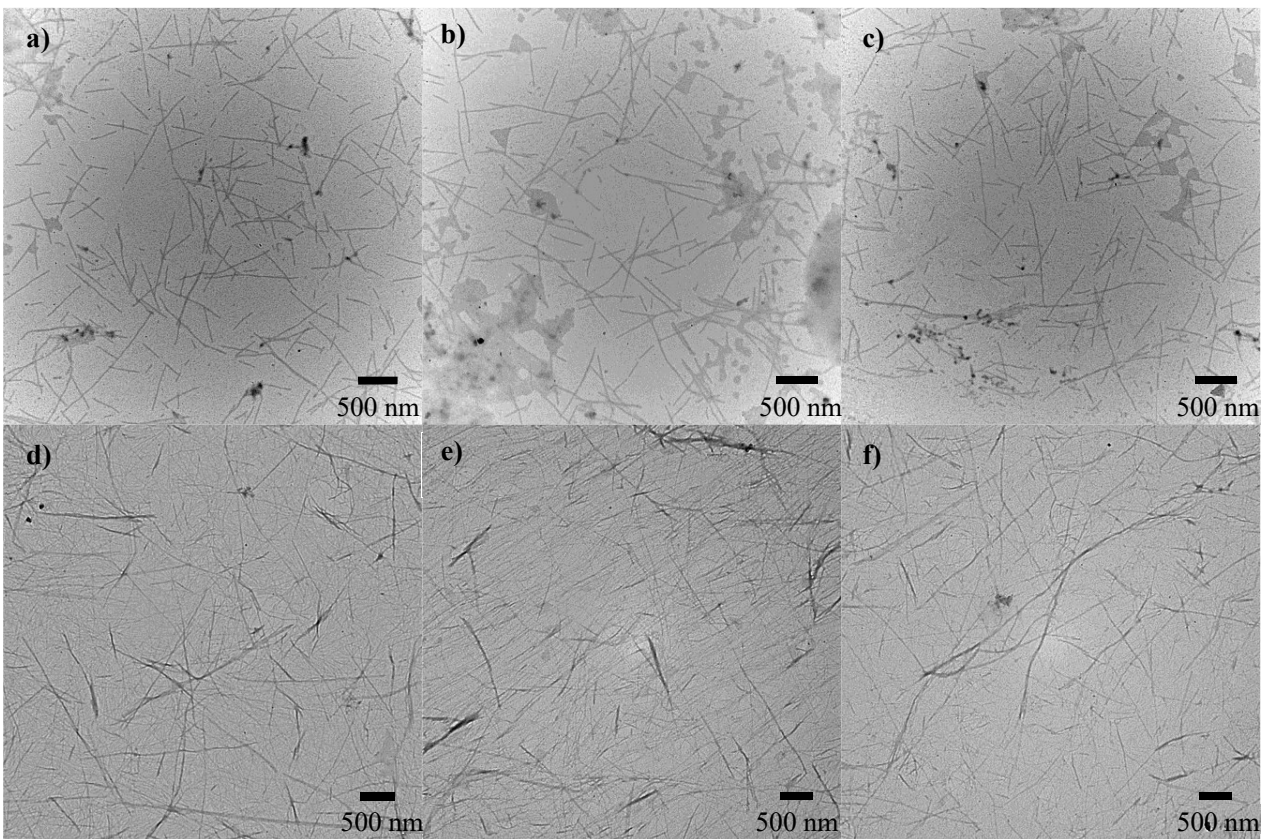


Figure 49 : Images TEM des peptides 35m1 (a, b, c) et 35m2 (d, e, f) préparés de la même manière. Incubation à 40°C, 1400 rpm pendant 60 h à 500  $\mu\text{M}$  de peptide, 50 mM d'HEPES  $\text{pH}=8.5$  avec agent de contraste (AcU).

Nous étions très surpris des fibres formées par le peptide 35m1 qui correspondent aux fibres attendues, à savoir, isolées, de taille homogène et très dispersée, avec peu d'agrégation latérale. Après discussion, et au vu de la technique utilisée, nous pensons qu'il s'agit d'objets avec une forte densité électronique, très compacts, qui pourraient presque s'apparenter à des nanocristaux plus qu'à des fibres, mais quoi qu'il en soit, les morphologies sont différentes. Ici, le 35m2 forme des fibres bien plus longues et isolées mais aussi en fagot à la différence de ce que nous avons observé précédemment. Ces observations vont dans le sens de la formation de fibres isolées pour pouvoir mener des analyses plus poussées sur celles-ci. Nous pouvons aussi remarquer une différence de diamètre de fibres lorsqu'on compare celles qui sont seules. En effet, nous nous attendons à un diamètre d'environ 4.5-5 nm, qui correspond à la section trimérique de l'assemblage. Mais la résolution du microscope ne nous a pas permis d'avoir une mesure de diamètre précise, qu'une mesure estimée de 10 à 15 nm pour le peptide 35m1 et 35m2. Cela peut sembler trop gros pour le trimère d'hélices mais la résolution n'étant pas assez bonne pour des objets aussi petits que les nôtres, nous ne pouvons pas être sûrs de la mesure.

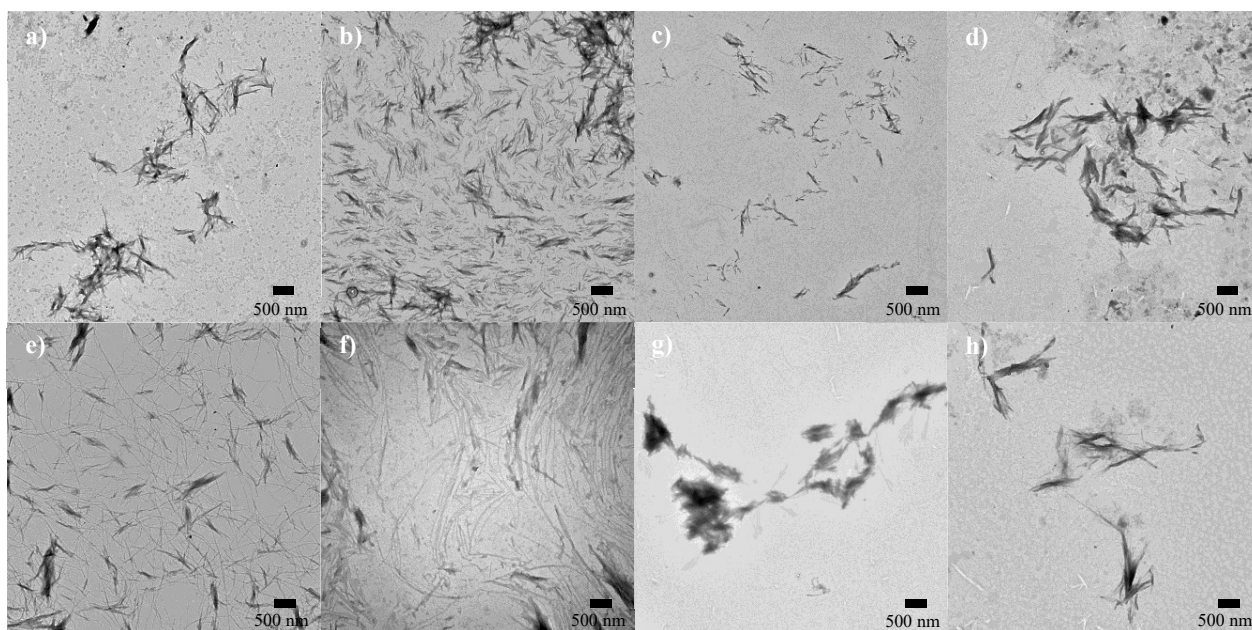


Figure 50 : Images TEM des 35m1 (haut) et 35m2 (bas) incubés à 40 °C, 1400 rpm, 500 μM pendant 24 h et redilués à 10 μM pour le dépôt sur grille a) et e), avec coloration négative (AcU) b) et f), pendant 1 h c) et g). d) et h) correspond aux solutions de fluorescence de la ThT à 50 μM et redilués à 10 μM pour le dépôt (incubation à 37 °C, 500 rpm pendant 72h).

N'étant pas totalement confiants de la reproductibilité des analyses, nous avons refait d'autres analyses en TEM à partir de stocks différents. À partir de peptides nouvellement synthétisés et purifiés, nous avons réalisé de nouvelles incubations à 40°C, 1200 rpm pendant 1 heure et 24 h afin de pouvoir regarder comment sont les fibres à ces temps-là, notamment pour le peptide 35m1 au vu des résultats de fluorescence de la ThT obtenus. Nous avons également pris les solutions de peptides issues des analyses de fluorescence de la ThT que nous avons déposées sur une grille de TEM afin de comparer entre les deux concentrations et conditions d'incubation. La concentration de tampon a été fixée à 50 mM dans toutes les analyses. Nous voulions aussi

voir si nous arrivions à voir une différence, sur le microscope de l'INCI, entre la coloration négative ou non. Les résultats de ces analyses sont exposés en Figure 50.

Nous n'observons aucune différence entre les grilles colorées négativement et non colorées signe que nous avons eu un problème de coloration lors de la préparation de l'échantillon (Figure 50 a, b, e, f). Nous pouvons cependant remarquer que les grilles sont très similaires, avec et sans coloration. Pour le peptide 35m1 les fibres sont plutôt courtes (quelques dizaines de nm de longueur) et agrégées alors que pour le peptide 35m2 elles sont très longues, fines, homogènes en diamètre et surtout non agrégées, même si on peut tout de même observer des paquets de fibres. Pour ce second peptide, le ratio de fibres seules par rapport aux fibres agrégées est assez élevé, ce qui correspond à ce qui est attendu (Figure 50 e et f). Des échantillons ont été incubés seulement 1 heure pour observer si la formation de fibres avait déjà lieu. En effet, les fibres sont déjà présentes, mais la morphologie est bien différente, plus courte pour le 35m1, et plus particulièrement pour le peptide 35m2 qui semble former des agrégats. Enfin, les grilles correspondantes aux dépôts à partir des solutions de fluorescence de la thioflavine-T, incubée à 37°C pendant 72 h nous montrent bien la formation de fibres mais de morphologies similaires, en fagots. Nous voulions voir si les fibres avaient la même morphologie que celles incubées dans le Thermomixer dans des tubes. Nous pouvons déjà dire que la concentration n'est pas la même : 10 fois plus dilué lors de l'incubation avec la ThT que lors des essais d'optimisation des fibres par la TEM. Cependant nous voyons une différence notable dans la morphologie du peptide 35m2 alors que le peptide 35m1 a une morphologie similaire. En effet, pour le peptide 35m2, seules des fibres agrégées ont pu être observées (Figure 50 d et h).

Sur la base de ces différentes observations, nous avons décidé, en collaboration avec les Dr. Pieri et Dr. Bressanelli, et le soutien du GDR MéDynA, de nous pencher sur le peptide 35m2 pour des études plus poussées en TEM mais aussi en Cryo-TEM.

### **3.4.2 Optimisation des conditions de fibrillation en vue d'une analyse en Cryo-TEM**

Jusqu'ici, l'optimisation des conditions de fibrillation pour les deux peptides nous a permis de voir quelles différences nous pouvions observer lors de la formation de fibres, juste avec 3 mutations sur les peptides. Cependant, pour la Cryo-TEM, les fibres doivent être homogènes, non agrégées, et longues. Le peptide s'en rapprochant le plus est le peptide 35m2 comme montrés en Figure 50 e et f. Des analyses ont alors été réalisées sur ce peptide afin d'identifier les conditions les plus propices à une analyse en Cryo-TEM. Lors de ces analyses, nous avons aussi chargé en surface (polarisé) les grilles de TEM, avant dépôt, afin de favoriser l'adsorption des peptides sur la grille. En effet, charger les grilles sous vapeur de plasma, appelé « glow discharge », permet de réduire l'hydrophobicité de la grille, ce qui favorise le contact et un dépôt homogène sur toute la surface.<sup>[46,47]</sup> La procédure réalisée est la même que précédemment mais la quantité de dépôt a été de 3  $\mu$ L et les temps pour contraster la grille ont été de 2 fois 40 secondes.

Nous avons alors regardé sa stabilité dans le temps, après fibrillation, pour observer une possible agrégation entre les fibres. Nous avons en parallèle testé deux contrastants différents, le phosphotungstate de sodium à pH=7.05 et l'acétate d'uranyle à pH 4, afin de visualiser les différences entre les deux, et de déterminer lequel était le meilleur pour notre système, c'est-à-dire qui ne re-solubilise pas les fibres et qui permet de voir les détails précisément. L'incubation puis le repos pendant quelque temps à température ambiante semblent favoriser la formation d'agrégats lorsque les fibres ne sont pas déposées tout de suite sur la grille. Il y a cependant des fibres seules qui sont homogènes en diamètres et longues mais des agrégats beaucoup plus larges et courts ont aussi été observés dans ces conditions (Figure 51 a et b).

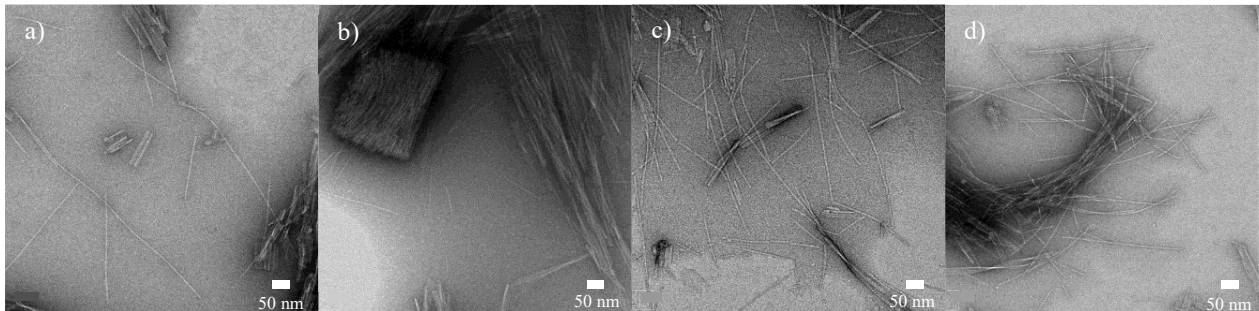


Figure 51 : Observations TEM en coloration négative de fibres formées par le peptide 35m2 incubé à 500  $\mu$ M dans HEPES pH=8.5 à 40 °C à 1200 rpm pendant 48 h et laissé à  $T_{amb}$  pendant 1 jour (a et b) et pendant 16 h (c et d). Les grilles correspondant aux images a), b) et d) ont été contrastées au phosphotungstate de sodium à pH=7.05 et C) contrastée avec de l'acétate d'uranyle à pH=4.

À l'inverse lorsque ce peptide a été incubé et déposé directement après incubation, nous avons observé une large majorité de fibres seules et longues, bien isolées, avec quelques agrégats certes mais en faible proportion (Figure 51 c et d). On peut aussi voir que l'image contrastée avec l'acétate d'uranyle permet de voir les fibres avec une meilleure résolution (Figure 51 c) que celles contrastées avec phosphotungstate de sodium (Figure 51 a, b, d). Par la suite, nous avons alors utilisé ce contrastant pour colorer nos fibres. Il est à noter que des conditions d'incubation avec du sulfate d'ammonium, du NaCl et du Laurylsulfate de sodium (SDS), à différentes concentrations de peptides, à différents temps et températures d'incubation ont été réalisés. Les fibres sont très régulières aussi en diamètre, et très longues dans chaque cas, mais le problème est la présence d'agrégats qui demeure. Une centrifugation contrôlée, après incubation, pourrait résoudre ce problème.

### 3.4.3 Suivi de la fibrillation par TEM

Dans le but de comprendre comment les fibres se forment, nous avons réalisé des dépôts en fonction du temps d'incubation. Pour se faire, et pour ne pas perturber l'incubation en fonction du temps, différents tubes ont été mis à incuber à 500  $\mu$ M, 40°C, 1200 rpm au Thermomixer dans l'HEPES (50 mM) à pH=8.5 et 7.4. Un dépôt a été réalisé à t=0 min afin de contrôler la présence de fibres ou non dans la solution de départ, à 30 min puis toutes les heures. Toutes les grilles ont été contrastées avec de l'AcU et analysées. Un contrôle a aussi été fait en incubant la solution de départ à 500  $\mu$ M sans tampon (pH $\approx$  3.5) pendant 5 heures.

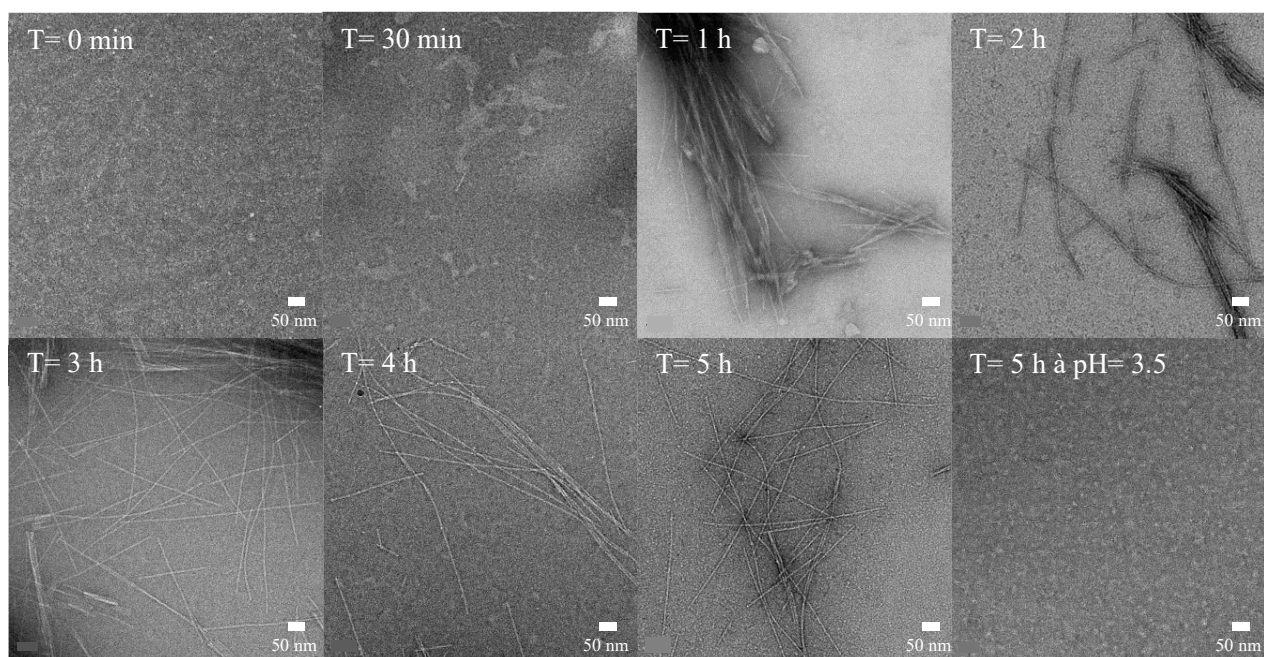


Figure 52 : suivi de formation de fibres du peptide 35m2 par TEM en coloration négative. Incubation à 500  $\mu$ M, 1200 rpm à 40°C dans l'HEPES pH=8.5 (50 mM).

La Figure 52 nous montre le résultat de ces analyses en fonction du temps de fibrillation. Nous pouvons confirmer qu'à  $t=0$ , il n'y a pas de fibres présentes dans le tube. Ce qui veut dire que la solution stock préparée auparavant est stable sur plusieurs semaines au moins. Ensuite, nous avons l'impression que les fibres ne se forment pas de manière isolée, mais qu'elles proviennent de la séparation de fagots qui se forment avant l'apparition des fibres seules. En effet, nous avons observé l'apparition de fibres qu'à partir d'une heure d'incubation, et celles-ci étaient principalement en fagots. Ensuite, au fur et à mesure du temps jusqu'à trois heures, les fibres ont l'air de se déliter petit à petit (i.e. apparition de fibres isolées).

Au bout de 3 heures d'incubation, les fibres semblent avoir atteint leurs formes finales puisque les analyses à 4 ou 5 heures ne montrent pas de changement. Cette observation de formation de faisceaux de fibres avant la formation de fibres seules va dans le sens de l'hypothèse du mécanisme explicité lors des analyses de fluorescence de la thioflavine-T. Nous pouvons aussi confirmer que même si nous ne voyons plus de fibres qui sont agrégées après repos, la morphologie des fibres seules est similaire, ce qui pourrait indiquer qu'à 48 heures, le système n'évolue pas et que nous n'observons pas de changement drastique de morphologie, si ce n'est de l'agrégation (Figure 51 a et b). Nous observons également, des plaques de fibres formant un réseau de fibres parallèles comme précédemment, mais en très faible proportion. Nous remarquons de petits objets qui peuvent être des agrégats ou précipités de peptides globulaires, déjà présents à 3 h d'incubation. Les analyses à pH=7.4 ne sont pas montrées ici car ce sont les mêmes que celles obtenues à pH=8.5. Nous avons aussi tenté de laisser les fibres après incubation pendant 5 heures (Figure 52) au repos au frigo à 4 °C afin de voir si nous avons toujours ce comportement d'agrégation des fibres. Après une nuit à 4°C, les fibres sont beaucoup plus agrégées que celles déposées directement après incubation, ce qui laisserait indiquer une affinité non-négligeable entre

les fibres (Figure 51 d). Ceci pourrait être dû à la sédimentation des fibres, qui une fois collées auraient du mal à se décoller, même après resuspension à la pipette et dilution. Enfin, nous pouvons aussi confirmer par cette technique que les fibres ne se forment pas à pH acide (Figure 51, 5h à pH 3,5).

Nous avons soniqué la suspension de fibres dans le but de savoir si cela allait permettre de séparer les fagots en fibres seules. Nous avons réalisé cela sur l'échantillon à 3 h d'incubation laissé au repos à  $T_{amb}$  pendant 2 heures où les fibres étaient seules mais avec quelques paquets de fibres. La comparaison est montrée en Figure 53 où les fibres seules et paquets de fibres longues sont présents avant sonication alors qu'après nous observons toujours des fibres seules mais beaucoup plus courtes, et surtout, il y a beaucoup plus d'agrégats et notamment de faisceaux de fibres. La sonication n'est donc pas une voie possible pour réduire la formation de fagots de fibres.

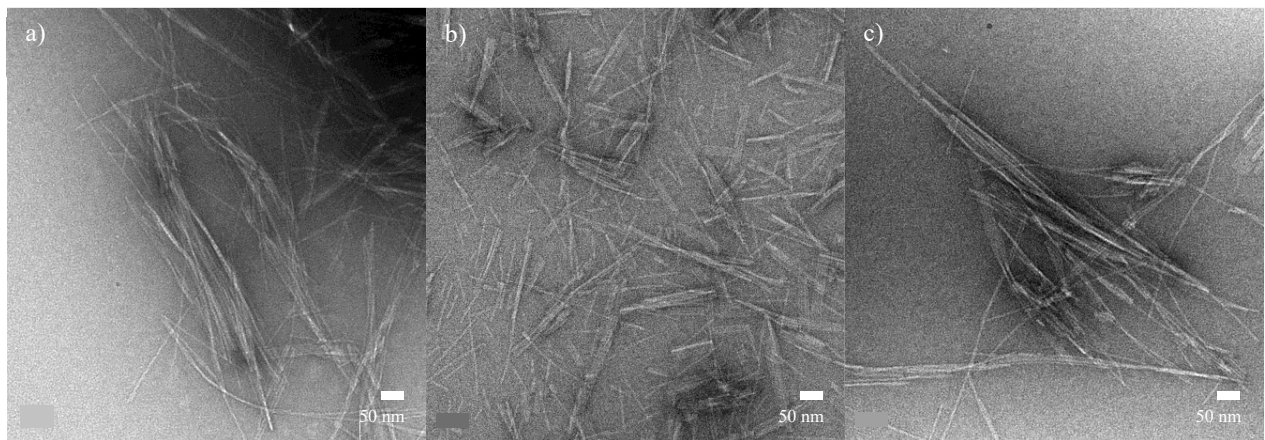


Figure 53 : Images TEM du 35m2 a) après 3 h d'incubation à 500  $\mu$ M, 40°C, 1200 rpm dans l'HEPES 50 mM pH=8.5 puis 2 h de repos à  $T_{amb}$ , b) après sonication pendant 30 s et c) après une nuit au frigo.

Nous avons alors choisi de prendre les conditions les plus favorables pour réaliser des essais en Cryo-TEM. Il s'agit des conditions d'incubation à 500  $\mu$ M dans l'HEPES pH=8.5 (50 mM) à 40°C à 1200 rpm pendant minimum 3 heures.

### 3.4.4 Analyses préliminaires en Cryo-TEM

Les observations par Cryo-TEM ont été réalisées à l'I2BC à la plateforme de Cryo-EM avec la Dr. Pieri.

#### 3.4.4.1 Préparation des grilles

Les premières analyses en Cryo-TEM ont été réalisées à partir de solutions fraîchement incubées dans les conditions les plus favorables déterminées précédemment. Une préparation rigoureuse et minutieuse a besoin d'être réalisée afin de déposer assez de matière, de ne pas tordre la grille, de ne pas avoir trop de glace

qui masquerait les fibres, etc...<sup>[48,49]</sup> Nous avons besoin de voir dans un premier temps si nous observons nos objets en conditions Cryo-TEM, dans lesquelles il n'y a pas de contrastant. Comme cette analyse se réalise à froid, nous devons vitrifier la grille afin d'emprisonner les fibres dans une matrice d'eau non-cristallisée. Cette étape peut s'avérer très difficile à réaliser car d'une part il faut absorber du liquide, comme en TEM, mais une quantité pas trop importante sinon l'analyse sera difficile, le contraste se faisant entre la matrice (l'eau) et les fibres déposées. Plus ce rapport est faible, moins les fibres seront visibles. Il est alors important de contrôler ce paramètre, appelé « temps de blot » qui peut être spécifique pour chaque échantillon en fonction de sa viscosité, concentration ou encore du solvant utilisé. La force du « blot » est elle aussi importante, elle correspond à la force mise contre la grille afin d'absorber le liquide.<sup>[50]</sup> Il faut ensuite congeler cette grille très rapidement dans l'éthane liquide. Le facteur limitant dans la vitrification est la conductivité thermique de l'eau, il faut alors un liquide en ayant une plus grande, c'est le cas de l'éthane, mais pas de l'azote liquide, c'est pour cette raison que malgré sa grande inflammabilité l'éthane est utilisé.<sup>[47,51-53]</sup> Cela peut être réalisé à la main, de la même manière qu'en TEM, ou via un robot, qui permet une reproductibilité plus importante. Dans notre cas, nous avons réalisé un essai « à la main » afin de voir si les fibres étaient visibles puis au robot (Vitrobot) pour préparer nos grilles. Les grilles utilisées ont été, dans un premier temps des grilles classiques Quantifoil R2.2, Cu-200 mais nous n'avons pas réussi à avoir une glace à assez bonne qualité pour pouvoir faire une bonne analyse. Nous sommes ensuite passés sur une grille Lacey300-Cu. La différence entre les deux grilles analysées par Cryo-TEM est montrée en Figure 54. La première grille a une taille de pore de 2  $\mu\text{m}$  séparés par 2  $\mu\text{m}$  de cuivre, ce qui résulte en beaucoup d'objets non observables car sur le cuivre. À l'inverse, la seconde ayant des tailles et formes de pores aléatoires, avec très peu de cuivre entre eux, permet d'observer beaucoup plus d'objets, notamment des fibres à cheval entre plusieurs pores.

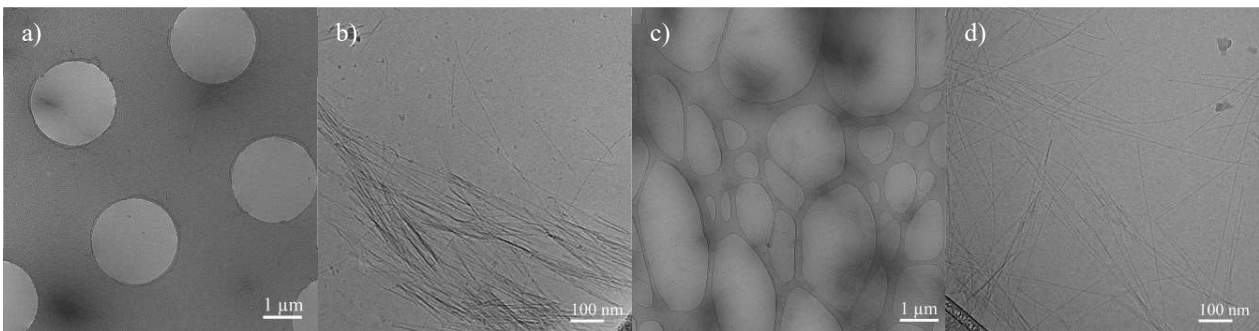


Figure 54 : Différences entre les deux grilles utilisées en Cryo-TEM pour observer les fibres formées par le peptide 35m2 incubé à 500  $\mu\text{M}$  40°C 1200 rpm pendant a) et b) 16 h sur grille Quantifoil R2.2 Cu-200 et 3 h sur grille Lacey300-Cu c) et d) et dilué à 250  $\mu\text{M}$  pour le dépôt.

Dans les deux cas, nous pouvons observer des fibres et notamment de gros fagots de fibres qui correspondent aux points plus foncés sur les images a) et c) de la Figure 54. Nous pouvons cependant confirmer, comme en TEM à coloration négative, le fait qu'à 3 ou 16 heures, la morphologie des fibres est la même, avec un diamètre moyen de 5 nm et des longueurs de l'ordre de quelques  $\mu\text{m}$ . Nous remarquons que les fibres paraissent moins agrégées dans le second cas, ce qui peut être dû au dépôt. Les petits points noirs observés sur les grilles



correspondent à de la glace cristallisée. Ici dans les deux cas, les grilles sont de bonnes qualités car la glace est bien vitrifiée avec très peu de cristaux ou de pore vides.

### 3.4.4.2 Résultats de l'analyse en Cryo-TEM du peptide 35m2

Après avoir trouvé des conditions acceptables pour des analyses préliminaires du peptide 35m2 en Cryo-TEM, nous avons réalisé des images à deux grossissements différents. Un grossissement à x14500 avec une taille de pixel de 2.5 Å et un à x 29000 avec une taille de pixel de 1.3 Å. La taille de pixel correspond à la taille que prend un pixel sur l'image, qui est un paramètre important à prendre en compte lors de l'analyse de clichés en Cryo-TEM.

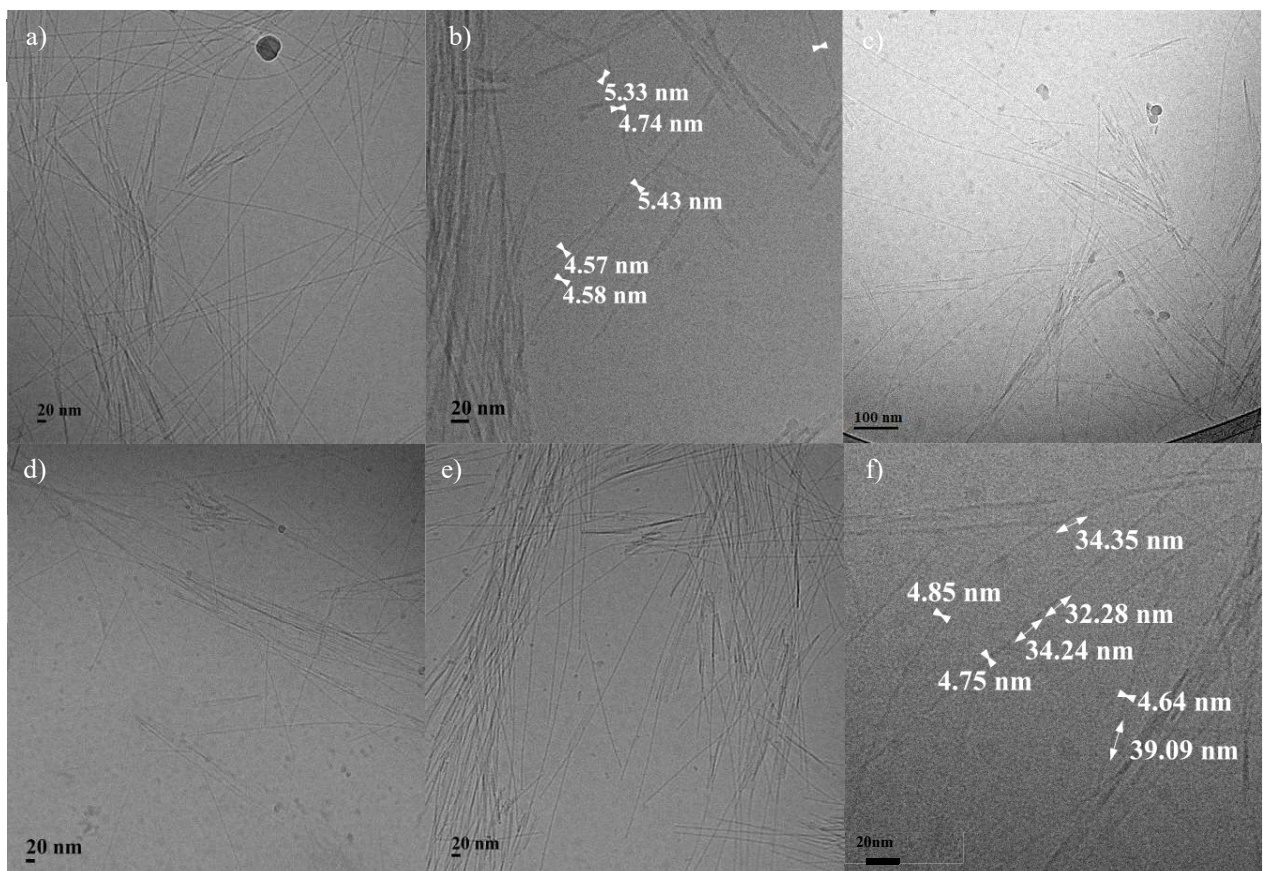


Figure 55 : Images Cryo-TEM obtenues pour le peptide 35m2 incubé 500  $\mu$ M 40°C 1200 rpm pendant 3 heures pour 2 grilles différentes. Un grossissement de x14500 est utilisé pour toutes les images sauf la b) qui a un grossissement de x29000. Les mesures de périodicité et de largeur de fibres sont montrées en b) et f).

La Figure 55 nous montre différents clichés obtenus pour le peptide 35m2. Nous voyons quand même que ces fibres sont homogènes en diamètre, lorsqu'elles sont seules, et que lorsqu'elles sont en fagots, elles sont bien parallèles. La largeur des fibres mesurée est d'environ 5 nm et on observe une périodicité d'environ 30-40 nm (mais ceci est à modérer car très difficile à voir au vu de la limite de résolution de l'appareil). Ces mesures sont

compatibles avec la formation de trimères d'hélice  $\beta$  gauche, les modèles prévoient une largeur de trimère d'environ 5 nm avec une périodicité (« pitch » en anglais) d'environ 35 nm. Peu de clichés à fort grossissement sont montrés car nous n'avons pas réussi à voir les fibres seules sur la grille à ce grossissement.

Nous avons réalisé des clichés pour voir si nous pouvions obtenir la structure à partir des fibres de peptide. Cependant, nos objets étant très petits en diamètres, ils demandent un appareil plus puissant afin d'accéder à de bonnes résolutions et de pouvoir reconstruire la structure des fibres. Nous avons besoin, pour cela, d'optimiser les conditions afin d'obtenir de meilleurs résultats qui sont encore à traiter pour pouvoir obtenir des informations structurales, sans remonter à la structure. L'analyse de plusieurs fibres longues et isolées, à partir de préparations différentes, nous permet d'obtenir leur taille de section, leur périodicité éventuelle mais aussi, par transformée de Fourier rapide (FFT) sur les fibres, de savoir si l'alignement des fibres est bien parallèle.<sup>[54–56]</sup> Avec une résolution assez bonne, nous pourrions déterminer l'orientation des feuillettes  $\beta$  à l'intérieur de la fibre afin de savoir si nous avons des feuillettes  $\beta$  qui sont parallèles ou perpendiculaires à l'axe de la fibre. Cette information permettrait d'attester que les fibres sont compatibles avec la formation d'hélices  $\beta$  gauche ou d'infirmer cela (des feuillettes  $\beta$  orientés parallèlement à l'axe de la fibre seraient incompatibles avec notre modèle).

D'après ces analyses, ces fibres pourraient correspondre à des fibres trimériques d'hélices  $\beta$  gauche. Cependant, comme nous pouvons le voir sur les images, nous avons tendance à former des paquets de fibres, qui dans certaines conditions sont difficiles à séparer du reste. Les conditions ont été améliorées afin de favoriser les fibres seules et isolées et nous savons que le temps optimal de fibrillation est supérieur à 3 heures à un pH supérieur à 6. Ici, les pH de 7.4 et 8.5 ont montré les mêmes morphologies mais à pH acide, aucune fibrillation n'a pu être observée. Les fibres étant de taille très petite (en diamètre), nous devons passer sur un microscope plus puissant pour permettre une reconstruction de la structure 3D de la fibre. Nous pouvons cependant dire avec certitude que les fibres sont homogènes en diamètre, et qu'elles forment des fagots mais aussi des fibres longues, et elles sont visibles en Cryo-TEM ce qui montre bien qu'elles sont présentes en solution, qu'elles ne sont pas qu'un artefact de séchage en TEM. Elles pourraient avoir une périodicité hélicoïdale mais cela reste à confirmer.

### 3.4.5 Informations structurales par dichroïsme circulaire

Cette technique spectroscopique permet de déterminer rapidement les structures secondaires que peuvent adopter les peptides ou protéines en observant les transitions  $n-\pi^*$  et  $\pi-\pi^*$  des liaisons amides.<sup>[30,57–59]</sup> Les protéines, en fonction de leur coefficient d'extinction molaire entre une lumière polarisée droite ou gauche, vont avoir un spectre caractéristique aux structures secondaires qui y sont présentes. Elle peut donc permettre de suivre des changements structuraux, comme un repliement/dépliement,<sup>[60]</sup> une cinétique de formation/dissociation de complexe, d'interaction avec des membranes ou autres.<sup>[31,61]</sup> C'est une technique bien connue aujourd'hui, et qui a permis de suivre la cinétique de formation de fibres avec différents systèmes comme l'A $\beta$ (1-40) via la formation de feuillettes  $\beta$ <sup>[62]</sup> ou encore l' $\alpha$ -synucléine<sup>[63]</sup>. Elle est complémentaire à

d'autres techniques comme la cristallographie, la Cryo-TEM ou encore la RMN du solide qui sont des techniques à un état solide et peut parfois servir à obtenir des informations structurales non accessibles par d'autres techniques.<sup>[64,65]</sup> Il a notamment été possible de mieux comprendre la structure de la  $\gamma$ D-crystalline responsable de différents types de cataracte congénitale par CD par observation de l'augmentation de la bande aux alentours de 218 nm permettant de faire une corrélation avec la perte de solubilité de la protéine.<sup>[65]</sup>

Évidemment, les systèmes comme les solénoïdes  $\beta$  ont été étudiés et montrent des signatures spectrales caractéristiques à ces repliements très similaires à ceux des amyloïdes  $\beta$ .<sup>[66]</sup> L'autoassemblage des protéines en solénoïdes  $\beta$  a aussi été étudié et notamment les hélices  $\beta$  gauche de type II<sup>[7]</sup> ou issues de bactériophage qui forment des hélices  $\beta$  gauche par association de 3 chaînes protéiques.<sup>[67]</sup> Une hélice  $\beta$  gauche de type II fibrillaire a déjà été analysée par CD (de 195 nm à 300 nm) (Figure 10 de l'Introduction).<sup>[7]</sup> Dans tous les cas, les signatures sont très similaires avec des minimaux aux alentours de 218 nm, qui sont caractéristiques des feuillets  $\beta$ .<sup>[68]</sup> Ces caractérisations sont compatibles avec les structures attendues des solénoïdes  $\beta$  mais l'orientation des feuillets n'est souvent pas accessible à cause des limites expérimentales (i.e. cuve, source de photons). Elles sont imposées par l'appareil mais aussi par les tampons et sels utilisés ainsi que les solvants qui ont une coupure de signal spécifique à chacun.<sup>[31,69,70]</sup> Cependant, il serait intéressant aujourd'hui de déterminer si la conformation des feuillets est parallèle ou anti-parallèle avec les informations obtenues à des longueurs d'onde inférieures à 200 nm même si la différence est moins marquée que par d'autres techniques.<sup>[71]</sup> La Figure 56 montre les différences de signal dues aux différentes conformations des feuillets  $\beta$ , que ce soit antiparallèle (Anti1 : torsion du feuillet vers la gauche, Anti2 : très légère torsion vers la droite, Anti3 : torsion du feuillet vers la droite) ou parallèle. En ce qui concerne les feuillets  $\beta$  parallèles, les signaux caractéristiques correspondent à une bande négative à environ 218 nm, une bande positive intense à environ 198 nm ainsi qu'une bande négative à partir de 187 nm. Il est important de noter ici que les bandes peuvent varier en fonction de l'angle des feuillets  $\beta$  ainsi que de leur environnement.

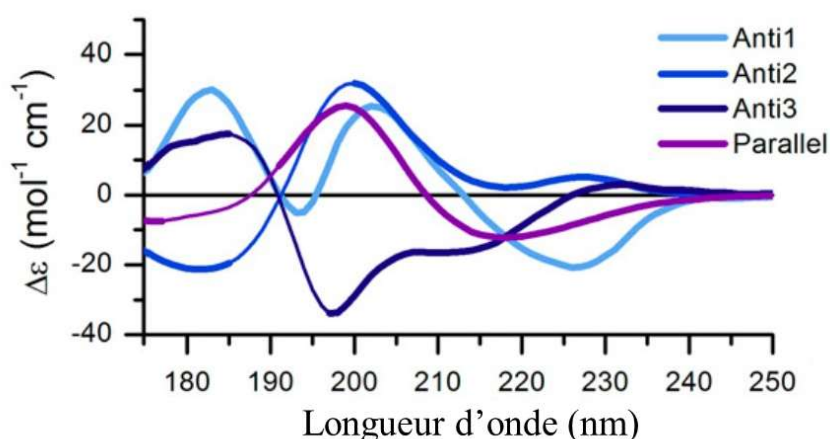


Figure 56 : Spectres CD calculés et optimisés utilisés comme références dans la base de données BestSel. (Reprise de la ref<sup>[71]</sup>)

Généralement, sur les spectromètres de paillasse, le trajet optique est assez important allant de la centaine de  $\mu\text{m}$  au cm. Plus il est long, plus il faut travailler en conditions diluées. En revanche, le dichroïsme circulaire

sur source synchrotron (SR-CD) permet de travailler à des concentrations bien plus élevées en utilisant des trajets optiques très courts qui sont de l'ordre de la dizaine de  $\mu\text{m}$ . La brillance, ou la quantité de photons du faisceau, permettent aussi de travailler en dessous des 190 nm (limite habituelle des spectromètres de paillasse) et d'avoir des temps d'acquisition de l'ordre de la dizaine de secondes au lieu de la dizaine de minutes en routine, tout cela menant à une résolution cinétique bien meilleure.<sup>[72]</sup> Nous pouvons obtenir beaucoup d'informations à des longueurs d'onde inférieures à 190 nm mais une limite est fixée à 170 nm environ à cause de l'eau qui tend à tronquer le signal à 168 nm, et sur des échantillons solides (i.e. secs) jusque 140 nm (la limite étant l'absorption du diazote).<sup>[73]</sup> Avoir accès à cette fenêtre spectrale permet d'obtenir plus d'informations structurales, puisqu'on y observe des signaux caractéristiques de l'arrangement des chaînes peptidiques. Par exemple, on peut différencier les feuillets  $\beta$  parallèles des antiparallèles, ou bien le type d'hélice  $\alpha$  ou encore les coudes  $\beta$ .<sup>[74]</sup> Il est de plus en plus possible de déterminer les différences de structures entre les différents feuillets  $\beta$  présents ainsi que leur angle de torsion par dichroïsme circulaire. Pour les bases de données existantes pour la prédiction de structures secondaires à partir du spectre CD, beaucoup de structures ont été enregistrées en n'allant que jusqu'à 190 nm, là où les différents types de feuillets ont des bandes similaires. Elles gagneraient en précisions si elles étaient alimentées avec plus de spectres incluant la fenêtre entre 200 et 170 nm. C'est notamment le cas pour SELCON<sup>[75]</sup>, CONTIN<sup>[76]</sup> et plus récemment BestSel<sup>[71]</sup>, qui utilisent des algorithmes couplés à des banques de données afin de prévoir la structure secondaire des protéines. Cependant, ces mesures sont faites avec des protéines qui ne sont quasiment jamais constituées de structures secondaires pures, c'est-à-dire qu'elles contiennent des parties globulaires avec des feuillets  $\beta$ , hélices  $\alpha$  ou pelotes statistiques, qui « polluent le signal », ce qui rend difficile l'attribution des signatures propres à chacune des structures secondaires.

Dans notre cas, nous nous attendons à un spectre pur d'une hélice  $\beta$  gauche seule, contenant des feuillets  $\beta$  parallèles mais aussi des coudes  $\beta$  qui serait similaire à celui-obtenu pour des fibres d'hélice  $\beta$  droite (Figure 10).<sup>[7]</sup>

En ce qui nous concerne, nous avons réalisé des expériences sur spectromètre de paillasse Jasco J-810 équipé d'une lampe Xenon 150 W. Ne pouvant aller en deçà des 200 nm à cause de la saturation du signal, nous avons aussi réalisé des expériences de SR-CD sur la ligne DISCO du synchrotron SOLEIL en collaboration avec le Dr. Wien nous permettant de mesurer des spectres allant jusque 170 nm en solution et d'acquérir des spectres très rapidement, ce qui permet de réaliser des cinétiques de repliement ou dépliement facilement.<sup>[77]</sup> Il a été aussi très intéressant de travailler avec le Dr. Wien qui nous a appris beaucoup sur les « bonnes pratiques en CD ». C'est-à-dire comment, par la suite des mesures, mettre les résultats normalisés à la quantité de peptide en utilisant l'absorbance mesurée à 205 nm en regard de la haute tension mesurée ainsi que la masse molaire pour obtenir la concentration exacte de protéine.<sup>[78,79]</sup> Aussi, la calibration avec l'acide D-10-camphorsulfonique (CSA) permet de s'assurer d'avoir des mesures reproductibles.<sup>[74,80]</sup>

### 3.4.5.1 Comparaison en CD et SR-CD des peptides fibrillés

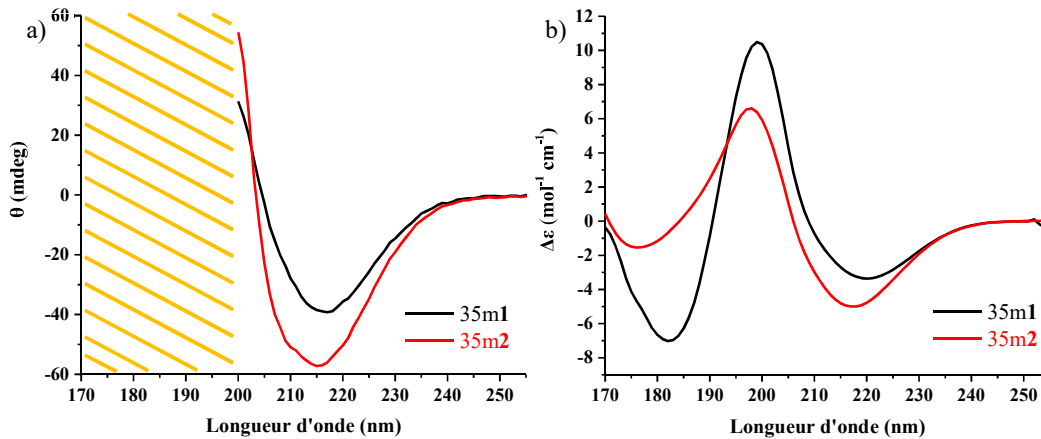


Figure 57 : Spectres CD obtenus pour les peptides 35m1 et 35m2 a) sur spectromètre de paillasse après incubation à 500  $\mu$ M HEPES pH=8.5 50 mM redilué à 60  $\mu$ M, trajet optique de 1 mm (la zone barrée en jaune symbolise la perte d'informations par rapport à une analyse SR-CD) et b) sur source synchrotron après incubation à 2mM HEPES pH=8.5 50 mM mesuré à 2 mM, trajet optique de 31  $\mu$ m.

Nous avons commencé par la mesure en CD classique pour voir si nous avions, à première vue, des feuillets  $\beta$  qui étaient présents dans les structures. Pour cela, nous avons incubé les peptides 35m1 et 35m2 dans les mêmes conditions que les expériences précédentes, à 500  $\mu$ M de peptide pour 50 mM de tampon HEPES à pH=8.5 à 40°C avec une agitation de 1200 rpm pendant 24 heures, afin de former les fibres de la même manière. Les suspensions ont ensuite été diluées à 60  $\mu$ M de peptide, et 20 mM de tampon pour l'analyse. Une cuve en quartz de 300  $\mu$ L a été utilisée avec un trajet optique de 1 mm. Une moyenne de 10 spectres a été enregistrée pour chaque peptide de 200 à 300 nm, avec des pas de 1 nm à 25 °C, à 1 seconde par point (16 minutes au total) qui est représentée en Figure 57 a. La différence d'intensité de la bande caractéristique des feuillets  $\beta$  pourrait être due à une erreur trop importante lors de la préparation de la solution stock, les concentrations étant estimées par pesée du peptide. Les valeurs des minima obtenus, sur spectromètre de paillasse, sont caractéristiques d'un repliement en feuillets  $\beta$  à 217 nm pour le peptide 35m1 et 215 nm pour le 35m2. Les spectres SR-CD des mêmes peptides enregistrés sur la ligne DISCO, moyenne de trois spectres sont montrés en Figure 57 b. Nous observons que les bandes sont un peu déplacées avec un minimum à 220 nm pour le peptide 35m1 et à 217 nm pour le peptide 35m2. Que ce soit sur paillasse ou au synchrotron, l'allure des spectres entre 200 et 250 nm est globalement la même, ce qui montre que les analyses sont reproductibles sur deux appareils différents. Pour le SR-CD, nous remarquons que l'intensité des bandes à 198 (35m1) et 199 nm (35m2) est différente mais aussi que le minima présent à 182 nm pour le 35m1 est intense alors que le 35m2 le minima présent autour de 176 nm est peu négatif. Cela pourrait être dû à la torsion des feuillets  $\beta$  dans la fibre finale qui est différente entre les deux peptides, un paramètre qui n'est généralement pas pris en compte.<sup>[71,81]</sup> Et évidemment cela peut être dû à des structures différentes, car le feuillet  $\beta$  parallèle est caractérisé par des bandes négatives à 218 nm et  $\approx$ 180 nm et positives à  $\approx$  198-200 nm. Pour les deux peptides, c'est le cas, mais nous voyons que les intensités sont différentes, ce qui traduit des structures fibrillaires légèrement différentes.

### 3.4.5.2 Différences entre les peptides en solution et fibrillés

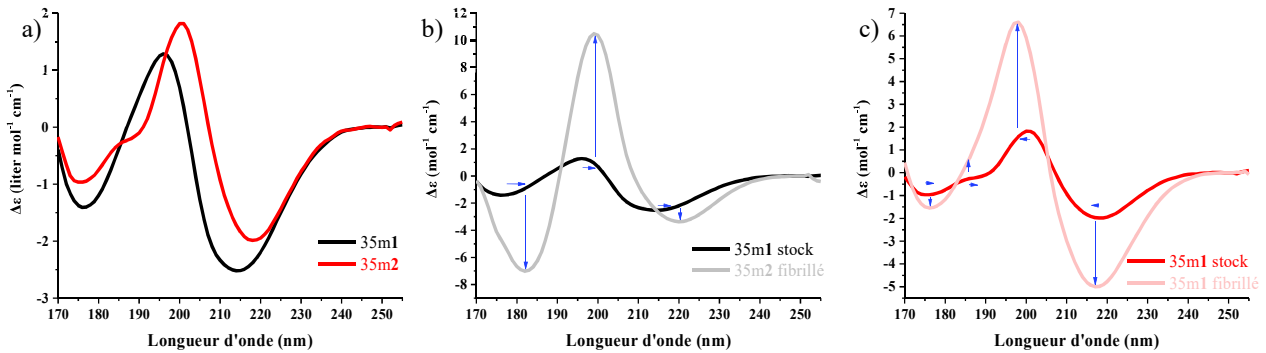


Figure 58 : Spectres SR-CD des peptides 35m1 et 35m2 à 2mM a) en solution; et différence entre la solution stock et fibrillée du peptide b) 35m1 et c) 35m2. Les flèches bleues représentent les déplacements des bandes des solutions vers celles des fibres.

Lors de la conception du peptide, nous partons du principe qu'à pH neutre, les peptides commencent à se replier sur eux-mêmes (pré-organisation) pour pouvoir former, par autoassemblage, des hélices  $\beta$  gauche de type I. Il serait alors logique d'avoir un spectre caractéristique des pelotes statistiques pour la solution stock de 35m1 ou 35m2 en conditions acides. C'est le cas lors de la solubilisation du peptide 35m2 dans l'eau, qui a une signature en pelote statistique en CD, et qui reste dans cette conformation même après une nuit au frigo en condition acide (i.e. spectres initiaux des cinétiques Figure 60). Cependant, au cours du temps, le peptide tend à se replier dans la solution stock, même avec la répulsion de charge, ce qui tend à montrer que la formation des feuillettes et du cœur hydrophobe est plus forte que la répulsion de charge (Figure 58). Mais nous n'avons pas accès à l'information sur la possible agrégation des peptides, s'ils sont toujours sous forme monomérique ou non. Les spectres CD de ces solutions stocks de peptides ont été comparés aux spectres obtenus après fibrillation (Figure 58 b, c). Le peptide 35m1 voit un déplacement bathochrome de ses bandes alors que le 35m2 voit des déplacements bathochromes en dessous de 190 nm et hypsochromes pour les bandes au-dessus de 190 nm. Il est aussi très intéressant de remarquer l'augmentation d'intensité significative entre les peptides en solution et fibrillés. Dans tous les cas, la proportion des bandes est augmentée avec la formation de fibres, très légèrement pour le 35m1 et de manière beaucoup plus significative pour le peptide 35m2 lorsqu'on regarde la bande caractéristique aux alentours de 218-220 nm. Pour chaque peptide, la bande positive présente à 198 et 199 nm est intense, ce qui est caractéristique des feuillettes  $\beta$ . Mais la différence majeure est à 182 nm pour le 35m1 avec une bande très prononcée caractéristique des feuillettes  $\beta$  parallèles<sup>[71]</sup> et déplacée à 176 nm pour le 35m2 et qui est moins intense. Cependant, nous ne pouvons pas dissocier les trois bandes car elles ne sont probablement pas indépendantes. Les différences pourraient être dues à un angle de torsion des feuillettes  $\beta$  différent en fonction des peptides. Il y a aussi de grandes chances que les parties terminales des peptides s'organisent afin de participer au réseau de liaisons hydrogène formé par les feuillettes mais aussi par chaînes latérales des asparagines (Figure 34). Nous nous attendons alors à un arrangement régulier des feuillettes  $\beta$  lorsque les fibres sont formées et moins régulier lorsque le peptide est en solution, l'effet template étant quasi nul et la flexibilité beaucoup plus importante.

Pour le 35m1, nous avons aussi réalisé des mesures de SR-CD, après incubation (i.e. formation des fibres) et centrifugation à 12000 rpm pendant 1 h à 20°C. Le surnageant et le culot ont été analysés séparément afin de voir s'il restait du peptide sous forme soluble dans la solution. La Figure 59 montre bien l'absence de signal dans le surnageant après la fibrillation, et au contraire, un spectre très intense dans l'échantillon correspondant au culot. Ceci suggère que tout le peptide présent en solution a été fibrillé (dans la limite de détection de l'appareil). L'analyse n'a pas été réalisée sur le peptide 35m2 par manque de temps, mais nous pouvons présumer que le comportement entre la solution et la formation de fibres se passe de manière similaire. Le déplacement des bandes pourrait être dû à la diffusion de la lumière par les fibres, présentes en grande quantité ici.

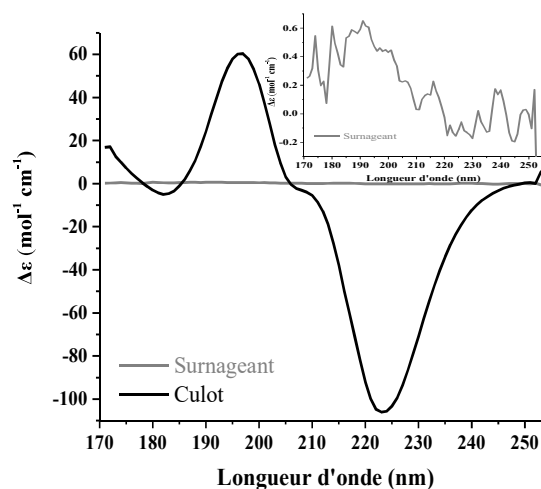


Figure 59 : Spectres SR-CD du peptide 35m1 après fibrillation et centrifugation montrant les différences entre le culot et le surnageant. En insert est montré un agrandissement du spectre obtenu pour le surnageant.

### 3.4.5.3 Suivi du repliement de 35m2 au cours du temps et en fonction de la température

Grâce à la rapidité d'acquisition des spectres en SR-CD, nous avons réalisé des cinétiques de repliement de nos peptides en fonction du temps à 40°C ; cela correspond aux conditions d'incubations les meilleures précédemment déterminées. Mais ici, l'incubation est réalisée directement dans la cuve de mesure en CaF<sub>2</sub> de 31 μm de trajet optique, sans possibilités d'agitation, ce qui marque une différence significative avec les conditions d'incubations que nous utilisons.

La Figure 60 a montre l'évolution du spectre SR-CD de 35m2 lors de l'incubation à 40 °C. Les spectres enregistrés toutes les 3 minutes, mais seuls ceux toutes les 30 minutes sont montrés ici pour faciliter la lecture. Mais nous observons une disparition rapide de la bande négative à 199 nm typique des pelotes statistiques (« random coils »). La bande à 218 nm était déjà présente à t = 0 h mais en très faible proportion, or lorsque le temps d'incubation augmente, le maximum est atteint au bout de 15 min pour diminuer par la suite, de même pour l'augmentation de la bande à autour des 200 nm. Une transition de 185 à 199 nm est aussi observée ce qui est compatible avec la formation de feuillet β. Mais même si l'apparition de cette bande indique l'apparition de feuillet β, l'incubation à 40°C sans agitation n'est pas suffisante, même après plus de six heures d'incubation nous n'observons pas de signal correspondant aux feuillets β purs.

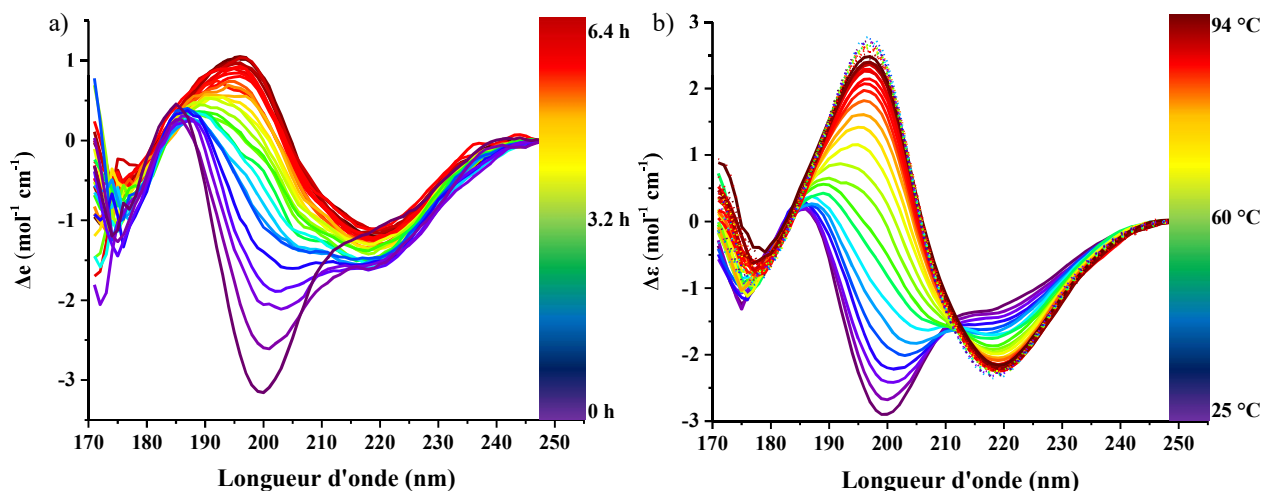


Figure 60 : a) Cinétique de repliement du peptide 35m2 à 2 mM à 40 °C et b) Cinétique de repliement suivie par SR-CD du peptide 35m2 à 2 mM dans HEPES à pH=8.5 (50 mM) en rampe de température croissante de 25 °C à 94 °C (traits pleins) puis en rampe descendante de 94 °C à 25 °C (traits pointillés).

Afin de vérifier si l'augmentation de température permettrait d'apporter l'énergie nécessaire au système pour se replier (i.e. compenser l'agitation mécanique appliquée lors de notre protocole habituel de formation des fibres), une rampe en température a été effectuée. Elle permettra aussi d'évaluer la stabilité de l'assemblage à haute température, mais aussi de voir sa possible réversibilité. Nous avons alors fait varier la température de la cuve de 25 °C à 94 °C, puis de 94 °C à 25 °C. Comme nous pouvons le voir en Figure 60 b, nous observons un point isobestique qui correspond à la transition entre une structure en pelote statistique (départ, violet) et en feuillettes  $\beta$  (état final, rouge). Le spectre final est identique aux spectres des fibres enregistrés précédemment. Ici, nous avons incubé le peptide tout fraîchement solubilisé, après sonication pendant quelques secondes, ce qui nous a permis de partir d'une solution désordonnée, sans préorganisation. Nous observons très vite que les feuillettes  $\beta$  se forment juste par augmentation de la température. Nous pouvons alors dire que le manque d'agitation de l'expérience précédente peut être compensé par une température plus élevée. Mais, à partir de ces mesures, une température seuil de formation complète et finale de la structure n'a pas pu être déterminée à cause de la vitesse d'analyse à chaque température, qui a pris environ 5 minutes pour chaque incrément de 3 °C, ce qui ne laisse pas le temps au système de s'équilibrer.

De plus, nous observons que ces fibres sont stables pour des températures allant jusque 94 °C. À noter que pour les fibres formées par 35m1, nous n'avons pas observé de changement significatif du spectre CD, même après 3h à 94°C. Cela montre le potentiel de ces systèmes pour des applications futures. Lors de la redescente en température, le spectre change très peu, indiquant une conservation de la structure, dont la flexibilité diminue comme peut l'attester l'augmentation de l'amplitude des bandes à 218 et 198 nm.

Constatant la stabilité en température de ses fibres, nous avons voulu tester l'effet d'un détergent sur la formation des fibres (voir si cela empêcherait le phénomène). Le peptide 35m2 a donc été mis à fibriller en présence de 0.5 % de SDS, dans les mêmes conditions que pour la Cryo-TEM. Nous voyons que nous retrouvons le spectre des fibres (Figure 61). Le signal correspondant aux feuillettes  $\beta$  est déjà important à 1h



d'incubation, signe que l'assemblage et le repliement du peptide se font, même en présence de 0,5 % de SDS. La baisse d'intensité pour les échantillons à 2 et 3 h est très probablement due à la préparation de la cuve avant analyse, mais nous pouvons voir que généralement, l'allure est similaire, ne montrant pas de dénaturation sur ce laps de temps.

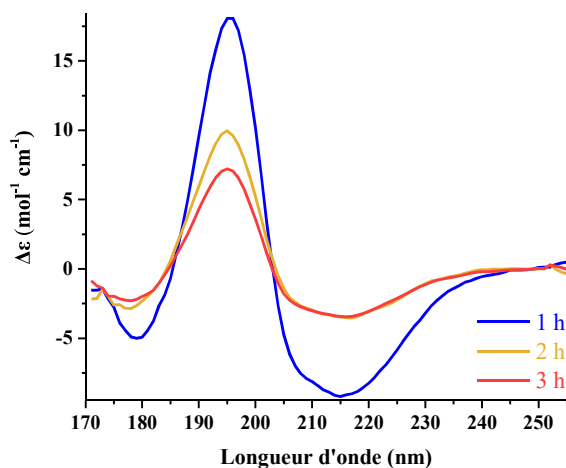


Figure 61 : Spectres SR-CD du peptide 35m2 à mis à incuber dans le SDS à 0.5 % à 500  $\mu$ M de peptide HEPES pH=8.5 50 mM, 40°C, 1200 rpm à 1h, 2h et 3h d'incubation. Incubation réalisée sur un agitateur rotatif différent de celui du laboratoire.

### 3.4.5.4 Alignement des fibres visible par dichroïsme linéaire

À l'inverse du dichroïsme circulaire, où nous regardons la différence d'absorbance de la molécule entre une lumière circulaire polarisée droite et gauche, le dichroïsme linéaire (LD) permet d'observer la différence d'absorption de lumière polarisée perpendiculairement à l'axe du faisceau (Figure 98 partie expérimentale). Cela nous informe notamment sur l'orientation des structures secondaires. Si nous connaissions l'orientation des fibres nous pourrions alors savoir comment les structures secondaires sont organisées dans celles-ci.<sup>[82-84]</sup> Mais le problème est que l'alignement des fibres n'est pas si facile à réaliser. En effet, sauf si les fibres forment spontanément des réseaux de fibres orientés de manière homogène dans l'espace, il est nécessaire de les aligner afin de pouvoir mesurer le dichroïsme linéaire. Il existe plusieurs techniques d'alignement mais dans notre cas, nous voulions voir si les fibres avaient déjà une orientation préférentielle dans l'échantillon. La préparation est la même que celle pour le CD, et nous avons regardé en dichroïsme linéaire s'il y avait alignement. Pour faire cela, à partir des échantillons préparés pour l'analyse en CD, le spectromètre a été passé en mode LD pour réaliser les mesures. Pour des orientations de 90° et 270° par rapport à la cuve analysée au départ, nous avons observé un minima et maxima aux alentours de 195 nm signe de l'alignement des fibres en feuillets  $\beta$ . Nous avons observé une différence avec le dépôt liquide (Figure 62 a) et séché, formant un film mince (Figure 62 b).

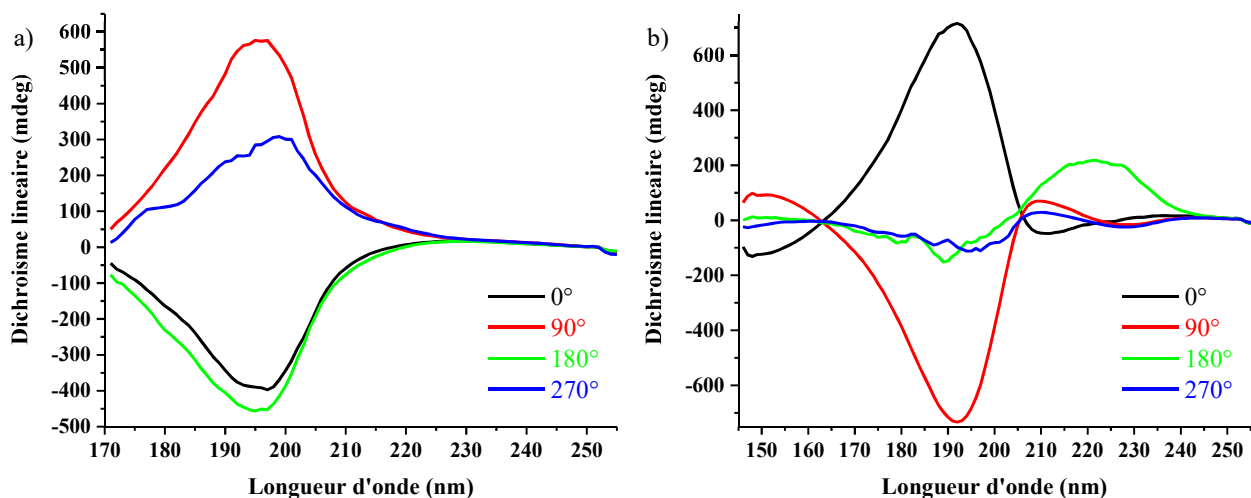


Figure 62 : Spectres de dichroïsme linéaire obtenus avec les fibres du peptide 35m2. a) Suspension de fibres ; b) Fibres déposées et séchées en film. Des rotations de 90° ont été effectuées afin de vérifier les orientations parallèles et perpendiculaires.

Nous observons une inversion du signal entre les orientations perpendiculaires (entre 0° et 90° et entre 180° et 270°) qui correspond aux orientations parallèles (positif) et perpendiculaires (négatif).<sup>[84]</sup> Ceci indique clairement un alignement (au moins partiel) des fibres en suspension dans la cuve de mesure. Cependant, nous avons eu de grandes difficultés à reproduire ces mesures, à re préparer des cuves avec du dichroïsme linéaire. Aussi, nous ne savons pas si cet alignement est présent en solution (suspension) ou bien s'il se fait lors du dépôt sur la cuve. Les spectres obtenus sont très similaires à ceux obtenus pour l'A $\beta$ <sub>1-42</sub> qui adopte également des structures en feuillets  $\beta$  parallèles.<sup>[85]</sup>

Nous avons aussi essayé d'aligner les fibres obtenues par séchage des suspensions. Le but est de former un film, voir si le séchage promeut l'alignement. Pour se faire, un dépôt de 4  $\mu$ L de suspension de fibres a été fait sur la cellule de mesure qui a ensuite été mise sous vide pendant 1 heure. Nous avons ensuite analysé ce film comme précédemment en SR-LD à différentes orientations pour voir si les fibres étaient alignées.<sup>[86]</sup> Il a été possible d'observer un certain alignement par cette technique comme peut l'attester la Figure 62 b où les orientations observées sont les mêmes que pour la suspension de fibres. Nous avons pu atteindre une longueur d'onde de 145 nm sur les mesures de film car l'eau n'est plus présente. La limite ici étant l'absorption du N<sub>2</sub>, qui se dissocie en dessous de cette longueur d'onde.<sup>[77]</sup>

### 3.4.5.5 Problèmes rencontrés lors des analyses SR-CD et SR-LC

Nous avons cependant rencontré des problèmes de reproductibilité lors de la fibrillation du peptide 35m2 signe que nous ne maîtrisons pas totalement son comportement. Ce manque de maîtrise est principalement dû à la différence entre les deux incubateurs utilisés, au laboratoire et au synchrotron. Au laboratoire, nous avons un Thermomixer avec un Thermotop ce qui permet d'éviter l'évaporation des échantillons, ce qui n'a pas été le cas au synchrotron où les évaporations n'étaient pas maîtrisées. Ces problèmes

ont permis de mettre en évidence que l'évaporation des échantillons est un paramètre très important, et que les fibres ne se forment pas juste par évaporation. Ainsi, nous pouvons dire que la préparation des échantillons joue un rôle prédominant, comme il a été montré auparavant avec la TEM par exemple. Nous avons observé des problèmes de stabilité du peptide monomérique solubilisé dans l'eau. Initialement, sa structure secondaire était en pelotes statistiques. Afin de garder le peptide déplié, nous avons vu qu'il est préférable de soniquer la solution pour avoir une solubilisation totale et de la garder en conditions acides, au frigo.

Notons finalement, que lors d'une mesure, nous avons observé un spectre de fibres de peptide 35m2 identique à celui des fibres du peptide 35m1, mais nous n'avons pas encore compris pourquoi.

### 3.4.6 Analyse des fibres par spectroscopie infrarouge

La dernière technique que nous avons utilisée pour caractériser les fibres peptidiques est la spectroscopie infrarouge (IR). Cette méthode d'analyse basée sur la vibration des liaisons présentes dans la molécule nous renseigne sur son repliement.<sup>[87,88]</sup> Nous nous focaliserons sur le mode de vibration des liaisons amides (chaîne principale), des déplacements d'onde caractéristiques des différentes structures secondaires.<sup>[89]</sup> Le mode de vibration le plus regardé pour observer une structure secondaire est le mode amide I. Il est caractérisé par l'élongation de la liaison carbonyle (C=O) de l'amide.<sup>[32]</sup> Dans notre cas, nous nous attendons à des feuillets  $\beta$  parallèles, caractérisés par une bande intense autour de  $1630\text{ cm}^{-1}$ . Les feuillets  $\beta$  antiparallèles étant caractérisés par deux maxima, un autour de  $1620\text{ cm}^{-1}$  et  $1690\text{ cm}^{-1}$  (Figure 63). C'est d'ailleurs ce qui a aidé à différencier les structures des amyloïdes  $\beta$  sous forme oligomérique (antiparallèle) et fibrillée (parallèle).<sup>[90]</sup> De plus, l'identification des coudes  $\beta$  est visible lorsqu'on observe la bande amide I, caractérisés par un maxima autour des  $1660\text{ cm}^{-1}$ , même si l'attribution des coudes  $\beta$  reste controversée.<sup>[91,92]</sup>

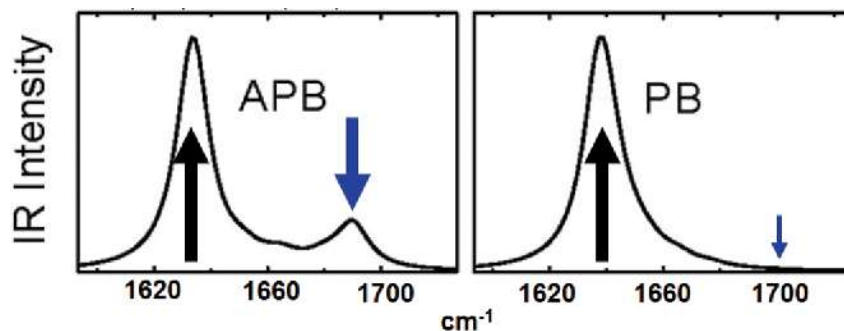


Figure 63 : Différence de bandes caractéristiques des groupements amides I entre les deux conformations de feuillets  $\beta$ , à droite antiparallèle et à gauche parallèle. (Figure reprise de la ref<sup>[32]</sup>)

Nous avons réalisé des analyses en ATR-IR qui permet d'analyser un échantillon fin déposé à la surface du diamant avec une sensibilité de surface accrue. L'avantage de l'ATR-IR ici est de vraiment analyser les fibres solides de peptides par séchage de la goutte déposée sur le diamant sous flux d'air. Cela ne requiert pas de préparation complexe. Il est très important de préciser ici que les contre-ions trifluoroacétates ( $\text{TFA}^{-}$ ) des

peptides ont été échangés contre des chlorures ( $\text{Cl}^{-}$ ) après purification. Cela a été fait car les TFA<sup>(-)</sup> ont une bande d'absorption importante dans la région des bandes amide I ( $1673 \text{ cm}^{-1}$ ).<sup>[93,94]</sup> Le TFA n'est pas non plus conseillé dans des applications biologiques du fait de sa toxicité.<sup>[95]</sup>

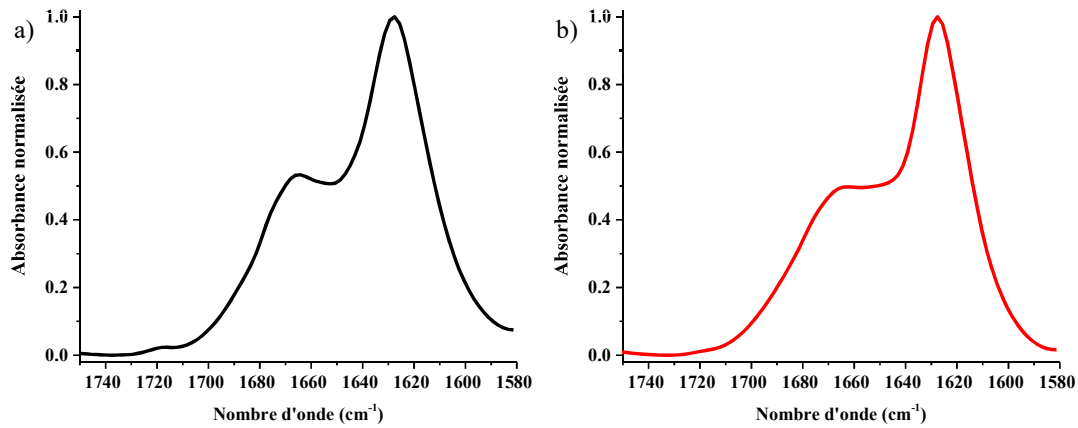


Figure 64: Spectres ATR-IR des peptides fibrillés dans les conditions déterminées précédemment a) 35m1 et b) 35m2.

Après fibrillation comme précédemment, nous avons centrifugé les suspensions pour avoir une concentration de peptide importante au fond des tubes puis nous avons déposé  $2 \mu\text{L}$  du culot sur le diamant de mesure. Les spectres réalisés sont montrés en Figure 64.

Les deux peptides donnent des spectres quasiment identiques avec des maxima à  $1629 \text{ cm}^{-1}$  (i.e. feuillets  $\beta$  parallèles) et à  $1666 \text{ cm}^{-1}$  (i.e. courbures  $\beta$ ). Sur le spectre du peptide 35m1, nous remarquons une bande de très faible intensité à  $1718 \text{ cm}^{-1}$  qui pourrait correspondre à un glutamate qui participe à une liaison hydrogène.<sup>[96,97]</sup>

La spectroscopie infrarouge nous montre donc clairement la présence de feuillets  $\beta$  en conformation parallèles ainsi que la présence de courbures  $\beta$ , confirmant ce que nous avons vu avant par dichroïsme circulaire, avec en plus l'attribution d'une bande appartenant aux courbures  $\beta$ .

### 3.5 Discussion et conclusion

Comme expliqué en introduction, l'objectif de ce chapitre était de concevoir un ou des peptides s'autoassemblant et se repliant en hélice  $\beta$  gauche de type I ( $L\beta H$ ). En se basant sur l'outil de conception décrit lors du premier chapitre, nous avons conçu deux séquences peptidiques susceptibles de se replier en  $L\beta H$ s. Cette conception prend en considération les différentes faces (non-équivalentes) du trimère et reprend tous les codes d'une  $L\beta H$  « classique » que l'on peut retrouver dans la nature. Comme il a été suggéré dans la littérature qu'une séquence correspondant à un tour d'hélice ne suffirait pas à avoir un autoassemblage stable, nous avons conçu des peptides correspondant à deux tours d'hélice, permettant de bénéficier de préorganisation du peptide.<sup>[9]</sup> Les séquences conçues sont de 35 acides aminés, ce qui correspond à deux tours moins 1 acide aminé. Ce résidu est retiré pour éviter le clash entre deux atomes et compensé par une formylation en  $N_{\text{ter}}$  et

une amidation en C<sub>ter</sub>. Des motifs structuraux caractéristiques afin de favoriser l'assemblage et la préorganisation ont été introduits : valines et leucines (pour former le cœur hydrophobe), des asparagines (connues pour former des échelles), des résidus de charge complémentaire (pour faire des ponts salins). Des histidines ont aussi été introduites de manière à ce qu'elles soient orientées au centre du trimère, afin de contrôler l'assemblage en fonction du pH.

Ces deux peptides ont été synthétisés et purifiés, puis étudiés par différentes techniques. Les résultats de fluorescence de la thioflavine-T nous indiquent que ces peptides forment des objets de type amyloïdes avec des feuillets  $\beta$ . Ces analyses nous montrent des comportements différents entre les deux peptides, le 35m1 ne montrant pas de phase latente, alors que le 35m2 en montre une. L'absence de phase latente suggère une structuration très rapide en fibres amyloïdes qui est ici probablement due à une préorganisation élevée du système, où les monomères forment des fibres avec une cinétique rapide. Les contrôles réalisés avec de différents temps d'ajout du tampon à pH=8.5 suggèrent qu'il n'y a pas de nucléi tant que le pH est maintenu acide. Aussi, l'incubation à pH différents nous montre qu'à un pH inférieur ou proche du pKa des histidines, la fibrillation n'a pas lieu, ce qui permet de contrôler le processus de fibrillation par le pH. Ceci est compatible avec l'hypothèse des histidines présentes à l'intérieur du trimère pour former les fibres mais aussi de la formation de ponts salins. L'autre peptide, le 35m2, a un comportement amyloïdique plus classique.<sup>[98]</sup> Les deux peptides auraient un mécanisme de formation où la nucléation secondaire multiétapes serait prépondérante, dont un beaucoup plus rapide que l'autre. Ce qui est compatible avec la formation de trimères d'hélices  $\beta$  gauche où une hélice catalyserait la formation d'autres hélices.

Les observations en microscopie électronique à transmission avec coloration négative nous ont permis d'identifier la morphologie des fibres formées par les deux peptides. Pour le peptide 35m2 les fibres individuelles ont des largeurs homogènes (env. 5 nm), ce qui est compatible avec des trimères de L $\beta$ Hs. Cela n'a pas été mesuré pour le peptide 35m1 au vu de l'empaquetage des fibres formées. Leurs morphologies sont cependant un peu différentes : le 35m1 a tendance à former des faisceaux/ réseaux de fibres courtes en paquets, le 35m2 forme plutôt des fibres très longues, homogènes en diamètre beaucoup moins agrégées (avec quand même la présence de fagot de fibres). Nous avons essayé d'optimiser les conditions d'incubation pour le peptide 35m2 qui nous paraissait le plus prometteur afin de n'avoir que des fibres non agrégées, sans arriver à ne plus avoir d'agrégats du tout.

Malgré la présence de paquets de fibres, nous avons réalisé des essais de Cryo-TEM pour voir si les fibres étaient visibles en solution, ce qui est le cas, et si nous pouvions déterminer la taille de celles-ci. Les fibres étaient très homogènes en diamètre, de l'ordre de 4-5 nm, ce qui est compatible avec la formation de fibres trimériques d'hélices  $\beta$  gauche. Une périodicité d'environ 35-40 nm a pu être observée dans certains cas, mais pas tout le temps, et celle-ci est aussi compatible avec la formation de fibres trimérique de L $\beta$ Hs. L'échantillon n'ayant pas assez de fibres isolées, et le microscope TEM pas assez puissant, il n'a pas été possible de déterminer la structure, et cela demande encore des étapes d'optimisation des conditions afin d'avoir un dépôt de fibres homogène. Une solution serait de centrifuger l'échantillon afin de faire « déposer » les paquets au fond et ne garder que les fibres longues et séparées dans le surnageant.

Les expériences de dichroïsme circulaire des fibres révèlent une structure secondaire en feuillets  $\beta$ , pour le spectromètre de paillasse, et en feuillets  $\beta$  parallèles pour les expériences de CD au synchrotron. Les ratios entre les bandes à 218 et 198 et 180 nm sont différents en fonction du peptide. La différence entre les spectres des deux peptides pourrait s'expliquer par un angle de torsion des feuillets  $\beta$  différent ainsi que la présence de coudes  $\beta$  soit de différents types soit dans différentes proportions. Après fibrillation et centrifugation, les spectres CD du surnageant et du culot indiquent clairement que les peptides sont présents presque que sous forme fibrillée et qu'il ne reste que très peu de peptide en solution (sous la limite de détection). Nous avons pu suivre le repliement du peptide 35m2 en fonction du temps et de la température. Nous avons observé un temps de repliement à 40 °C bien supérieur à 3 heures qui peut s'expliquer par l'absence d'agitation, et qui peut être compensée par une augmentation de la température. L'incubation des fibres formées, pour le 35m2, après 3-4 heures à 94°C montre la robustesse et la stabilité des fibres, n'observant pas de changement dans le spectre. La stabilité des celles-ci s'illustre aussi par leur fibrillation en présence de SDS et ne montrant pas de dénaturation. Enfin le SR-LD nous a permis de voir que nos peptides peuvent s'aligner et les bandes caractéristiques d'un alignement parallèle et perpendiculaire de feuillets  $\beta$  ont été identifiées. Ces résultats sont similaires à ceux décrits auparavant pour l'A $\beta$ .

La mesure des spectres infrarouges des deux peptides nous montre des signatures identiques lorsqu'on regarde la bande amide I. Elle confirme la présence de feuillets  $\beta$  parallèles. En plus, l'IR nous renseigne sur la présence de coudes  $\beta$  dans les structures fibrillaires.

La structure attendue étant composée de feuillets  $\beta$  parallèles et de coudes  $\beta$ , ces mesures sont compatibles avec le repliement des peptides en hélice  $\beta$  gauche. Aucune autre structure secondaire n'a été observée, ce qui permet de confirmer que la fibrillation est homogène et qu'elles sont composées seulement de feuillets  $\beta$  parallèles et de coudes  $\beta$ .

Des études de stabilité pour une L $\beta$ H de type II (SBAFP-m1) ont été réalisées par l'équipe de Cox, montrant des résultats remarquables.<sup>[7,99]</sup> Dans leur cas, le spectre CD reste le même jusqu'à environ 80°C, alors que nos fibres ne changent pas de spectre même après 3h à 94°C. Nous avons, par exemple, déjà vérifié que les fibres se forment aussi en présence de 0,5% SDS. Ces résultats sont donc encourageants, et appellent à tester la stabilité de nos fibres dans une gamme plus large de conditions (i.e. en présence de trypsine, d'urée ou de chlorure de guanidinium).

Le Tableau 12 dresse un résumé des résultats obtenus pour chaque technique. Ils sont compatibles avec la formation d'objets en hélice  $\beta$  gauche, ce qui validerait la conception du peptide. Cependant, la structure n'a pas pu être obtenue, et aucune technique spectroscopique ne fournit de preuves qu'il s'agit de ce repliement-là précisément. Pour cela, nous devons optimiser les conditions en TEM afin d'avoir, pour sûr, de fibres homogènes et bien séparées, afin de réaliser des analyses en Cryo-TEM et déterminer la structure de nos fibres. Grâce à cela et à toutes les techniques déjà utilisées, nous espérons pouvoir dresser une « carte d'identité » spectroscopique de ce type de structure. Ceci qui permettrait, plus tard, d'identifier plus facilement la formation d'une structure en L $\beta$ H par des techniques accessibles.

Tableau 12 : Résumé des informations structurales apportées par chaque technique, pour chaque peptide.

Technique de caractérisation	Information apportée par peptide fibrillé	
	35m1	35m2
Fluorescence de la ThT	Fibres amyloïdes, mécanisme dominé par la nucléation secondaire multiétapes lorsque le pH est augmenté au-dessus de 7 ; pas d'assemblage à pH acide	
	Cinétique rapide, pas de phase de latence	Cinétique lente, phase de latence
TEM	Fibres courtes et agrégées	Fibres longues, homogènes et seules à pH > 7.5
Cryo-TEM	<del> </del>	Diamètre de 4-5 nm + périodicité de 30-40 nm
CD	Feuillets $\beta$	Feuillets $\beta$
SR-CD	Feuillets $\beta$ parallèles + ?	Feuillets $\beta$ parallèles + ?
SR-LD	<del> </del>	Alignement possible des fibres
ATR-IR	Feuillets $\beta$ parallèles + coudes $\beta$	Feuillets $\beta$ parallèles + coudes $\beta$

### 3.6 Références bibliographiques

- [1] K. Basu, R. L. Campbell, S. Guo, T. Sun, P. L. Davies, *Protein Sci.* **2016**, *25*, 946–958.
- [2] Z. Peng, A. S. Parker, M. D. R. R. Peralta, K. M. Ravikumar, D. L. Cox, M. D. Toney, *Biophys. J.* **2017**, *113*, 1945–1955.
- [3] S. I. Stupp, R. H. Zha, L. C. Palmer, H. Cui, R. Bitton, *Faraday Discuss.* **2013**, *166*, 9.
- [4] C. Valéry, F. Artzner, M. Paternostre, *Soft Matter* **2011**, *7*, 9583–9594.
- [5] K. Papanikolopoulou, G. Schoehn, V. Forge, V. T. Forsyth, C. Riekel, J. F. Hernandez, R. W. H. Ruigrok, A. Mitraki, *J. Biol. Chem.* **2005**, *280*, 2481–2490.
- [6] L. Adler-Abramovich, E. Gazit, *Chem. Soc. Rev.* **2014**, *43*, 6881–6893.
- [7] M. D. R. R. Peralta, A. Karsai, A. Ngo, C. Sierra, K. T. Fong, N. R. Hayre, N. Mirzaee, K. M. Ravikumar, A. J. Kluber, X. Chen, et al., *ACS Nano* **2015**, *9*, 449–463.
- [8] J. K. Myers, T. G. Oas, *Nat. Struct. Biol.* **2001**, *8*, 552–558.
- [9] J. H. Choi, C. Govaerts, B. C. H. H. May, F. E. Cohen, *Proteins Struct. Funct. Bioinforma.* **2008**, *73*, 150–160.
- [10] A. T. Hagler, L. Leiserowitz, *J. Am. Chem. Soc.* **1978**, *100*, 5879–5887.
- [11] P.-N. Cheng, J. D. Pham, J. S. Nowick, *J. Am. Chem. Soc.* **2013**, *135*, 5477–5492.

- [12] G. F. Fabiola, S. Krishnaswamy, V. Nagarajan, V. Pattabhi, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53*, 316–320.
- [13] P. Iengar, N. V. Joshi, P. Balaram, *Structure* **2006**, *14*, 529–542.
- [14] J. H. Choi, C. Govaerts, B. C. H. May, F. E. Cohen, *Proteins Struct. Funct. Genet.* **2008**, *73*, 150–160.
- [15] J. H. Choi, B. C. H. H. May, C. Govaerts, F. E. Cohen, *Structure* **2009**, *17*, 1014–1023.
- [16] D. Baker, A. K. Shiau, D. A. Agard, *Curr. Opin. Cell Biol.* **1993**, *5*, 966–970.
- [17] Q. Shao, K. M. Wong, D. T. Seroski, Y. Wang, R. Liu, A. K. Paravastu, G. A. Hudalla, C. K. Hall, *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 4710–4717.
- [18] Y. Yang, *Side Reactions in Peptide Synthesis*, Elsevier Science, **2015**.
- [19] Y. Wang, Q. Shao, C. K. Hall, *J. Biol. Chem.* **2016**, *291*, 22093–22105.
- [20] G. J. Kleywegt, T. A. Jones, *Structure* **1995**, *3*, 535–540.
- [21] A. Martínez, *Amino Acids* **1995**, *9*, 285–292.
- [22] J. P. Gallivan, D. A. Dougherty, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9459–9464.
- [23] C. A. Rohl, C. E. M. Strauss, K. M. S. Misura, D. Baker, *Methods Enzymol.* **2004**, *383*, 66–93.
- [24] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al., *Nature* **2021**, *596*, 583–589.
- [25] D. E. Kim, D. Chivian, D. Baker, *Nucleic Acids Res.* **2004**, *32*, 526–531.
- [26] M. Mirdita, S. Ovchinnikov, M. Steinegger, *bioRxiv* **2021**, 2021.08.15.456425.
- [27] M. Amblard, J.-A. A. Fehrentz, J. Martinez, G. Subra, *Mol. Biotechnol.* **2006**, *33*, 239–254.
- [28] S. I. A. Cohen, M. Vendruscolo, C. M. Dobson, T. P. J. Knowles, *J. Mol. Biol.* **2012**, *421*, 160–171.
- [29] X. Deng, A. Gonzalez Llamazares, J. M. Wagstaff, V. L. Hale, G. Cannone, S. H. McLaughlin, D. Kureisaite-Ciziene, J. Löwe, *Nat. Microbiol.* **2019**, *4*, 2357–2368.
- [30] N. Greenfield, G. D. Fasman, *Biochemistry* **1969**, *8*, 4108–4116.
- [31] N. J. Greenfield, *Nat. Protoc.* **2007**, *1*, 2876–2890.
- [32] S. D. Moran, M. T. Zanni, *J. Phys. Chem. Lett.* **2014**, *5*, 1984–1993.
- [33] R. Khurana, A. L. Fink, *Biophys. J.* **2000**, *78*, 994–1000.
- [34] M. Biancalana, S. Koide, *Biochim. Biophys. Acta - Proteins Proteomics* **2010**, *1804*, 1405–1412.
- [35] G. Musteikyte, M. Ziaunys, V. Smirnovas, *PeerJ* **2020**, *8*, e9719.
- [36] G. V. De Ferrari, W. D. Mallender, N. C. Inestrosa, T. L. Rosenberry, *J. Biol. Chem.* **2001**, *276*, 23282–23287.
- [37] G. Meisl, J. B. Kirkegaard, P. Arosio, T. C. T. Michaels, M. Vendruscolo, C. M. Dobson, S. Linse, T. P. J. Knowles, *Nat. Protoc.* **2016**, *11*, 252–272.
- [38] M. Törnquist, T. C. T. Michaels, K. Sanagavarapu, X. Yang, G. Meisl, S. I. A. Cohen, T. P. J. Knowles, S. Linse, *Chem. Commun.* **2018**, *54*, 8667–8684.
- [39] V. Foderà, M. Groenning, V. Vetri, F. Librizzi, S. Spagnolo, C. Cornett, L. Olsen, M. Van De Weert, M. Leone, *J. Phys. Chem. B* **2008**, *112*, 15174–15181.
- [40] E. V. Hackl, J. Darkwah, G. Smith, I. Ermolina, *Eur. Biophys. J.* **2015**, *44*, 249–261.
- [41] S. L. Gras, L. J. Waddington, K. N. Goldie, in *Protein Folding, Misfolding, and Disease*, Humana Press, Totowa, NJ, **2011**, pp. 197–214.
- [42] J. R. Glover, A. S. Kowal, E. C. Schirmer, M. M. Patino, J. J. Liu, S. Lindquist, *Cell* **1997**, *89*, 811–819.



- [43] C. Liu, M. R. Sawaya, P.-N. Cheng, J. Zheng, J. S. Nowick, D. Eisenberg, *J. Am. Chem. Soc.* **2011**, *133*, 6736–6744.
- [44] Q. Cao, D. R. Boyer, M. R. Sawaya, R. Abskharon, L. Saelices, B. A. Nguyen, J. Lu, K. A. Murray, F. Kandeel, D. S. Eisenberg, *Nat. Struct. Mol. Biol.* **2021**, *28*, 724–730.
- [45] T. Dokland, M. M. L. Ng, *Tech. Microsc. Biomed. Appl.* **2006**, 153–211.
- [46] U. Aebi, T. D. Pollard, *J. Electron Microsc. Tech.* **1987**, *7*, 29–33.
- [47] J. Dubochet, M. Ducommun, M. Zollinger, E. Kellenberger, *J. Ultrastructure Res.* **1971**, *35*, 147–167.
- [48] R. F. Thompson, M. Walker, C. A. Siebert, S. P. Muench, N. A. Ranson, *Methods* **2016**, *100*, 3–15.
- [49] L. A. Passmore, C. J. Russo, *Methods Enzymol.* **2016**, 51–86.
- [50] G. Weissenberger, R. J. M. Henderikx, P. J. Peters, *Nat. Methods* **2021**, *18*, 463–471.
- [51] M. J. Dobro, L. A. Melanson, G. J. Jensen, A. W. McDowell, *Plunge Freezing for Electron Cryomicroscopy*, Elsevier Inc., **2010**.
- [52] P. Schultz, J. Dubochet, M. Adrian, J.-J. Chang, J.-C. Homo, J. Lepault, A. W. McDowell, P. Schultz, *Q. Rev. Biophys.* **1988**, *21*, 129–228.
- [53] J. Dubochet, A. W. McDowell, *J. Microsc.* **1981**, *124*, 3–4.
- [54] A. Hoffmann, V. Perrier, S. Grudinin, *J. Appl. Crystallogr.* **2017**, *50*, 1036–1047.
- [55] D. Jeulin, M. Moreaud, *Image Anal. Stereol.* **2008**, *27*, 183–192.
- [56] E. D. Sone, S. Weiner, L. Addadi, *J. Struct. Biol.* **2007**, *158*, 428–444.
- [57] S. M. Kelly, T. J. Jess, N. C. Price, *Biochim. Biophys. Acta - Proteins Proteomics* **2005**, *1751*, 119–139.
- [58] L. Whitmore, B. A. Wallace, *Biopolymers* **2008**, *89*, 392–400.
- [59] A. T. B. Gilbert, J. D. Hirst, *J. Mol. Struct. THEOCHEM* **2004**, *675*, 53–60.
- [60] M. G. H. María Florencia Pignataro, *Molecules* **2020**, *25*, 121–140.
- [61] A. J. Miles, B. A. Wallace, *Chem. Soc. Rev.* **2016**, *45*, 4859–4872.
- [62] E. Y. Hayden, K. K. Hoi, J. Lopez, M. Inayathullah, M. M. Condron, D. B. Teplow, *Sci. Rep.* **2017**, *7*, 1–16.
- [63] J. Meuis, M. Gerard, L. Desender, V. Baekelandt, Y. Engelborghs, *Biochemistry* **2010**, *49*, 9345–9352.
- [64] H. D. Flack, G. Bernardinelli, *Chirality* **2008**, *20*, 681–690.
- [65] P. Evans, K. Wyatt, G. J. Wistow, O. A. Bateman, B. A. Wallace, C. Slingsby, *J. Mol. Biol.* **2004**, *343*, 435–444.
- [66] P. L. Davies, *Trends Biochem. Sci.* **2014**, *39*, 548–555.
- [67] S. A. Buth, L. Menin, M. M. Shneider, J. Engel, S. P. Boudko, P. G. Leiman, *Viruses* **2015**, *7*, 4676–4706.
- [68] R. W. Woody, in (Ed.: G.D. Fasman), Springer US, Boston, MA, **1996**, pp. 25–67.
- [69] A. J. Miles, R. W. Janes, B. A. Wallace, *Chem. Soc. Rev.* **2021**, *50*, 8400–8413.
- [70] R. Acharya, M. Gupta, S. Ramakumar, U. A. Ramagopal, V. S. Chauhan, *BMC Struct. Biol.* **2007**, *7*, 1–9.
- [71] A. Micsonai, F. Wien, L. Kernya, Y.-H. H. Lee, Y. Goto, M. Réfrégiers, J. Kardos, *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, E3095–E3103.
- [72] D. T. Clarke, G. Jones, *J. Synchrotron Radiat.* **2004**, *11*, 142–149.
- [73] B. M. Bulheller, A. J. Miles, B. A. Wallace, J. D. Hirst, *J. Phys. Chem. B* **2008**, *112*, 1866–1874.
- [74] A. J. Miles, B. A. Wallace, *Chem. Soc. Rev.* **2006**, *35*, 39–51.

- [75] N. Sreerama, R. W. Woody, *Anal. Biochem.* **1993**, *209*, 32–44.
- [76] S. W. Provencher, J. Gloeckner, *Biochemistry* **1981**, *20*, 33–37.
- [77] M. Réfrégiers, F. Wien, H.-P. Ta, L. Premvardhan, S. Bac, F. Jamme, V. Rouam, B. Lagarde, F. Polack, J.-L. Giorgetta, et al., *J. Synchrotron Radiat.* **2012**, *19*, 831–835.
- [78] R. W. Woody, in *Circular Dichroism and the Conformational Analysis of Biomolecules*, Springer US, Boston, MA, 1996, pp. 25–67.
- [79] N. J. Anthis, G. M. Clore, *Protein Sci.* **2013**, *22*, 851–858.
- [80] A. J. Miles, F. Wien, J. G. Lees, A. Rodger, R. W. Janes, B. A. Wallace, *Spectroscopy* **2003**, *17*, 653–661.
- [81] M. C. Manning, M. Illangasekare, R. W. Woody, *Biophys. Chem.* **1988**, *31*, 77–86.
- [82] B. M. Bulheller, A. Rodger, J. D. Hirst, *Phys. Chem. Chem. Phys.* **2007**, *9*, 2020–2035.
- [83] T. R. Dafforn, J. Rajendra, D. J. Halsall, L. C. Serpell, A. Rodger, *Biophys. J.* **2004**, *86*, 404–410.
- [84] B. M. Bulheller, A. Rodger, M. R. Hicks, T. R. Dafforn, L. C. Serpell, K. E. Marshall, E. H. C. Bromley, P. J. S. King, K. J. Channon, D. N. Woolfson, et al., *J. Am. Chem. Soc.* **2009**, *131*, 13305–13314.
- [85] T. L. Benzinger, D. M. Gregory, T. S. Burkoth, H. Miller-Auer, D. G. Lynn, R. E. Botto, S. C. Meredith, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 13407–13412.
- [86] M. R. Hicks, T. R. Dafforn, A. Damianoglou, P. Wormell, A. Rodger, S. V. Hoffmann, *Analyst* **2009**, *134*, 1623–1628.
- [87] A. Barth, C. Zscherp, *Q. Rev. Biophys.* **2002**, *35*, 369–430.
- [88] A. Barth, *Biochim. Biophys. Acta - Bioenerg.* **2007**, *1767*, 1073–1101.
- [89] J. L. R. Arrondo, A. Muga, J. Castresana, F. M. Goñi, *Prog. Biophys. Mol. Biol.* **1993**, *59*, 23–56.
- [90] E. Cerf, R. Sarroukh, S. Tamamizu-Kato, L. Breydo, S. Derclaye, Y. F. Dufrière, V. Narayanaswami, E. Goormaghtigh, J.-M. Ruyschaert, V. Raussens, *Biochem. J.* **2009**, *421*, 415–423.
- [91] J. Wang, *Phys. Chem. Chem. Phys.* **2009**, *11*, 5310–5322.
- [92] Y. El Khoury, G. Le Breton, A. V. Cunha, T. L. C. Jansen, L. J. G. W. Van Wilderen, J. Bredenbeck, *J. Chem. Phys.* **2021**, *154*.
- [93] W. K. Surewicz, H. H. Mantsch, D. Chapman, **1993**, 389–394.
- [94] L. E. Valenti, M. B. Paci, C. P. De Pauli, C. E. Giacomelli, *Anal. Biochem.* **2011**, *410*, 118–123.
- [95] J. Cornish, K. E. Callon, C. Q. X. Lin, C. L. Xiao, T. B. Mulvey, G. J. S. Cooper, I. R. Reid, *Am. J. Physiol. - Endocrinol. Metab.* **1999**, *277*, 779–783.
- [96] S. J. Parikh, J. D. Kubicki, C. M. Jonsson, C. L. Jonsson, R. M. Hazen, D. A. Sverjensky, D. L. Sparks, *Langmuir* **2011**, *27*, 1778–1787.
- [97] H. Tai, K. Nishikawa, Y. Higuchi, Z. wan Mao, S. Hirota, *Angew. Chemie - Int. Ed.* **2019**, *58*, 13285–13290.
- [98] A. K. Buell, *Biochem. J.* **2019**, *476*, 2677–2703.
- [99] Z. Peng, M. D. R. Peralta, M. D. Toney, *Biochemistry* **2017**, *56*, 6041–6050.

## **4 Séquence peptidique minimale afin de former une hélice $\beta$ gauche**

## 4.1 Introduction

Dans ce chapitre, un peu dans le sens opposé à la stratégie employée précédemment, nous avons voulu concevoir des peptides de taille minimale permettant de former des hélices  $\beta$  gauche, ou du moins des assemblages qui en sont proches. Ceci reposerait sur la formation de réseaux de liaisons hydrogène entre les briques moléculaires permettant de former des édifices plus importants. Avec des peptides très courts, le contrôle d'assemblages supramoléculaires et la formation de cristaux ne sont pas aussi simples qu'avec certaines molécules organiques par exemple.<sup>[1]</sup> Ceci est notamment dû aux différentes conformations que peuvent adopter les acides aminés, et au manque de préorganisation (i.e. interactions intramoléculaires) des peptides courts.<sup>[2]</sup> Aujourd'hui, les avancées technologiques permettent de mieux comprendre ces interactions, notamment grâce à des résolutions bien meilleures.<sup>[3,4]</sup>

En effet, les peptides et protéines sont composés d'un enchaînement de différents acides aminés, constituant un polymère naturel via la formation de liaison amide entre les différents résidus. Cet enchaînement constitue le codage de l'information d'une protéine, et s'appelle la séquence primaire. Cette séquence conditionne les interactions intra-moléculaires que la protéine va pouvoir former, et donc, la formation de structures secondaires. On y retrouve les hélices  $\alpha$  et les feuillets  $\beta$  constituant des structures secondaires très fréquentes dans les protéines naturelles. En ce qui concerne l'assemblage de solénoïdes  $\beta$  à partir de chaînes polypeptidiques, les interactions principales correspondent à celles entre les fonctions amides.<sup>[5,6]</sup> Nous nous intéresserons plus particulièrement aux feuillets  $\beta$  qui sont présents en majorité (avec les coudes  $\beta$ ) dans les objets que nous visons à concevoir. Il est important de noter ici que les feuillets  $\beta$  peuvent adopter deux conformations principales étant parallèles ou antiparallèles, les plus présentes dans la Nature étant l'antiparallèle. Comme ce que nous souhaitons construire est des  $\beta$  solénoïdes, constitués de feuillets  $\beta$  parallèles et de coudes  $\beta$ , nous nous attendons à des signatures spectroscopiques caractéristiques de ces types de repliement secondaire.

À travers ce projet de conception d'hélice  $\beta$  gauche par le biais de peptide de taille réduite, nous visons à déterminer la taille de séquence minimale permettant de coder l'information pour permettre la formation de cette structure.

## 4.2 Contexte

Afin de former des édifices de cette manière, à partir de petites séquences peptidiques, nous avons besoin de définir quelques hypothèses de travail. En effet, avec de courtes séquences, la préorganisation due aux interactions intra-moléculaires (du monomère) n'a plus lieu. Cela implique alors de prendre en compte ce paramètre lors de la conception de ces peptides. Il faut en conséquence choisir les résidus afin de maximiser leurs interactions intermoléculaires, mais aussi, les quelques interactions intramoléculaires auxquelles on peut

s'attendre. Les contributions intermoléculaires permettent de maintenir un réseau de liaisons hydrogène au travers des feuillets  $\beta$  mais aussi, de former le cœur hydrophobe de l'hélice par association de plusieurs molécules. Un pas d'hélice est constitué de 18 résidus répartis sur trois faces, donc six par face. Les travaux d'Iconomidou et collaborateurs montrent que, pour des séquences de six acides aminés, il était très difficile, voire impossible de former des hélices  $\beta$  gauche qui soient vraiment fidèles à celles retrouvées dans la Nature.<sup>[7,8]</sup> Même si des travaux montrent que ces peptides ont des signatures spectroscopiques allant dans le sens de la formation de L $\beta$ Hs, rien n'indique la formation de celles-ci.<sup>[7-9]</sup>

### 4.3 Conception des séquences

L'hélice faisant 18 acides aminés de section, il existe de nombreuses possibilités de division de cette dernière. La taille de séquence ne pourra pas être, dans tous les cas, inférieure à 1/3 de pas d'hélice, car, l'assemblage étant de section triangulaire, 1/3 de pas constitue l'unité minimale possible pour former un triangle. Cependant, il faut garder en tête que nous parlons de séquences étant de cinq résidus, ce qui sous-entend une très faible préorganisation, voire nulle, et donc une flexibilité conformationnelle très importante. Nous devons alors choisir des peptides qui permettront d'optimiser la formation de mimes de L $\beta$ Hs. Nous sommes partis de l'idée que si l'on conçoit des peptides dont la séquence correspond aux feuillets  $\beta$  de l'hélice, ces peptides vont très probablement s'assembler en feuillets orientés dans la même direction plutôt que formant un triangle. Il nous a alors paru plus judicieux de centrer la séquence sur les parties contenant des coudes  $\beta$  (Figure 65). Ces coudes pouvant comporter une proline, en position T<sub>1</sub>, comme l'a été reporté précédemment (voir chapitre 1).<sup>[10]</sup>

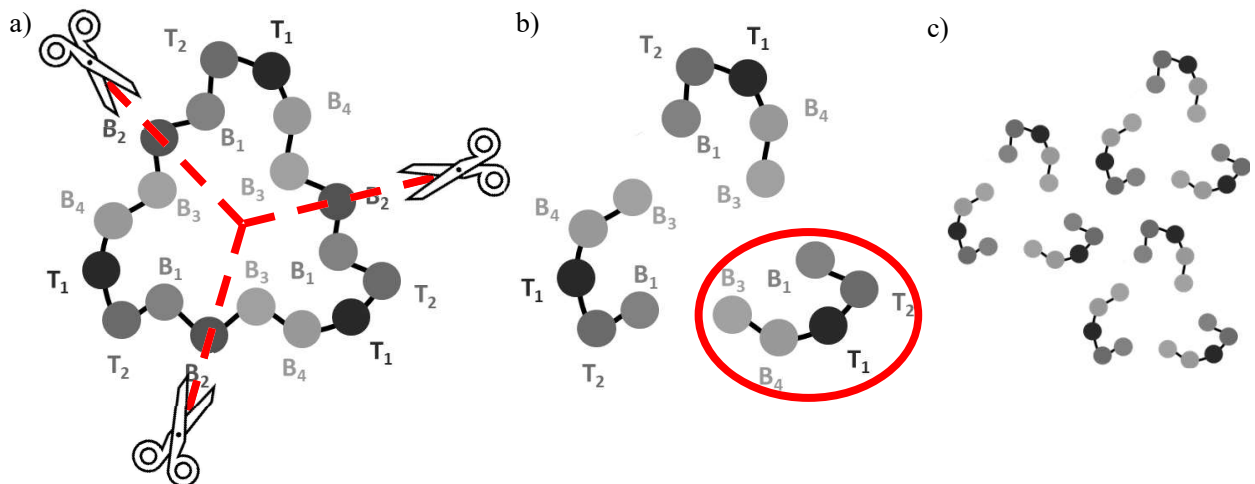


Figure 65 : Pas d'une hélice  $\beta$  gauche de type I représenté avec les positions correspondantes a) avant découpage en 3 parties distinctes, b) après, la séquence minimale étant entourée en rouge, c) arrangement hypothétique des peptides dans les axes x et y, l'axe z étant le plus évident en termes d'assemblage (empilement), il n'est pas montré par souci de clarté.

Nous avons alors décidé de centrer nos séquences autour de cette proline qui apporte une contrainte structurale à la chaîne principale. Ce qui résulte donc d'une séquence générique de type : **f-XXPXX-Am**. Ce qui veut dire que la position supprimée correspond à B<sub>2</sub> qui, nous pensons, est la position la moins impactante lors de la suppression pour ces petites unités peptidiques.

Nous avons décidé de choisir une séquence de cinq résidus, ce qui correspond à une face d'hélice moins un A.A. Ce choix s'explique par le fait que, naturellement, les hélices β gauche sont formées d'une seule et unique chaîne peptidique qui s'enroule sur elle-même en section triangulaire (Figure 14 de l'intro). Ici, le but étant de former des mimes de ce type d'édifice, il y aura plusieurs chaînes interagissant entre-elles. C'est pour cela que l'amidation et la formylation ont été réalisées comme précédemment (Figure 29). Nous avons alors choisi d'appliquer ces modifications à tous nos peptides synthétisés dans cette partie afin d'avoir une séquence de type Formyle(f)-peptide-Amide(Am).

Nous savons aussi que le cœur hydrophobe est très important, donc nous conservons les positions qui y participent, c'est-à-dire B<sub>1</sub> et B<sub>3</sub> (Figure 12 de l'introduction). Il est connu que la thréonine participe à une liaison hydrogène (avec son groupement hydroxyle et le carbonyle de la chaîne principale du résidu en position i+15 et l'amide i+18 après la thréonine), mais aussi au cœur hydrophobe par son groupe méthyle en position T<sub>1</sub>. Il reste alors la valine, isoleucine, alanine, et phénylalanine pour la position B<sub>3</sub>, et nous avons choisi une valine car, elle est retrouvée en majorité, et elle n'est pas trop encombrante pour l'intérieur de l'hélice. Ce qui donne, pour commencer, une séquence de type **f-VXPXT-Am**.

Il reste maintenant à trouver les résidus appropriés pour les positions restantes, à savoir B<sub>4</sub> et T<sub>2</sub>. Il est évident de placer une glycine en B<sub>4</sub>, qui se retrouve à plus de 74 % dans les différentes structures (voir chapitre 1). Dans cette séquence, c'est la position T<sub>2</sub> qui semble la plus libre, et nous allons la faire varier. Ce choix est supporté par le fait que les résidus à cette position se retrouvent dans les coudes et pointent entre les hélices, ce qui permet de favoriser, selon les résidus, les interactions entre les peptides.

Nous avons alors une séquence finale générique **f-VGPXT-Am** avec en X (T<sub>2</sub>) des résidus tels que L, D, N, E, H, F, Q. Dans l'hypothèse où ces peptides s'assemblent en mimes d'hélice β gauche, la répétition de la séquence conçue aura pour conséquence d'apporter une forte symétrie au système. Nous nous attendons alors à une symétrie dans les trois directions de l'espace, ce qui revient à la formation d'un cristal. C'est pour ces raisons que pour les étudier, nous avons décidé de les cristalliser. Les peptides seront nommés par la suite 5M<sub>X</sub> pour pentamère avec X correspondant au résidu en position T<sub>2</sub>. Une séquence différente a été conçue afin de comprendre l'influence de la proline en position T<sub>1</sub>.

## 4.4 Propriétés des peptides

### 4.4.1 Synthèses

Les séquences étant très petites, il est alors très facile de synthétiser ces peptides par synthèse en phase solide classique Fmoc/tBu manuelle, et cela sur des échelles importantes au laboratoire. La majorité des synthèses effectuées sur ces peptides a été réalisée sur des échelles de synthèse de 400  $\mu\text{mol}$  avec des rendements finaux de 35 à 75 %. Cela a permis d'avoir de la matière pour réaliser les analyses par la suite. Ces peptides ont été analysés par HPLC analytique, et purifiés par HPLC semie-préparative en phase inverse. Des solutions stocks ont été préparées à 10 et 5 mg/ml de peptide dans l'eau milliQ.

### 4.4.2 Solubilité

La série de séquences synthétisées a permis de « screener » les résidus à la position X ( $T_2$ ) afin d'observer les différences de propriétés et structures en fonction de l'acide aminé à cette position. L'une de ces propriétés est la solubilité. Par exemple, la présence d'un acide aspartique à cette position conduit à un peptide étant peu soluble dans l'eau MilliQ ou dans le tampon A, acide. Dans le but d'en préparer une solution, nous avons dû augmenter le pH au-dessus de 6,5 avec de la soude (i.e. au-dessus du pKa du résidu), qui lorsqu'il est déprotoné permet sa solubilisation dans l'eau. Un autre peptide a posé des problèmes de solubilités, il s'agit de celui comportant une glutamine et une asparagine (f-VGNQT-Am), qui après déprotection et lyophilisation formait un hydrogel assez dense, même en présence d'acétonitrile comme co-solvant. Nous pensons que cela est dû à la formation d'échelles de glutamines et d'asparagines qui formeraient des réseaux tridimensionnels stables.<sup>[11]</sup> Concernant ce peptide, la purification n'a pas été réalisée et aucune autre analyse que HPLC et LC-MS a été réalisée.

## 4.5 Vers une structure en $L\beta H$ ?

S'il s'assemble tel qu'escompté, le système conçu devrait être très symétrique : les trois coins de l'hélice étant composés de la même séquence. Ainsi, chaque coin peut être vu comme celui au cœur du trimer d'hélice (Figure 65). Donc, nous nous attendons à ce que ces peptides s'assemblent dans les trois différentes directions de l'espace, c'est-à-dire, à la formation d'un cristal. La diffraction des rayons X (DRX) nous permettra d'obtenir la structure de ces édifices. Une autre technique, à ne pas négliger, est la résonance magnétique nucléaire (RMN) qui peut nous permettre de résoudre leur structure sous forme cristallisée. Elle peut aussi, confirmer les résultats obtenus par DRX, ou inversement.



## 4.5.1 Dichroïsme circulaire

Les différents peptides n'ont pas tous été analysés par des techniques de spectroscopie en solution. Seulement un, le dérivé avec la leucine (5M<sub>L</sub>), a été analysé par dichroïsme circulaire. Or, de toute la série, c'est ce peptide qui s'assemble (cristallise) le mieux, donc nous nous attendions à ce que ce soit, de toute la série, celui qui est le plus probablement partiellement replié et/ou assemblé en solution. Ce n'est pas le cas, donc les autres peptides n'ont pas été analysés. Le spectre en Figure 66 nous permet d'observer un minima à 205 nm pour des pH de 7.4 ou 9. Dans les conditions que nous avons testées, pour ce peptide, nous l'avons toujours observé non-structuré (Figure 66). Des concentrations élevées, jusqu'à 250 µM de peptide dans du tampon HEPES 500 mM à pH=7.4 ainsi que l'utilisation de sel (sulfate d'ammonium) à concentration élevée (500mM) n'a pas permis l'observation d'une structure secondaire définie.

Nous avons aussi réalisé, en phase solide cette fois-ci, des spectres de dichroïsme circulaire pour déterminer si, avec les cristaux de peptides obtenus, nous observons des signaux caractéristiques à l'objet recherché. Ces expériences ont été réalisées sur la ligne DISCO au synchrotron Soleil.

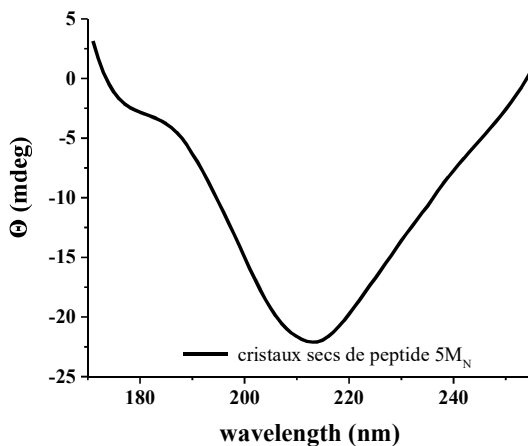


Figure 67: Spectre SR-CD (lissé) du peptide 5M<sub>N</sub> cristallisé.

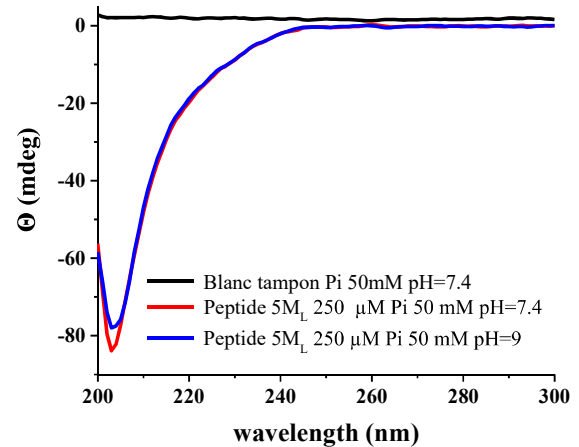


Figure 66: Spectre CD du peptide 5M<sub>L</sub> mettant en évidence la désorganisation de ce dernier en solution.

Pour pouvoir enregistrer des spectres CD de nos cristaux (du peptide 5M<sub>N</sub> de séquence f-VGPNT-Am), nous en avons étalé sur la cellule de mesure en CaF<sub>2</sub> de manière la plus homogène possible, à l'aide d'une pointe de pipette scellée préalablement. Ensuite la cellule de mesure a été montée sur l'appareil et trois spectres ont été enregistrés de 170 à 255 nm. Une moyenne de ces trois spectres a été réalisée afin d'avoir un meilleur rapport signal sur bruit. Nous pouvons observer la présence d'une bande caractéristique des feuillettes β avec un minimum à 218 nm (Figure 67), mais pas de signal positif à 200 nm. Il est important de rappeler ici que l'échantillon n'est pas en solution, mais solide, ce qui peut expliquer les différences possibles avec une structure secondaire de type feuillettes β en solution.



## 4.5.2 Spectroscopie Infra-rouge

Dans la même optique que dans le Chapitre 2 de cette thèse, de la spectroscopie infra-rouge a été réalisée pour savoir si les peptides dans les cristaux formés adoptent une conformation en feuillets  $\beta$  parallèles ou non, et si nous avons la présence de coudes  $\beta$ . Pour cela, nous avons déposé les cristaux de peptide 5M<sub>L</sub> directement sur le diamant de l'appareil et nous avons mesuré le spectre ATR-IR. Nous avons dû attendre quelques minutes le temps que les cristaux sèchent totalement, et afin d'avoir seulement le signal du peptide cristallisé. Les signaux obtenus sont caractéristiques des feuillets  $\beta$ , mais antiparallèles avec des maxima caractéristiques à 1680 et 1631  $\text{cm}^{-1}$  (Figure 68). Un épaulement à 1650  $\text{cm}^{-1}$  suggère la présence de structures désordonnée, pouvant être due au peu de peptide restant en solution. Afin de pouvoir confirmer cela, et même d'avoir des informations structurales plus précises, nous nous sommes tournés vers d'autres techniques comme la diffraction des rayons X ou la RMN du solide.

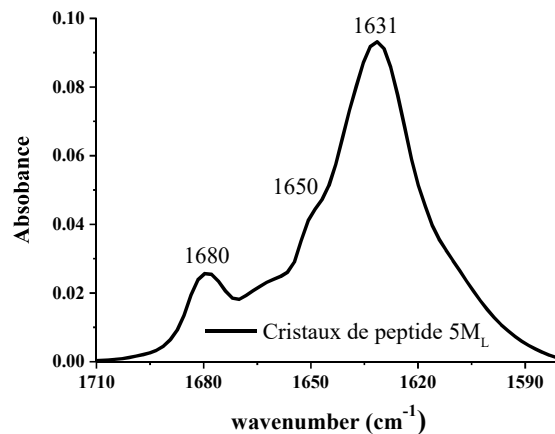


Figure 68: ATR-IR de cristaux composés du peptide 5M<sub>L</sub>.

## 4.5.3 Diffraction des rayons X

La partie de préparation et d'optimisation des conditions de cristallisation en vue d'analyse en DRX a demandé beaucoup de temps. Au total, plusieurs milliers de conditions de cristallisations différentes ont été testées avec différents peptides et avec des techniques de cristallisation différentes. Dans cette partie, nous allons voir comment il a été possible de transférer les conditions d'une technique à une autre, dans le but d'avoir des cristaux de meilleure qualité, et en plus grande quantité.

### 4.5.3.1 Cristallogénèse en goutte assise

Il s'agit-là de la méthode la plus commune pour la cristallographie des protéines. Cette technique de cristallisation permet de tester rapidement différentes conditions en même temps car cela est réalisé sur des plaques multipuits. Les solutions peuvent d'ailleurs être achetées déjà préparées (commerciales). Elle permet

d'identifier rapidement les conditions où les cristaux/précipités/séparations de phase, etc. se produisent et d'affiner les tests autour les conditions les plus intéressantes. Étant novices dans ce domaine, nous avons dû nous renseigner, apprendre, et acheter le matériel nécessaire pour pouvoir effectuer ces essais. Nous avons alors acheté des screens commerciaux dans un premier temps (Hampton

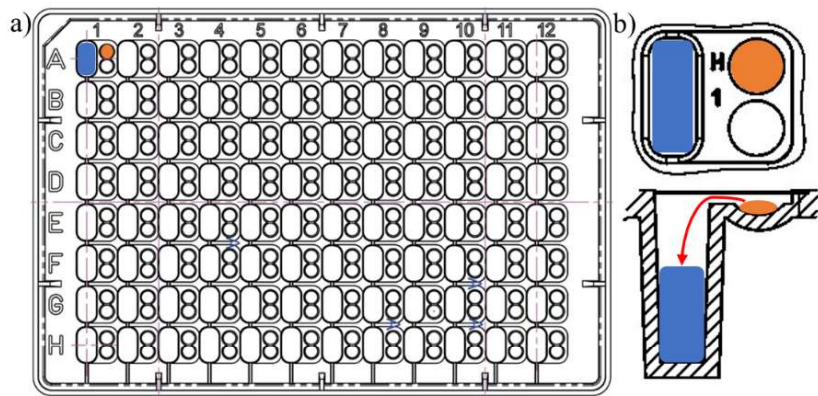


Figure 69 : Schéma (du dessus) de la plaque de cristallisation 96 puits utilisée (Hampton research) avec la solution du réservoir en bleu et de la goutte de cristallisation en orange, b) vue de face et profil et chemin de diffusion de la vapeur, globalement de la goutte vers le puits.

Research), ainsi que des plaques de cristallisation permettant de tester 192 conditions sur une seule plaque (Figure 69 a). L'avantage de commencer avec un screen commercial (i.e. Index, Crystal Screen) est de pouvoir regarder, dans un large spectre de conditions, le comportement du peptide lors de la concentration de la goutte. Les plaques de cristallisation que nous avons utilisées sont des plaques permettant de faire de la cristallisation en goutte assise. Ici, la goutte contenant le peptide repose sur un support (Figure 69 b), et est en équilibre de vapeur avec la solution du puits, contenant le même tampon et sel que la goutte, mais plus concentrés. Pour ces plaques-ci, nous avons déposé, dans le creux de la goutte, 1  $\mu$ L de solution de peptide (typiquement à 5 ou 10 mg/mL) et 1  $\mu$ L de la solution du réservoir de cristallisation (50  $\mu$ L). Le réservoir est notamment constitué de sels, tampons et autres additifs tels que des polyéthylènes glycols (PEGs) à concentrations et pH différents. La concentration de ces derniers est donc deux fois plus importante dans le réservoir que dans la goutte. Une fois le système isolé (par un film étanche), les gouttes et le réservoir tendent vers un équilibre de vapeur. C'est comme cela que, petit à petit, la goutte va se concentrer en peptide, en tampon et en additifs, afin de dépasser la limite de solubilité du peptide et permettre l'apparition d'un cristal dans les meilleurs cas. Évidemment, le risque est la diffusion trop rapide de l'eau, qui pourrait amener à la précipitation du peptide, et inversement, si elle est trop lente, nous pourrions ne pas observer de cristallisation. En fonction de la température, du volume de la goutte, du réservoir, ainsi que les concentrations en équilibre, nous pouvons contrôler l'équilibre de vapeur au fil du temps et ainsi faire cristalliser le peptide lorsque les conditions permettant la formation de cristaux sont réunies.<sup>[12]</sup>

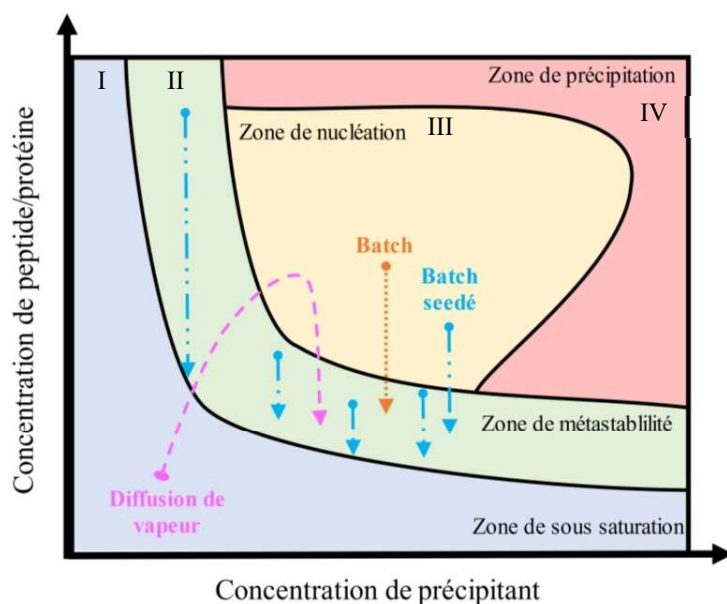


Figure 70: Diagramme de cristallisation d'une protéine/peptide en fonction de la technique utilisée (adapté de la ref<sup>[13]</sup>).

pour former des nuclei (Figure 70, flèche rose), et à mesure que les cristaux grandissent, la concentration diminue, pour retourner dans la zone de métastabilité, voire de sous saturation. Néanmoins, s'il n'y a pas de nucleus de cristal qui se forme en (III), et si la goutte continue de se concentrer, l'état de la goutte peut aller dans la zone de précipitation (IV) où la concentration de peptide va finir par diminuer par précipitation.<sup>[13,16,17]</sup>

Cette technique permet donc de trouver assez rapidement les conditions optimales de concentration de peptides, et de sels, tampons ou autres composés présents dans la goutte et le réservoir. C'est le principal avantage de la cristallisation en plaque par diffusion de vapeur, selon le type de plaque utilisée, nous pouvons avoir, pour un même réservoir, une, deux ou trois gouttes avec des concentrations en peptides différentes et tester des concentrations de peptides différentes en parallèle.

Tous les peptides cités précédemment ont été testés sur plaques multipuits avec le test commercial « Index screen », afin d'identifier les conditions dans lesquelles ils formaient des cristaux dans ces conditions. Cette expérience permet de tester, qu'avec une seule plaque, un large spectre de conditions, afin de pouvoir déterminer celles qui sont les meilleures. Elle est donc très souvent utilisée comme point de départ lors de la cristallisation en goutte assise ou suspendue.<sup>[18]</sup> Nous avons premièrement commencé les essais sur le peptide 5M<sub>L</sub>, de séquence f-VGPLT-Am. À partir des solutions de ce peptide à 10 et 5 mg/mL, nous avons réalisé une plaque de cristallisation avec la solution de cristallisation du test commercial « Index screen » (Hampton Research) (voir les volumes indiqués précédemment). Ainsi, pour chaque condition, nous avons déposé une goutte à 2.5 et une autre à 5 mg/mL de peptide. La plaque est ensuite scellée avec un film pour plaque multipuits.

Après s'être bien assuré que la plaque est bien hermétiquement scellée et que toutes les conditions le soient aussi, la plaque a été mise à l'abri de la lumière à température ambiante. Les premiers cristaux sont apparus au bout de 2 jours. Ils ressemblaient à des oursins plutôt qu'à des monocristaux (Figure 71 a). Les conditions pour lesquelles nous avons obtenu des oursins sont multiples comme il est montré dans Tableau 13, où seulement les cristaux les plus prometteurs sont

Tableau 13 : Conditions pour lesquelles, des cristaux ont été obtenus pour le Peptide 5M<sub>L</sub> f-VGPLT-Am à 5 et 2.5 mg/mL.

0.05 M Citric acid pH 3.5	<b>1.0 M Ammonium sulfate</b>
0.05 M Sodium acetate trihydrate pH 4.5	
0.05 M BIS-TRIS pH 5.5	
0.05 M BIS-TRIS pH 6.5	
0.05 M HEPES pH 7.5	
0.05 M Tris pH 8.5	

choisis. Nous remarquons que peu importe le pH, le tampon ou la concentration de peptide utilisée, des cristaux ont principalement été obtenus avec une concentration à 1 M de sulfate d'ammonium.<sup>[19]</sup> Nous avons alors continué avec le tampon qui donnait les cristaux les plus prometteurs. Nous avons aussi refait une plaque avec les mêmes conditions de cristallisation que nous avons mise à 4°C après la préparation. Cependant, contrairement à la plaque réalisée à T<sub>amb</sub>, aucun cristal n'a pu être observé après 2 mois à 4°C. Nous l'avons alors sortie et après seulement 3-4 jours à T<sub>amb</sub>, des cristaux sont apparus, semblant plus denses que ceux observés dans le cas de la plaque « initiale » (à T<sub>amb</sub> seulement) (Figure 71 b).<sup>[20]</sup>

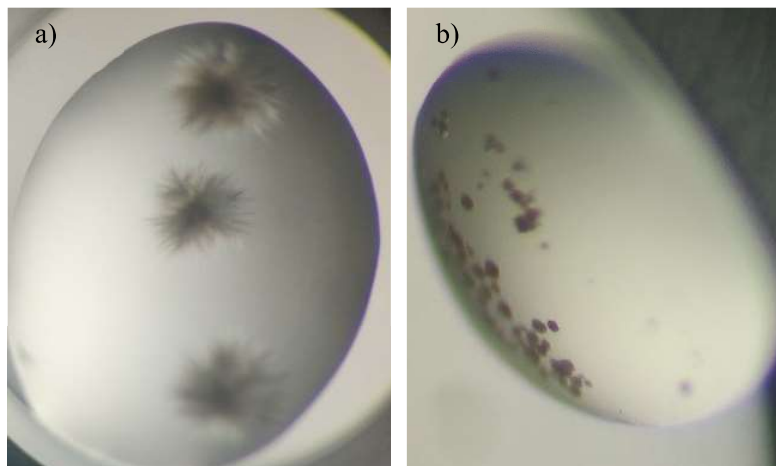


Figure 71 : Cristaux de peptide 5M<sub>L</sub> a) 5mg/mL, 50 mM HEPES pH=7.5, 1M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> à T<sub>amb</sub> après 2 semaines, b) 5mg/mL, 10 mM Chlorure de Calcium dihydrate, 10 mM Chlorure de Cadmium hydrate, 10 mM Chlorure de Cobalt (II) hexahydrate, 10% w/v Polyéthylène glycol 3,0 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> à 4°C pendant 2 mois puis T<sub>amb</sub> après 1 mois.

À partir de là, nous avons continué d'affiner les conditions de cristallisation, dans le but d'obtenir des monocristaux plutôt que des oursins. Nous avons essayé d'autres screens comme le « PEG-ION screen » par exemple mais la formation de cristaux était moins évidente. Nous avons alors choisi de partir de la condition qui nous paraissait la meilleure et nous avons l'avons reproduite « manuellement » au laboratoire et avons fait varier les paramètres autour de cette condition. Il s'agit de la condition avec 2.5 mg/mL de peptide, 50 mM d'HEPES à pH=7.5 et 1M de sulfate d'ammonium dans la goutte. Autour de cette condition, les concentrations de peptide et de sels ont été variées plus finement. Une plaque a été réalisée avec une concentration de peptide

allant de 0.5 à 3.5 mg/mL dans la goutte, pour des concentrations de sulfate d'ammonium de 0.25 à 1 M dans la goutte, tout en gardant la concentration de tampon HEPES à pH=7.4 fixée à 50 mM. Avec cette plaque, nous avons réussi à mettre en évidence la rapidité de formation des cristaux, qui ont encore été des oursins, en fonction de la concentration de chacun. Les concentrations les plus élevées en sels et peptides formaient des cristaux les plus rapidement, en 2-3 jours, tandis qu'il fallait jusqu'à 1 mois pour les concentrations les moins élevées. Cependant, dans ce cas, les cristaux étaient plus denses. Les autres peptides ont été testés de la même manière afin de trouver des conditions optimales de cristallisation pour chacun (voir plus bas).

Cependant, la plaque concernant le peptide 5M<sub>L</sub> ayant été faite en juin 2020, la température ambiante a posé un réel problème lors de sa préparation. Il a fallu être très rapide afin de préparer la plaque (même avec une pipette multicanaux) car les premières gouttes avaient commencé à s'évaporer avant la fin, et donc à se concentrer. Ce qui a amené à observer la formation de cristaux en à peine une heure. Encore une fois, les oursins formés étaient très denses et petits. Nous avons aussi observé l'apparition de très petites aiguilles, très fines et isolées. La plaque a quand même été mise à incuber à 25°C dans un incubateur pour plaque (merci à E. Ennifar, de l'IBMC). Nous avons regardé la plaque toutes les semaines afin de pouvoir suivre la formation des cristaux. Après 3 semaines, nous avons identifié une condition où les cristaux étaient plus épais que ceux obtenus précédemment, et aussi plus isolés. Cette condition correspond à celle déjà identifiée précédemment, mais avec une température contrôlée à 25 °C (Figure 72).



Figure 72: Cristaux de peptide 5M<sub>L</sub> en oursins avec aiguilles longues et isolées (conditions de cristallisation : 2.5 mg/mL, 50 mM d'HEPES à pH=7.4 et 1M de sulfate d'ammonium)

Nous avons essayé de résoudre ce problème de différentes manières. La première a été de diminuer encore les concentrations de peptides dans les conditions déjà testées pour ne pas favoriser la formation d'oursins. La deuxième a été l'ensemencement et l'utilisation de gel d'agarose.

#### 4.5.3.1.1 L'ensemencement : apport de nucléi pour optimiser la cristallisation

L'ensemencement correspond à l'ajout de cristaux préalablement obtenus puis broyés dans la solution de cristallisation fraîchement préparée.<sup>[21]</sup> L'intérêt de cette technique est multiple. Premièrement, elle permet de se passer de l'étape la plus difficile lorsque l'on veut cristalliser des protéines, à savoir l'étape de nucléation. En effet, l'aspect stochastique de la nucléation n'est plus présent grâce à la présence de nucléi.<sup>[16,22,23]</sup> Ensuite, comme vu dans le schéma de la Figure 70, nous n'avons pas à passer par la zone de nucléation, les cristaux poussent directement à partir de la zone métastable (II) (à des concentrations plus faibles) ce qui devrait mener à des cristaux de meilleure qualité. Les cristaux ici grandiront directement sur les nuclei apportés (communément avec un poil, un crin de cheval par exemple) via l'ensemencement et cela permet d'avoir des cristaux, plus volumineux, et moins polydispersés. Nous avons réalisé ceci en broyant des cristaux que nous avons obtenus dans une goutte, avec une pointe de pipette scellée, puis nous y avons trempé le crin, qui a

ensuite été trempé dans les gouttes à cristalliser (préparées dans les mêmes conditions que précédemment). Pour cette technique, nous observons des temps d'apparition de cristaux plus courts qu'habituellement. Par exemple, dans la condition de référence (2.5 mg/mL, 50 mM d'HEPES à pH=7.4 et 1M de sulfate d'ammonium) les cristaux sont apparus après un jour alors que sans ensemencement, ils sont seulement apparus après cinq jours en goutte assise. Nous avons alors réalisé une autre plaque en testant de 2 à 3.1 mg/mL en peptide 5M<sub>L</sub> avec des concentrations de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> comprises entre 0.25 et 2 M, et une concentration de tampon HEPES à pH=7.4 de 50 mM. L'ensemencement a permis de former des cristaux de peptide rapidement, en trois jours pour les conditions les plus diluées, et en quelques heures seulement, pour les conditions plus concentrées. Malheureusement, la cristallisation s'est quand même faite en oursins.

#### 4.5.3.1.2 Influence de l'agarose

C'est pour cette raison que nous nous sommes tournés vers l'agarose, qui permet la croissance des cristaux sans mouvements de convection, juste de diffusion, rappelant les conditions rencontrées en microgravité.<sup>[13,24]</sup> C'est-à-dire qu'ici, qu'une fois que le premier nucleus sera formé (ou apporté), celui-ci reste immobile, et les autres molécules de peptides vont venir, par diffusion à travers le gel, interagir avec le cristal en croissance ce qui a pour but de réduire la mosaïcité (Figure 70). Nous pouvons alors jouer sur la concentration du gel d'agarose pour jouer sur la densité du gel et donc modifier la diffusion des peptides.<sup>[25]</sup> Ici, l'étape d'ensemencement a été réalisée afin de combiner l'avantage des deux méthodes. Pour préparer ce type d'échantillon, nous faisons exactement la même procédure que pour l'étape de préparation avec ensemencement, mais l'agarose est ajoutée juste après dans la goutte, encore « chaud » pour qu'il soit liquide. Comme nous avons utilisé de l'agarose à bas point de fusion, qui fond à 28°C, l'ajout se faisait lorsqu'il était à 35 °C. La contribution de l'agarose n'était pas prise en compte lors du calcul des concentrations de peptides dans la goutte car celui-ci n'apportait que de l'eau qui s'évapore par la suite. Nous avons remarqué que pour des concentrations de d'agarose de 0.33 % les cristaux mettaient plus de temps à pousser passant de quelques jours en moyenne à plusieurs semaines, illustrant l'utilité de cette technique. Mais encore une fois, nous avons obtenu des cristaux en oursins et pas de cristaux isolés (Figure 72).

### 4.5.3.1.3 Applications à d'autres peptides

Les conditions de cristallisation des peptides par lesquelles nous avons obtenu des cristaux par diffusion de vapeur en goutte assise sont résumées par la Figure 73 que ce soit avec ou sans ensemencement. Il est important de préciser ici que les tampons utilisés sont les mêmes pour tous les peptides, 50 mM HEPES pH=7.4, sauf pour le f-VGPDT-Am qui a été réalisé avec 50 mM d'acide citrique à pH=3.5 afin de favoriser la protonation du résidu acide aspartique. D'autres peptides ont été testés comme les dérivés f-VGPXT-Am avec X= H, E, Q mais aucun cristal n'a pu être obtenu pour ces composés. Les conditions optimales pour la diffusion de vapeur par goutte assise sont montrées en Tableau 14 pour chaque peptide testé.

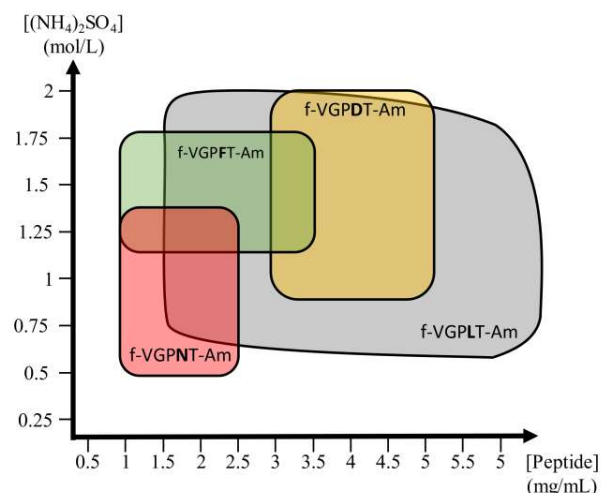


Figure 73: Conditions de cristallisation de différents peptides illustrant leurs concentrations en fonction de la concentration de précipitant au bout d'un mois d'incubation à T=30°C.

Tableau 14 : Conditions optimales de cristallisation pour les 4 peptides les plus étudiés.

Peptide f-VGPXT-Am	[Peptide] (mg/mL)	[(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ] (mol/L)	Tampon (mol/L)
5M <sub>L</sub> : X= <u>L</u>	2.5-3.0	0.75-1	HEPES pH=7.4 50 mM
5M <sub>N</sub> : X= <u>N</u>	1.5-2.0	0.5-0.75	
5M <sub>F</sub> : X= <u>F</u>	1-1.5	1-1.25	
5M <sub>D</sub> : X= <u>D</u>	3-5	1.5-1.75	Acide citrique pH=3.5 50 mM

La forme des cristaux était différente suivant les peptides, le peptide 5M<sub>L</sub> a tendance à former des oursins mais avec des aiguilles très longues et assez séparées. Le peptide 5M<sub>N</sub> forme des oursins compacts de couleur orangée mais aussi des aiguilles plus larges que le précédent. Les objets orange sont attribués à une formation trop rapide des cristaux les amenant à être trop denses. Le peptide 5M<sub>F</sub> forme des cristaux courbés et non réguliers, qui ont l'air assez amorphes. Enfin, le peptide 5M<sub>D</sub> forme quant à lui seulement des oursins très denses en épis. Des photos représentatives sont illustrées en Figure 74.

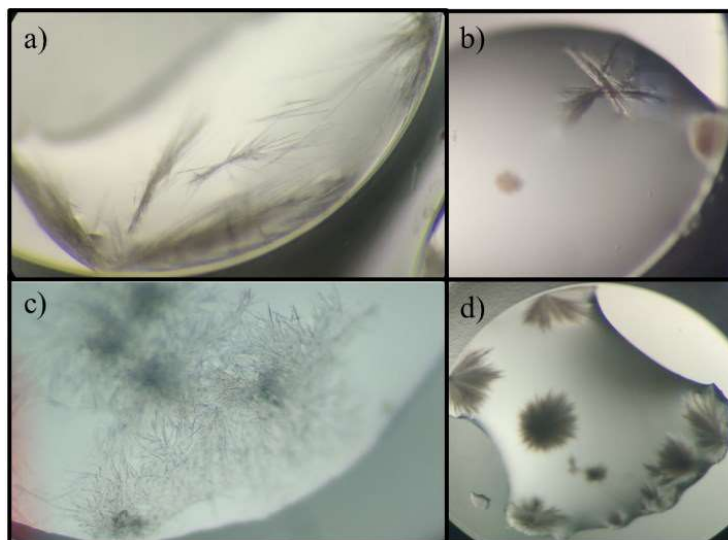


Figure 74 : Cristaux des peptides 5ML (a), 5MN (b), 5MF (c), et 5Md (d) obtenus suivant les meilleures conditions énumérées en Tableau 14 (tous les cristaux dans ces intervalles de conditions se ressemblent).

Malgré les toutes les conditions testées et le recentrage autour de celles qui étaient les plus prometteuses, nous n'avons pas réussi à obtenir des monocristaux qui soient isolés. Dans nos cas, les peptides formaient trop rapidement des cristaux, ce qui résultait en la formation d'oursins ou d'aiguilles collées. Cela peut s'expliquer par le fait que la concentration en précipitant requise pour la croissance des cristaux est plus faible que celle pour la nucléation.<sup>[26]</sup> Il y a aussi le changement des conditions entre le début de la diffusion de vapeur jusqu'à la formation de cristaux. Comme montré dans la Figure 70, avec l'évaporation de l'eau de la goutte, la concentration de peptide augmente progressivement avec celle de précipitant jusqu'à atteindre un seuil où le peptide va soit précipiter, si les concentrations sont trop élevées, soit cristalliser lorsque les concentrations le permettent.

Nous avons alors voulu essayer de contrôler un peu mieux la formation de ces cristaux en nous tournant vers une autre technique, la cristallographie en batch où les concentrations varient très peu.<sup>[27]</sup>

#### 4.5.3.2 Diffraction des rayons X sur plaques

Il a donc été possible de « pêcher » les cristaux de la plaque à l'aide d'une boucle en nylon et ensuite, de les tremper dans le Cryo-protectant, qui dans notre cas correspond à la même condition que la goutte initiale plus 20 % de glycérol<sup>[28]</sup>. Le glycérol empêche l'eau du cristal de former de la glace lors de la congélation, ce qui endommagerait le cristal. Après avoir été trempé dans cette solution, le cristal est congelé dans l'éthane liquide puis stocké dans l'azote liquide.<sup>[29]</sup> L'avantage de passer par la congélation à l'éthane est d'éviter la formation d'une couche de glace à la surface du cristal, ce qui le rendrait très difficile à analyser. Un premier essai de DRX a été réalisé par E. Ennifar (IBMC), mais aucune tâche de diffraction n'a été détectée.



Plus tard, en collaboration avec C. Sauter (IBMC), la qualité des cristaux a pu être vérifiée directement sur plaque de cristallisation sur la ligne de lumière Proxima-2A (synchrotron SOLEIL). Cela a pu être réalisé grâce au CRIBLEUR développé sur la ligne et qui permet de faire de la diffraction directement (*in situ*) sur les plaques de cristallisation.<sup>[30]</sup> On nous alors proposé d'essayer cette technique et de participer au développement de la ligne, leur but étant de déterminer la structure des cristaux de protéines directement sur plaque. Nos cristaux étant présents en quantités mais de tailles très limitées, la technique s'y prêtait bien, surtout pour pouvoir observer quels étaient les cristaux qui diffractaient.

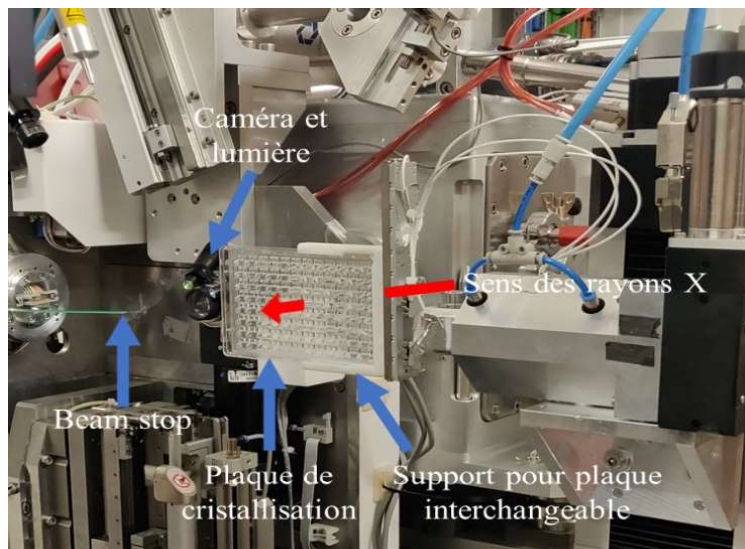


Figure 75 : Montage du cribleur pour la diffraction des rayons X *in situ* en plaque.

Les cristaux de protéines pouvant être très difficiles à obtenir, surtout de bonne qualité et de taille convenable pour réaliser une DRX dessus, le cribleur peut aider en cela. Cette technique a aussi pour avantage de se passer de l'étape de pêche et de congélation des cristaux, lorsque ceux-ci résistent à température ambiante aux rayons X.

Le montage est très simple, la plaque est placée sur un porte-plaque, porté par un bras robotisé qui remplace le goniomètre classique (Figure 75).<sup>[31]</sup> Nous avons irradié aux rayons X nos cristaux à travers la plaque, pour les conditions où ils étaient les plus homogènes et isolés. Nous avons observé que certaines aiguilles ne diffractaient pas ou très peu (Figure 76).<sup>[32]</sup> Cela est probablement dû à différents paramètres qui sont la qualité du cristal, avec des défauts présents ou non, la taille de celui-ci et l'angle d'analyse. En effet, lors d'une analyse classique en cristallographie, le cristal est placé sur un goniomètre qui va permettre de faire des rotations sur le cristal, pour collecter des tâches de diffraction.<sup>[33]</sup> Or, ici la rotation est limitée par la plaque elle-même à 45° d'amplitude, nous ne pouvions donc pas voir les autres plans comme en analyse classique avec un goniomètre (Figure 75).

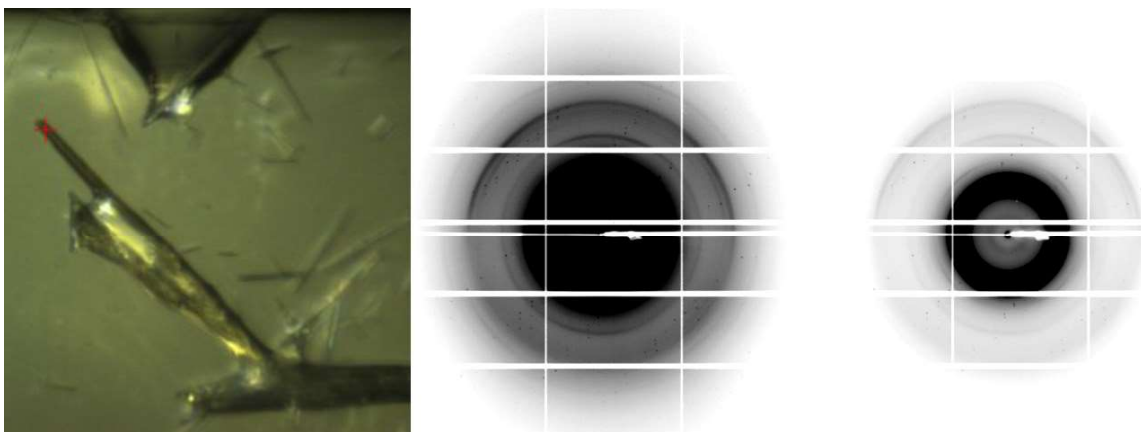


Figure 76 : Gauche : photo d'une aiguille de peptide 5ML et endroit où les rayons X ont été tirés (croix rouge). Milieu et droite : clichés de diffractions correspondants montrant la mosaïcité.

Lorsque nous avons passé un cristal en analyse avec le cribleur à température ambiante, nous avons observé des tâches de diffraction selon l'orientation de celui-ci. L'axe principal de diffraction a toujours été dans la direction perpendiculaire à l'axe de l'aiguille, indiquant un réseau cristallin dans la longueur du cristal. Cependant, le gros problème que nous avons eu est la mosaïcité. Les grains du cristal ne sont pas tous orientés dans la même direction, ce qui est dû à des défauts entre ces grains causant des orientations différentes, ce qui donne une superposition de sets de tâches de diffractions. À partir de cela, il est très difficile d'indexer les points pour en obtenir la structure car en réalité il y en a deux au minimum (Figure 76).<sup>[34]</sup>

#### 4.5.3.3 Cristallogenèse en micro-batch : à la recherche de cristaux de meilleure qualité

Ce type de technique de cristallogenèse permet de cultiver des quantités de cristaux plus importantes qu'avec la technique en goutte assise. Elle permet aussi d'obtenir des cristaux plus gros qu'en goutte assise, qui souvent est limitée par le volume (1 à 2  $\mu\text{L}$ ) alors qu'en batch on peut atteindre, dans notre cas 6  $\mu\text{L}$ .<sup>[26,35,36]</sup> Dans ces plaques, on prépare des réservoirs où les peptides, tampons, sels/précipitants sont directement mélangés, puis recouverts d'huile de paraffine (Figure 77). Le volume reste fixe tout au long de l'expérience, la concentration de peptide ou sel ne change pas jusqu'à la cristallisation (Figure 70). Ce qui mènerait à la formation de monocristaux, plus gros que nos cristaux obtenus par diffusion de vapeur. Cependant, cela implique de connaître un minimum les conditions de cristallisation afin de se mettre dans des conditions de sursaturation de l'objet à cristalliser. Enfin, après avoir trouvé les bonnes conditions de cristallisation de nos

peptides, nous pouvons facilement faire des quantités beaucoup plus importantes par transfert des conditions du réservoir de cristallisation (généralement, 6-7  $\mu\text{L}$  de solution) à des réservoirs comme des tubes allant de plusieurs centaines de  $\mu\text{L}$  à 1 mL. Pour finir, l'étape d'ensemencement est ici beaucoup plus facile et nous avons la possibilité de réaliser de « l'ensemencement en série » qui correspond à ensemencer des réservoirs successivement avec le même crin. Le grand avantage de faire ceci est de diluer le nombre de nuclei apportés dans chaque réservoir, dans le but d'éviter une cristallisation trop rapide et la formation d'oursins.<sup>[37]</sup>

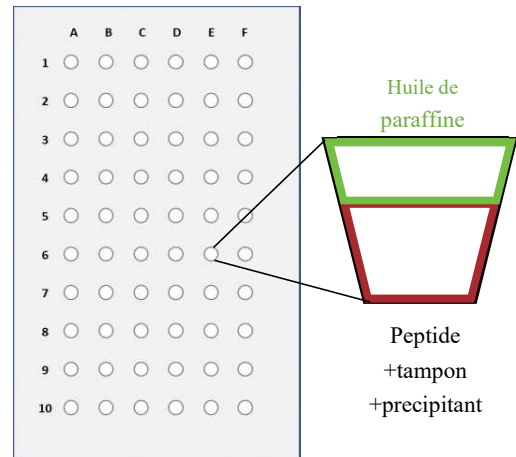


Figure 77: plaque de cristallisation batch 60 puits et compositions utilisées en sursaturation.

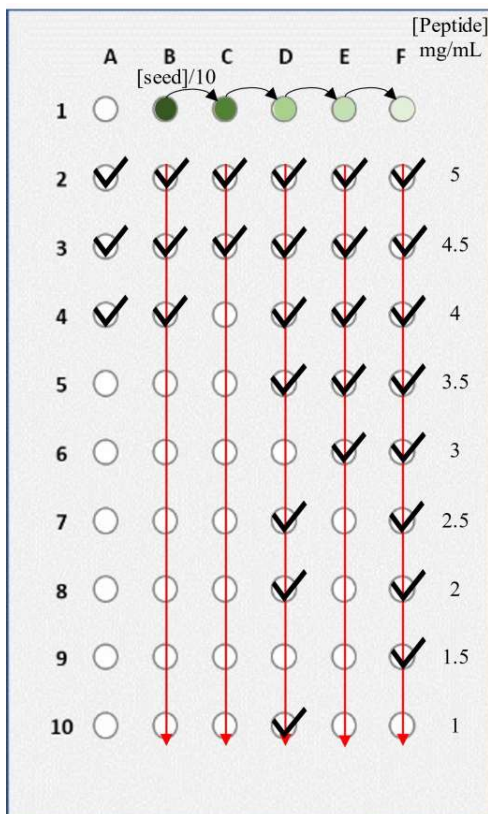


Figure 78: Schéma d'une plaque de cristallisation en batch réalisée avec des ensemencements en série, à partir de cristaux broyés et dilués en cascade. Le sens de « l'ensemencement » est indiqué par une flèche rouge et au bout de 21 jours les conditions cristallisées sont montrées par une marque noire. Les réservoirs contenaient des mélanges de 1 à 5 mg/mL de peptide 5M<sub>L</sub>, 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 50 mM HEPES pH=7.4 à 30°C.

Le protocole que nous avons appliqué est très simple. Nous avons préparé les solutions stocks de peptides, de tampons, et de sels, que nous avons mixés dans le réservoir afin d'obtenir un volume total de 6  $\mu\text{L}$ . Nous avons ensuite ensemencé, ou non, les réservoirs puis les avons recouverts d'huile de paraffine (4  $\mu\text{L}$ ) afin d'éviter l'évaporation de l'eau. Le volume ajouté correspond au volume nécessaire d'huile afin de ne pas avoir d'évaporation de l'eau, le puits pouvant contenir 10  $\mu\text{L}$  au maximum. Nous avons utilisé cette technique sur le peptide 5M<sub>L</sub> dans le but d'identifier les conditions optimales où le peptide cristallise de manière régulière, idéalement sans former d'oursins. Dans la plaque réalisée en Figure 78, les conditions utilisées sont proches de celles identifiées précédemment. Des cristaux ont été obtenus pour la première plaque réalisée mais en oursins, comme précédemment. Nous avons alors diminué les concentrations de peptide et ensemencé les gouttes pour réduire la formation d'oursins, mais nous en avons quand même observé (dans un laps de temps plus long). De plus, le suivi au cours du temps nous a permis de savoir quand la croissance s'arrête, ce qui veut dire que la concentration en peptide restante en solution a diminué jusqu'à atteindre la concentration de saturation. En Figure 79, est montrée l'évolution du contenu du puits F4 (plaque en Figure 78). Nous avons observé que la croissance des cristaux s'est stoppée au bout de 13 jours. Encore une fois, on observe d'abord la formation d'aiguilles longues et, petit à petit, d'autres viennent

se « coller » à elles, ou poussent directement à partir de celles-ci, formant des épis. Par cette technique il est alors difficile de contrôler la formation d'aiguilles seules, sauf, si avant la formation d'oursins, les cristaux sont récupérés pour analyse. Nous avons alors tenté les conditions de cristallisation avec du gel d'agarose à 0.33 % comme précédemment. Pour les conditions où nous réussissions à obtenir des cristaux, ils avaient toujours la même forme, mais leurs apparitions étaient retardées.

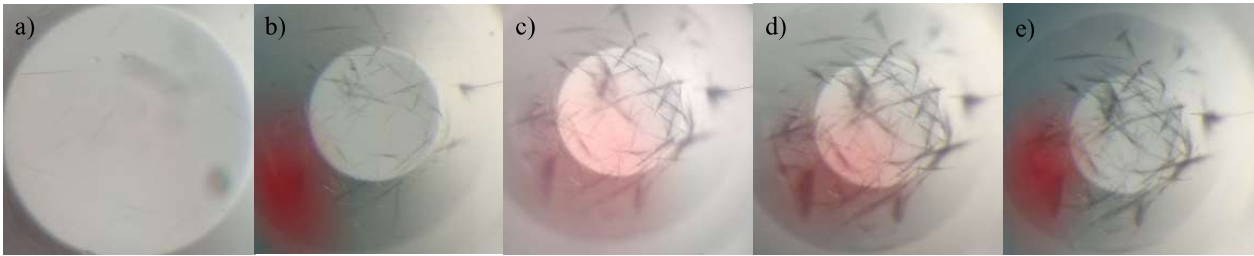


Figure 79 : Avancement au cours du temps, de la formation de cristaux en batch pour le puits F4 reporté dans la Figure 78 à a) 8 ; b) 10 ; c) 11 ; d) 13 ; e) 21 jours.

À partir de ces résultats, nous avons voulu essayer de transférer les conditions où les aiguilles obtenues étaient les plus régulières, volumineuses, par la cristallisation en grande échelle, directement en tube dans le but de les analyser par RMN du solide.

#### 4.5.3.4 Cristallogenèse en tube

##### 4.5.3.4.1 Production de cristaux en grandes quantités

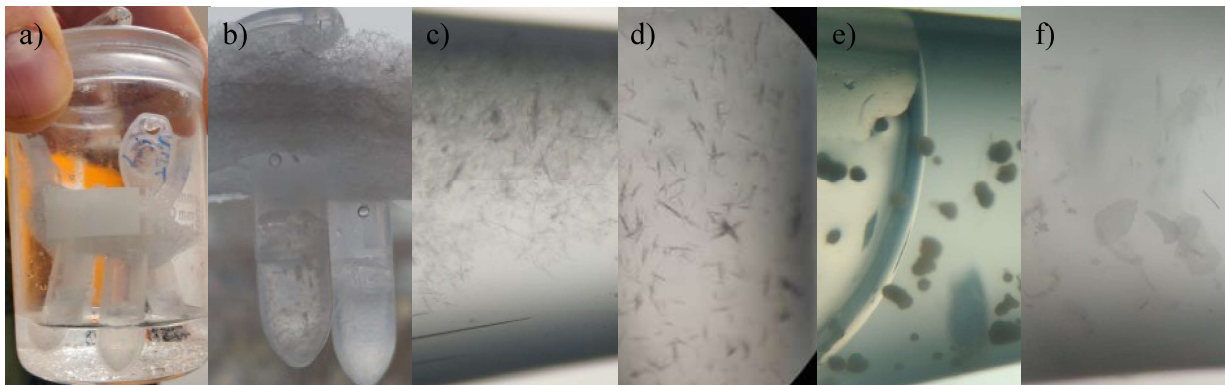


Figure 80 : a) Montage expérimental de diffusion de vapeur pour de grandes quantités de peptides ; b), c) Tubes contenant des cristaux après 2 semaines d'incubation pour le peptide 5ML ; d) 5MN ; 1 mois pour le peptide e) 5MF et f) 5MD. Tous les tubes ont été ensemencés et les conditions sont résumées en Tableau 15.

L'analyse par RMN du solide demandant de grandes quantités de cristaux, la technique en batch n'est plus suffisante, nous nous sommes donc dirigés vers la cristallisation en tube. Cette technique est proche de celle de la cristallisation en goutte assise car la solution contenue dans le tube sera en « équilibre » de vapeur avec une solution de sels. Cependant, il y a plusieurs différences. Les quantités de peptides utilisées sont

beaucoup plus importantes, pour en avoir en quantité (de 1 à 5 mg de peptide pour chaque tube). La solution de sels utilisée dans le réservoir, contrairement aux plaques, est une solution saturée de sulfate d'ammonium, connue pour son hygroscopicité importante.<sup>[19]</sup> Cela nécessite de vérifier tous les jours l'évolution de la cristallisation. Les conditions de départ utilisées pour chaque peptide sont les meilleures précédemment identifiées en batch. Le montage est très simple à mettre en place, mais ne peut servir que lorsque les conditions ont déjà été optimisées. Il consiste à mettre un ou plusieurs tubes de 1.5 à 2 mL contenant la solution de peptide, sel, tampon et ensemencement éventuel qui sera concentrée progressivement grâce à une solution de sulfate d'ammonium saturée (voir Figure 80). Les tubes sont tenus avec un portoir en polystyrène au milieu de la solution saturée de sulfate d'ammonium et mis à incuber dans une cuve classique pour chromatographie sur couche mince (CCM) à 30°C pour observer des cristaux en deux semaines. Ici, tous les tubes ont été ensemencés avec des cristaux broyés de chaque peptide. Via cette technique de culture de cristaux en grande quantité dont les conditions sont résumées en Tableau 15.

Tableau 15 : Conditions optimisées pour la croissance de peptides en grandes quantités.

Peptide	[Peptide] (mg/mL)	[(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ] (mol/L)	Tampon (mol/L)
5M <sub>L</sub> : f-VG <u>P</u> <u>L</u> T-Am	5	1	HEPES pH=7.4 50 mM
5M <sub>N</sub> : f-VG <u>P</u> <u>N</u> T-Am	1.5	1	
5M <sub>F</sub> : f-VG <u>P</u> <u>F</u> T-Am	1	1.2	
5M <sub>D</sub> : f-VG <u>P</u> <u>D</u> T-Am	4.6	1.5	Acide citrique pH=3.5 100 mM

#### 4.5.3.4.2 Préparation de cristaux optimisés pour la DRX

Ici, nous cherchons à avoir des cristaux de bonne qualité pour la diffraction de rayons X. Ceci requiert de bien connaître les conditions de cristallisation via diffusion de vapeur. L'ensemencement par des cristaux de très petite taille en suspension très diluée (Figure 78) est très important. Cela a été fait seulement pour le peptide 5M<sub>L</sub>, c'est celui que nous avons le plus étudié et où l'on connaît le mieux son comportement. À partir des conditions optimales reportées en Tableau 14, pour 100 µL, nous avons choisi une concentration de peptide initiale de 3 mg/mL, 50 mM HEPES pH=7.4 et 1 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> pour la solution dans le tube et deux fois plus concentrée en tampon et sel dans le pilulier (1mL). Le tube a été ensemencé avec des cristaux broyés et dilués précédemment et le pilulier a été scellé avec un bouchon et du parafilm puis incubé à 30°C. En deux semaines, les premières aiguilles sont apparues longues et isolées (ci-dessous).

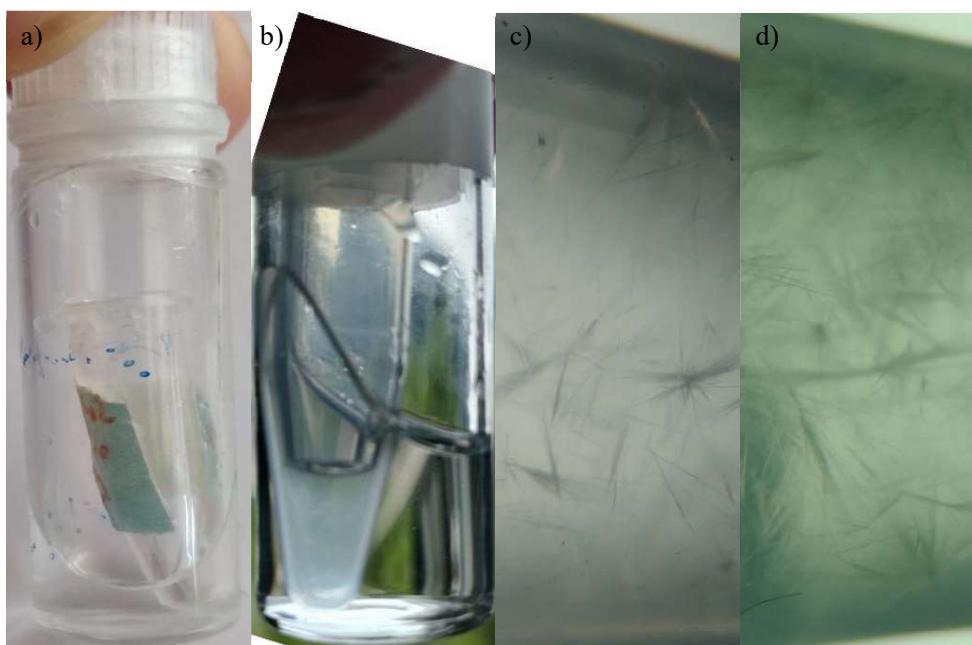


Figure 81 : a) Montage expérimental ; b) après 1 semaine ; cristaux de peptide 5M<sub>L</sub> après c) 2 semaines et après d) 5 semaines.

C'est à ce moment-là que nous avons sorti le tube et nous l'avons fermé pour ralentir la croissance et favoriser la formation de longues aiguilles, sans concentrer plus pour éviter la formation d'oursins (Figure 81).

Pour les conditions résumées dans le Tableau 13, l'influence de l'ensemencement, de la présence d'agarose mais aussi de la séquence a changé le temps de cristallisation (Tableau 16).

Tableau 16 : Tableau récapitulatif des conditions testées pour le peptide 5M<sub>N</sub> en tube,

Peptide	[Peptide] (mg/mL)	[(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ] <sub>tube</sub> (mol/L)	[(NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub> ] <sub>pilulier</sub> (mol/L)	Tampon <sub>tube</sub> (mol/L)	Tampon <sub>pilulier</sub> (mol/L)	Agarose <sub>pilulier</sub> (%)	Temps de cristallisation (semaines)
5M <sub>L</sub> : f-VG <u>P</u> L <u>T</u> -Am	3.0	1	2	HEPES pH=7.4 50 mM	HEPES pH=7.4 100 mM	<del>0.6</del>	2
	3.0	1	2			0.6	2
	1.5	1	2			<del>0.6</del>	6

#### 4.5.3.5 Résultats de diffraction

La collecte des données de diffraction ainsi que le traitement des données ont principalement été faits par Claude Sauter (IBMC, Strasbourg) et ses étudiants, Kevin Rollet et Petr Pachl, sur la ligne Proxima-2A du Synchrotron SOLEIL. Les structures ont été obtenues pour le peptide f-VGPLT-Am (5M<sub>L</sub>), à partir des cristaux avec les conditions de la ligne 1 ci-dessus et le peptide f-VGPNT-Am (5M<sub>N</sub>), à partir des cristaux de la condition de la ligne 2 du Tableau 15, avec des résolutions respectives de 0.95 Å et 0.87 Å (voir ci-dessous). Plus tard, nous avons obtenu des cristaux, plus gros en aiguilles (Tableau 16, ligne 3 et Tableau 15), qui ont été

analysés en République Tchèque. Les structures obtenues ont été les mêmes que celles obtenues juste avant, avec de meilleures résolutions.

Il est à noter que nos cristaux, dans leur plus petite dimension (section), sont de quelques  $\mu\text{m}$  seulement. Ce qui fait que la ligne Proxima-2A (synchrotron Soleil) était particulièrement adaptée à leur analyse puisque la taille du faisceau y est de  $10 \times 5 \mu\text{m}$ .

Les données concernant les paramètres cristallographiques sont données en Annexe 9.

#### 4.5.3.5.1 Structures obtenues

Les structures obtenues pour les cristaux, différentes d'un peptide à l'autre, ne révèlent pas d'assemblage rappelant les hélices  $\beta$ . Ceci permet déjà de dire que le design de 5-mers n'est pas suffisant pour former des assemblages moléculaires mimant le repliement voulu (Figure 82). Les deux structures sont certes différentes, mais forment toutes les deux des feuillettes  $\beta$  parallèles, orientés dans le sens perpendiculaire des aiguilles.

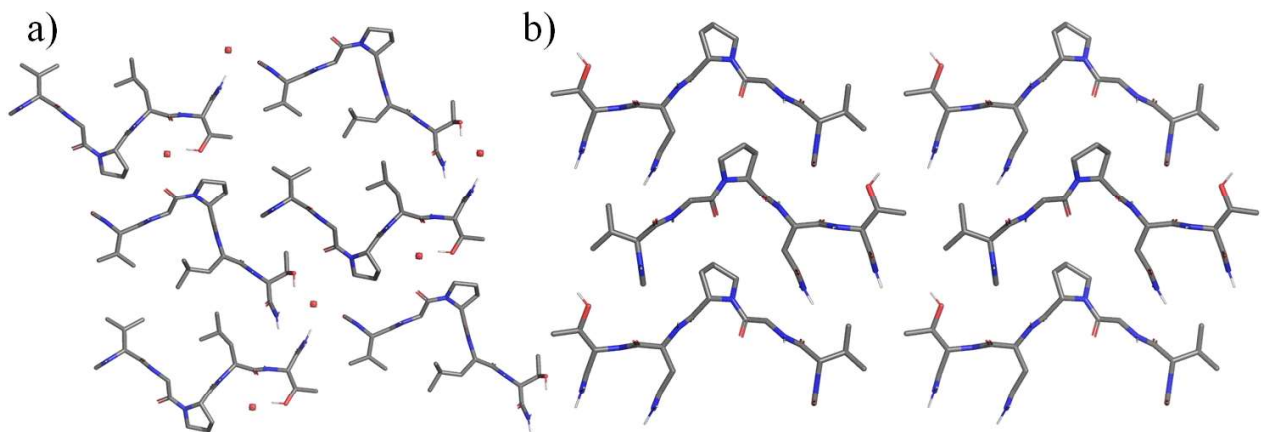


Figure 82 : Arrangements des peptides dans l'espace pour a) f-VGPLT-Am et b) f-VGPNT-Am résolus par DRX (vue de dessus).

#### 4.5.3.5.2 f-VGPLT-Am

Le cristal formé par ce peptide contient une molécule d'eau par peptide, ce qui est très peu comparé à la quantité d'eau dans les cristaux de protéines en général, et qui peut comporter jusqu'à 70 % d'eau (Figure 82 a).<sup>[38]</sup> Ici, l'eau représente environ 3.6 % du cristal, et est expliquée par un assemblage très compact du peptide. Les groupements formyle (en N-ter) et amide (en C-ter) sont retrouvés, comme attendu, impliqués dans des liaisons hydrogène chaîne principale-chaîne principale (Figure 83). La conformation qu'adoptent les peptides afin de diminuer l'énergie du système est de former un cœur hydrophobe entre 2 unités identiques orienté en sens opposé, ce qui ne correspond pas à ce qui était initialement souhaité. Nous pensons ici que

l'isoleucine dans la séquence est un résidu avec une hydrophobicité trop élevée pour être compatible, dans notre design, avec la formation d'assemblage de trois peptides s'autoassemblant. L'angle formé par la proline ici est d'environ  $88^\circ$  et adopte une conformation *cis* alors que les structures naturelles d'hélices  $\beta$  gauche de type I adoptent une conformation *trans*.<sup>[5]</sup> Nous remarquons également que les superpositions de feuillets  $\beta$  ne sont pas équivalentes suivant l'orientation des chaînes peptidiques. Il y a 2 conformations présentes dans le cristal, la I lorsque les chaînes latérales des valines et leucines pointent vers le haut dans la Figure 82, et la II lorsqu'elles sont orientées vers le bas.

#### 4.5.3.5.3 f-VGPNT-Am

Contrairement au peptide précédent, celui-ci ne contient aucune molécule d'eau dans sa structure cristalline (Figure 82 b). Ici aussi on retrouve le formyle et l'amide C-terminal dans les interactions chaîne principale-chaîne principale (Figure 84). Nous avons aussi remarqué que la chaîne latérale de l'asparagine fait une liaison hydrogène avec un amide d'une chaîne principale en face, ce qui renforce la cohésion de l'assemblage des peptides. L'autre différence, par rapport à l'autre peptide, est l'angle formé par la proline d'environ  $105^\circ$  et est de configuration *trans*, ce qui est compatible avec le coude dans les hélices  $\beta$  gauche.

#### 4.5.4 RMN du solide

Cette technique est très avantageuse car elle nous permet de corrélérer directement, par observation des signaux du carbone, les interactions entre les chaînes principales (i.e. amides) mais aussi entre les chaînes latérales du cœur hydrophobe par exemple par corrélation entre les protons et carbones.<sup>[39,40]</sup> Ayant des microcristaux, l'échantillon est homogène (i.e. ils sont orientés dans toutes les directions) et les peptides ont tous la même conformation dans l'espace (puisque ce sont des cristaux). Afin de s'affranchir des contributions anisotropiques de l'échantillon, l'analyse est réalisée en rotation à l'angle magique (MAS). Cette expérience

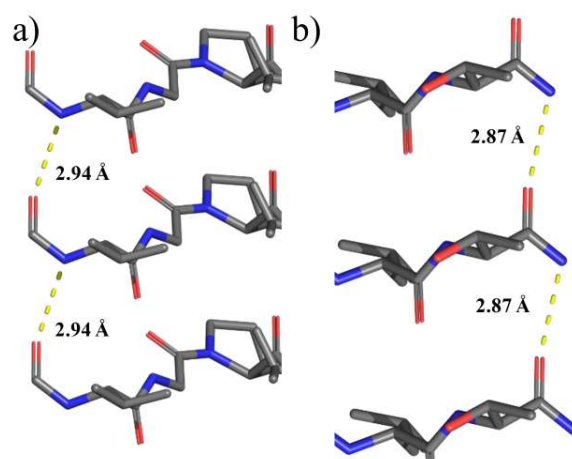


Figure 83: Réseau de liaisons hydrogène formé par a) le formyle et b) l'amide.

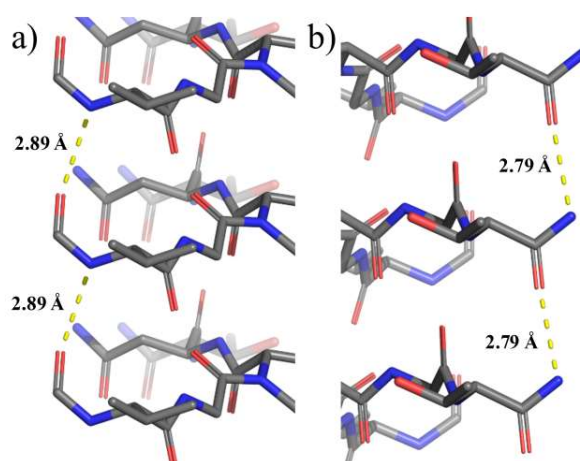


Figure 84 : Réseau de liaisons hydrogène formé par a) le formyle et b) l'amide.



demande une quantité assez importante de peptide cristallisé (cf. 4.5.3.4.1) : plusieurs dizaines mg de cristaux de peptides. Il est important d'en avoir une quantité suffisante afin de remplir le rotor (pour avoir un signal intense). Et afin de ne pas en avoir en solution, qui seraient dans des conformations libres et non verrouillées comme dans les cristaux, ce qui polluerait le signal, les cristaux ont été rincés (aussi pour enlever les sels) puis séchés par capillarité avec un papier absorbant. Ces analyses ont été effectuées en collaboration avec P. Bertani du Laboratoire de biophysique des membranes et RMN.

La résolution des signaux en analyse de polarisation croisée couplée à la rotation à l'angle magique (CP-MAS) s'est avérée très bonne avec des pics bien définis (Figure 85 a). Les résultats montrent un arrangement des peptides en feuillettes  $\beta$  avec les déplacements chimiques des carbones  $^{13}\text{C}$  appartenant aux groupes amides. Une analyse en *FSLG-HETCOR*  $^1\text{H}$ - $^{13}\text{C}$  (Frequency switch Lee-Goldburg HETeronuclear chemical shift CORrelation) a été réalisée afin de voir les corrélations entre les résidus (Figure 85 b). Ce travail préliminaire montre qu'il est possible d'observer nos peptides sous forme cristallisée et que leurs structures sont homogènes. Malgré une bonne résolution en CP-MAS, le marquage des peptides  $^{13}\text{C}$  et  $^{15}\text{N}$  sera nécessaire afin de pouvoir corréler les résidus entre eux. C'est particulièrement le cas pour les chaînes latérales, qui nous permettraient de déterminer les distances entre les différents résidus du cœur hydrophobe par exemple ou encore en fonction de leurs expositions au solvant.<sup>[41,42]</sup>

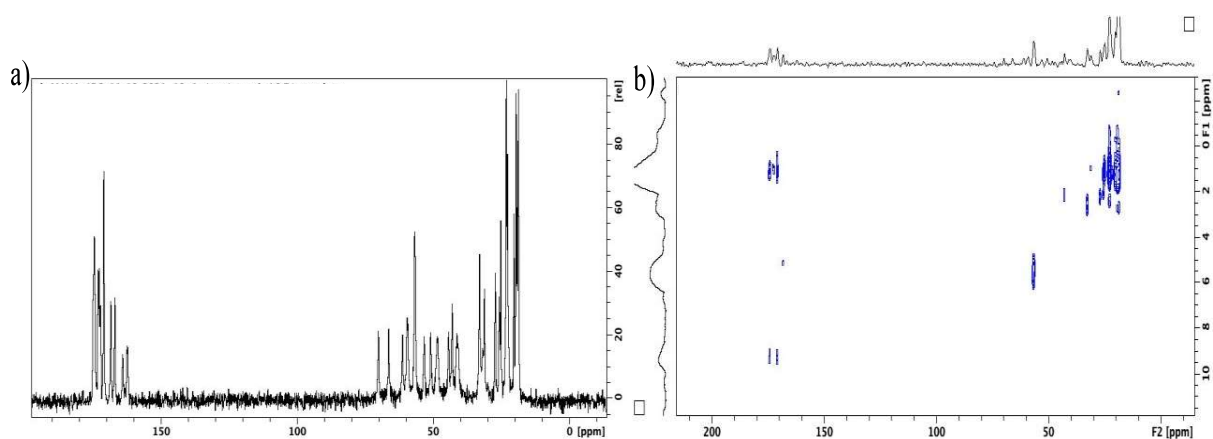


Figure 85 : a) Spectre CP-MAS, b) Spectre *FSLG-HETCOR*  $^1\text{H}$ - $^{13}\text{C}$  (Abondance naturelle) du peptide 5M<sub>L</sub>.

## 4.6 Conclusion

L'objectif de cette partie a été de tester s'il était possible de former des mimes de L $\beta$ Hs par autoassemblage d'unités peptidiques de taille très réduite (Figure 65). Le design de séquences de pentamères n'a pas permis d'encoder la formation de mimes d'hélice  $\beta$  gauche. Le criblage et l'optimisation des conditions, via différentes méthodes, a permis d'obtenir deux structures cristallographiques, pour les peptides 5M<sub>L</sub> et 5M<sub>N</sub> de séquence générique de type f-VGPXT-Am (X=L, N) nous permettant de confirmer l'hypothèse faite lors de la conception concernant le rôle du formyle. Son rôle comme contribution supplémentaire permettant la

continuité d'un réseau de liaisons hydrogène chaîne principale-chaîne principale a été confirmée. Néanmoins, au vu de l'assemblage obtenu pour les deux peptides, nous n'avons pas pu attester de la pertinence de choisir une séquence de 5 A.As qui correspondrait à 1/3 de pas d'hélice, puisque ces peptides n'ont pas formé d'assemblage rappelant ces L $\beta$ Hs. Nous voyons également que notre design permet de former, sous forme cristalline, des feuillets  $\beta$  parallèles qui sont moins stables que les antiparallèles.<sup>[43]</sup> Le coude  $\beta$  est bien présent avec la proline en milieu de séquence même si la conformation *cis* n'est pas celle attendue. Les deux résidus en B<sub>3</sub> (valine) et B<sub>1</sub> (thréonine), censés être orientés du même côté afin de former le cœur hydrophobe avec 2 autres peptides, se retrouvent en fait orientés de manière opposée. Un cœur hydrophobe est formé pour le peptide 5M<sub>L</sub> mais entre 2 unités, et non 3, de plus, la leucine, en position T<sub>2</sub> est censée se retrouver vers l'extérieur et non vers l'intérieur de l'hélice. La thréonine ne forme aucune interaction avec d'autres peptides, elle forme, à la place, une liaison hydrogène avec une molécule d'eau.

Mais en ce qui concerne les autres peptides, nous n'avons pas pu avoir de bons cristaux qui puissent être exploités convenablement. Nous pouvons mentionner que des clichés de poudre ont été réalisés pour les peptides 5M<sub>L</sub>, 5M<sub>N</sub>, 5M<sub>D</sub> (Figure 86). Ces clichés pourront nous permettre par la suite de vérifier la structure du peptide 5M<sub>D</sub> en le corrélant avec les autres obtenus via une collaboration avec des spécialistes du domaine.

Les résultats en RMN du solide décrivent également une structure en feuillets  $\beta$ , ce qui confirme la structure obtenue en DRX. Par la suite, cette technique peut permettre, si les cristaux obtenus ne sont pas de qualité suffisante pour la diffraction des rayons X, de déterminer la structure des peptides obtenus que ce soit en solution ou sous forme cristalline. Il serait judicieux de réaliser une expérience de cristallisation *in situ* afin de voir la variation des signaux.

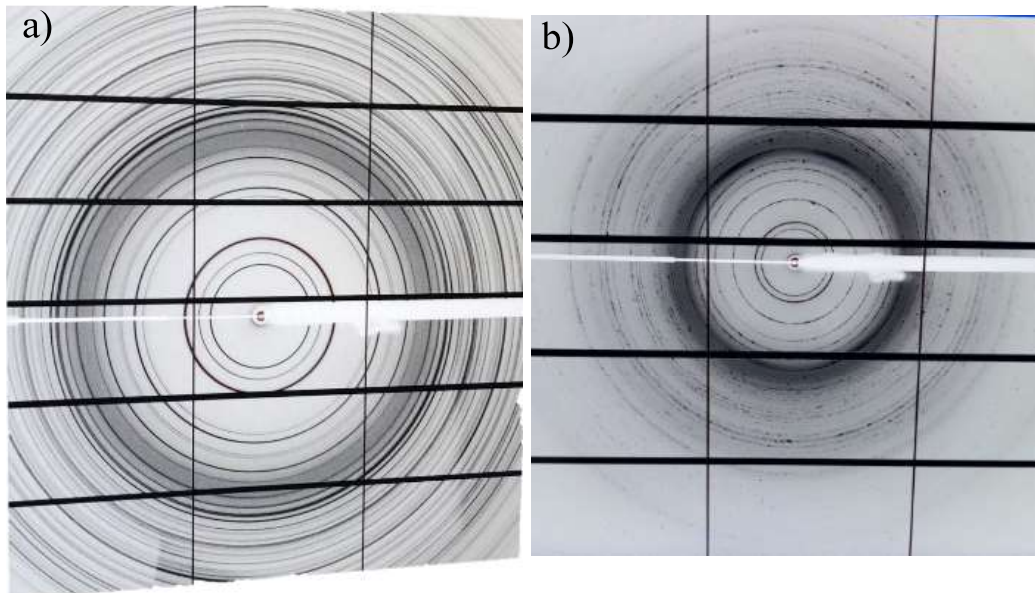


Figure 86 : Clichés de poudre du peptide 5M<sub>D</sub> (a) et 5M<sub>L</sub> (b).

## 4.7 Références bibliographiques

- [1] G. Marinescu, S. Ferlay, N. Kyritsakas, M. W. Hosseini, *Chem. Commun.* **2013**, 49, 11209–11211.
- [2] A. Mohan, V. N. Uversky, P. Radivojac, *PLoS Comput. Biol.* **2009**, 5, DOI 10.1371/journal.pcbi.1000497.
- [3] R. Giegé, C. Sauter, *HFSP J.* **2010**, 4, 109–121.
- [4] R. Giegé, *IUCrJ* **2017**, 4, 340–349.
- [5] L. Pauling, R. B. Corey, *Proc. Natl. Acad. Sci.* **1951**, 37, 729–740.
- [6] L. Pauling, R. B. Corey, H. R. Branson, *Proc. Natl. Acad. Sci.* **1951**, 37, 205.
- [7] P. L. Tsiolaki, N. N. Louros, V. A. Iconomidou, *J. Mol. Biol.* **2018**, 430, 3774–3783.
- [8] V. A. Iconomidou, G. D. Chryssikos, V. Gionis, A. S. Galanis, P. Cordopatis, A. Hoenger, S. J. Hamodrakas, *J. Struct. Biol.* **2006**, 156, 480–488.
- [9] N. N. Louros, F. A. Baltoumas, S. J. Hamodrakas, V. A. Iconomidou, *J. Comput. Aided. Mol. Des.* **2016**, 30, 153–164.
- [10] J. H. Choi, C. Govaerts, B. C. H. H. May, F. E. Cohen, *Proteins Struct. Funct. Bioinforma.* **2008**, 73, 150–160.
- [11] M. F. Perutz, J. T. Finch, J. Berriman, A. Lesk, *Proc. Natl. Acad. Sci.* **2002**, 99, 5591–5595.
- [12] V. Mikol, J. L. Rodeau, R. Giegé, *Anal. Biochem.* **1990**, 186, 332–339.
- [13] G. Ruyters, C. Betzel, **2017**, 11–26.
- [14] G. T. DeTitta, J. R. Luft, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1995**, 51, 786–791.
- [15] M. A. Dessau, Y. Modis, *J. Vis. Exp.* **2010**, 9, 1–6.
- [16] C. Sauter, B. Lorber, A. McPherson, R. Giegé, *Int. Tables Crystallogr.* **2012**, F, 99–121.
- [17] R. de Wijn, K. Rollet, S. Engilberge, A. G. McEwen, O. Hennig, H. Betat, M. Mörl, F. Riobé, O. Maury, E. Girard, et al., *bioRxiv* **2020**, 1–13.
- [18] R. K. Spencer, J. S. Nowick, *Isr. J. Chem.* **2015**, 55, 698–710.
- [19] P. Wingfield, *Curr. Protoc. Protein Sci.* **1998**, A.3F.1–A.3F.8.
- [20] I. V. Kurinov, R. W. Harrison, *Acta Crystallogr. - Sect. D Biol. Crystallogr.* **1995**, 51, 98–109.
- [21] A. Watts, *Crystallization of Nucleic Acids and Proteins. A Practical Approach*, **1993**.
- [22] A. D'Arcy, T. Bergfors, S. W. Cowan-Jacob, M. Marsh, *Acta Crystallogr. Sect. Struct. Biol. Commun.* **2014**, 70, 1117–1126.
- [23] T. Bergfors, *J. Struct. Biol.* **2003**, 142, 66–76.
- [24] E. H. Snell, J. R. Helliwell, *Reports Prog. Phys.* **2005**, 68, 799–853.
- [25] F. Artusio, A. Castellví, A. Sacristán, R. Pisano, J. A. Gavira, *Cryst. Growth Des.* **2020**, 20, 5564–5571.
- [26] I. Rayment, *Structure* **2002**, 10, 147–151.
- [27] R. Giegé, *FEBS J.* **2013**, 280, 6456–6497.
- [28] E. F. Garman, R. L. Owen, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2006**, 62, 32–47.
- [29] S. Kriminski, M. Kazmierczak, R. E. Thorne, *Acta Crystallogr. Sect. D* **2003**, 59, 697–708.
- [30] L. Jacquamet, J. Ohana, J. Joly, F. Borel, M. Pirocchi, P. Charrault, A. Bertoni, P. Israel-Gouy, P. Carpentier, F. Kozielski, et al., *Structure* **2004**, 12, 1219–1225.

- [31] D. Jeangerard, S. Sirigu, A. Haouz, P. Weber, L. Ciccone, E. D. E. Zitter, S. Jaho, M. Spano, K. Rollet, M. Naude, et al., *The CRIBLEUR: An Any Format Screener for in Situ X-Ray Diffraction Experiments on PROXIMA 2A A*, **2021**.
- [32] A. E. Goeta, J. A. K. Howard, *Chem. Soc. Rev.* **2004**, *33*, 490–500.
- [33] Z. Dauter, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1999**, *55*, 1703–1717.
- [34] H. D. Bellamy, E. H. Snell, J. Lovelace, M. Pokross, G. E. O. Borgstahl, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2000**, *56*, 986–995.
- [35] J. H. Beale, R. Bolton, S. A. Marshall, E. V. Beale, S. B. Carr, A. Ebrahim, T. Moreno-Chicano, M. A. Hough, J. A. R. Worrall, I. Tews, et al., *J. Appl. Crystallogr.* **2019**, *52*, 1385–1396.
- [36] A. McPherson, J. A. Gavira, *Acta Crystallogr. Sect. F Structural Biol. Commun.* **2014**, *70*, 2–20.
- [37] E. A. Stura, I. A. Wilson, *J. Cryst. Growth* **1991**, *110*, 270–282.
- [38] M. Frey, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1994**, *50*, 663–666.
- [39] H. Van Melckebeke, C. Wasmer, A. Lange, E. Ab, A. Loquet, A. Böckmann, B. H. Meier, *J. Am. Chem. Soc.* **2010**, *132*, 13765–13775.
- [40] C. Wasmer, A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek, B. H. Meier, *Science (80-. )*. **2008**, *319*, 1523–1526.
- [41] E. D. BECKER, *High Resolut. NMR* **1969**, *54*, 59–85.
- [42] F. Avbelj, D. Kocjan, R. L. Baldwin, *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 17394–17397.
- [43] A. Perczel, Z. Gáspári, I. G. Csizmadia, *J. Comput. Chem.* **2005**, *26*, 1155–1168.

## **5 Conclusion**

Le projet de thèse, dans son ensemble, a pour but de concevoir des peptides pouvant se replier et s'autoassembler en hélice  $\beta$  gauche de type I (L $\beta$ Hs). Ils pourront plus tard servir de « briques élémentaires » pour la conception de métalloenzymes artificielles multisites. Pour se faire, ce projet s'est scindé en trois parties. Le chapitre 1 traitant de l'analyse des protéines adoptant un repliement en hélice  $\beta$  gauche de type I, un type particulier de solénoïde  $\beta$ , afin de mieux comprendre la relation séquence-structure de ce repliement. Le second chapitre est axé sur la conception de peptides, à partir des résultats du chapitre 1, comme briques de base pouvant former des fibres de L $\beta$ Hs par autoassemblage afin de former des édifices stables. Enfin, le chapitre final de ma thèse s'articule autour de la recherche de la séquence peptidique minimale permettant de former une structure rappelant l'hélice  $\beta$  gauche.

Afin de concevoir des peptides susceptibles de se replier en L $\beta$ Hs, nous avons étudié, dans un premier chapitre, les structures cristallographiques des protéines adoptant ce repliement. Celui-ci a déjà fait l'objet d'études, mais avec des limites que nous pensons importantes pour sa compréhension : notamment le fait de considérer chaque hélice comme un objet isolé. En effet, la particularité de ce repliement est son autoassemblage sous forme de trimère, qui est retrouvé dans quasi toutes les structures naturelles. Nous nous sommes alors proposé d'étudier les L $\beta$ Hs non plus comme objets seuls mais comme objets trimériques, ce qui a pour conséquences de considérer les trois faces (hexades) de chaque hélice comme différentes. Nous nous sommes aussi rendu compte, lors de l'étude, qu'il y avait différentes parties au repliement. Nous avons identifié trois différentes parties qui nous ont semblé importantes, le premier tour après le bouchon du N<sub>ter</sub>, les tours canoniques ainsi que le dernier tour avant le bouchon C<sub>ter</sub>.

Les trois différentes hexades du repliement ont permis de mettre en évidence le fait que, en plus de l'importance du réseau de liaisons hydrogène formé par les feuilletts  $\beta$  et le cœur hydrophobe, la distribution est différente selon la face sur laquelle les résidus sont. En effet, une est exposée au solvant (II) alors que les deux autres faces se font mutuellement face (I) et (III). Ce qui entraîne une répartition de résidus plutôt polaires sur la première, comparée aux deux autres.

Les trois parties du repliement ont permis de mettre en lumière le caractère nucléant du premier tour après le bouchon N<sub>ter</sub>. Dans ce premier tour, nous avons identifié des motifs renforçant le repliement par interactions entre les différentes chaînes latérales et principales, comme la séquence aspartate/histidine-proline-sérine/thréonine dans les coudes  $\beta$ . C'est aussi dans ce premier tour que la proline, identifiée précédemment comme importante au repliement, est présente en majorité comparé aux autres parties. Les tours les plus réguliers ont des motifs structuraux permettant la consolidation du repliement, par l'intermédiaire d'échelles d'asparagines par exemple. C'est aussi dans ces tours que l'interface entre les hélices est la plus importante. Elle est caractérisée par des interactions au centre du trimère d'hélices, comme des empilements  $\pi$  ou des échelles d'asparagines, ayant pour rôle de garder l'assemblage trimérique. La dernière partie est le tour avant le bouchon C<sub>ter</sub>. Ce tour a pour rôle de stopper le repliement en L $\beta$ H, tout en ne déstabilisant pas le reste de la structure. Pour cela, il repose sur des suppressions de résidus afin de stopper la continuité du réseau de liaisons

H mais aussi sur la présence de proline à une position qui avait été reportée comme non-compatible avec le repliement en L $\beta$ H, la position B<sub>4</sub>. Celle-ci, étant perpendiculaire à la chaîne principale, prévient l'empilement d'autres feuillets. Nous avons aussi remarqué que la grande majorité des structures étudiées se terminaient par un bouchon formant une épingle  $\beta$  commençant toujours par la même hexade. Ce motif a été étudié pour comprendre son rôle de masque du cœur hydrophobe, gardant le repliement solidaire.

À partir des analyses effectuées, des motifs structuraux de nucléation et de terminaison ont été identifiés en fonction des différentes faces des hélices, suivant leur position dans le trimère, et des parties. Tout cela a servi à construire un tableau de matrice 3 x 3 (hexades) dont le but est de servir comme outil pour la conception de protéines et protéines se repliant en L $\beta$ H.

Ce tableau a d'ailleurs servi à concevoir un peptide long de 59 acides aminés. Ce peptide correspond à trois pas d'hélice, un tour N<sub>ter</sub>, un tour canonique et un tour C<sub>ter</sub> auxquels il faut ajouter des motifs de bouchon, notamment celui en épingle  $\beta$ . La structure de ce peptide a été modélisée et le repliement est bien celui d'une L $\beta$ H. Mais nous n'avons pas encore obtenu la structure expérimentale. De plus, à cause de sa longueur, ce peptide n'a pas pu être obtenu avec une pureté satisfaisante. Des essais de cristallisation ont été réalisés, mais aucun n'a mené à des cristaux pour le moment. Nous avons cependant pu apprendre, avec les analyses en dichroïsme circulaire au synchrotron, que la structure secondaire passe de pelotes statistiques à feuillets  $\beta$  lorsque l'on augmente la température, et cette structure reste stable dans le temps.

Le second chapitre de ma thèse est axé sur la conception rationnelle de peptides se repliant et s'assemblant en fibres d'hélices  $\beta$  gauche. En se basant sur l'outil développé lors du premier chapitre, deux séquences ont été conçues. Basée sur les travaux précédents, la taille des séquences conçues correspond à deux pas d'hélice afin d'avoir une certaine préorganisation du peptide, facilitant son repliement et son assemblage. Ensuite, la séquence a été optimisée pour la formation d'hélice  $\beta$  gauche, et surtout des assemblages trimériques, par différenciation des trois faces de l'hélice. La formylation et l'amidation des extrémités du peptide ont été faites dans le but de garder un réseau de liaison hydrogène régulier malgré la suppression d'un acide aminé. Enfin, nous avons pensé la séquence du peptide de manière à contrôler son autoassemblage par le pH. C'est pour cette raison que deux histidines ont été placées en position T<sub>2</sub> de la face (III) de chaque hélice, pointant vers l'intérieur du trimère. À partir de cela, deux peptides ont été conçus, 35m1 et 35m2. Ce dernier est un mutant du premier, avec trois mutations, possédant une charge globale plus élevée.

Ceux-ci ont été produits classiquement (SPPS) et caractérisés par différentes techniques. Cette approche multitechniques nous a permis d'avoir une corrélation entre les différentes analyses et de déterminer s'il y a ou non des incompatibilités entre elles, qui seraient signe que le repliement attendu pourrait ne pas être le bon.

Le premier moyen de caractérisation de l'autoassemblage formé, accessible au laboratoire, est la fluorescence de la thioflavine-T. Ce marqueur de fibres amyloïdes nous a permis d'étudier, pour les deux peptides, les mécanismes d'assemblages mais aussi de déterminer que les autoassemblages sont de type amyloïde. Cela a

été réalisé à différentes concentrations et pH, et nous avons pu remarquer que les cinétiques d'assemblages des deux peptides 35m1 et 35m2 sont bien différentes. Le peptide 35m1 ne montre pas de phase de latence, alors que 35m2 si. Cette phase de latence correspond généralement au temps nécessaire à la formation des nucléi. Il semble que pour 35m1, ce phénomène soit très rapide, peut-être grâce à la forte préorganisation du peptide. Nous avons confirmé qu'à pH acide, il n'y a pas d'assemblage, et que dès que le pH est monté, au-dessus de 7.5, l'autoassemblage se fait. Ces résultats sont compatibles avec la conception initiale, notamment avec les histidines, qui déprotonées, permettent l'autoassemblage. Nous avons pu aussi émettre l'hypothèse que l'autoassemblage de ces deux peptides se fait par un mécanisme dominé par une nucléation secondaire multiétapes, ce qui serait compatible avec la formation de trimères d'hélices.

Nous avons observé la morphologie des deux peptides par microscopie électronique à transmission afin de savoir si nous avions bien des fibres et si leurs tailles et diamètres étaient compatibles avec la structure trimérique attendue des fibres. Les deux peptides ont des morphologies différentes avec des fibres plutôt courtes et agrégées pour le 35m1 et plutôt longues et isolées (même s'il y a quelques agrégats) pour le 35m2. C'est pour cela que nous nous sommes concentrés sur le peptide 35m2 optimisé pour lequel nous avons optimisé les conditions d'incubation afin de pouvoir réaliser des grilles en Cryo-TEM pour avoir des informations structurales (et par la suite la structure). Nous avons pu déterminer que les fibres seules formées sont compatibles avec la formation de trimères d'hélices par leur diamètre. Les conditions d'incubations restent à optimiser pour des analyses plus poussées en Cryo-TEM, mais nous avons vu que les trois mutations jouent un rôle crucial dans l'assemblage et la conception est compatible avec le fait que les charges apportées repoussent les autres hélices, les empêchant de se « coller ».

Afin d'investiguer la structure secondaire des peptides dans les fibres, nous avons réalisé des analyses de dichroïsme circulaire (SR-CD). Nous avons observé, pour les deux peptides, des spectres caractéristiques de structures secondaires en feuillets  $\beta$  (parallèles) ce qui est compatible avec le repliement et l'autoassemblage attendu. Nous avons remarqué des différences dans les spectres pouvant être dues à des angles de torsions des feuillets  $\beta$  différents. Nous n'avons pas observé de signal correspondant à la présence de coudes  $\beta$  par cette technique. La contribution au spectre CD des fibres a pu être vérifiée après fibrillation par centrifugation des fibres, montrant que les fibres contribuent, en très large majorité au signal alors que la quantité de peptide en solution est négligeable. Des incubations à des températures identiques à celle de la TEM ont été réalisées directement dans la cuve démontrant le repliement/assemblage partiel des peptides, probablement dû au manque d'agitation, qui peut être compensé par une élévation de la température. Nous avons pu confirmer la robustesse de l'assemblage formé par incubation des fibres à des températures élevées, sans changement du spectre CD. Comme en TEM, des incubations dans le SDS et en présence de sels montrent la formation de feuillets  $\beta$ . Le dichroïsme linéaire (SR-LD) a révélé un alignement des fibres possible dans l'échantillon.

La dernière technique que nous avons utilisée pour caractériser nos assemblages est la spectroscopie InfraRouge (ATR-IR). Toujours dans l'optique de confirmer les résultats des autres techniques, nous avons pu confirmer la présence de feuillets  $\beta$  parallèles en addition aux coudes  $\beta$ , que nous n'avons pas observés en CD.



Les deux peptides 35m1 et 35m2 montrent des signatures identiques, signe que les structures secondaires sont les mêmes, pouvant indiquer un repliement des peptides égaux.

Les caractérisations réalisées sont toutes compatibles avec le repliement en L $\beta$ H mais sans aucune certitude, c'est pour cela que des analyses supplémentaires doivent être réalisées afin d'être sûr de la structure.

Le dernier chapitre de ma thèse correspond à la recherche de la brique minimale pouvant former des mimes d'hélice  $\beta$  gauche. Pour cela plusieurs séquences de cinq acides aminés ont été conçues (5M<sub>x</sub>) et produites. Nous sommes partis d'un pas d'hélice que nous avons coupé en trois parties puisque l'hélice est de section triangulaire. Les peptides conçus correspondent à une partie de feuillet  $\beta$  ainsi que le coude  $\beta$  que nous pensons important de former avec un peu de préorganisation comme c'est le cas avec la proline qui induit un angle de torsion. Nous avons cristallisé ces peptides en vue d'analyse en DRX pour obtenir la structure et observer comment ces peptides s'autoassemblent. Plusieurs techniques de cristallisation ont été utilisées comme la goutte assise, la microbatch avec et sans ensemencement/agarose. Nous avons aussi développé une méthode de cristallisation pour avoir des cristaux en quantités importantes pour des analyses en RMN du solide et des cristaux les plus gros possibles pour la DRX. À partir de cristaux en oursins, après plusieurs étapes d'optimisation, nous avons pu obtenir des cristaux assez gros et homogènes pour en obtenir la structure. Les structures ont été résolues pour deux peptides, montrant l'autoassemblage en feuillets  $\beta$  parallèles, ne formant donc pas de L $\beta$ Hs comme attendu. Les briques n'étant sûrement pas suffisantes pour encoder l'information nécessaire à la formation de ce type de structure. Nous avons cependant pu confirmer la contribution du formyle en N<sub>ter</sub> et de l'amide en C<sub>ter</sub> comme participant au réseau de liaisons hydrogène impliquant les chaînes principales, confirmant par la même occasion l'intérêt de ces modifications pour les peptides 35m1 et 35m2. En plus de cela, des analyses ATR-IR, CD et RMN du solide préliminaires sur ces peptides ont été réalisés pour avoir une confirmation des structures formées.

Globalement, ce travail de thèse a posé la première brique d'un projet à long terme : la conception de protéines artificielles avec un repliement en hélice  $\beta$  gauche de type I.



# 6 Perspectives



## 6.1 Perspectives à court terme : suite de la caractérisation

Comme nous l'avons vu tout au long de la thèse, caractériser une structure protéique peut s'avérer difficile et fastidieux, surtout lorsque celle-ci ne possède pas de signature spectroscopique spécifique. Dans le but d'améliorer la compréhension de ce type d'autoassemblage adoptant ce repliement, il est nécessaire de réaliser le plus d'analyses possible afin de les caractériser. Pour cela, des analyses pourront être réalisées notamment en microscopie à force atomique (AFM) pour confirmer et avoir une meilleure résolution sur la taille des fibres formées. Des techniques comme la titration calorimétrique isotherme (ITC) et la calorimétrie différentielle à balayage (DSC) pourront nous renseigner sur le mécanisme d'autoassemblage, sur le repliement de nos objets ainsi que leur stabilité. Les études de stabilité des fibres plus poussées ont besoin d'être réalisées pour déterminer les limites d'utilisations des fibres formées, c'est aussi le cas pour d'éventuels essais mécaniques de torsion, élongation, etc... Des analyses en DRX, diffusion des rayons X aux grands angles (WAXS) et diffusion des neutrons (DRN) nous renseigneraient sur les interactions, les structures secondaires et les structures en général des fibres. Malgré toutes les analyses réalisées, la structure est nécessaire afin d'être sûr du repliement et de l'autoassemblage des peptides conçus. Les conditions de fibrillation restant à être optimisée pour la Cryo-TEM, la résonance magnétique nucléaire du solide (ssNMR) peut s'affranchir de ce problème et nous permettre de déterminer la structure des fibres. Le seul bémol étant la nécessité de marquer certains A.As au carbone 13 et à l'azote 15. La Figure 87 résume les techniques ayant été ou qui seront probablement utilisées pour construire la « carte d'identité » de nos fibres. Tout cela ayant pour but d'avoir une structure connue, en L $\beta$ H, stable et robuste dans différentes conditions pour ensuite concevoir différents matériaux.

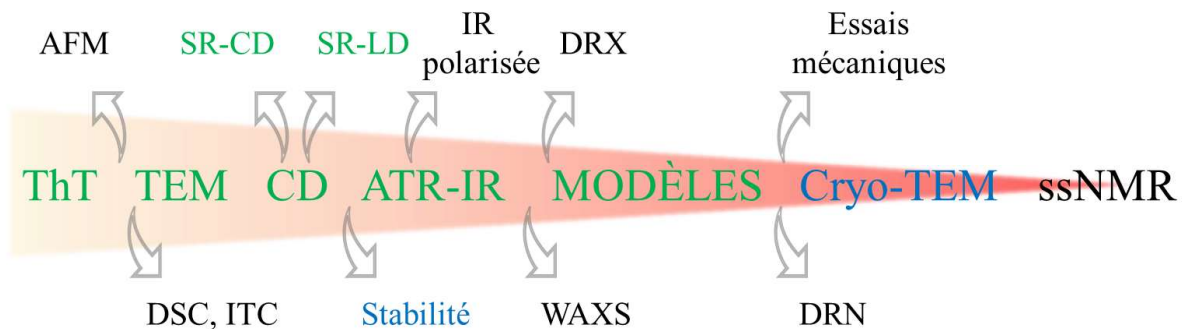


Figure 87 : Résumé des différentes analyses effectuées en vert, des celles partiellement faites ou en cours en bleu et celle que nous pouvons réaliser en noir pour déterminer les propriétés des fibres.

## 6.2 Vers la conception rationnelle d'ArMs

Ne perdant pas de vue l'objectif de concevoir des métalloenzymes artificielles à partir de briques peptidiques de base, nous avons muté le peptide 35m2 afin de lui donner des propriétés catalytiques. Ce travail a principalement été réalisé par deux stagiaires L. Escudero (L3, ENS Saclay) et S. Behr (M2, université de Strasbourg). Toujours à partir de l'outil du chapitre 1, nous avons conçu deux peptides mutants du 35m2 appelés 34mHIS et 34mMET. Ils s'inspirent directement de l'anhydrase carbonique  $\gamma$ , qui est une enzyme de type hélice  $\beta$  gauche. Son site catalytique à  $Zn^{II}$  est situé à l'interface entre deux hélices (Figure 88 a, b).

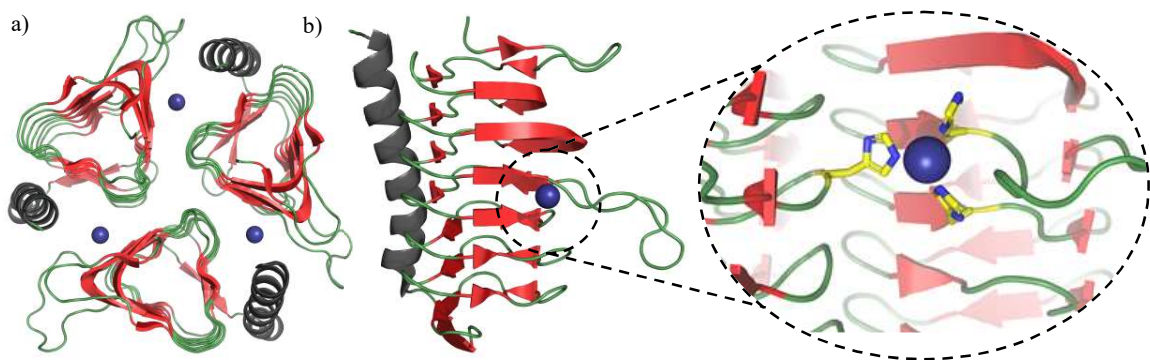


Figure 88 : Métalloenzyme naturelle ayant une structure en hélice  $\beta$  gauche de type I : l'anhydrase carbonique  $\gamma$ , vu du dessus (trimère) (a) ou bien de profil (monomère seul) (b), mettant en évidence le site à zinc (2 His +1 His) (PDB : 3TIO).

Le site catalytique est formé de deux histidines situées en position B<sub>4</sub> de la face (I) et d'une autre placée dans le coude de l'hexade (III) à la position T<sub>1</sub> d'une deuxième hélice. Ces deux positions se faisant face, un site de coordination à trois azotes (3N) est formé pour coordiner un atome de zinc (Figure 88 c). Une molécule d'eau complète la sphère de coordination du zinc pour former un site à  $Zn^{II}$  (3N1O).

Les deux peptides conçus, dont les séquences sont détaillées en Figure 89 reprennent alors la séquence du peptide 35m2 avec quelques mutations dans le but d'adapter la brique de base à la coordination de métaux.

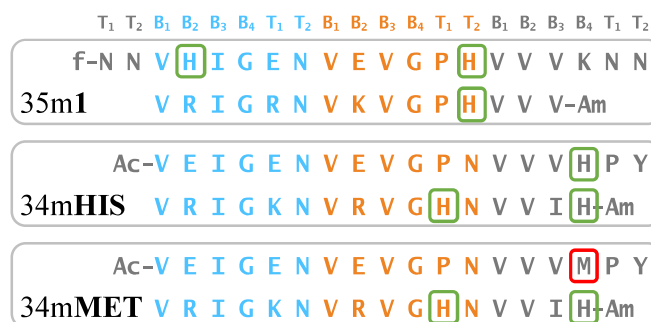


Figure 89: Comparaison des séquences conçues pour coordonner des métaux, les résidus entourés représentent les mutations pour permettre la coordination des métaux.

Dans un premier temps, la coupure s'est faite au niveau d'un des coudes  $\beta$  exposés au solvant. Pour la simple raison d'apporter de la flexibilité aux deux histidines de la face (I) afin de former le site de coordination pour le zinc (34mHIS). Un autre mutant a été conçu en remplaçant une histidine par une méthionine afin de favoriser la coordination du Cuivre ainsi que son cycle d'oxydation/réduction  $\text{Cu}^{\text{I}}/\text{Cu}^{\text{II}}$  (34mMET). Une tyrosine a été incorporée dans la séquence afin de pouvoir connaître la quantité exacte de peptide en solution (dosage par UV-Visible). D'autres mutations ont été faites mais ne seront pas détaillées ici. Le but étant de concevoir des briques identiques, mais avec des sites catalytiques différents permettant de créer, par autoassemblage, des fibres catalytiques.

La conception, appuyée par la modélisation, nous a permis de confirmer que le repliement pourrait être compatible avec les séquences choisies, et cela malgré les mutations faites (Figure 90). Leurs synthèses et purifications ont été réalisées comme décrites précédemment. Et quelques difficultés au niveau de la solubilité ont été néanmoins rencontrées, ces deux mutants étant moins solubles que le peptide parent 35m2. Cette différence est probablement due au fait que les histidines au centre du trimère ne sont plus présentes. Le 34mMET ayant une solubilité moins importante que son homologue avec l'histidine, sûrement dû à la différence de charges.

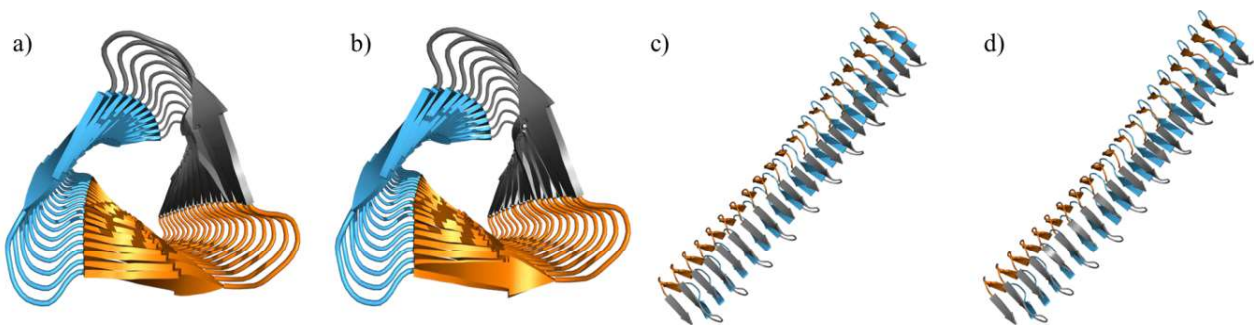


Figure 90 : Modélisation AlphaFold des autoassemblages pour 10 monomères de peptides a, c) 34mHIS et b, d) 34mMET.

Nous les avons aussi caractérisés par les techniques citées précédemment, et ceux-ci montrent des signatures spectrales et microscopiques très proches de celle du 35m2. Il ne s'agit ici que d'analyses préliminaires qui restent à confirmer et surtout à comprendre certains résultats. Les préparations et analyses ont été effectuées comme décrites précédemment.

Les cinétiques de fluorescence de la ThT pour ces peptides nous ont posé quelques problèmes, cela à cause de leurs solubilités. Ceux-ci ont dû être stockés dans une solution à 60% de DMSO et dilués à 20 % lors de la mesure afin de pouvoir observer une augmentation de la fluorescence. Il se peut aussi que la ThT ne se lie pas bien aux structures formées, que ce soit à cause du métal, ou non. L'assemblage semble être stable une fois formé, signe de la robustesse du repliement.

Des analyses en TEM pour le peptide 34mHIS, ont été faites, en fonction du métal mis lors de l'incubation. Des fibres du peptide seul, et en présence de  $\text{Zn}^{\text{II}}$ ,  $\text{Cu}^{\text{II}}$ , et  $\text{Ag}^{\text{I}}$  ont été observés pour déterminer l'influence des métaux sur l'autoassemblage. Comme nous pouvons le voir sur les clichés (Figure 91), le métal joue une

influence sur l'assemblage. Comme nous le voyons en Figure 91, les métaux, notamment le zinc et l'argent permettent de former des fibres plutôt homogènes en diamètres et relativement longues pour les fibres en présence d'argent. Le peptide seul est plutôt très agrégé, formant des paquets de fibres reflétant la tendance de celui-ci à se replier rapidement.

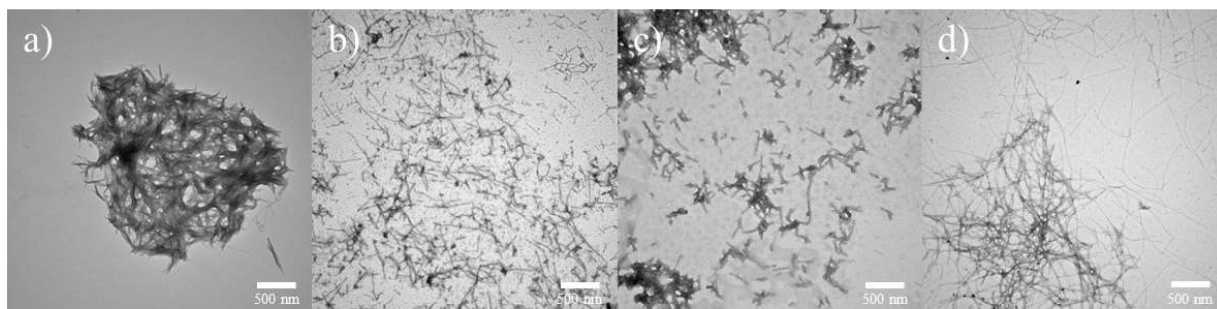


Figure 91 : Influence de la présence de métal sur la fibrillation du peptide 34mHIS à 500  $\mu$ M, 50 mM HEPES pH=7.4, 450  $\mu$ M métal à 45 °C, 1400 rpm pendant 24h, a) seul et en présence de b) Zn<sup>II</sup>, c) Cu<sup>II</sup>, et d) Ag<sup>I</sup>.

Les analyses en CD et ATR-IR nous montrent des signatures similaires au peptide parent, ce qui reste compatible avec la formation du repliement en L $\beta$ H, cela en présence de métal ou non. Ceci semble indiquer que le métal n'a pas d'influence sur la structure secondaire des peptides, ce qui contraste avec les observations que nous avons faites en TEM (morphologie des fibres).

Des tests de catalyse ont été réalisés en présence de Zn, Cu, Ag pour les deux peptides, 34mHIS et 34mMET. Ceux-ci, notamment en présence de Zn<sup>II</sup>, pour des réactions d'hydrolyses (avec l'acétate de paranitrophénol), et de Cu<sup>II</sup> pour des réactions d'oxydoréductions (oxydation de l'ascorbate et de la phénazine). Ces tests préliminaires n'ont pas révélé d'activité particulièrement intéressante. Néanmoins ils nécessitent d'être reproduits, et étendus à d'autres types de réactions. Nous pourrions par exemple envisager de réaliser des réactions de Diels-Alder avec les fibres peptidiques chargées en Cu<sup>II</sup>.

Nous devons cependant, avant de continuer les tests, optimiser les séquences afin de rendre ces peptides plus solubles, car les problèmes de solubilité rencontrés jusque-là ont sévèrement limité les expériences. Cela peut être réalisé en mutant des résidus asparagines de la position T<sub>2</sub> de la face (II) par des résidus basiques comme des lysines, ou arginines par exemple, afin de donner une charge globale plus importante au peptide. Malgré cela, les résultats des caractérisations effectuées sont en accord avec la conception initiale, ce qui montre l'intérêt de la conception de briques peptidiques se repliant en L $\beta$ Hs formant des trimères d'hélices.

Par la suite nous pourrions aussi essayer d'incorporer d'autres types de sites, accommodant d'autres métaux différents, nous menant toujours un peu plus vers la conception rationnelle de métalloenzymes artificielles à partir de briques peptidiques.





# 7 Partie expérimentale

## 7.1 Matériel

**Synthèse :** Tous les acides aminés protégés Fmoc ont été achetés chez NovaBiochem. Les résines Fmoc-Rink-Amide-TentaGel XV (XV30023, Iris Biotech GMBH, 0.2-0.4 mmol/g, 100–200  $\mu\text{m}$ ) et Fmoc-Rink Amide AM (Br-1330 Iris Biotech GMBH, 0.74 mmol/g, 100–200 mesh) utilisées ainsi que le N,N'-Diisopropylcarbodiimide (DIC) ont été achetés chez IRIS Biotech. L'acide trifluoroacétique (TFA), pipéridine, Oxyma, Triisopropyl silane (TIS), Thianisole et N,N-Diisopropylethylamine(DIEA) ont été achetés chez Merck. L'anhydride acétique ( $\text{Ac}_2\text{O}$ ) et l'éther diéthylique ( $\text{Et}_2\text{O}$ ) ainsi que la N-Méthyl-2-pyrrolidone (NMP) ont été achetés chez Sigma-Aldrich. Les solvants, dichlorométhane (DCM), acétonitrile (MeCN) et diméthyle formamide (DMF) ont été achetés chez Fisher Chemical et Carlo Erba.

Un agitateur oscillant et un robot de synthèse Initiator Alstra+ de chez Biotage ont été utilisés pour la synthèse.

Un spectromètre Agilent Cary-60 UV-Vis a été utilisé, notamment pour quantifier le peptide couplé par dosage des groupements Fmoc.

**HPLC analytique :** HPLC analytique de type Hitachi Primade avec détecteur Primaide Diode Array Detector instrument avec détection UV à 214 nm, 254 nm et 280 nm.

**Colonne a :** colonne Waters XBridge peptide BEH C18, 300 Å, 3.5  $\mu\text{m}$ , 4.6 mm  $\times$  150 mm.

**Colonne b :** colonne Waters XBridge peptide BEH C4, 300 Å, 3.5  $\mu\text{m}$ , 4.6 mm  $\times$  150 mm.

**Colonne c :** colonne HALO C4 bioclass, 1000 Å, 2.7  $\mu\text{m}$ , 4.6 x 150 mm.

**HPLC semi-préparative :** HPLC semi-préparative de type VWR LaPrep + four Jasco CO-4060

**Colonne d :** colonne Waters XBridge protein BEH C18, OBD, Prep Column, 130 Å, 5  $\mu\text{m}$ , 19 mm  $\times$  250 mm.

**Colonne e :** colonne Waters XBridge protein BEH C4, OBD, Prep Column, 300 Å, 5  $\mu\text{m}$ , 19 mm  $\times$  250 mm.

**Colonne f :** colonne Waters XBridge peptide BEH C18, OBD, Prep Column, 300 Å, 5  $\mu\text{m}$ , 10 mm  $\times$  250 mm.

**Chromatographie Liquide couplé à la Spectrométrie Masse (LC-MS) :** Thermo-Scientific LCQ Fleet Ion Trap LC-MS & Vanquish UHPLC avec détecteur Primaide Diode Array Detector instrument avec détection UV à 214 nm et 280 nm.

**Colonne g :** colonne Waters XBridge peptide BEH C18, OBD, Prep Column, 300 Å, 5  $\mu\text{m}$ , 10 mm  $\times$  250 mm.

**Analyses :** Les peptides utilisés ont été synthétisés et purifiés classiquement. L'eau ultrapure a été générée au laboratoire avec un appareil à eau Milli-Q® Direct. Les sels et Tampons ont été achetés chez Sigma-Aldrich. La Thioflavine-T provient de chez Acros.

Les tubes de 1.5, 0.75 mL et 100  $\mu\text{L}$  ont été achetés chez Eppendorf.



Pour l'incubation, un thermomixer Eppendorf C a été utilisé avec un portoir pour tube de 1.5 mL avec ThermoTop®.

Le spectromètre UV-visible utilisé est un Agilent Cary-60 UV.

Le logiciel OriginPro a été utilisé pour générer et mettre en forme les résultats et générer les figures.

**Fluorescence de la ThT** : Les plaques 96 puits GREINER 96 F-BOTTOM  $\mu$ Clear® Black viennent de chez Dutscher et les films pour microplaque BRAND® ont été achetés chez Sigma-Aldrich.

Le lecteur de plaques multipuits pour le suivi de fluorescence est un CLARIOstar Plus de chez BMG Labtech piloté par les logiciels CLARIOstar et MARS.

**Microscopie électronique** : Les solutions d'acétate d'uranyle à pH=4 ou 5 à 2 % et de phosphotungstate de sodium à pH=7.05 à 2% ont été fournies par nos collaborateurs.

Les grilles de microscopie électronique utilisées ont été des FF300-Cu (film Formvar) de chez Electron Microscopy Sciences, des S160 carbon film 200 mesh copper fournies par nos collaborateurs.

Les grilles de TEM ont été chargées en utilisant un ELMO Glow Discharge Cleaning system.

Les grilles utilisées pour la Cryo-TEM sont des QUANTIFOIL® R2/2 Cu200 de chez Quantifoil et des LC300-Cu de chez Electron Microscopy Sciences. Elles ont été vitrifiées à la main dans de l'éthane liquide et en utilisant un FEI Vitrobot™ Mark IV (Thermo Fisher, New York, NY, USA).

Le premier microscope utilisé est un FEI Tecnai G2 200kV FEG (Thermo Fisher, New York, NY, USA) en collaboration avec le Dr. Schmutz (PLAMICS, ICS, Strasbourg). Le second microscope est un Hitachi H 7500 équipé d'une caméra numérique en collaboration avec la Dr. Demais (UPS 3156, INCI, Strasbourg). Ensuite un microscope FEI Tecnai Spirit 120kV LaB6 (Thermo Fisher, New York, NY, USA) équipé d'une caméra DDC K2 Base 4 k  $\times$  4 k camera (Gatan, Pleasanton, CA, USA) et pour la Cryo-TEM, FEI Tecnai G2 200kV FEG équipé d'une caméra DDC K2 Summit en collaboration avec la Dr. Pieri et le Dr. Bressanelli (I2BC, Cryo-EM, Gif-sur-Yvette).

**Dichroïsme circulaire et linéaire** : L'incubation au synchrotron a été réalisée dans un agitateur pour tube chauffant couvert avec du papier aluminium pour conserver la chaleur.

Le dichroïsme circulaire a été mesuré à la plateforme d'analyse de l'institut de chimie de la faculté de Strasbourg sur un spectromètre de paillasse Jasco J-810 équipé d'une lampe Xenon 150 W. Les cuvettes utilisées ont été des cuvettes en quartz de 1 mm (300  $\mu$ L) et cuve « sandwich » de 0.1 mm (30  $\mu$ L) de trajet optique.

Le dichroïsme circulaire et linéaire sur source synchrotron a été mesuré sur la ligne DISCO du synchrotron SOLEIL (Gif-sur-Yvette) en collaboration avec le Dr. Wien (Scientifique Synchrotron SOLEIL, Gif-sur-Yvette). Des cuves en CaF<sub>2</sub> (Hellma Ltd) de 31  $\mu$ m (2  $\mu$ L), pour les suspensions, et 12  $\mu$ m, pour les films, de

trajet optique ont été utilisées. Les films ont été préparés en appliquant une dépression dans une chambre d'incubation en plastique.

Une centrifugeuse Tube centrifuge 5804R a été utilisée pour centrifuger les tubes de peptides fibrillés.

**Infrarouge** : Les mesures et interprétations des spectres infrarouges ont été réalisées en collaboration avec le Dr. El Khoury (LBS, institut LeBel, Strasbourg).

Le spectromètre infrarouge utilisé est un Vertex 70 (Bruker Optics, Karlsruhe) équipé d'un détecteur MCT et d'un module ATR diamant avec une résolution spectrale de  $4\text{ cm}^{-1}$ . Le logiciel d'acquisition OPUS 7.2 est celui fourni avec le spectromètre.

**Cristallisation** : Les plaques 96 puits ont été achetées chez Hampton Research pour la MRC 2 Well UVP et chez Corning pour la 96-well COC Protein Crystallization Microplate. Les plaques de cristallisation CrystalQuick X Plate, et In Situ-1 Crystallization Plate provenant de chez MiTeGen et les plaques 60 well Greiner HLA Terasaki multiwell plates ont été données par notre collaborateur C. Sauter. Les films BRAND® sealing film for microplates pour couvrir les plaques ont été achetées chez Sigma-Aldrich. L'huile de paraffine utilisée provient de chez Sigma-Aldrich. L'agarose à bas point de fusion (mp.  $28^{\circ}\text{C}$ ) provient de chez Roth.

Les screens commerciaux, Index HT, et Crystal Screen HT ainsi que le l'outil d'ensemencement proviennent de chez Hampton Research.

La loupe binoculaire à zoom MOTIC SMZ161-BLED équipée d'une lentille de grossissement x2 a été achetée chez MOTIC.

Les photos des cristaux ont été prises avec un téléphone équipé d'une caméra optique.

Les MicroMounts™ et MicroLoops™ proviennent de chez MiTeGen.

**Diffraction des rayons X** : Les données de diffraction du peptide  $5M_L$  (f-VGPLT-Am) ont été recueillies avec un système Rigaku HF-007 équipé d'un détecteur HyPix6000, d'un monochromateur multicouche et d'une anode tournante en cuivre ( $\lambda = 1.54178\text{ \AA}$ ) à 100 K comme source d'électron.

Les données de diffraction du peptide  $5M_N$  (f-VGPNT-Am) ont été recueillies sur la ligne Proxima-2A Synchrotron Soleil, France équipée d'un détecteur Eiger 9M. L'appareil est contrôlé depuis l'extérieur de la cabine par un ordinateur avec le logiciel ALBULA (Dectris, Switzerland).

Le traitement des données collectées a été réalisé par nos collaborateurs C. Sauter, K. Rollet et P. Pachl (ARN, IBMC, Strasbourg). Les logiciels ADXV et ALBULA ont été utilisés pour collecter et visualiser les clichés de diffraction. Les logiciels CrysAlisPro, Olex, AutoPROC, XDS et SHELXT et SHELXL pour traiter les données.

**RMN du solide** : Les cristaux de peptides utilisés ont été préparés comme décrits en 7.2.1.3.1

Le rotor RMN utilisé est un rotor pour MAS de 1.3 mm de diamètre en oxyde de zirconium (ZrO<sub>2</sub>) fermé avec des bouchons en Vespel®.

Les spectres 1D et 2D en RMN du solide ont été réalisés sur un spectromètre Bruker™ AVANCE 500 MHz à grand diamètre fonctionnant à une fréquence de 500.03 MHz pour 1H (125.7 MHz pour 13C) et à une fréquence de rotation de 22.5 kHz avec le tétraméthylsilane (TMS) comme référence interne et un échantillon d'adamantane comme référence secondaire.

## 7.1.1 Analyse bio-informatique des structures en hélice $\beta$ gauche naturelles

### 7.1.2 Logiciels et bases de données

La banque de données sur les protéines PDB (<https://www.rcsb.org/>) ainsi que la banque de données DbStRiP (<https://bioinf.iiit.ac.in/dbstrips/>) ont permis une recherche de structures en hélices  $\beta$  gauche de manière précise jusqu'au 15/08/2021.

Toutes les figures de structures de protéine ont été générées via PyMol.<sup>[1]</sup>

### 7.1.3 Démarche d'analyse

#### 7.1.3.1 Visualisation

À partir d'une structure d'hélice  $\beta$  gauche trouvée sur la PDB, téléchargée au format .pdb et ouverte avec PyMol, nous avons identifié le motif structural canonique de l'hélice de section triangulaire. Les molécules d'eau de la structure ainsi que les métaux ou ligands ont été retirés. Lorsque ce n'était pas le cas, les structures trimériques ont été générées à partir des opérateurs de symétrie obtenus grâce à la maille cristalline élémentaire via la commande « symexp sym,XXX,(XXX),4 » où XXX représente le nom de la structure PDB et 4 le rayon de génération de structures en Angströms.

Les structures ne contribuant pas à l'hélice (boucles, motifs de terminaison, etc...) ont été masquées pour ne visualiser que les structures canoniques. Ce qui résulte en un trimère d'hélices  $\beta$  gauche régulier avec des longueurs de 4 à 11 pas d'hélice.

Les couleurs ont ensuite pu être attribuées par face en fonction de leur position dans le trimère en N<sub>ter</sub> vers le C<sub>ter</sub>. La face exposée à une autre hélice et au solvant, notée **I**, a été colorée en orange. L'hexade faisant face au coin de l'hélice opposée, notée **II**, est colorée en gris. Celle la plus exposée au solvant est notée **III**, et est colorée en bleue.



### 7.1.3.2 Report et tri dans une feuille de calcul

Les couleurs aidant à se repérer quant aux résidus appartenant aux différentes faces, il faut maintenant les reporter dans une table. Cela a été réalisé sur Excel, avec une feuille de calcul par partie, c'est-à-dire, une pour le tour de nucléation (vert), identifié comme le premier tour après le motif de terminaison en N<sub>ter</sub>, une pour les tours canoniques (noir) et une pour le tour de terminaison juste avant le C<sub>ter</sub> (violet). Le report a été réalisé dans le sens N<sub>ter</sub> vers le C<sub>ter</sub> pour chaque partie. Chaque résidu a été reporté suivant sa position, et sa face, ce qui donne une ligne d'un tableau à 3 x 6 colonnes pour chaque pas d'hélice. Tous les acides aminés ont ensuite été comptés par position, en fonction des faces dans lesquels ils sont et reportés dans une table Excel. Au total, 1377 résidus ont été recensés pour les tours de nucléation, 7377 pour les tours canoniques et 1388 pour les tours de terminaison.

Les acides aminés ont ensuite été triés, pour chaque position en utilisant la fonction « NB.SI » du tableur. Pour les 18 positions, les 20 acides aminés ont été triés automatiquement via cette fonction, à cela il faut ajouter les boucles et les résidus manquants. Chaque résidu a ensuite été pondéré afin d'obtenir la fréquence d'apparition de celui-ci pour chacune des positions. Les résultats de fréquences, en %, ont été compilés dans un histogramme à 18 entrées correspondant à chaque position. Le code à une lettre des résidus a été ajouté à la main et la taille correspondante est celle de la fréquence d'apparition.

Cela résulte en 3 histogrammes à 18 entrées correspondant aux différents tours.

Les motifs en épingle  $\beta$  (après le dernier tour en C<sub>ter</sub> de l'hélice) ont été recensés en suivant la même procédure.

Le peptide de 59 résidus de long a été synthétisé et « purifié » comme décrit dans la partie suivante 7.1.4.

## 7.1.4 Conception de peptides de novo s'autoassemblant en hélice $\beta$ gauche

### 7.1.5 Méthodes

#### Solvants utilisés en Chromatographie Liquide à Haute Performance (HPLC) :

Les solvants ont été préparés à l'aide d'éprouvettes graduées de 1000 et 500 mL. Le solvant A est constitué d'eau avec 0,1% de TFA (pH = 2). Le solvant B est composé d'un mélange d'acétonitrile à 90% et 10% d'eau, avec 0,1% TFA. Pour des raisons de simplicités, les gradients de solvants seront toujours écrits comme suit : A/B où A et B représentent les pourcentages de solvants respectifs.

#### Solvants utilisés en Chromatographie Liquide couplé à la Spectrométrie Masse (LC-MS) :



Les solvants ont été préparés minutieusement à l'aide d'éprouvettes graduées de 1000 et 500 millilitres. Le solvant C, est constitué d'eau à 0,1% de FA. Le solvant D, est composé d'un mélange de MeCN à 90% et 10% d'eau, le tout à 0,1% FA. Pour des raisons de simplicités, les gradients de solvants seront toujours écrits comme suit : C/D où C et D représentent les pourcentages de solvants respectifs.

#### **Méthodes utilisées en HPLC analytique :**

**Méthode A.** Gradient 95(A)/5(B) à 50/50 en 15 minutes, 1mL/min, à 37°C.

**Méthode B.** Analyse avec gradient 95/5 à 0/100 en 30 minutes, 1mL/min, à 37°C.

**Méthode C.** Analyse avec gradient 80/20 à 70/30 en 30 minutes, 1mL/min, à 37°C.

**Méthode C'.** Analyse avec gradient 80/20 à 70/30 en 30 minutes, 1mL/min, à 60°C.

#### **Méthodes utilisées en HPLC semi-préparative :**

**Méthode 1 :** gradient 95/5 à 50/50 en 20 min, 14mL/min, T<sub>amb.</sub>

**Méthode 2 :** gradient 80/20 à 35/65 en 30 min, 14mL/min, 60°C.

**Méthode 3 :** gradient 80/20 à 30/70 en 30 min, 14mL/min, 60°C.

**Méthode 3' :** gradient 78/22 à 25/75 en 30 min, 14mL/min, 60°C.

#### **Méthodes utilisées en LC-MS :**

**Méthode  $\alpha$  :** gradient 98(A)/2(B) à 50/50 en 15 minutes, 40 $\mu$ L/min, 40°C.

**Méthode  $\gamma$  :** gradient 98(A)/2(B) à 0/100 en 15 minutes, 40 $\mu$ L/min, 40°C.

**Méthode  $\eta$  :** gradient 98(A)/2(B) à 0/100 en 30 minutes, 40 $\mu$ L/min, 40°C.

## **7.1.6 Synthèse Peptidique en Phase Solide (SPPS), stratégie**

### **Fmoc/tBu :<sup>[2,3]</sup>**

Les peptides les plus longs, de 35 acides aminés (35m1 et 35m2) ainsi que celui de 59 résidus de long (59m), ont été synthétisés avec un synthétiseur automatique de peptide Biotage Initiator+ Alstra en utilisant un nombre de moles de départ de 0,05 ou 0,1 mmol, avec une résine Fmoc-Rink-Amide-TentaGel XV (Iris Biotech) de charge 0.74 mmol.g<sup>-1</sup>. Une perte d'environ 25 % de la charge de la résine a été constatée par la suite, donc, certaines synthèses ont en fait été réalisées avec des charges 25 % plus faibles. Les solutions d'acides aminés préparées pour les synthèses utilisant le robot avaient une concentration de 0,15 mol/L, les solutions de DIC et d'Oxyma de 0,5 mol/L, les solutions d'Ac<sub>2</sub>O de 5 mol/L, et de pipéridine à 20% (v/v) dans de la DMF. Seules les solutions de DIEA ont été préparées à 2 mol/L dans de la (NMP) pour des raisons de solubilité et stabilité.

La résine est directement pesée dans le réacteur de synthèse de 10 mL, rincée une fois et mise pour solvation pendant 30 min à l'agitateur oscillant dans environ 8 mL de DCM. Ensuite, la résine a été solvatée dans 8 mL de DMF pendant 15 min. **Deprotection des Fmoc.** L'étape de deprotection des groupements « Fmoc » est réalisée sous agitation pendant 3 minutes dans un mélange pipéridine/DMF 1:4 (3 fois). Les Fmoc issus de la deprotection de la résine ont été dosés (voire « dosage des Fmoc » plus bas), afin de déterminer la charge exacte de la résine.

La résine est ensuite lavée 3 fois avec 4 mL de DMF avant d'être mise dans le robot pour synthèse. **Couplage.** Des doubles couplages de 30 min ont été réalisés pour tous les acides aminés, avec 4 équivalents d'acides aminés, 3,8 équivalents de DIC/Oxyma à 70°C (micro-ondes) sous agitation. Les résidus non compatibles avec les micro-ondes présents dans la séquence (i.e. histidine, acide aspartique, cystéine, arginine) ont été couplés 2 fois pendant 1h à température ambiante (avec les mêmes proportions que les autres). Pour certains couplages tels que R-N (en position 11), un double couplage d'1h est réalisé à  $T_{amb}$  avec une préactivation (16 eq. de DIEA, 3.8 eq. d'HCTU et 4 eq. d'arginine). Une étape de lavage avec 1 x 4.5 mL de DMF est réalisée entre chaque étape de couplage. **Acétylation.** Après chaque couplage d'A.A, une acétylation (« capping ») est réalisée, à l'aide d'une solution d'anhydride acétique ( $Ac_2O$ ) à 5 %, de 10 % de DIEA dans de la DMF, sous agitation pendant 1 minute. La résine est ensuite lavée 4 fois avec 4.5 mL de DMF avant d'effectuer l'étape de deprotection des Fmoc suivante.

Ce cycle (déprotection des Fmoc – couplage – acétylation) est répété jusqu'à obtenir le peptide de longueur voulue.

La **formylation** du résidu situé en  $N_{ter}$  s'effectue en préactivant, pendant 1 minute sous vortex, 10 équivalents d'acide formique (FA), dans un premier temps dissout dans environ 10 mL de DMF, et 10 équivalents de DIC. La solution est ensuite transférée dans le réacteur de synthèse et mise sous agitation pendant 30 minutes en dégazant le réacteur toutes les 5 minutes. Cette étape est réalisée 2 fois afin de s'assurer de la bonne formylation du  $N_{ter}$ . Puis, la résine est lavée 3 fois avec 5 mL de DMF. Ensuite, un test avec du 2,4,6 - Trinitrobenzenesulfonic Acid (TNBS) est réalisé, sur une pointe de spatule de résine, en mélangeant 2 gouttes de solution de TNBS à 1% avec 2 gouttes d'une solution de DIEA à 10% dans de la DMF. Cela permet de révéler s'il reste ou non des amines primaires libres (coloration rouge de la résine). Dans notre cas, si la formylation est totale, les billes restent non colorées.

Les étapes d'**acétylation** éventuelles du  $N_{ter}$  ont été réalisées à la main avec les mêmes proportions de réactifs utilisées lors du capping pendant 2 x 10 min. La résine est lavée avec 3 fois avec 5 mL de DMF et le test TNBS réalisé comme précédemment pour s'assurer de la bonne fonctionnalisation.

Le **clivage final** a été réalisé à l'aide d'une solution d'acide trifluoroacétique (TFA)/Triisopropyl silane (TIS)/ $H_2O$  (95/2,5/2,5) pendant 60 min puis 2 fois 30 min. Durant cette étape, le peptide est clivé de la résine, les groupements protecteurs des chaînes latérales sont aussi enlevés lors de cette étape. Le peptide brut ainsi obtenu, soluble dans le TFA, est précipité par ajout d'éther diétylique ( $Et_2O$ ) glacé. Il en résulte une suspension blanchâtre, que l'on centrifuge à 4000 rpm pendant 3 min. Le surnageant est jeté, et le culot est repris (rincé)



dans de l'Et<sub>2</sub>O glacé. L'opération est répétée une troisième fois. Le culot finalement obtenu est ensuite dissout dans une solution d'eau à 0,1% de TFA (A), et, si besoin, d'acétonitrile/eau/0,1% TFA, dans des proportions permettant la solubilisation. Un peu de solution est analysée en HPLC analytique afin d'évaluer la pureté du peptide brut, ainsi que le pourcentage de solvants A et B nécessaires pour éluer le peptide de la colonne. Le reste est lyophilisé dans le but d'obtenir une poudre homogène en vue de la purification.

Un schéma global de synthèse SPPS est montré en Figure 92.

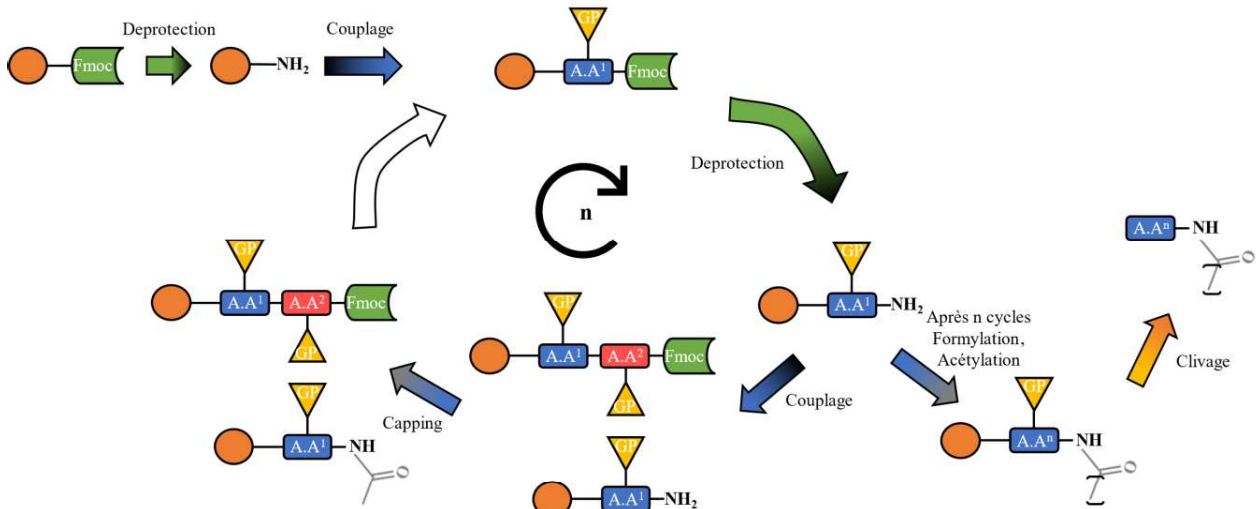


Figure 92 : Schéma de synthèse pour la SPPS.

- **Dosage des Fmoc :**

Le dosage des Fmoc permet de déterminer, par dosage, un rendement grossier de la synthèse réalisée. Pour cela, les solutions ont été mises dans une fiole jaugée de 100 mL puis complétées au trait de jauge avec du méthanol. Après avoir mesuré le spectre d'absorbance entre 400 et 200 nm sur spectrophotomètre UV-Vis Agilent Cary 60 du méthanol seul, comme ligne de base (2,8 mL dans une cuve de 3 mL), 0,2 mL de solution de déprotection diluée a été ajouté et l'absorbance remesurée. Le coefficient d'extinction molaire de la molécule issue de la déprotection, un adduit pipéridine-dibenzofulvène (pip-DBF), est de  $\epsilon = 7800 \text{ L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$  à 301 nm (Figure 93).<sup>[4]</sup> Ce qui nous permet de suivre et de déterminer la quantité exacte de groupements Fmoc déprotégés par la loi de Beer-Lambert appliquée à notre cas, en considérant la double dilution (Équation 1).

$$\text{(Équation 1)} \quad n = \frac{A}{7800} \times \frac{V_{\text{sol}}}{V_{\text{cuve}}} \times V_{\text{tot}}$$

où n représente la charge de la résine en  $\mu\text{mol}$ , A, l'absorbance de la solution à 301 nm, 7800 est le coefficient d'absorption molaire à 301 nm,  $V_{\text{cuve}}$  est le volume de la cuve (3mL),  $V_{\text{sol}}$  correspond au volume de solution de déprotection diluée 1 fois ajouté (0,2 mL) et  $V_{\text{tot}}$  est le volume total de la fiole (100mL). On peut alors simplifier l'Équation 1 comme suit :

$$\text{(Équation 2)} \quad n = \frac{A}{7800} * 1,5$$

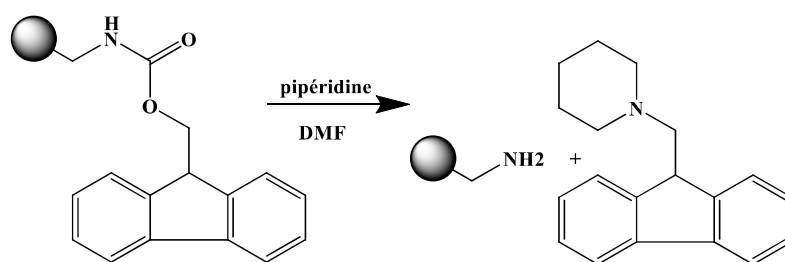


Figure 93 : Réaction de deprotection des groupements Fmoc et formation des adduits pipéridines-dibenzofulvènes.

## 7.1.7 Purification

### 7.1.7.1 Protocole

La purification s'effectue en dissolvant le peptide brut lyophilisé dans du tampon A, à environ 5 mg/mL. Avant injection, le peptide est passé sur filtre 220  $\mu\text{m}$  dans le but d'enlever toutes les poussières et impuretés nuisibles à la colonne. La solution est ensuite injectée en HPLC semie-préparative. Les fractions collectées ont été analysées par HPLC analytique, LC-MS avant d'être réunies, congelées dans la glace carbonique puis lyophilisées (Annexe 3 et Annexe 4). Les peptides sont ensuite stockés au congélateur à  $-18^{\circ}\text{C}$ . Il en résulte une poudre de peptide pure. Les rendements totaux obtenus sont d'environ 4 à 6.5%. La purification des peptides se faisant à 0.1 % de TFA, nous devons considérer, pour chaque peptide, le nombre de contre-ions trifluoroacétate. Un résumé des conditions de purification et d'analyse des deux 35-mers est reporté en Tableau 17.

Tableau 17 : Tableau récapitulatif de synthèse et purification des 35 mers conçus.

Peptide	Masse Molaire (g/mol)	Masse Molaire avec contre-ions TFA (g/mol)	condition purification; colonne	Condition HPLC Analytique; colonne	Nombre de mole initial ( $\mu\text{mol}$ )	Nombre de mole final ( $\mu\text{mol}$ )	rendement Fmoc (%)	masse peptide pur (mg)	rendement (%)
35m1	3806.4	4492.6	Méthode 2; e	Méthode C,C'; a, b	75.0	3.7	4.9	15.5	4.6
35m2	3862.6	4744.8	Méthode 3;3'; e	Méthode C,C'; a, b	75.0	5.1	6.8	23.0	6.5

Pour chaque peptide pur, une partie de la poudre est séparée du reste afin d'échanger les contre-ions TFA<sup>(-)</sup> par des contre-ions chlorure (Cl<sup>(-)</sup>). Le peptide est dissout dans une solution de HCl 2 mM, puis est lyophilisé. Cette opération est répétée 3 fois pour s'assurer du bon échange de contre-ions, en vue d'analyse en spectroscopie infrarouge à réflectance totale atténuée (ATR-IR).<sup>[5,6]</sup>

## 7.1.8 Suivi cinétique à l'aide de la thioflavine-T

### 7.1.8.1 Préparation de la solution de ThT

La solution stock de thioflavine-T a été préparée par dissolution de 15.9 mg de la poudre commerciale ( $M=318.86$  g/mol,  $49,8$   $\mu\text{mol}$ ) dans 10 mL d'eau MilliQ dans un tube falcone de 10 mL. **Dosage de la solution de ThT** : 1  $\mu\text{L}$  de cette solution est mis dans une cuve UV et elle est complétée jusqu'à 1 mL. La molécule de ThT libre a un coefficient d'absorption molaire à 412 nm  $\epsilon_{412}= 36000$   $\text{L}\cdot\text{mol}^{-1}\cdot\text{cm}^{-1}$ .<sup>[7]</sup> Après agitation, l'absorbance à 412 nm est mesurée et la concentration calculée par la loi de Beer-Lambert.

À partir de la solution dont la concentration a été déterminée précédemment, une solution stock à 2 mM est préparée. La solution stock est ensuite aliquotée dans des tubes et stockée au congélateur à l'abri de la lumière.

La concentration dans les plaques est de 200  $\mu\text{M}$ , donc 10  $\mu\text{L}$  dans 100  $\mu\text{L}$  finaux. Il est important de toujours garder la solution stock à l'abri de la lumière, c'est pour cela que les tubes utilisés ont été entourés de papier aluminium et sortis du congélateur environ 20 minutes avant utilisation pour qu'ils reviennent à température ambiante.

### 7.1.8.2 Préparation des tampons

Les tampons utilisés, d'un pH=3 à un pH=10 ont été préparés, pour beaucoup d'entre eux, par dissolution d'une poudre commerciale dans de l'eau MilliQ. Les sels sont dissouts dans une quantité minimale d'eau, puis le pH est ajusté à l'aide de solutions concentrées de NaOH et de HCl.

En Tableau 18 sont résumés les tampons préparés, et à quel pH, ainsi que les sels commerciaux utilisés.

Tableau 18 : Tampons utilisés pour quelle gamme de pH.

Tampon	Phosphate	Acide citrique	Acide acétique	MES	HEPES	CHES
Sel de départ	$\text{KH}_2\text{PO}_4/\text{K}_2\text{HPO}_4$	$\text{HOC}(\text{COOH})$ $(\text{CH}_2\text{COOH})_2 \cdot \text{H}_2\text{O}$	$\text{CH}_3\text{COOH}$	$\text{C}_6\text{H}_{12}\text{NO}_4\text{S}\cdot\text{Na}$	$\text{C}_8\text{H}_{18}\text{N}_2\text{O}_4\text{S}$	$\text{C}_8\text{H}_{17}\text{NO}_3\text{S}$
pH de tampon	3	3.5	4 à 5	5.5 à 6.5	7 à 8.5	9 à 10

### **7.1.8.3 Préparation des plaques pour la mesure de la fluorescence**

#### **7.1.8.3.1 Paramètres de mesures**

Les paramètres du lecteur de plaques utilisés sont adaptés à la plaque utilisée et la longueur d'onde d'excitation est centrée à 450 ( $\pm 15$ ) nm alors que celle d'émission est centrée sur 480 ( $\pm 20$ ) nm. La température est fixée à 37°C et la valeur de fluorescence mesurée est une moyenne de 106 points (flash) par puits avec une mesure en spirale.<sup>[7]</sup> La valeur du gain est fixée soit manuellement à 900 soit avec la mesure de la valeur sur un puits de référence. L'agitation de la plaque entre les mesures est fixée à 500 rpm. Enfin, un point est pris toutes les 5 minutes afin de suivre la cinétique de fibrillation des peptides.

#### **7.1.8.3.2 Préparation des plaques**

Les plaques multipuits sont préparées toutes de la même manière à partir des solutions à température ambiante. Ensuite, la plaque est « lestée » en ajoutant de l'eau entre les puits, jusqu'à avoir un volume atteignant les  $\frac{3}{4}$  de la hauteur des puits. Cela permet aussi d'éviter une évaporation trop importante des solutions. Des poids de 20 grammes sont aussi ajoutés à l'intérieur des deux bords courts de la plaque afin d'éviter qu'elle ne se déloge de son emplacement lors de l'agitation. Enfin, les solutions sont ajoutées à la pipette multicanaux dans l'ordre suivant : eau, ThT, peptide et enfin tampon. Une fois ajoutées, les solutions des puits sont homogénéisées à la pipette multicanaux (en évitant de faire des bulles) et la plaque est scellée en collant le film sur le dessus. À partir de l'étape d'ajout du tampon, il est nécessaire d'être rapide dans le but de perdre le moins d'information sur la cinétique d'assemblage (notamment pour le 35m1). Après ajout, la plaque peut être mise à incuber dans le lecteur de plaque à 37°C et le suivi cinétique réalisé. Lorsque des ajouts à des temps différents de tampons ont dû être faits, nous avons mis en pause l'analyse et sorti la plaque, décollé le film et ajouté le tampon, homogénéisé et le film soit changé, soit recollé. La plaque est réintroduite dans le lecteur de plaque.

#### **7.1.8.3.3 Collecte et traitement des données**

La collecte des données se fait après la mesure via le logiciel CLARIOstar par le logiciel MARS en exportant sous forme de fichier .xlsx qui peut être directement ouvert sur Excel ou Origin. Les courbes sont ajustées au niveau de la fluorescence la plus stable, lorsque les valeurs sont constantes dans le temps, en début de mesure avant d'ajouter le tampon par exemple.

## **7.1.9 Microscopie électronique**

### **7.1.9.1 Incubation des solutions pour fibrillation**

Les solutions de tampons et d'eau ont été filtrées sur filtre 200  $\mu\text{m}$  afin d'enlever toutes les fibres et poussières présentes qui ne seraient pas du peptide. Les solutions stocks de peptide (35m1 ou 35m2) à 2 mM ont été obtenues par dissolution des peptides purifiés dans de l'eau ultrapure puis soniquées pendant 10 min. À partir de ces solutions, ont été préparées, par dilution, des solutions à 500, 250 ou 100  $\mu\text{M}$  de peptide pour incubation. Les solutions tampons d'HEPES (pH=7.4 et 8.5) à 500 mM de concentration stock ont été ajoutées aux solutions de peptides pour obtenir une concentration finale de 50 mM. Le volume d'eau ultrapure nécessaire a été complété pour obtenir les concentrations voulues. Toutes les incubations ont été réalisées dans des tubes de 100  $\mu\text{L}$  imbriqués dans des tubes de 750  $\mu\text{L}$  et 1.5 mL pour pouvoir s'adapter sur l'incubateur. Ils ont ensuite été entourés de parafilm autour de leurs capuchons et mis à incuber dans le Thermomixer® équipé d'un ThermoTop®.

Plusieurs conditions ont été testées et les meilleures conditions déterminées sont données en dans le chapitre 2 et sont une incubation de minimum 3 heures, à 40 °C à 500  $\mu\text{M}$  de peptide dans l'HEPES pH=8.5 à 1200 rpm.

### **7.1.9.2 Préparation des grilles de microscopie électronique à transmission (TEM) et Cryo-TEM**

#### **7.1.9.2.1 Grilles de TEM**

Les dépôts sur grilles de TEM ont été aussi optimisés afin d'avoir des fibres pas trop denses sur la grille, au risque de n'avoir qu'un tapis et de ne rien voir. Après incubation, nous avons alors trouvé qu'une dilution des suspensions à 100  $\mu\text{M}$  pour un dépôt de 3  $\mu\text{L}$  sur grille correspond aux meilleures conditions.

En collaboration avec la Dr. Pieri à l'I2BC, les grilles ont reçu un traitement préalable pour les charger en surface, afin d'augmenter le contact avec les gouttes. Le traitement a été réalisé sous vide à 30  $\mu\text{A}$  pendant 10 secondes pour les grilles de TEM ou à 10  $\mu\text{A}$  pendant 1 min. Pour les autres analyses, les grilles n'ont pas reçu ce traitement.

Ensuite, 3  $\mu\text{L}$  de suspension de peptide dilué sont déposés sur grille pendant 1 min puis le liquide a été absorbé avec un papier filtre plié. La grille est ensuite retournée sur une goutte de contrastant et est laissée pendant 2 fois 45 secondes. Le liquide est ensuite absorbé avec du papier absorbant et la grille est ensuite laissée à sécher pendant quelque temps. Par la suite, les grilles sont ensuite montées sur le porte échantillon et analysées au microscope électronique.

### 7.1.9.2.2 Grille de Cryo-TEM

La préparation des grilles pour la Cryo-TEM se fait de manière identique à celles de la TEM avec charge de la grille et dépôt de 3  $\mu\text{L}$  de solution diluée à 100 ou 250  $\mu\text{M}$  de peptide pendant 1 min. Ensuite, à la main, le liquide a été absorbé précautionneusement avec du papier absorbant et très rapidement plongé dans l'éthane liquide refroidi dans l'azote liquide. Avec le Vitrobot, les conditions les meilleures ont été une force d'absorption (force de serrage de la grille entre les deux disques de papier absorbant) de 0, pendant 6 secondes, avec une humidité de 100 % et à température ambiante dans la chambre où a lieu le dépôt et l'absorption (Figure 94). Après absorption, la grille est plongée très rapidement dans de l'éthane liquide pour permettre une vitrification de l'eau et non une cristallisation.<sup>[8]</sup> Après quelques minutes, la grille peut être transférée dans de l'azote liquide et stockée dans un dewar en attendant d'être transférée dans le porte-échantillon pour la Cryo-TEM.<sup>[8]</sup>

Pour l'observation des grilles, le porte-échantillon doit être refroidi pendant 1 heure à l'azote liquide jusqu'à stabilisation de la température puis transférer dans le microscope pour être refroidit jusqu'à ce que la température se stabilise et ensuite peut être sortie et rapidement la grille est mise dessus. Il faut faire très attention à ne pas tordre la grille ou ne pas la faire décongeler lors de cette étape. Ensuite le porte échantillon est remis et le vide est fait. L'analyse par Cryo-TEM peut alors être réalisée.

Lors de l'analyse, il est important de choisir une zone non-intéressante pour faire un ajustement du focus puis de faire une acquisition d'image sur la zone d'intérêt. Très souvent, comme les électrons ont une énergie cinétique importante, nous ne pouvons faire qu'une acquisition sur une zone de la grille. En effet, cette énergie est transférée à la glace qui fond et rend toute nouvelle observation dans cette zone impossible.

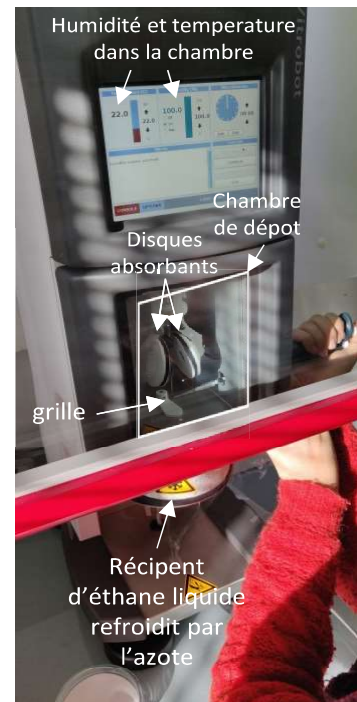


Figure 94 : Dépôt d'une grille de Cryo-TEM au Vitrobot.

## 7.1.10 Dichroïsme circulaire et linéaire

### 7.1.10.1 Préparation des cuves/cellules

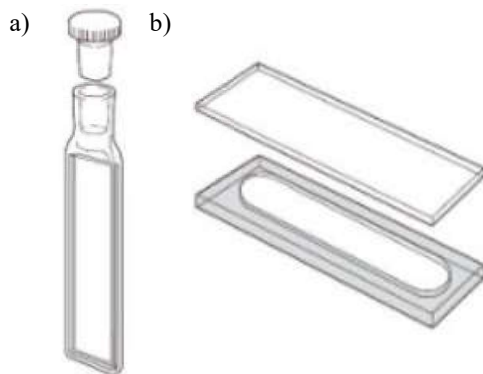


Figure 95 : Cuves utilisées a) de 300  $\mu\text{L}$ , 1 mm de trajet optique et b) 30  $\mu\text{L}$ , 0.1 mm de trajet optique.

En ce qui concerne le spectromètre de paillasse, deux types de cuves ont été utilisés, une fermée avec un bouchon, et une cuve/cellule dit « sandwich » fermée en collant deux lames. La première a été remplie avec 300  $\mu\text{L}$  de peptide dilué à 60  $\mu\text{M}$ . La seconde correspond à deux plaques, dont une creusée permettant de mettre 30  $\mu\text{L}$  de peptide 500  $\mu\text{M}$  qui est refermée avec une plaque en quartz (Figure 95). Les mêmes procédures ont été réalisées pour les tampons qui ont servi de blanc. Entre chaque mesure, les cuves ont été lavées avec, successivement, de l'eau ultrapure, de l'Hellmanex® à 2%, de l'eau ultrapure et de l'éthanol à 96%. Elles ont ensuite été séchées avec, pour la cuve de 300  $\mu\text{L}$  de l'air comprimé, et pour la cellule de 30  $\mu\text{L}$  du papier

kimtech. L'extérieur des cellules étant aussi lavées avec de l'éthanol et séché pour enlever les traces éventuelles.

Pour les mesures de spectres SR-CD et SR-LD au synchrotron, nous avons utilisé deux cuves/cellules avec des trajets optiques de 31  $\mu\text{m}$  pour les mesures de solutions et suspensions de fibres et 12  $\mu\text{m}$  pour les films (Figure 96). Le dépôt s'est fait comme pour la préparation des cuves pour le spectromètre de paillasse mais avec 2  $\mu\text{L}$  comme dépôt de liquide à des concentrations de 2 mM ou 500  $\mu\text{M}$ . Les cuves ne sont pas rectangulaires, mais circulaires. Pour le SR-LD, le film a été fait en faisant le vide dans la chambre où a été mis le côté creusé de la cuve, cela pendant 1 heure. Ensuite, ce côté de la cuve est mis dans le porte-échantillon, le film orienté vers le faisceau et l'analyse est lancée.

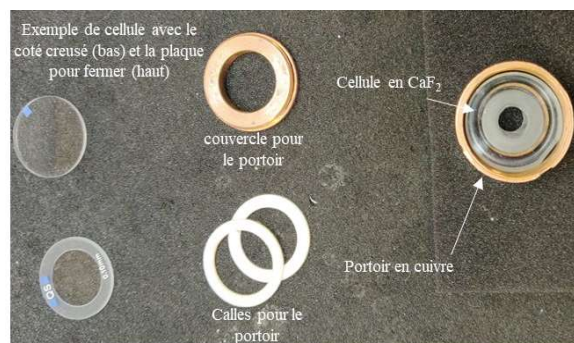


Figure 96 : Différentes parties composant le portoir afin de monter la cellule dans la chambre.

### 7.1.10.2 Acquisition et traitement des spectres

Le dichroïsme circulaire (CD) a été mesuré sur spectrophotomètre de paillasse à des longueurs d'onde allant de 200 à 300 nm avec des pas de 1 nm dans une chambre de mesure sous flux d'azote à température constante de 25 °C contrôlés par module Peltier. Les spectres ont été moyennés sur 10 scans afin d'avoir un

ratio signal sur bruit correct et soustrait de la ligne de base. Le logiciel utilisé pour la collecte de données et la moyenne est le logiciel fourni avec le spectromètre et la mise en forme réalisée avec le logiciel OriginPro.

Les spectres SR-CD (la ligne DISCO du synchrotron SOLEIL (St-Aubin, France)) ont été collectés avec un pas de 1 nm pour des longueurs d'onde de 170 à 255 nm à 25°C (sauf lors des analyses en températures) maintenue par module Peltier. Les cuves ont été placées dans un portoir en cuivre pour faciliter la conduction de la chaleur et la chambre de mesure a été mise sous un flux d'azote constant. L'iris présent sur le portoir a été ouvert à 7 mm de diamètre. Le logiciel CDToolX a permis de collecter et de normaliser les spectres, et le logiciel A205 a permis de déterminer la concentration de peptide lors de la mesure.<sup>[9-11]</sup> Les spectres (hors scans en température) ont été moyennés sur 3 scans, soustraits de la ligne de base, lissés, calibrés avec le CSA et normalisés en  $\Delta\epsilon$  avec la concentration à 205 nm. <sup>[12]</sup> Pour les cinétiques à température constante, les spectres ont été mesurés toutes les 3 minutes. Pour les suivis en rampe de température, les spectres ont été mesurés toutes les 5 min, avec des incréments de 3°C, et normalisés comme décrit précédemment. Le montage expérimental est montré en Figure 97.

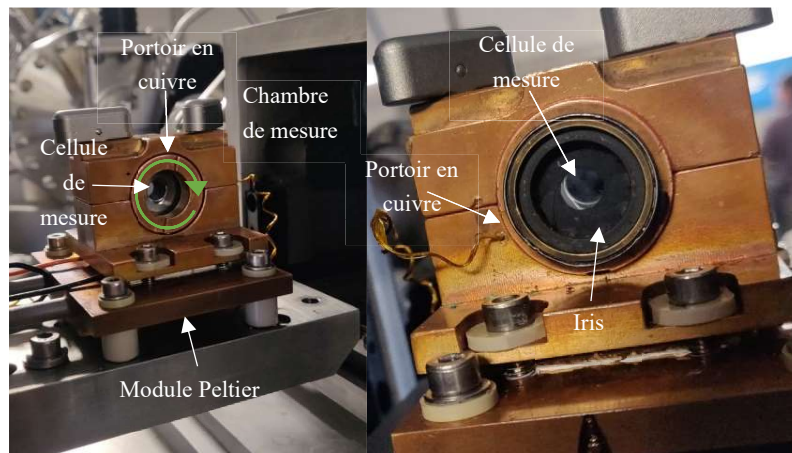


Figure 97 : Montage expérimental utilisé en SR-CD et SR-LD.

Le dichroïsme linéaire (LD) sur source synchrotron a été réalisé sur la ligne DISCO de la même manière que précédemment, en changeant de mode du CD vers le LD (switch), avec des longueurs d'onde allant de 145 à 255 pour les films et de 170 à 255 pour les solutions. L'iris présent sur le portoir a été ouvert à 7 et 12 mm de diamètre. Les spectres ont été moyennés sur 3 scans, soustraits de la ligne de base, lissés, calibrés avec le CSA mais sans normalisation. La rotation de la cuve a été réalisée à la main sur le module Peltier dans le sens perpendiculaire au faisceau.

Un schéma de fonctionnement de la ligne DISCO du synchrotron SOLEIL est montré en Figure 98.



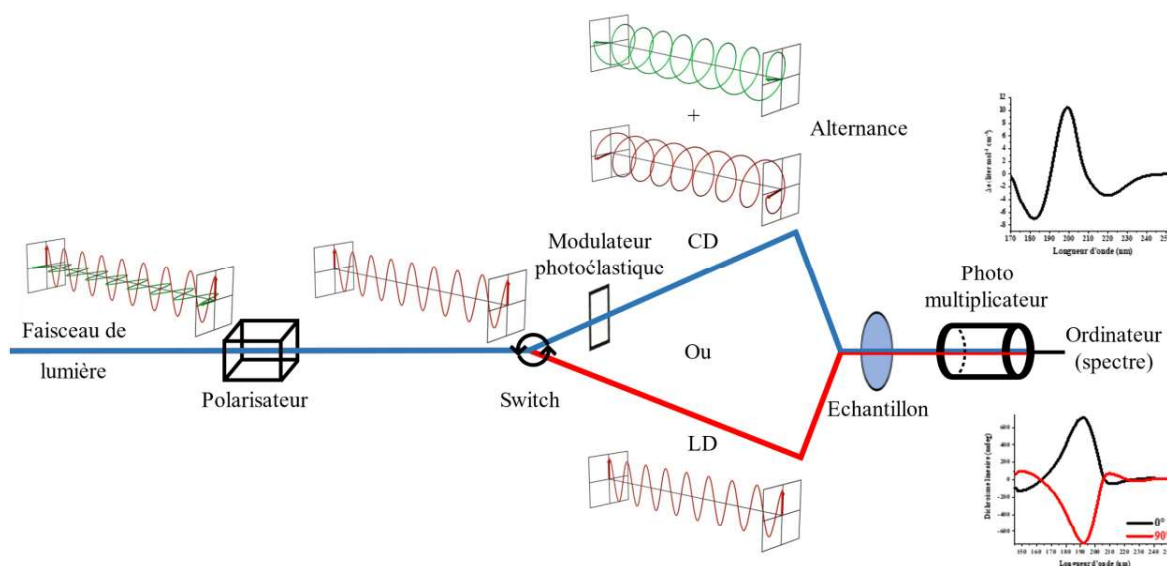


Figure 98 : Schéma de fonctionnement des expériences de CD et LD réalisées au synchrotron.

## 7.1.11 Spectroscopie infrarouge

### 7.1.11.1 Incubation et centrifugation des peptides

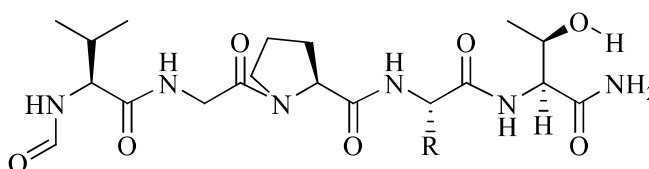
Les peptides, dont les contre-ions ont été échangés, ont été mis à incuber dans l'incubateur pour tube de la même manière que celle réalisée précédemment. Ensuite, les tubes contenant les suspensions ont été mis à centrifuger pendant 10 minutes à 5000 rpm à température ambiante.

### 7.1.11.2 Acquisition des spectres infrarouges

Les spectres IR acquis avec un module ATR ont été faits en déposant 1  $\mu$ L du culot de la suspension de fibres peptidiques précédemment centrifugée. La goutte est laissée sécher directement sur le module (diamant) pendant jusqu'à ce que la bande d'absorption de l'eau disparaisse et que le signal soit stable. Chaque spectre est une moyenne de 64 acquisitions.

## 7.2 Motif minimal permettant la formation d'une hélice $\beta$ gauche

Les peptides f-VGPXT-NH<sub>2</sub> (Figure 99) ont été synthétisés par synthèse peptidique en phase solide (SPPS) en suivant la stratégie Fmoc/tBu, manuellement, en utilisant une résine Fmoc-Rink Amide AM (Br-1330 Iris Biotech GMBH, 0.74 mmol/g, 100–200 mesh). Les synthèses ont été faites sur 0,4 mmol (540,5 mg de résine), avec X=L, N, D. Le peptide avec X=F avait déjà été synthétisé. D'autres dérivés avec des X=H, E, Q ont été synthétisés, purifiés selon les mêmes méthodes mais les résultats de cristallisation ne sont pas concluants, et donc ne sont pas montrés.



f V G P X T Am  
Figure 99 : Séquence générique des peptides conçus de 5 A.As (5m<sub>x</sub>) avec X = L, N, D, F.

La méthode de synthèse est la même que celle réalisée précédemment pour les peptides de 5 résidus s. La masse de résine a directement été pesée dans le réacteur de synthèse de 30 mL et après solvatation, elle est déprotégée et les Fmoc sont dosés.

Des double couplages (30 min) à température ambiante ont été réalisés pour tous les acides aminés couplés à l'aide de 4 équivalents de Fmoc-L-acide aminé (A.A), 3,8 équivalents de O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU) et 10 équivalents de N,N-Diisopropylethylamine (DIEA) à température ambiante sous agitation. Après chaque couplage d'A.A, une acétylation est réalisée sous agitation pendant 1 minute. Les étapes 1 à 3 ont été répétées autant de fois que nécessaire afin d'obtenir la séquence finale de 5 acides aminés. L'étape de formylation a été réalisée comme précédemment.

Les étapes suivantes comme le clivage, la lyophilisation et la purification sont fait comme décrit précédemment.

Après purification des poudres de peptides pures ayant des rendements de 49 à 75 % ont été obtenus (voir Tableau 19). Leurs analyses HPLC et LC-MS sont montrées en Annexe 5 à Annexe 8. La purification des peptides se faisant à 0.1 % de TFA, nous devons considérer, pour chaque peptide, le nombre de contre-ions TFA. Remarque, ici, un seul des peptides, le 5M<sub>D</sub>, a été purifié dans un tampon NH<sub>3</sub>/HCOOH à pH= 9. Après lyophilisation des fractions pures, les peptides sont stockés au congélateur à -18°C.

Tableau 19 : Tableau récapitulatif de synthèse et purification des 5 mers conçus.

Peptide	Masse Molaire (g/mol)	Condition purification; colonne	Condition HPLC Analytique; colonne	Nombre de mole initial ( $\mu$ mol)	Nombre de mole final ( $\mu$ mol)	Rendement Fmoc (%)	Masse peptide pur (mg)	Rendement (%)
5M <sub>L</sub>	512.61	Methode 1; e	Methode A; a	360.00	185.30	51.47	95.00	51.5
5M <sub>D</sub>	514.54 (531.57 avec NH <sub>4</sub> <sup>+</sup> )	Methode 1; d	Methode A; a	346.40	260.40	75.17	143.50	75

5M <sub>N</sub>	513.55	Methode 1; d	Methode A; a	351.00	172.00	49.00	88.30	47.76
-----------------	--------	--------------	--------------	--------	--------	-------	-------	-------

L'échange des contre-ions TFA<sup>(-)</sup> en Cl<sup>(-)</sup> a été réalisé comme précédemment en vue d'analyses en ATR-IR.

## 7.2.1 Cristallisation

### 7.2.1.1 Cristallisation en goutte assise

Deux types de plaques 96 puits ont été utilisées : *MRC 2 Well UVP (2 gouttes)* et *96-well COC Protein Crystallization Microplate (3 gouttes)* (Figure 100 a et b). Dans les deux cas, le même protocole a été suivi. Les peptides 5M<sub>L</sub>, 5M<sub>N</sub>, 5M<sub>D</sub> après purification et lyophilisation ont été dissouts afin d'obtenir des solutions stocks à 10 mg/mL à partir desquelles des solutions à 5 mg/mL (ou moins) ont été préparées par dilution. Le peptide 5M<sub>D</sub> a dû être préparé à pH environ 6 du fait de la présence d'acide aspartique, qui le rend insoluble à pH acide. Pour cela, le pH de la solution stock du peptide a été monté avec une solution de soude. Le peptide 5M<sub>F</sub> a été directement dissout à partir d'un stock de poudre déjà présent et des solutions ont été préparées comme pour les autres.

Les solutions stocks de tampons HEPES (pH=7.4) et d'acide citrique (pH=3.5) à 500 mM ont été préparées par dissolution de poudre dans de l'eau MilliQ et leur pH a été ajusté avec NaOH et HCl. À partir de ces stocks, des sous-stocks à 100 mM ont été préparés par dilution avec de l'eau ultra pure.

Une solution stock d'agarose à 2 % a été préparée par dissolution d'agarose à bas point de fusion dans un tube mis à incuber dans un thermomixer C tube à 1600 rpm, 60°C pendant 1 h. Ensuite des aliquots à 1 % et 0.66 % ont été préparés par dilution de cette dernière.

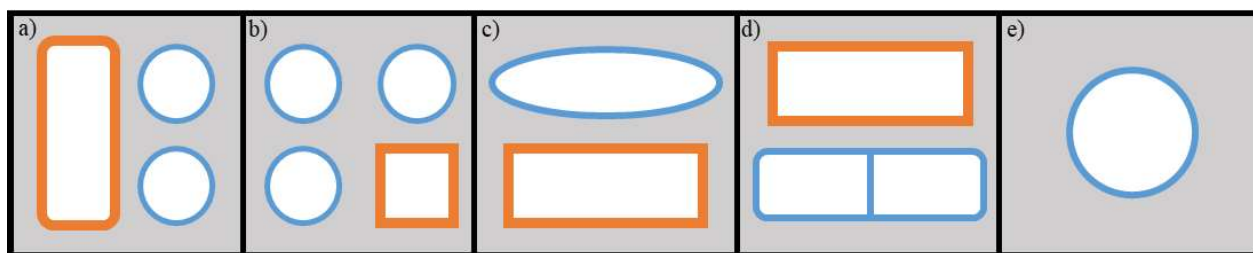


Figure 100 : Schéma des plaques utilisées avec en orange les réservoirs et en bleu les puits. A) MRC 2 Well UVP, b) 96-well COC Protein Crystallization Microplate, c) In Situ-1 Crystallization Plates, d) CrystalQuick X Plate et e) 60 well Greiner HLA Terasaki multiwell plates.

#### 7.2.1.1.1 Protocole

Une fois les solutions de peptide prêtes, nous avons déposé un volume de 52  $\mu$ L, pour la plaque *MRC 2 Well UVP* et 53  $\mu$ L pour la plaque *96-well COC Protein Crystallization Microplate* de la solution de screen (commercial ou préparé au laboratoire) dans le réservoir. Lors de l'utilisation de screens commerciaux, une

pipette multicanaux a été utilisée pour déposer plus rapidement les solutions. Ensuite 1  $\mu$ L des solutions des réservoirs ont été déposés à chaque emplacement de goutte, toujours à la pipette multicanaux. Les solutions stocks de peptides ont été transférées dans des plaques 96 puits classiques afin de pouvoir les prélever à la pipette multicanaux. Puis 1  $\mu$ L de solutions de peptides, à concentrations voulues ont été ajoutés dans les gouttes et mélangées à la solution de réservoir déjà présente par aspiration/dépôt avec la pipette. Lorsque cela était nécessaire, 1  $\mu$ L d'agarose à 37°C a été ajouté à la pipette et homogénéisé rapidement pour éviter la formation du gel trop rapidement.<sup>[13]</sup> L'ensemencement a été réalisé en trempant le l'outil d'ensemencement dans la suspension de nucléi préparée avec des cristaux broyés, utilisée comme cela ou diluée.<sup>[14]</sup> À l'œil, la présence de bulles a été vérifiée pour chaque puits, et percée si besoin avec une pointe de pipette scellée. Pour finir, le film imperméable a été collé sur la plaque et tous les puits ont été vérifiés pour être sûrs d'être fermés de manière étanche. Un appui pas trop important à la main avec un papier absorbant convient parfaitement. La plupart des plaques ont ensuite été placées dans un tiroir pour incuber à l'abri de la lumière et à  $T_{amb}$ . Ce n'est que récemment que le laboratoire s'est équipé d'un incubateur pour plaque avec température régulée à 30°C. Tous les jours pendant une semaine, la présence de cristaux a été vérifiée à la loupe binoculaire. Il est important, lors de cette étape d'être rapide afin d'éviter que la lampe ne chauffe de trop la plaque, ce qui peut altérer grandement la cristallisation et qui forme de la condensation sur le film, empêchant l'observation. Des photos ont été prises lorsque les cristaux ont commencé à pousser et lors du processus de cristallisation lorsque c'était possible. Lors du confinement, plusieurs plaques avaient été réalisées avant et n'ont pas pu être suivies pendant 2 mois, ce qui se traduit par une perte d'information au niveau du comportement des cristaux.

En vue de l'utilisation du Cribleur sur la ligne Proxima-2A du synchrotron SOLEIL, des cristaux ont aussi été cultivés en plaques mieux adaptées pour la DRX sur plaque comme les *In Situ-1 Crystallization Plates* et *CrystalQuick X Plates* qui sont faites pour la cristallisation et l'analyse In-situ des cristaux (Figure 100 c et d).

Le transport des plaques juste qu'au synchrotron, en bus et train, s'est fait en les emballant bien dans du papier bulle, et dans un sac glacière non réfrigéré. Le but ici est de les maintenir bien à plat, avec le minimum de perturbation pour ne pas risquer de redissoudre les cristaux. Il est arrivé que lors du transport de cristaux soit redissout mais ceci représente une minorité.

### 7.2.1.2 Cristallisation en microbatch

Ce type de cristallisation a été réalisé dans des plaques *60 well Greiner HLA Terasaki multiwell plates* (Figure 100 e). Seulement réalisé pour le peptide 5M<sub>L</sub>, f-VGPLT-Am, afin d'obtenir des cristaux plus gros. Les solutions stocks utilisées en peptide, sels et tampons l'agarose et les solutions de nucléi sont les mêmes que précédemment. À la différence de la cristallisation en équilibre de vapeur, ce type de méthode se réalise à partir de solution saturée sans équilibre de vapeur.<sup>[15]</sup>

### 7.2.1.2.1 Protocole

La préparation de ce type de plaque demande plus de quantité de peptide que pour la technique précédente, ce qui se traduit par un volume plus important. 3  $\mu\text{L}$  de solution stock de peptide a été mélangé à 3  $\mu\text{L}$  de solution de tampon et sel afin d'avoir un mélange 1:1 comme précédemment. Les 6  $\mu\text{L}$  ont ensuite été homogénéisés à la pipette etensemencés si besoin. Des dilutions en cascade de nucléi ont aussi été réalisées en partant de la solution stock puis en la diluant 10 fois, puis encore 10 fois, cela réalisé 5 fois afin d'obtenir une suspension 10000 fois moins concentrée en nucléi que la première. À partir de stock ou des dilutions en cascade réalisées, l'ensemencement en série a été fait en trempant le l'outil d'ensemencement dans le puits contenant les nucléi puis glissé dans un puits contenant la solution à cristalliser. Ensuite, celui-ci a été directement glissé dans le puits du dessous et ainsi de suite (Figure 78). Lorsque l'ajout d'agarose a été réalisé, celui-ci a été fait de la même manière que précédemment, mais ici, la solution de peptide et sel a été de 4  $\mu\text{L}$  pour 2  $\mu\text{L}$  d'agarose. Pour finir, 4  $\mu\text{L}$  d'huile de paraffine a été ajouté afin de couvrir le puits et éviter toute évaporation qui sécherait la condition et ferait cristalliser le sel en priorité. Enfin, la plaque a été mise à incuber comme précédemment et vérifiée tous les jours Figure 79.

### 7.2.1.3 Cristallisation en Tube

#### 7.2.1.3.1 Pour obtenir des cristaux en quantité

Les solutions utilisées sont les mêmes que précédemment. La cristallisation a été réalisée en tube de 1 mL en équilibre de vapeur dans une cuve CCM avec un fond d'environ 1-2 cm de solution saturée en sulfate d'ammonium (Figure 101).

La préparation de la cuve CCM pour la pousse des cristaux s'est faite en ajoutant environ 5 à 10 mL d'eau à environ 2 grosses spatules de sulfate d'ammonium. Le tout a été agité puis laissé au bain à ultrasons pendant 30 minutes pour s'assurer que la solution préparée est encore saturée comme l'atteste la présence de cristaux après sonication. Il est important aussi de graisser la connexion entre le couvercle et la cuve pour ne pas la bloquer (sinon, une goutte de solution saline peut sécher là, scellant le couvercle). Cela a été fait avec de la graisse à vide classique. Un support pour tube pouvant soutenir jusqu'à 4 tubes a été fabriqué avec du polystyrène. Cela permet de maintenir droits les tubes dans la cuve.

Les tubes de cristallisation ont été préparés en mélangeant le stock de peptide avec les sels et tampons afin d'obtenir les concentrations reportées en Tableau 15. Puis ont étéensemencés comme décrit précédemment, et mis en cuve, dans le porte-tube.

La cuve a ensuite été fermée et mise à incuber dans l'incubateur à 30 °C. L'apparition des cristaux a été suivie à l'œil pour éviter d'ouvrir la cuve et perturber le processus. Après l'apparition des cristaux, en général 2

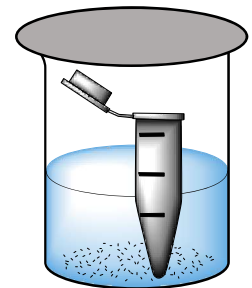


Figure 101 : montage expérimental de cristallisation en tube avec le  $(\text{NH}_4)_2\text{SO}_4$  saturé en bleu, la cuve étant scellée avec le couvercle.

semaines (Figure 81), les tubes ont été sortis et refermés pour stopper le processus de concentration du tube (par évaporation) puis gardés dans l'incubateur pour réaliser, par la suite, la diffraction des rayons X dessus.

Lorsque l'analyse ATR-IR ou RMN du solide a été réalisée, les tubes ont été centrifugés, le surnageant retiré et lavé 3 fois avec de l'eau ultra-pure glacée afin d'enlever le sel tout en évitant la solubilisation trop importante des cristaux. Après ceux-ci ont été séchés par capillarité avec du papier absorbant. Les cristaux « secs » ont ensuite été mis au frigo à 4°C dans l'attente d'être analysés.

### 7.2.1.3.2 Pour obtenir des cristaux de qualité

Le même protocole est repris, mais cela se fait sur de plus petites quantités. Des solutions (500 µL) où les cristaux étaient les meilleurs (par cristallisation sur plaque en goutte assise) ont été mises à incuber dans des tubes de 750 µL. Les solutions correspondantes aux réservoirs (1 mL) ont été mises dans un pilulier de 4 mL dans lequel le tube est ouvert ; il a ensuite été refermé et scellé avec du parafilm. Cela a seulement été fait pour le peptide le plus étudié, f-VGPLT-Am (5M<sub>L</sub>). Comme précédemment dit, une solution de peptide est mélangée à une solution de sels + tampon dans l'éppendorf pour obtenir les concentrations détaillées en Tableau 15. L'agarose a été ajoutée chaud à la solution précédente, et l'ensemencement réalisé selon le protocole déjà décrit. La surveillance d'apparition des cristaux peut se faire directement dans le pilulier passé sous la loupe binoculaire. Après les temps de cristallisation reportés en Tableau 15, correspondant à l'apparition des premiers cristaux, les tubes ont été sortis et refermés afin de « stopper » ou ralentir la croissance cristalline. Les tubes ont ensuite été conservés à l'incubateur à 30°C en attendant d'être pêchés pour réaliser de la diffraction des rayons X.

## 7.2.2 Diffraction des Rayons X

### 7.2.2.1 Méthodes

La diffraction *in-situ* a été réalisée grâce au montage expérimental mis en place par D. Jeangerard et W. Shepard permettant de mesurer la diffraction des rayons X de cristaux directement dans la plaque de cristallisation sans devoir les pêcher et les congeler. Il sert aussi à pré-sélectionner les cristaux de bonne qualité, en utilisant le « grid mode » ou en tirant directement sur des cristaux. Ce montage est le CRIBLEUR, montré Figure 75. La plaque multipuits est montée côté film vers le détecteur sur le portoir.<sup>[16]</sup>

Les cristaux ont été pêchés avec des MicroMounts™ et MicroLoops™. Le cristal est pêché directement dans la goutte de cristallisation ou transféré dans un puits plus large (en ce qui concerne la cristallisation tube). Des cristaux ont directement été placés sous le faisceau de rayons X et analysés mais la diffraction n'était pas optimale. D'autres ont été congelés en les trempant dans une solution de Cryoprotection, pendant 15-30s, qui correspond à la condition de cristallisation + 20 % de glycérol afin d'éviter l'explosion, due à la glace du cristal.

Ensuite, ils ont été congelés dans l'azote liquide en les trempant dedans puis en les gardant dans un puck stocké dans l'azote liquide. La diffraction sur boucle s'est faite sur un montage avec un goniomètre classique.<sup>[17]</sup>

### 7.2.2.2 Collecte et traitement des données

Le logiciel CrysAlisPro a été utilisé pour la collecte de donnée, le réajustement de la maille cristalline et la réduction des données. La structure a été résolue par méthode directe avec le logiciel Olex et affinée par la matrice des moindres carrés avec le même logiciel.

Le logiciel XDS a été utilisé pour l'ajustement de la maille cristalline et l'intégration. La mise à l'échelle et la réduction des données a été réalisées via AutoPROC. La méthode directe avec SHELXT a permis de résoudre la structure, l'affinement par la matrice des moindres carrés a été réalisé avec SHELXL et OLEX utilisé comme interface graphique. Les paramètres de position et d'anisotropie thermique de tous les atomes non-hydrogènes ont été affinés. Les atomes d'hydrogène étant localisés dans une carte de Fourier, ils ont pu être repositionnés géométriquement. Les détails expérimentaux sur les diffractions cristallographiques sont résumés en Annexe 9.

Le CRIBLEUR permet la collecte de diffraction RX de cristaux directement dans la plaque mutlipuits. Il est possible de collecter à différents angles, mais seulement de  $\pm 45^\circ$ . Dans notre cas, cela n'a pas suffi à résoudre la structure, à la fois à cause de la mosaïcité de nos cristaux mais aussi car ils étaient de symétrie P1.

En ce qui concerne la collecte des données sur le goniomètre classique, en mode Cryo ou non, une fois la boucle montée, un bras robotique permet de bouger le cristal comme souhaité et de l'orienter dans toutes les directions. Il faut alors centrer la rotation sur l'endroit où on veut effectuer la rotation, et ensuite effectuer une collecte en rotation autour du cristal. Cela permet d'obtenir des informations dans toutes les directions de l'espace et de remonter, si les informations sont suffisantes à la maille cristalline, mais aussi à la structure.

Le traitement des données cristallographiques a été réalisé par nos collaborateurs de l'IBMC.

## 7.2.3 RMN du solide

### 7.2.3.1 Préparation de l'expérience RMN du solide

L'analyse des cristaux a été réalisée en collaboration avec P. Bertani de l'équipe biophysique des membranes et RMN. Les cristaux, après avoir été lavés et séchés, ont été transférés dans le rotor RMN. Les cristaux ont été transférés avec une petite spatule et tassés avec le « sample packer » adapté au rotor. En tout, environ 20 mg de cristaux de peptides ont été transférés afin de pouvoir remplir le rotor. Une fois rempli avec les cristaux, celui-ci a été scellé avec les bouchons adaptés. Le rotor a ensuite été placé dans le spectromètre RMN.

### 7.2.3.2 Méthodes d'analyse en RMN du solide

1H-13C 1D CP-MAS : Les expériences 1D ont été enregistrées en utilisant un temps de contact de 2ms avec une modulation tangentielle sur chaque fréquence. Le pulse du proton à 90° a été de 3 µs et la fréquence du champ de découplage spinal de 80 kHz.

1H-13C 2D CP-MAS : L'expérience 1H-29Si Frequency switch Lee-Goldburg HETeronuclear chemical shift CORrelation (FSLG HETCOR) a été acquise dans les mêmes conditions que précédemment en utilisant la FSLG pour améliorer la résolution dans la dimension du proton. 64 scans ont été enregistrés pour chaque expérience à 256 t1.

## 7.3 Références bibliographiques

- [1] Schrödinger, LLC, *The {PyMOL} Molecular Graphics System, Version~1.8*, **2015**.
- [2] M. Amblard, J.-A. A. Fehrentz, J. Martinez, G. Subra, *Mol. Biotechnol.* **2006**, *33*, 239–254.
- [3] J. Alsina, F. Albericio, *Biopolymers* **2003**, *71*, 454–477.
- [4] W. S. Newcomb, T. L. Deegan, W. Miller, J. A. Porco, *Biotechnol. Bioeng.* **1998**, *61*, 55–60.
- [5] J. Cornish, K. E. Callon, C. Q. X. Lin, C. L. Xiao, T. B. Mulvey, G. J. S. Cooper, I. R. Reid, *Am. J. Physiol. - Endocrinol. Metab.* **1999**, *277*, 779–783.
- [6] V. V. Andrushchenko, H. J. Vogel, E. J. Prenner, *J. Pept. Sci.* **2007**, *13*, 37–43.
- [7] M. Groenning, L. Olsen, M. van de Weert, J. M. Flink, S. Frokjaer, F. S. Jørgensen, *J. Struct. Biol.* **2007**, *158*, 358–369.
- [8] M. J. Dobro, L. A. Melanson, G. J. Jensen, A. W. McDowall, *Plunge Freezing for Electron Cryomicroscopy*, Elsevier Inc., **2010**.
- [9] N. J. Anthis, G. M. Clore, *Protein Sci.* **2013**, *22*, 851–858.
- [10] J. Sutherland, in (Ed.: G.D. Fasman), Springer US, Boston, MA, **1996**, pp. 25–67.
- [11] J. G. Lees, B. R. Smith, F. Wien, A. J. Miles, B. A. Wallace, *Anal. Biochem.* **2004**, *332*, 285–289.
- [12] A. J. Miles, F. Wien, J. G. Lees, A. Rodger, R. W. Janes, B. A. Wallace, *Spectroscopy* **2003**, *17*, 653–661.
- [13] B. Lorber, C. Sauter, A. Théobald-Dietrich, A. Moreno, P. Schellenberger, M. C. Robert, B. Capelle, S. Sanglier, N. Potier, R. Giegé, *Prog. Biophys. Mol. Biol.* **2009**, *101*, 13–25.
- [14] T. Bergfors, *J. Struct. Biol.* **2003**, *142*, 66–76.
- [15] A. D'Arcy, C. Elmore, M. Stihle, J. E. Johnston, *J. Cryst. Growth* **1996**, *168*, 175–180.
- [16] D. Jeangerard, M. Savko, L. Ciccone, K. Desjardins, A. Le Jollec, A. Haouz, W. Shepard, *Proc. MEDSI'18* **2018**, 283–285.
- [17] L. Jacquamet, J. Ohana, J. Joly, F. Borel, M. Pirocchi, P. Charrault, A. Bertoni, P. Israel-Gouy, P. Carpentier, F. Kozielski, et al., *Structure* **2004**, *12*, 1219–1225.



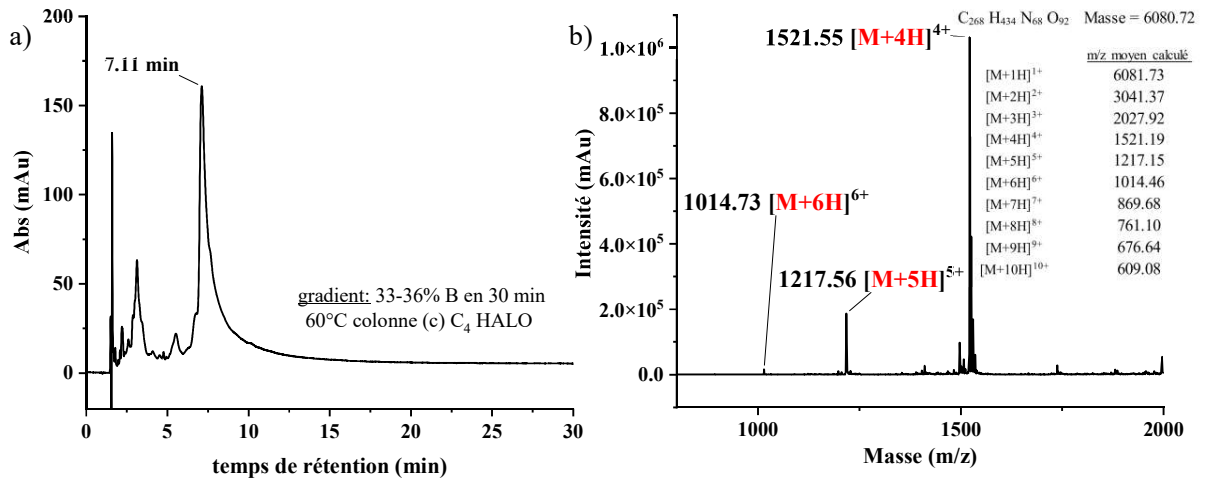


## Annexes

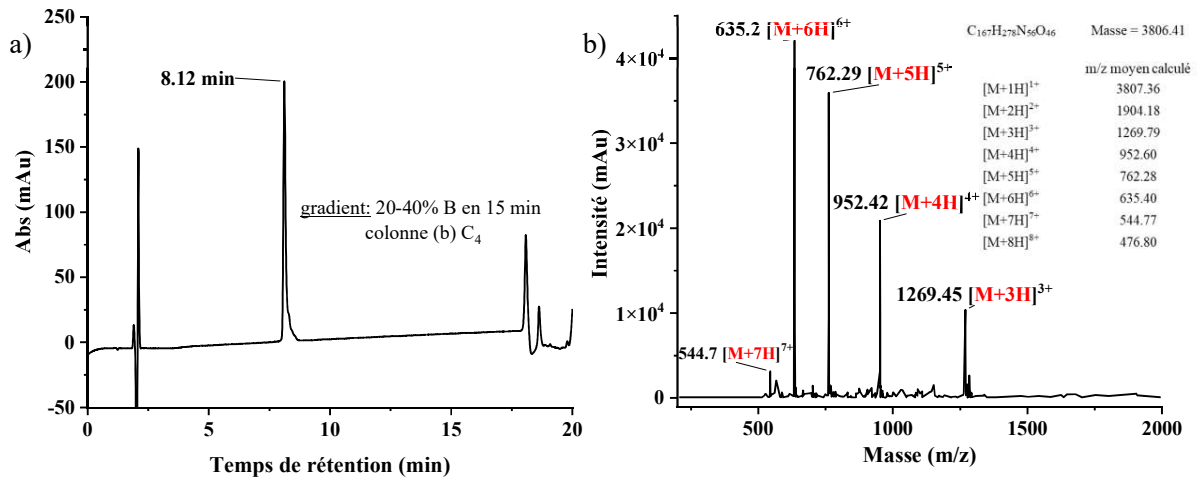
Annexe 1 : Liste des structures cristallographiques étudiées classées par code PDB.

Code PDB	Topologie (principalement $\beta$ ) CATH	Famille	Espèce (organisme)	Nb de tours	Resolution (Å)
1FWY	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Escherichia coli	4	2.3
1G97	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Streptococcus pneumoniae	10	1.96
1HM0	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Streptococcus pneumoniae	10	2.3
1JZ2	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Helicobacter pylori	9	2.1
1KRR	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Escherichia coli	6	2.5
1LXA	UDP N-Acetylglucosamine Acyltransferase	ACYLTRANSFERASE	Escherichia coli K-12	10	2.6
1OCX	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Escherichia coli	6	2.15
1SSM	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Haemophilus influenzae	5	2.15
1T3D	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Escherichia coli	5	2.2
1V3W	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Pyrococcus horikoshii OT3	7	1.5
1XHD	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Bacillus cereus ATCC 14579	7	1.9
2GGO	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Sulfurisphaera tokodaii	8	1.8
2IU8	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Chlamydia trachomatis	10	2.2
2P2O	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Geobacillus kaustophilus HTA426	6	1.74
2RIJ	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Campylobacter jejuni	6	1.9
2V0J	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Haemophilus influenzae	10	2
2XAT	UDP N-Acetylglucosamine Acyltransferase	ACETYLTRANSFERASE	Pseudomonas aeruginosa	5	3.2
3BSY	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Campylobacter jejuni	6	1.8
3BXY	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Escherichia coli O157:H7	6	2
3CJ8	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Enterococcus faecalis	6	1.95
3DHO	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Enterococcus faecium	5	1.8
3EEV	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Brucella abortus 2313	5	2.61
3EG4	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Brucella suis	6	1.87
3F1X	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Bacteroides vulgatus ATCC 8482	5	2
3FS8	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Thermoanaerobacterium thermosaccharolyticum	11	1.7
3FSX	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Mycobacterium tuberculosis H37Rv	6	2.15
3GOS	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Yersinia pestis	6	1.8
3GVD	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Yersinia pestis	5	2.4
3HSQ	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Leptospira interrogans	9	2.1
3IGJ	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Bacillus anthracis	6	2.6
3IXC	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Anaplasma phagocytophilum str. HZ	7	1.61
3KWE	UDP N-Acetylglucosamine Acyltransferase	LYASE	Thermosynechococcus elongatus BP-1	7	1.1
3MC4	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Brucella abortus 2308	5	1.95
3MQH	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Bordetella petrii DSM 12804	7	1.43
3OTM	UDP N-Acetylglucosamine Acyltransferase	LYASE	Methanosarcina thermophila	7	1.5
3PMO	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Pseudomonas aeruginosa	10	1.3
3Q1X	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Entamoeba histolytica	5	1.59
3R0S	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Campylobacter jejuni subsp. jejuni NCTC 11168 = ATCC 700819	9	2.3
3R3R	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	Salmonella enterica subsp. enterica serovar Typhimurium	7	1.2

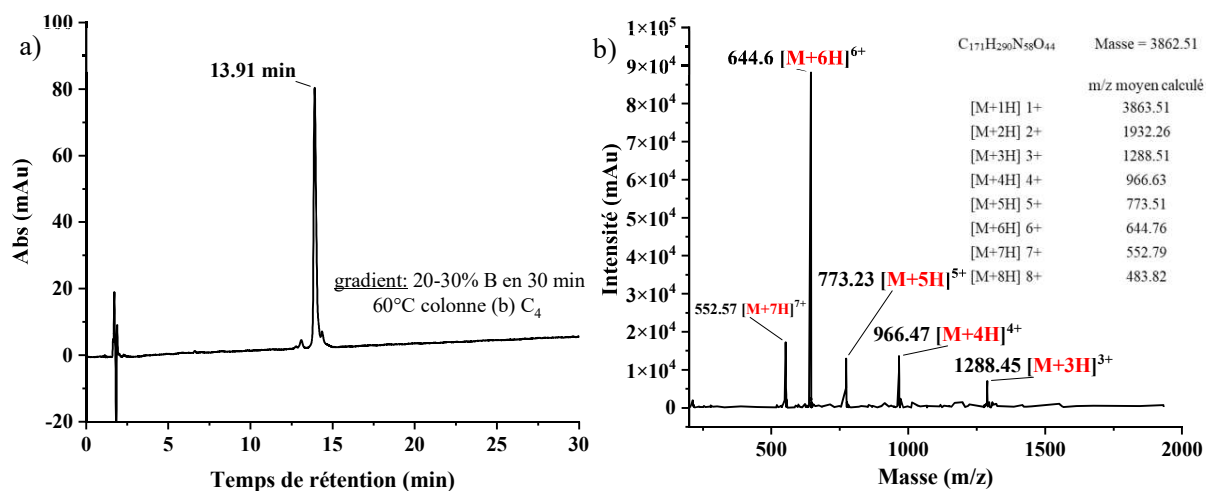
Code PDB	Topologie (principalement $\beta$ ) CATH	Famille	Espèce (organisme)	Nb de tours	Resolution (Å)
3R5D	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Pseudomonas aeruginosa</i>	6	1.8
3R8Y	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Bacillus anthracis</i>	6	1.7
3SRT	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Clostridioides difficile</i> 630	6	2.5
3T57	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Arabidopsis thaliana</i>	10	2.1
3TDT	UDP N-Acetylglucosamine Acyltransferase	ACYLTRANSFERASE	<i>Mycobacterium tuberculosis</i> variant bovis	6	2
3TIO	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Escherichia coli</i> K-12	7	1.41
3TK8	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	3TK11Burkholderia pseudomallei 1710b	6	1.8
3VBP	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Bacillus cereus</i> SJ1	7	2.3
3VNP	UDP N-Acetylglucosamine Acyltransferase	STRUCTURAL GENOMICS, gamma-carbonic anhydrase-like	<i>Geobacillus kaustophilus</i> HTA426	7	2.4
4E79	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Acinetobacter baumannii</i>	10	2.66
4E6U	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Acinetobacter baumannii</i>	9	1.41
4EA9	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Caulobacter vibrioides</i>	6	0.9
4EQY	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Burkholderia thailandensis</i> E264	9	1.8
4FCE	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Yersinia pestis</i> CO92	10	1.96
4H7O	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Vibrio cholerae</i> O1 biovar El Tor str. N16961	5	2.17
4HHF	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE/LIPID BINDING PROTEIN	<i>Escherichia coli</i> K-12, <i>Escherichia coli</i> str. 'clone D i14	10	2.1
4M98	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Neisseria gonorrhoeae</i> FA 1090	6	1.67
4M9C	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Acinetobacter baumannii</i>	6	2.1
4MFG	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Clostridioides difficile</i> 630	7	2
4MZU	UDP N-Acetylglucosamine Acyltransferase	ISOMERASE, TRANSFERASE	<i>Shewanella denitrificans</i> OS217	7	2.2
4N27	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Brucella abortus</i> 2308	7	2.73
4N6A	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Glycine max</i>	5	1.75
4R36	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Bacteroides fragilis</i> NCTC 9343	9	1.9
5B04	Initiation factor	TRANSLATION	<i>Schizosaccharomyces pombe</i> 972h-	6	2.99
5DEP	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Pseudomonas aeruginosa</i> PA7	9	2.16
5E3R	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Corynebacterium glutamicum</i> ATCC 13032	6	1.85
5F42	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Francisella tularensis</i> subsp. novicida U112	9	2.06
5JXX	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Moraxella catarrhalis</i> BBH18	9	3
5U2K	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Staphylococcus aureus</i> subsp. aureus COL	5	1.38
5UX9	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Aliivibrio fischeri</i> ES114	5	2.7
5VMK	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Acinetobacter baumannii</i>	10	2.55
6AMZ	Acyltransferase	TRANSFERASE	<i>Acinetobacter baumannii</i>	6	2.5
6CKT	Acyltransferase	TRANSFERASE	<i>Legionella pneumophila</i> subsp. pneumophila str. Philadelphia 1	6	1.8
6GE9	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Mycobacterium tuberculosis</i> H37Ra	10	2.26
6IVE	UDP N-Acetylglucosamine Acyltransferase	METAL BINDING PROTEIN	<i>Thermus thermophilus</i> HB8	7	2.3
6JVU	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Klebsiella pneumoniae</i> MGH 78578	5	2.8
6MFK	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Elizabethkingia anophelis</i>	5	1.65
6OSS	UDP N-Acetylglucosamine Acyltransferase	TRANSFERASE	<i>Proteus mirabilis</i> HI4320	9	2.19
6SC4	UDP N-Acetylglucosamine Acyltransferase	METAL BINDING PROTEIN	6sc7 candidate division MSBL1 archaeon SCGC-AAA259109	7	2.6



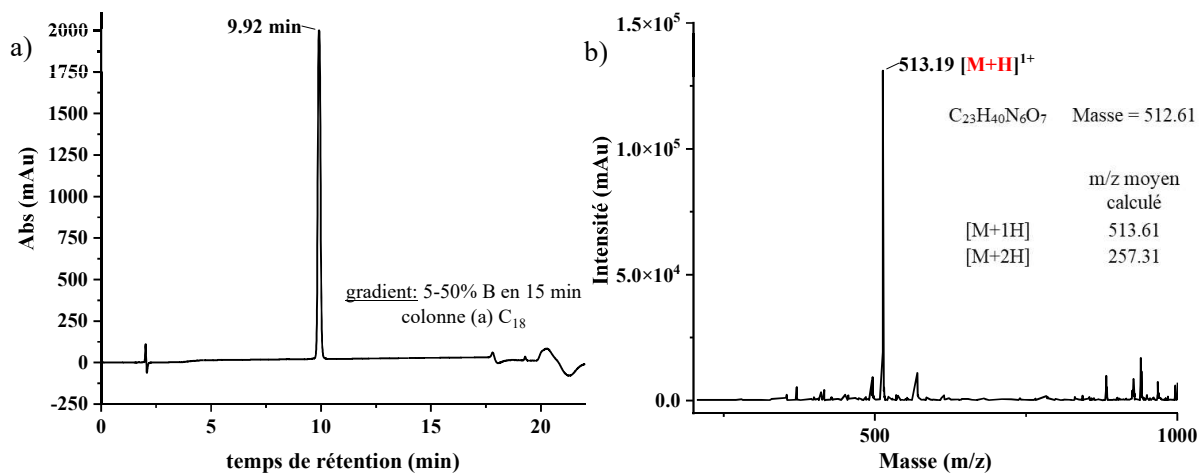
Annexe 2 : a) Chromatogramme HPLC du peptide 59m purifié une fois, et b) LCMS montrant que le peptide est présent.



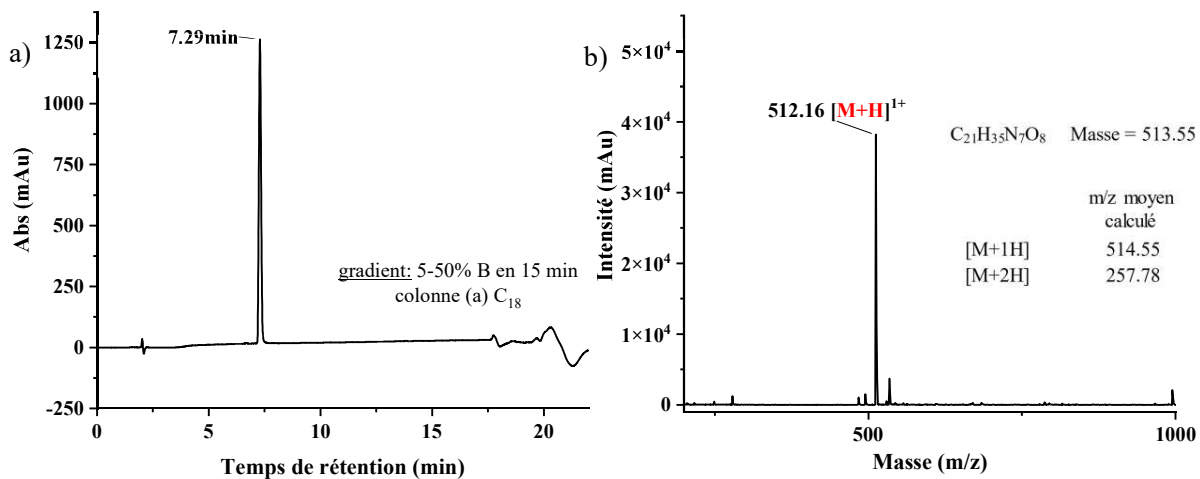
Annexe 3 : a) Chromatogramme HPLC du peptide 35m1 pur, et b) analyse LCMS correspondante.



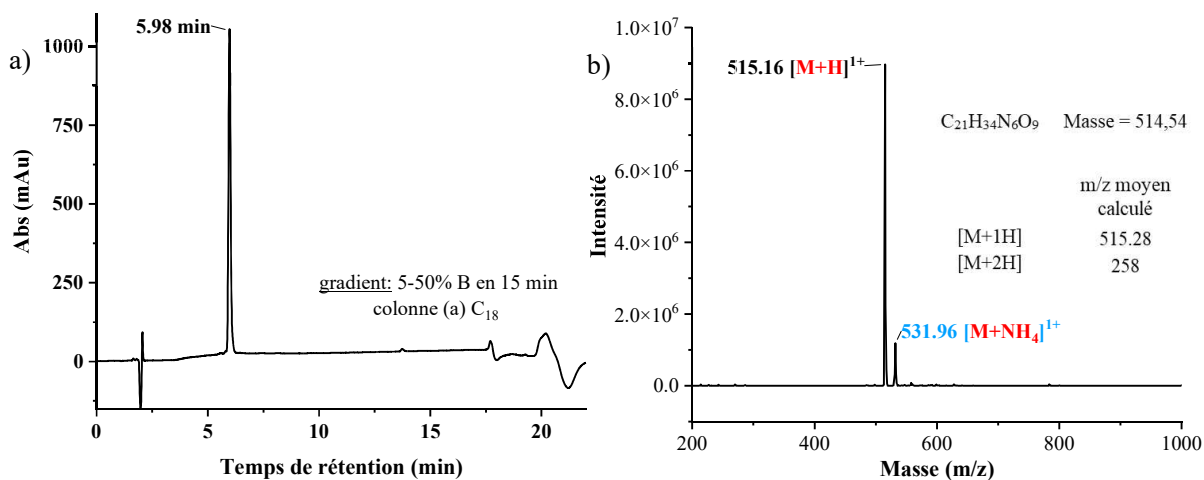
Annexe 4 : a) Chromatogramme HPLC du peptide 35m<sub>2</sub> pur, et b) analyse LCMS correspondante.



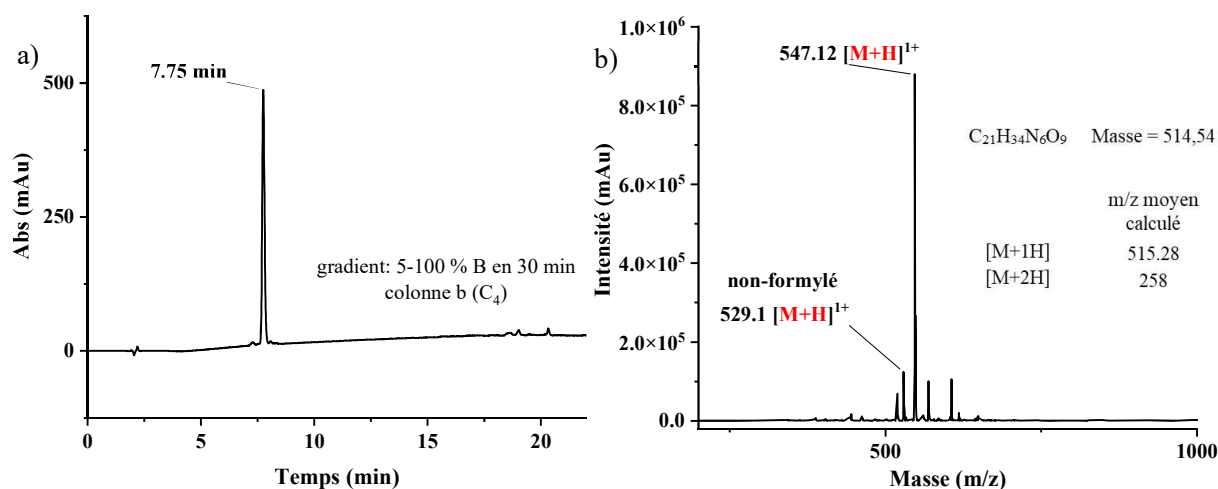
Annexe 5 : a) Chromatogramme HPLC du peptide 5m<sub>L</sub> pur, et b) analyse LCMS correspondante.



Annexe 6 : a) Chromatogramme HPLC du peptide 5m<sub>N</sub> pur, et b) analyse LCMS correspondante.



Annexe 7 : a) Chromatogramme HPLC du peptide 5m<sub>D</sub> pur, et b) analyse LCMS correspondante.



Annexe 8 : a) Chromatogramme HPLC du peptide 5m<sub>F</sub> pur, et b) analyse LCMS correspondante.

Annexe 9 : Tableau résumant les différents paramètres cristallographiques des deux peptides 5M<sub>L</sub> et 5M<sub>N</sub>.

Paramètre/ Peptide	f-VGPLT-Am (5M <sub>L</sub> )	f-VGPNT-Am (5M <sub>N</sub> )
Formule brute	C <sub>48</sub> H <sub>58</sub> N <sub>10</sub> O <sub>14</sub>	C <sub>48</sub> H <sub>58</sub> N <sub>10</sub> O <sub>14</sub>
Masse molaire, g/mol	512.30	513,55
Dimensions cristallines, mm	0.006 x 0.06 x 0.030	
Système cristallin	Triclinic	Monoclinic
Groupe d'espace	<i>P</i> 1	<i>P</i> 2 <sub>1</sub>
Dimensions de la cellule unitaire :		
a, Å	4.833985	4.805
b, Å	16.159267	13.471
c, Å	19.291795	19.849
α, °	66.1567	90.0
β, °	85.3626	94.011
γ, °	87.2601	90.0
T, K	100	100
Réflexions indépendantes	4154	1260
Limites de résolution	17.61 – 0.95 (0.98 – 0.95)	19.80 – 0.87 (0.98-0.87)
Réflexions observées <i>I</i> > 2σ( <i>I</i> )	2989	1241
Paramètres	676	327
R <sub>1</sub>	0.0957	0.0535
wR <sub>2</sub>	0.1741	0.1361



**Maxime NAUDÉ**  
**Conception de peptides *de novo***  
**s'autoassemblant en hélice  $\beta$  gauche**

## Résumé

Les métalloenzymes artificielles représentent une option viable pour concevoir des catalyseurs combinant les avantages de la catalyse homogène et enzymatique afin de remplacer les catalyseurs utilisés de nos jours, souvent à base de métaux rares et issus de la pétrochimie. Pour se faire, une des manières est de concevoir rationnellement de petites unités peptidiques, appelées briques, s'autoassemblant pour former des fibres stables et robustes. Un des repliements identifiés pour cette application est l'hélice  $\beta$  gauche de type I (L $\beta$ H)(e.g. plusieurs exemples avec des métaux). Cette thèse présente l'analyse de séquences naturelles contenant des L $\beta$ Hs mettant en évidence les résidus importants pour le repliement. Ensuite, à partir de cette analyse, des peptides formant des fibres très susceptibles de se replier en L $\beta$ Hs ont été conçus et analysés par spectroscopies IR et CD, fluorescence, TEM, modélisation... tout étant en accord avec le repliement attendu. Nous nous sommes également intéressés à la conception d'un motif minimal pouvant permettre la formation de mimes de L $\beta$ Hs étudiés par DRX. Malheureusement, les structures ne sont pas en accord avec le repliement attendu.

**Mots clés:** Conception rationnelle, métalloenzymes artificielles, peptides, autoassemblages, fibres, cristaux

## Résumé en anglais

Artificial metalloenzymes represent a viable option to design catalysts combining the advantages of homogeneous and enzymatic catalysis to replace today's rare metal and petrochemical based catalysts. One way to do this is to rationally design small peptide units, called bricks, which self-assemble to form stable and robust fibers. One of the folds identified and promising for this application is the type I left-handed  $\beta$  helix (L $\beta$ H) (e.g. several examples with metals). This thesis presents the analysis of natural sequences containing L $\beta$ Hs highlighting important residues for folding. Then, based on this analysis, fiber-forming peptides highly susceptible to fold into L $\beta$ Hs were designed and analysed by IR and CD spectroscopies, fluorescence, TEM, modeling... all in agreement with the expected folding. We were also interested in the design of a minimal motif allowing the formation of L $\beta$ Hs mimics that could be studied by XRD. Unfortunately, the structures are not in agreement with the expected folding.

**Keywords:** Rational design, artificial metalloenzymes, peptides, self-assembly, fibers, crystals