



**HAL**  
open science

# Synthèse et étude structurale multi-échelle de peptides mimes de collagène

Anaïs Terrien

► **To cite this version:**

Anaïs Terrien. Synthèse et étude structurale multi-échelle de peptides mimes de collagène. Biophysique. Université Pierre et Marie Curie - Paris VI, 2015. Français. NNT : 2015PA066647 . tel-01359905

**HAL Id: tel-01359905**

**<https://theses.hal.science/tel-01359905>**

Submitted on 5 Sep 2016

**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.

## UNIVERSITÉ PIERRE ET MARIE CURIE

ÉCOLE DOCTORALE: Chimie physique et chimie analytique de Paris Centre

SPÉCIALITÉ : Biophysique

Présentée par

**Anaïs TERRIEN**

POUR OBTENIR LE GRADE DE

DOCTEUR DE L'UNIVERSITE DE PIERRE ET MARIE CURIE

Sujet de la thèse :

# Synthèse et étude structurale multi-échelle de peptides mimes de collagène

Dirigée par Emeric Miclet

Présentée et soutenue publiquement le 17 décembre 2015

Devant un jury composé de :

Mme Florine CAVELIER	Directeur de Recherche, CNRS, Montpellier	Rapporteur
Mme Céline LANDON	Chargée de Recherche, CNRS, Orléans	Rapporteur
Mme Maïté PATERNOSTRE	Directeur de Recherche, CNRS, Paris Saclay	Examineur
Mr Xavier PANNECOUCKE	Professeur des universités, Rouen	Examineur
Mr Thierry AZAIS	Maitre de conférences, Paris 6	Examineur
Mr Emeric MICLET	Maitre de conférences, Paris 6	Directeur de thèse



*A mon papa,*



## Remerciements

Et voilà, nous y sommes ... après 3 années de travail, je m'attaque à la partie qui me paraît la plus difficile de la rédaction de thèse, celle des remerciements ! Eh oui ... ce travail ne serait pas le même et n'aurait pas abouti sans l'aide de toutes les personnes citées ci-dessous. Je m'excuse par avance auprès des personnes que j'aurais oubliées.

Les premières personnes que je vais remercier sont les membres du jury. Je remercie Céline Landon et Florine Cavelier d'avoir accepté d'être rapporteurs de cette thèse. Je vous remercie pour votre lecture attentive et vos remarques précieuses. Je remercie Xavier Pannecoucke d'avoir accepté d'être président du jury. Je remercie également Thierry Azaïs et Maïté Paternostre d'avoir participé à ce jury. Je vous remercie tous pour la discussion enrichissante qui a suivi la présentation de ce travail.

Je tiens à remercier la personne qui m'a permis de montrer mes capacités à mener un travail de recherche, mon directeur de thèse, Emeric Miclet. Je suis ravie d'avoir pu partager ta passion pour la recherche et garde d'excellents souvenirs de ces 3 années. Années que j'ai passées au sein d'une équipe nommée « Structure et Dynamique des Biomolécules », je remercie donc tous les membres de cette équipe et plus particulièrement Olivier Lequin, Jean-Jacques Lacapère, Ludovic Carlier (merci pour ta relecture de dernière minute de mon résumé de thèse), et Lucie Khemtémourian pour votre aide et les discussions animées qui ont eu lieu lors de nos « Lunch Meeting ». Bien que l'établissement de ces réunions ait été très controversé, elles m'ont beaucoup aidée à mener à bien ce projet. Cette équipe fait partie du Laboratoire des BioMolécules, laboratoire dans lequel cohabitent des personnes aux compétences variées et complémentaires et de ce fait, je remercie toutes les personnes de ce laboratoire et particulièrement Sandrine Sagan qui m'a intégré dans ce laboratoire. And particularly, Cillian, for his help in synthesis. Merci à Solange Lavielle pour ses remarques et conseils lors de la préparation à ma soutenance de thèse. Merci à Eliane et Stéphane pour la gestion de ce laboratoire. Eh oui, sans eux ce serait le chaos !

Ce travail fut le fruit de nombreuses collaborations unifiées par un projet ANR qui a regroupé de nombreuses personnes (12 pour être plus précise) et auxquelles s'en sont associées de nouvelles. Je commence donc par remercier Gervaise Mosser pour son investissement dans ce projet (oui, parce que les microcellules made by Anaïs sont

## REMERCIEMENTS

extraordinaires !) mais également pour son soutien, sa bonne humeur et ses relectures. Merci à François et Aurélien, mes formateurs « microcellules », avec qui j'ai passé de bons moments (culture walt disney au top). Je tiens également à remercier Thibaud Coradin, Francisco Fernandes, Claire Aimé, Claire Teulon et Marie-Claire Schanne-Klein qui m'ont accueillie au sein de leurs laboratoires et qui ont participé à ce travail. Je remercie les cergysois Grégory Chaume et Olivier Monasson pour votre accueil. Merci de m'avoir fait partager vos connaissances et votre savoir-faire en synthèse. Je remercie également Nathalie Lensen et Thierry Brigaud pour votre accueil chaleureux à chaque fois que j'ai sollicité votre aide. Mes remerciements vont également vers Yves Jacquot et Rodrigue Marquant pour m'avoir prodigué maints conseils en matière de synthèse peptidique. Je souhaite également exprimer toute ma gratitude à Luminita Duma pour son investissement et l'intérêt qu'elle a manifesté à l'égard de ce projet. Merci d'avoir fait les déplacements pour venir discuter, merci pour le travail que tu as réalisé à distance et merci pour ta disponibilité. Merci à Claire Loison pour son excellente contribution à ce projet. Je tiens également à remercier Valérie Alezra et Isabelle Correia pour avoir assisté aux réunions annuelles ANR. Bien que nous n'ayons pas eu l'occasion de travailler ensemble, je vous remercie pour vos conseils et remarques lors de ces réunions.

Merci aux nombreux stagiaires qui ont participé à ce projet : Emelyne, Keyvan, Mathilde, Lauriane et Félix.

Et que serait la thèse sans les autres doctorants ? Joker. C'est pourquoi, je tiens à tous les remercier : les garçons du 4<sup>e</sup> (le Grand Medhi, Sébastien : merci pour le petit déjeuner, Frédéric, Simon et Didier), Claudia, Benjamin, Mathilde (merci de m'avoir accueillie dans ton chez toi durant les dernières semaines de ma thèse). Merci, à vous autres gens de l'ENS, Sam et Cyril, pour votre aide précieuse sur Mathematica et particulièrement à toi, Samuel, pour ce joli bouquet de fleurs apporté à un moment, comment dirais-je, un peu tendu. Merci à vous les filles du bureau 516. Je vais commencer par la grande, la folle, la joyeuse Aude. Je te remercie pour ces moments de rigolade et de papotage. Merci à toi ma petite Lucile pour les moments riches en émotions qui ont rendus cette thèse bien plus qu'agréable. Tu es géniale, j'espère que tu le sais et si tu ne le sais pas et bien je te le dis. Nos repas « chez Ann » vont me manquer et qui va pouvoir soutenir mes envies de sushis ??? Et je ne t'oublie pas Damien, seul garçon de ce bureau. Merci pour ta présence à tous les instants. Je tiens également à remercier les anciens doctorants et donc jeunes docteurs Anaïs, Tom, Sarah et Pierre pour nos discussions lors de la fête de la science et/ou autour d'un verre. Je te remercie Lucie pour ton

## REMERCIEMENTS

soutien, tes relectures attentives et les supers moments passés ensemble (le bal des JA, l'alsace, le pays angevin ...). Que dire de plus, que des rencontres formidables !

Maintenant, me voilà arrivée aux remerciements de ma famille et amies de toujours. Merci à mes coupines Mélanie, Lucile et Noémie, vous êtes toujours là et c'est ça qui compte ! Merci à toute ma graaaande famille qui a toujours été présente dans les bons comme les mauvais moments. Sans vous, je n'y serai pas arrivée. Je remercie particulièrement ma maman, Loïc, mon frère, ma sœur, mon beau-frère, ma belle-sœur et la plus belle (désolée pour les autres) ma petite Lilou pour tous les moments passés à vos côtés et qui rendent cette vie merveilleuse. Merci à toi mamie, je suis tellement contente que tu sois venue me soutenir le jour ma thèse.

Gracias a mi amiga Mariana por su paciencia que me permitió mejorar mi castellano durante los momentos agradables donde hablamos, reímos, bebimos y comimos (es muy importante). Gracias a tí, Joana por tu presencia a mi « pot de thèse ». Y lo mejor para el final, gracias a ti Adrián por todo, todo, todo y más.





# Table des matières

Abréviations .....	11
Introduction.....	13
<b>Références de l'introduction</b> .....	16
<b>Chapitre 1: Etude des relations monomères / triples hélices</b> .....	19
<b>1.1 Peptide modèle de collagène CMP</b> .....	20
1.1.1 Le collagène : ubiquitaire, protéiforme mais difficilement exploitable .....	20
1.1.2 Nature chimique des molécules CMP .....	25
1.1.3 La stabilisation de la triple hélice .....	26
<b>1.2 Synthèse et caractérisation de CMP</b> .....	27
1.2.1 Synthèse.....	27
1.2.2 Caractérisation par dichroïsme circulaire .....	31
<b>1.3 Etude des populations visibles par RMN liquide</b> .....	39
1.3.1 Attribution des résonances .....	39
1.3.2 Etudes des populations monomériques et triples hélices .....	43
<b>1.4 Marquage spécifique des résidus glycines</b> .....	46
<b>1.5 Etude structurale des CMP modèles</b> .....	51
1.5.1 Utilisation de l'expérience TROSY pour accéder à des informations locales .....	51
1.5.2 Dynamique moléculaire .....	63
<b>Conclusions et perspectives</b> .....	71
<b>Partie expérimentale</b> .....	72
A. Tests colorés.....	72
A1. Test Kaiser.....	72
A2. Test Chloranil.....	72
B. Préparation des échantillons de CD et traitement des données .....	72
C. SPPS pour les peptides marqués $^{13}\text{C}$ et $^{15}\text{N}$ .....	74
D. Calculs des coefficients de température des populations majoritaires de $(\text{POG})_7$ .....	75
E. Calculs des constantes de couplages $^3J_{\text{H}\alpha\text{-NH}}$ .....	76
F. Valeurs des constantes de couplages obtenues sur les expériences 3D.....	77

G. La dynamique moléculaire sur [(POG) <sub>10</sub> ] <sub>3</sub> .....	80
G1. Les paramètres de simulation .....	80
G2. Données issues de la simulation.....	81
<b>Références du chapitre 1</b> .....	84
ANNEXE 1: Classement des collagènes suivant leur appartenance, leur localisation et leur composition .....	90
ANNEXE 2 : Séquence des brins $\alpha 1$ et $\alpha 2$ du collagène de type I .....	91
<b>Chapitre 2 : Caractérisation d'assemblages supramoléculaires de CMP</b>	93
<b>3.1 Mise en évidence de la présence d'assemblages supramoléculaires en solution</b> .....	94
2.1.1 La diffusion dynamique de la lumière.....	94
2.1.2 Observations des assemblages supramoléculaires par microscopie électronique .....	100
<b>2.2 Analyse par RMN liquide</b> .....	103
2.2.1 Mise en évidence d'assemblages supramoléculaires par des expériences STD.....	103
2.2.2 Etudes cinétiques .....	106
<b>2.3 Mise en évidence et étude d'assemblages supramoléculaires non solubles</b> .....	114
2.3.1 Mise en évidence d'assemblages supramoléculaires non solubles .....	114
2.3.2 Caractérisation des assemblages supramoléculaires non solubles .....	115
<b>2.4 Etude de la modification des phases denses de collagène par dopage avec des peptidomimétiques</b> .....	123
2.4.1 Préparation des solutions de collagène.....	123
2.4.2 Les phases cristallines liquides.....	125
2.4.3 Dopage du collagène par des peptidomimétiques .....	127
<b>Conclusions et perspectives</b> .....	132
<b>Partie expérimentale</b> .....	133
A. Etude DLS .....	133
B. Préparation des échantillons pour la microscopie électronique .....	135
B1. Protocole de fixation pour la microscopie électronique.....	135
B2. Préparation des échantillons pour la cryomicroscopie .....	136
B3. Préparation des échantillons pour l'ombrage .....	137
C. Préparation des échantillons pour les études par microscopie optique .....	138

D. Etudes cinétiques.....	138
D1.Préparation des échantillons .....	138
D2.Expériences RMN et traitement des données .....	139
E. RMN solide .....	140
F. Protocole d'extraction du collagène.....	141
G. Détermination de la concentration d'hydroxyproline .....	142
H. Préparation des gels d'électrophorèse .....	144
<b>Références du chapitre 2.....</b>	<b>147</b>
ANNEXE 3 : Estimation du nombre de peptide par oligomère.....	150
ANNEXE 4 : Exemple de code Mathematica utilisé pour ajuster les courbes expérimentales des 3 états (monomériques, triples hélices et oligomériques) aux quatre températures .....	151
<b>Chapitre 3 : Vers de nouveaux peptidomimétiques du collagène.....</b>	<b>155</b>
<b>3.1 Contexte.....</b>	<b>156</b>
3.1.1 La proline et ses analogues.....	156
3.1.2 Pourquoi le fluor ?.....	160
3.1.3 Les pseudo-prolines trifluorométhylés.....	161
<b>3.2 Synthèse et caractérisation des triplets.....</b>	<b>165</b>
3.2.1 Synthèse.....	165
3.2.2 Caractérisation par RMN.....	170
<b>3.3 Etudes de peptides fluorés.....</b>	<b>180</b>
3.3.1 Synthèse.....	180
3.3.2 Analyse structurale .....	184
<b>Conclusions et perspectives .....</b>	<b>187</b>
<b>Partie expérimentale .....</b>	<b>188</b>
A. Synthèse des peptides P1, P2 , P3 et P4.....	189
A1.Synthèse sur support solide .....	189
A2.Synthèse en phase liquide .....	192
B. Calculs des populations up/down.....	198
C. Synthèse sur support solide des CMPF.....	199
C1.Synthèse avec le triplet PΨPG.....	199

C2.Synthèse avec le triplet $\Psi\Psi\Psi$ PG.....	201
<b>Références du chapitre 3</b> .....	205
ANNEXE 5 : Synthèse du peptide modèle $P3$ Ac-Pro-(2 <i>R</i> ,4 <i>S</i> )-Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro)-Gly-NH <sub>2</sub> en stratégie ester méthylique.....	205
ANNEXE 6 : Optimisation du couplage G- $\Psi$ Pro .....	212
<b>Conclusions et perspectives</b> .....	215
<b>Références de la conclusion et perspectives</b> .....	219





## Abréviations

**AA** : Acide aminé  
**ACN** : Acétonitrile  
**CCM** : Chromatographie sur couche mince  
**CD** : Circular Dichroïm  
**Clp** : (2*S*,4*R*)-4-chloroproline  
**clp** : (2*S*,4*S*)-4-chloroproline  
**DCM** : Dichlorométhane  
**DFT** : Density Functional Theory  
**DIC** : 1,3-Diisopropylcarbodiimide  
**DIEA** : *N,N*-Diisopropylethylamine  
**DLS** : Dynamic Light Scattering  
**DSS** : Acide 4,4-diméthyl-4-silapentane-1-sulfonique  
**Flp** : (2*S*,4*R*)-4-fluoroproline  
**flp** : (2*S*,4*S*)-4-fluoroproline  
**Gly (G)** : Glycine  
**HATU** : 1-[Bis(diméthylamino)méthylène]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate  
**HBTU** : *N,N,N',N'*-Tetraméthyl-*O*-(1*H*-benzotriazol-1-yl)uronium hexafluorophosphate  
**HFIP** : Hexafluoroisopropanol  
**HPLC** : High performance liquid chromatography  
**HSQC** : Heteronuclear Single Quantum Coherence  
**Hyp (O)** : 4*R*-Hydroxyproline  
**INEPT** : Insensitive Nuclei Enhanced by Polarization Transfer  
**M** : Monomère  
**Mep** : (2*S*,4*R*)-4-méthylproline  
**mep** : (2*S*,4*S*)-4-méthylproline  
**NMP** : *N*-méthyl-2-pyrrolidone  
**NOE** : Nuclear Overhauser Effect  
**NOESY** : Nuclear Overhauser Effect Spectroscopy  
**O** : nanoObjet  
**PBS** : Phosphate Buffer Saline  
**PDB** : Protein Data Bank  
**PPI** : PolyProline I  
**PPII** : PolyProline II  
**Pseudo-Pro (ΨP)** : Pseudo-proline  
**Pro (P)** : Proline  
**RDC** : Residual Dipolar Coupling  
**RMN** : Résonance Magnétique Nucléaire  
**ROESY** : Rotating-frame nuclear Overhauser Effect Spectroscopy  
**SHG** : Second Harmonique Generation  
**SPPS** : Synthèse peptidique sur phase solide  
**STD** : Saturation Transfer Difference  
**TEM** : Transmission Electron Microscopy  
**TFA** : Trifluoroacetic acid  
**TH** : Triple hélice  
**TIS** : Triisopropylsilane  
**T<sub>m</sub>** : Température de fusion  
**TROSY** : Transverse relaxation optimized spectroscopy





# Introduction

Une protéine est une macromolécule biologique caractérisée par un enchainement d'acides aminés (structure primaire) et repliée dans l'espace (structure secondaire et tertiaire). Sa structure tridimensionnelle, éventuellement associée à d'autres sous-unités protéiques (structure quaternaire) est à l'origine de son activité biologique. Les processus de repliement et d'auto-assemblage sont essentiellement guidés par des interactions non covalentes (liaisons hydrogènes, van der Waals, électrostatiques) et/ou des modifications post-traductionnelles (méthylation, hydroxylation, acétylation, glycosylations, phosphorylations...). L'isolement et l'étude de protéines en vue de relier leurs propriétés structurales à leur(s) fonction(s) biologique(s) constituent un champ important de recherche scientifique.

Lors de ce travail, nous nous sommes intéressés à la synthèse et la caractérisation structurale de peptides mimes de collagène, seuls ou en présence de collagène naturel. Chez les vertébrés, celui-ci représente un quart des protéines de l'organisme et chez l'homme un tiers des protéines exprimées. On le rencontre également chez les invertébrés : insectes ainsi que dans certains organismes marins.<sup>1</sup> Sans sa présence, nous ne serions qu'une « flaque d'eau ». En effet, son rôle, en coopération avec l'élastine ou d'autres molécules associées à sa structure quaternaire, est d'assurer la cohésion des tissus et de l'ensemble des organes. En fonction de son organisation structurale, il leur confère des propriétés de résistance (os), de souplesse (peau), ou de transparence (cornée). Selon l'expression du Dr P. Bououyrie, le collagène est à l'être ce que l'acier est au béton armé. A poids égal, la résistance des fibrilles de collagène surpasse en effet celle de l'acier. De façon générale, les collagènes désignent des protéines constituées de domaines hélicoïdaux résultant de l'enroulement de trois chaînes  $\alpha$ -peptidiques. On en recense, aujourd'hui, 28 types chez les vertébrés.

Nous nous sommes particulièrement intéressés au collagène de type I, fibrillaire, qui est extrêmement abondant puisqu'il représente 90% des collagènes. Chez l'homme, il est constitué de trois chaînes comptant chacune plus d'un millier d'acides aminés et associées au sein d'une triple hélice. Il est alors aisé de comprendre que les propriétés mécaniques du collagène sont étroitement liées à la stabilité de la triple hélice. Une première condition

nécessaire à sa formation est d'avoir un nombre important de résidus cycliques de type pyrrolidine stabilisant une structure secondaire polyproline II (PPII) et la présence systématique d'une glycine tous les trois résidus dans sa séquence primaire. Au-delà de ces contraintes de structure primaire, *Raines et al.*<sup>2</sup> ont proposé un « code » qui repose sur la nature des plissements des cycles pyrrolidines et qui sera présenté en détail dans le chapitre 3.<sup>3-5</sup>

De nombreuses équipes de recherche ont tenté d'étudier et/ou de mimer la triple hélice de collagène pour mieux contrôler sa stabilité. L'omniprésence du collagène dans le corps humain et les nombreuses pathologies qui peuvent subvenir suite à une anomalie lors de sa biosynthèse ou à des défauts de repliement expliquent qu'il constitue un objet d'étude de premier plan. Parmi les maladies les plus connues, on peut citer le scorbut, l'ostéogénèse imparfaite (ou maladie des « os de verre ») ou encore les syndromes d'Ehler-Danlos qui se signalent par une hyper-élasticité des tissus (d'où le syndrome de « l'homme caoutchouc »). Outre son intérêt d'un point de vue médical, il est également très étudié pour la conception de nouveaux biomatériaux. En effet, le collagène naturel est le principal composant des sutures biodégradables, des valves cardiaques artificielles et des implants pour l'incontinence urinaire. Ce marché est en nette augmentation en raison du vieillissement de la population. Cependant, il existe des facteurs limitant l'utilisation de collagène naturel. Une première raison vient de son origine animale, principalement bovine, à l'origine d'allergies ainsi que de problèmes d'immunogénicité. Par ailleurs, il est difficile d'obtenir du collagène naturel sans dégrader ses propriétés structurales et mécaniques car les méthodes d'extraction et de purification employées sont longues et fastidieuses, souvent incompatibles avec les conditions de préservation du collagène natif. C'est pourquoi des molécules synthétiques ont été développées pour mimer le collagène. Ces molécules nouvelles ont été dénommées CMP, l'acronyme anglais de « Collagen Model Peptide ». Le peptide (Pro-Hyp-Gly)<sub>n</sub> s'est illustré comme le CMP de référence puisqu'il correspond à la répétition du triplet le plus présent dans la séquence de collagène. Différentes modifications ont alors été envisagées, basées sur cette séquence de référence : utilisation d'acides aminés naturels non cycliques à des positions stratégiques, insertion d'analogues de prolines variés,<sup>1,6</sup> création de liens covalents entre chaînes  $\alpha$ -peptidiques.<sup>7</sup> Certains de ces CMP sont déjà de bons candidats pour la fabrication de biomatériaux soit par dopage du collagène naturel,<sup>8</sup> soit par greffage sur des polymères synthétiques, soit par addition à des matériaux inorganiques.<sup>9,10</sup> Certains d'entre eux forment

des fibres, des gels, mais ne sont pas encore assez stables d'un point de vue thermique, manquent de résistance mécanique, ou ne sont pas biocompatibles.<sup>11,12</sup>

Les études que nous avons menées ont visé à analyser en détail les relations entre une triple hélice modèle et les différentes espèces monomériques présentes en solution, identifiées grâce à des développements d'outils originaux. Une approche multi-échelle a été privilégiée pour appréhender les différents niveaux de structuration de peptides mimes de collagène : conformation locale, structures secondaires, assemblages trimériques et supramoléculaires. Le travail est ici présenté en trois grands chapitres :

Dans un premier chapitre, après une présentation détaillée du collagène naturel, nous aborderons les synthèses peptidiques menées et nous nous focaliserons sur leur étude structurale en solution. Nous appréhenderons la stabilité de la triple hélice de collagène et sa caractérisation structurale grâce à l'utilisation de peptides modèles marqués. Il contiendra un premier aperçu des techniques utilisées durant ce travail : dichroïsme circulaire, Résonance Magnétique Nucléaire et dynamique moléculaire.

Dans un deuxième chapitre, nous étudierons les assemblages supramoléculaires solubles et insolubles de ces peptides modèles mais également le dopage des phases denses de collagène par ces molécules peptidomimétiques. Les techniques mises en œuvre pour réaliser ces études sont nombreuses : microscopies optiques/électroniques, RMN liquide/solide, diffusion de lumière (Diffusion Light-Scattering).

Enfin dans un dernier chapitre, nous proposerons une nouvelle famille de molécules peptidomimétiques. Les synthèses et les analyses structurales de triplets fluorés seront détaillées et les premières études de peptides contenant des groupes triflorométhylés seront évoquées. Ceux-ci nous apparaissaient prometteurs dans le cadre de la conception de nouveaux CMP, en vue d'améliorer la stabilité de la triple hélice et de promouvoir des assemblages supramoléculaires supérieurs.

## Références de l'introduction

- 1 Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annual review of biochemistry* **78**, 929-958, doi:10.1146/annurev.biochem.77.032207.120833 (2009).
- 2 Holmgren, S. K., Taylor, K. M., Bretscher, L. E. & Raines, R. T. Code for collagen's stability deciphered. *Nature* **392** (1998).
- 3 Yang, W., Chan, V. C., Kirkpatrick, A., Ramshaw, J. A. M. & Brodsky, B. Gly-Pro-Arg Confers Stability Similar to Gly-Pro-Hyp in the Collagen Triple-helix of Host-Guest Peptides. *The Journal of biological chemistry* **272**, 28837-28840 (1997).
- 4 Ackerman, M. S. *et al.* Sequence Dependence of the Folding of Collagen-like Peptides. *J Bio Chem* **274**, 7668-7673 (1999).
- 5 Horng, J. C., Kotch, F. W. & Raines, R. T. Is glycine a surrogate for a D-amino acid in the collagen triple helix? *Protein science : a publication of the Protein Society* **16**, 208-215, doi:10.1110/ps.062560107 (2007).
- 6 Kar, K., Wang, Y. H. & Brodsky, B. Sequence dependence of kinetics and morphology of collagen model peptide self-assembly into higher order structures. *Protein science : a publication of the Protein Society* **17**, 1086-1095, doi:10.1110/ps.083441308 (2008).
- 7 Kotch, F. W. & Raines, R. T. Self-assembly of synthetic collagen triple helices. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3028-3033, doi:10.1073/pnas.0508783103 (2006).
- 8 Wang, A. Y., Mo, X., Chen, C. S. & Yu, S. M. Facile Modification of Collagen Directed by Collagen Mimetic Peptides. *Journal of the American Chemical Society* **127**, 4130-4131 (2005).
- 9 Suehiro, T., Kojima, C., Tsumura, S., Harada, A. & Kono, K. Higher order structure of short collagen model peptides attached to dendrimers and linear polymers. *Biopolymers* **93**, 640-648, doi:10.1002/bip.21413 (2010).
- 10 Desimone, M. F. *et al.* Silica-collagen bionanocomposites as three-dimensional scaffolds for fibroblast immobilization. *Acta Biomater* **6**, 3998-4004, doi:10.1016/j.actbio.2010.05.014 (2010).
- 11 O'Leary, L. E., Fallas, J. A., Bakota, E. L., Kang, M. K. & Hartgerink, J. D. Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel. *Nature chemistry* **3**, 821-828, doi:10.1038/nchem.1123 (2011).
- 12 Przybyla, D. E. & Chmielewski, J. Higher-order assembly of collagen peptides into nano- and microscale materials. *Biochemistry* **49**, 4411-4419, doi:10.1021/bi902129p (2010).





# Chapitre 1

## Etude des relations

### monomères / triples hélices

C'est en 1803 que la notion d'équilibre chimique est pour la première fois évoquée par le chimiste Claude Louis Berthollet. Les peptides que nous introduisons dans ce chapitre, les CMP, constituent un bel exemple d'équilibre chimique. Ils ne font pas intervenir de réaction chimique mais des échanges conformationnels entre des espèces monomériques et des assemblages de trois monomères, structurés en triple hélice.

Nous évoquerons dans un premier temps le contexte qui nous a amené à nous intéresser à ces molécules CMP. Nous montrerons, après une présentation détaillée du collagène naturel, leurs intérêts ainsi que les stratégies développées pour construire des assemblages en triple hélice stables. Dans une seconde partie, nous présenterons la synthèse et la caractérisation de ces peptides. Ensuite, nous étudierons par RMN les différentes populations présentes à l'état liquide. Certaines de ces études reposent sur des développements méthodologiques nouveaux utilisant un marquage isotopique spécifique des glycines. La dernière partie est enfin consacrée à des études de dynamique moléculaire. Les paramètres structuraux obtenus, caractéristiques des édifices moléculaires, seront comparés aux observations faites par RMN en solution.



## 1.1. Peptides modèles de collagène CMP

### 1.1.1 Le collagène naturel : ubiquitaire, protéiforme, mais difficilement exploitable

Le collagène est la protéine structurale la plus abondante du règne animal. Le terme « *collagène* » est un néologisme de la langue française (du grec : *kolla* – colle et *genno* – donner la vie) apparu au cours du XIX<sup>e</sup> siècle par les histologistes pour désigner le constituant des tissus conjonctifs.<sup>1</sup> Le collagène, comme élément de structure de base pour la plupart des tissus conjonctifs, joue un rôle important dans le maintien de l'intégrité structurale et biologique de l'architecture de la matrice extracellulaire. De façon générale, les collagènes désignent des protéines constituées de domaines hélicoïdaux résultant de l'enroulement de trois chaînes  $\alpha$ -peptidiques. On recense, aujourd'hui, 28 types de collagène chez les vertébrés que l'on peut classer en 5 catégories : fibrillaire, en réseau, FACIT (Collagène Associé en Fibrilles avec Triple hélice Interrompue), MACIT (Collagène transMembranaire avec des interruptions de triple hélice) et MULTIPLEXIN (domaine multiple de triple hélice et interruption).<sup>2,3</sup> Un tableau présentant les différents types de collagène ainsi que leur localisation et les pathologies liées à des altérations de leur structure est présenté en ANNEXE 1.

*La biosynthèse du collagène fibrillaire* – Les fibres de collagène se forment à partir de chaînes polypeptidiques synthétisées par les ribosomes liés à la membrane du réticulum endoplasmique et sécrétés dans le lumen de celui-ci. Les étapes principales dans la formation de fibres de collagène sont : (i) hydroxylation de certaines prolines et lysines en 4-hydroxyproline, 3-hydroxyproline et hydroxylysine, (ii) glycoslation de certaines hydroxylysines en galactosylhydroxylysine et glucogalactosylhydroxylysine, (iii) glycosylation de certaines asparagines sur les extrémités *N*- et *C*-terminales (télopeptides) (iv) rapprochement de trois chaînes peptidiques au niveau de leurs extrémités *C*-terminales (2 chaînes  $\alpha$ 1, 1 chaîne  $\alpha$ 2) (v) formation de ponts disulfures intra et intermoléculaires stabilisant le point de nucléation *C*-terminal (Figure 1). On parle ensuite de processus zipper-like : à partir du point de nucléation il y a une propagation de l'assemblage des trois chaînes peptidiques du *C*-terminal vers le *N*-terminal pour former le procollagène.<sup>4</sup> Les molécules de procollagène sont ensuite transportées dans l'appareil de Golgi où elles s'associent latéralement pour former des vésicules de sécrétion. Au moment de leur

excrétion de la cellule, les télépeptides du procollagène sont clivés par des propeptidases *N*- et *C*-terminales au niveau de la membrane cellulaire. L'étape suivante est l'auto-assemblage spontané des molécules de collagène résultant à la formation de fibrilles. De courts segments, non associés en triple hélice sont présents aux extrémités des chaînes  $\alpha$  et sont importants pour former les fibrilles de collagène. Les amines de certaines lysines et hydroxylysines présentes sur ces courts segments vont être oxydées et transformées en aldéhydes réactifs, ce qui va permettre la création de liaisons covalentes entre les fibrilles et ainsi les rendre plus rigides (fibres).

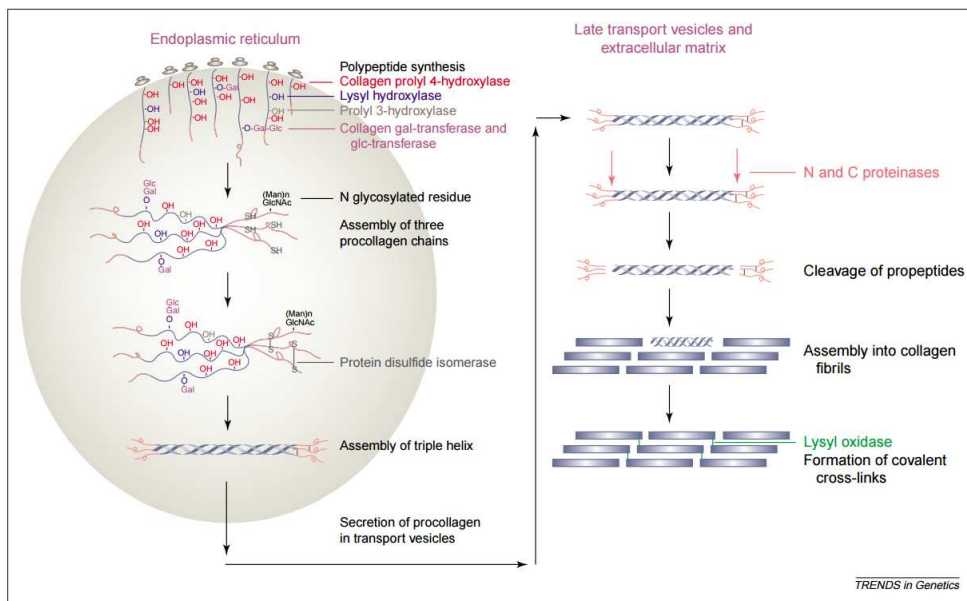


Figure 1: Schéma issu de la publication de Kivirikko *et al.*<sup>5</sup> présentant la formation de la fibre de collagène

La Figure 2 permet d'avoir une idée de la taille relative des assemblages aux différentes étapes de formation de la fibre de collagène.

Il est surprenant de savoir que le collagène, bien qu'il soit le constituant majeur de la plupart des tissus vivants, est instable à 37°C.<sup>6,7</sup> Les chaînes  $\alpha$  adoptent alors une conformation en pelote statistique et la structure en triple hélice est perdue de manière irréversible, ce qui se traduit, à l'échelle macroscopique, par de la gélatine liquide à température ambiante (20-25°C) et par l'apparition d'un gel quand elle est refroidie. Ce comportement est défini à l'échelle moléculaire comme une transition pelote-hélice.<sup>8</sup> La gélatine, bien que moins structurée que le collagène, possède des propriétés relativement intéressantes, très utilisées dans des domaines aussi variés que la photographie et l'agroalimentaire.

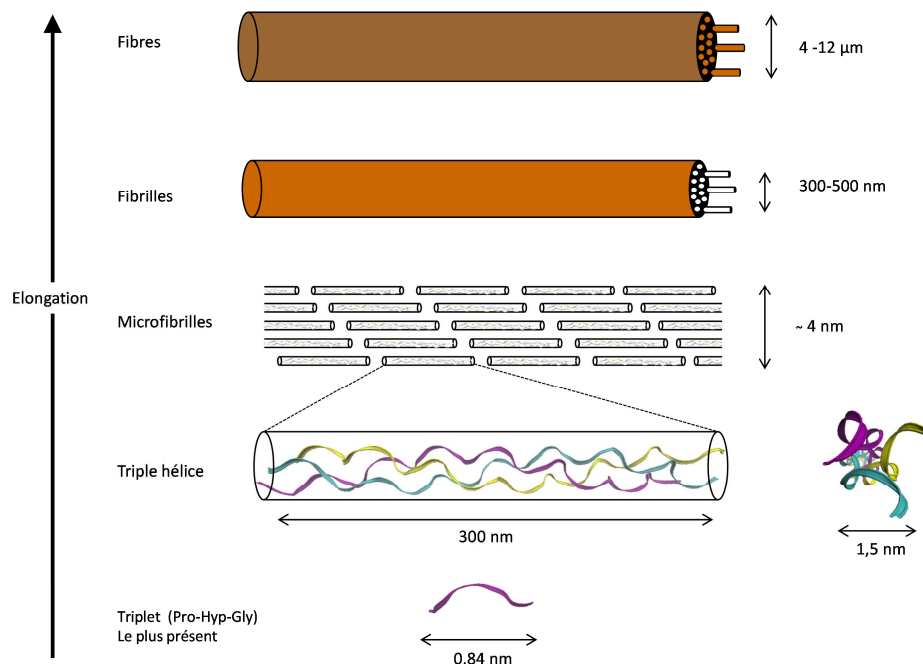


Figure 2: Les différents niveaux d'assemblages au sein des fibres de collagène

On voit ainsi que la triple hélice représente la « brique élémentaire » pour la formation d'édifices stables, et constitue à ce titre un pilier dans la structure du collagène. Ceci explique pourquoi de nombreuses études ont été réalisées pour comprendre sa stabilité. Par la suite, nous nous concentrerons sur le collagène fibrillaire de type I puisqu'il représente 90% des collagènes.

*Structure primaire, secondaire/tertiaire et quaternaire* – Chez l'Homme, il est constitué de deux chaînes ( $\alpha 1$ ) composées de 1464 acides aminés et d'une chaîne ( $\alpha 2$ ) de 1366 acides aminés (cf. ANNEXE 2), associées au sein d'une triple hélice.<sup>9</sup> Le résidu **glycine** apparaît périodiquement tous les trois résidus dans la structure primaire, la séquence du collagène pouvant alors se noter : **(Xaa-Yaa-Gly)<sub>n</sub>**. Le résidu **proline** occupe la position Xaa dans 28% des cas, et après hydroxylation enzymatique, le résidu **4R-hydroxyproline** occupe la position Yaa à 38%. Avec une occurrence de 10.5%, le triplet (Pro-Hyp-Gly) est donc le plus représenté dans le collagène de type I.<sup>10,11</sup> Les contraintes structurales apportées par les nombreux cycles pyrrolidines permettent de stabiliser une structure secondaire particulière appelée polyproline II (PPII). Cette conformation correspond à une hélice gauche, caractérisée par des angles diédraux  $\omega$ ,  $\phi$  et  $\psi$  de  $180^\circ$ ,  $-75^\circ$  et  $150^\circ$ , respectivement.<sup>12</sup> La séquence polyproline ne présentant pas de protons amides, aucune

liaison hydrogène intramoléculaire ne participe à la stabilisation de cette structure secondaire. En revanche, à l'image des brins  $\beta$  qui s'assemblent en feuillet  $\beta$  par liaisons hydrogène inter-brins, les protons amides des résidus glycine présents sur la séquence de collagène permettent aux hélices PPII de s'associer en assemblages supramoléculaires. Des liaisons hydrogène sont en effet observées sur les structures RX obtenues pour des modèles de triple hélice (PDB 1CAG).<sup>10</sup> Celles-ci impliquent chaque HN amide des résidus glycine comme donneur de liaison H et chaque carboxyle CO des résidus proline comme accepteur (Figure 3A, 3B). La structure quaternaire du collagène consiste ainsi en un superenroulement de trois hélices PPII. Les deux premières conditions nécessaires à la formation de la triple hélice sont d'avoir une séquence comportant un nombre important de résidus cycliques de type pyrrolidine et la présence de la glycine tous les trois résidus. L'absence périodique de chaîne latérale sur le squelette peptidique (résidu glycine) est en effet requis pour que les trois hélices PPII puissent s'associer en trimères. Au-delà de ces contraintes de structure primaire, *Raines et al.*<sup>13</sup> ont proposé un « code » expliquant pourquoi les résidus proline et 4*R*-hydroxyproline étaient autant représentés dans les séquences de collagène. Ce « code » repose sur la nature des plissements des cycles pyrrolidines. Des analyses thermodynamiques fines ont en effet montré qu'un plissement de cycle *C<sup>γ</sup>-endo* (*down puckering*) était nécessaire en lieu et place de la proline (Xaa) alors qu'un plissement de cycle *C<sup>γ</sup>-exo* (*up puckering*) était stabilisant en position Yaa (Figure 3C). En outre, la présence du substituant hydroxyle en position 4 sur le résidu hydroxyproline stabilise fortement le plissement de cycle *C<sup>γ</sup>-exo* par effet stéréoelectronique et inductif. De nombreuses études ont démontré qu'une mauvaise alternance de ces plissements de cycle engendrait une déstabilisation de la triple hélice, tel que cela est détaillé dans le chapitre 3.<sup>14-16</sup> Par ailleurs, il est important de préciser qu'une conformation *trans* de toutes les liaisons peptidiques est nécessaire pour maintenir la conformation PPII au sein de la triple hélice. Cette condition est largement vérifiée dans le cas des liaisons peptidique Xaa-nonPro (conformation *trans* à près de 99.97%), alors que ~ 5.2% des liaisons Xaa-Pro sont en conformation *cis*.<sup>17</sup> L'isomérisation *cis-trans* peut ainsi être une étape limitante dans la formation de la triple hélice.

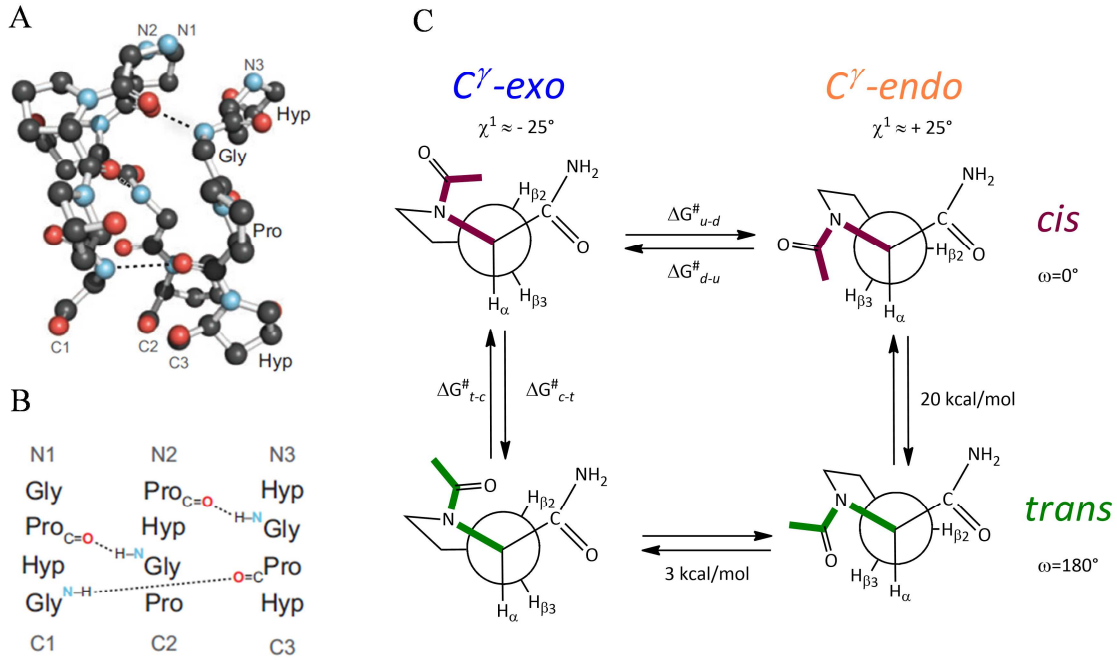


Figure 3: Prérequis structural pour former la triple hélice de collagène. A, B : liaisons hydrogène représentées entre les trois chaînes peptidiques. C : équilibres conformationnels présents au niveau des résidus prolines. Les cycles sont représentés en projection de Newman par rapport à l'angle dièdre  $\chi^1$ .

*Intérêt de l'étude du collagène* - Le collagène se présente sous des morphologies différentes suivant les tissus et se réorganise constamment afin de remplir des fonctions variées.<sup>18</sup> Dans les os, il fournit des tissus rigides et résistants aux chocs grâce à son organisation particulière avec l'hydroxyapatite (phase minérale à base de calcium).<sup>19</sup> Il peut se comporter comme un élastomère avec une faible rigidité et une forte tolérance à la déformation comme on peut le voir dans le tendon.<sup>20</sup> Il montre également des propriétés optiques telles que la transparence dans la cornée.<sup>21</sup> Cette polyvalence a inspiré de nombreux chercheurs qui se sont dédiés au développement de nouveaux biomatériaux basés principalement sur le collagène de type I qui est fibrillaire et représente 90% des collagènes. Cependant, les applications actuelles sont encore peu nombreuses et se limitent principalement au domaine des cosmétiques et à la mise au point de « pansements actifs » (Catrix®, Opraskin®). Des études en cours sur des hydrolysats de collagène de type II (gélatine issue du collagène de type II) montrent que l'administration de 0.1mg par jour diminuerait significativement les douleurs liées à l'arthrose.<sup>22</sup> Il a été également publié que la prise d'un supplément de collagène de type I améliorerait la guérison des ulcères de pression (escarres).<sup>23,24</sup>

La méthode principale pour la fabrication de ces biomatériaux est l'utilisation de collagène animal du fait de son abondance. Il peut être extrait à partir de presque tous les animaux

vivants, néanmoins, les sources communes de collagène pour les applications d'ingénierie tissulaire sont la peau et les tendons de bovins, la peau de porc, de poissons, et la queue de rat. Deux problèmes majeurs sont cependant rencontrés et nécessitent de nouvelles méthodes de production. L'intégrité structurale du collagène n'est en effet pas conservée durant les processus d'extraction et de purification. Par ailleurs, l'emploi de collagène d'origine animale peut déclencher des réponses immunitaires compromettant son utilisation.

Une alternative est développée depuis une vingtaine d'année, à la fois dans la recherche académique et industrielle, et consiste à synthétiser des molécules capables de mimer ce polymère aux propriétés spectaculaires, ces molécules portent le nom CMP pour Collagen Model Peptide.

### 1.1.2 Nature chimique des molécules CMP

Les CMP sont définis comme étant des peptides ou peptidomimétiques synthétiques ayant la capacité d'adopter en solution la conformation triple hélice, caractéristique du collagène. Les peptides mimes de collagène sont généralement composés de la répétition du triplet (Xaa-Yaa-Gly), où Xaa et Yaa sont souvent un résidu Proline ou un de ses analogues. Pour chacun de ces peptides CMP, une température de fusion ( $T_m$ ) peut être déterminée, celle-ci correspondant à la température critique de dénaturation de la triple hélice. Cette température  $T_m$  est une mesure directe de la stabilité des assemblages non covalents, comme cela est couramment fait pour mesurer les stabilités des doubles hélices d'ADN. La répétition du triplet (Pro-Hyp-Gly) permet d'obtenir des triples hélices plus stables, en accord avec les occurrences très importantes des résidus Proline et 4(R)-Hydroxyproline aux positions Xaa et Yaa, respectivement. Cinq répétitions du triplet suffisent à former la triple hélice, (Pro-Hyp-Gly)<sub>5</sub> ayant une  $T_m$  égale à 5°C.<sup>25</sup> Lorsque le peptide CMP est allongé, sa température de fusion augmente, (Pro-Hyp-Gly)<sub>10</sub> est très stable avec une  $T_m=61$ °C.<sup>10</sup> Le modèle de *Brodsky et al.* basé sur les systèmes « hôte-invité » a été le plus largement utilisé pour la conception de mimes de collagène dédiés à des applications biologiques.<sup>26,27</sup> Dans cette approche, le triplet d'intérêt Gly-Xaa-Yaa, dit séquence "invitée", est entouré par plusieurs répétitions de « l'hôte » (Gly-Pro-Hyp) qui est le triplet le plus commun dans la séquence de collagène et donc connu pour améliorer la stabilité de la triple hélice. Une formule générique de peptide hôte-invité est alors Ac-(Gly-Pro-Hyp)<sub>3</sub>-Gly-Xaa-Yaa-(Gly-Pro-Hyp)<sub>4</sub>-Gly-NH<sub>2</sub>. La mesure des températures de fusion

des peptides avec diverses combinaisons de Xaa et Yaa fournit des informations pour prédire la stabilité thermique des peptides mimes de collagène. Les analyses suggèrent que la nature du résidu en position de Xaa, que ce soit une proline, un résidu hydrophobe ou chargé, a relativement peu d'effet sur la stabilité de la triple hélice, peut-être en raison de la nature exposée du site Xaa. En revanche, la nature du résidu en position Yaa influence fortement la stabilité de la triple hélice. Brodsky *et al.* ont pu montrer que la présence de résidus hydrophobes, à la position Yaa, même de faible encombrement stérique, est déstabilisante pour la triple hélice. Il en a été conclu une absence d'interactions hydrophobes à cette position. Il est d'ailleurs trouvé que l'invité Gly-Pro-Arg stabilise tout autant la triple hélice que Gly-Pro-Hyp,<sup>14</sup> ce qui s'explique par la présence de 5 sites potentiels permettant d'établir des liaisons hydrogène et de surcroît stabiliser le réseau d'hydratation de la triple hélice. L'ensemble des données collectées avec différents triplets Gly-Xaa-Yaa invités<sup>28,29</sup> a permis de mettre au point un programme informatique pouvant prédire, pour une séquence aléatoire, la stabilité de la triple hélice.<sup>30</sup>

Au-delà des modifications de la séquence primaire (Xaa-Yaa-Gly)<sub>n</sub> avec des acides aminés naturels, de nombreux analogues ont été synthétisés utilisant des acides aminés non naturels, notamment des analogues de proline qui font l'objet d'une revue présentée dans le chapitre 3.

### 1.1.3 Les méthodes utilisées pour stabiliser la triple hélice

Différentes stratégies ont été abordées pour stabiliser la triple hélice. L'une d'entre elles consiste en l'utilisation de petites molécules organiques conjuguées aux peptides et capables de les relier par liaison covalente. Ces petites molécules permettent de forcer la trimérisation de l'hélice en jouant sur un gain entropique. Les molécules ayant trois groupements réactifs tels que l'acide *cis,cis*-1,3,5-triméthylcyclohexane-1,3,5-tricarboxylique (KTA)<sup>31,32</sup>, la tris(2-aminoéthyl)amine (TREN) couplée à l'acide succinique<sup>33</sup>, la tris(hydroxyméthyl) aminométhane (TRIS)<sup>34,35</sup> ainsi que la triserine lactone (TSL)<sup>36</sup> ont été utilisées pour maintenir les chaînes peptidiques proches les unes des autres et promouvoir leur assemblage en triples hélices. Il a été montré qu'une telle approche permettait d'incorporer des peptoïdes au sein des triples hélices. La N-isobutyglycine a en effet été placée avec succès en position Xaa et Yaa d'une triple hélice stabilisée par le KTA.<sup>37-39</sup>

Une autre approche est la création de liaisons entre les trois brins monomériques par incorporation d'acides aminés particuliers comme la cystéine ou la lysine.<sup>40,41</sup> *Raines et al.* ont synthétisé deux chaînes (Pro-Yaa-Gly)<sub>8</sub>-Cys-Gly et une chaîne (Pro-Yaa-Gly)<sub>3</sub>-Cys-Cys-Gly-(Pro-Yaa-Gly)<sub>5</sub> afin de pouvoir lier les trois chaînes entre elles, de manière décalée, par l'établissement de ponts disulfures.<sup>42</sup> Une température de fusion  $T_m$  de 47°C a été trouvée pour cette hélice, qui, à une échelle supérieure, peut former des assemblages supramoléculaires plus stables. Il a également été montré que l'ajout de fonctions spécifiques permettant de chélater un ion métallique, tel que le catéchol ou l'acide hydroxamique, en C-terminal de chaque monomère augmentait remarquablement la stabilité de la triple hélice ( $T_m + 22^\circ\text{C}$ ).<sup>43,44</sup> Enfin, l'ajout de chaînes alkyles en N-terminal a permis d'augmenter sensiblement cette température de fusion.<sup>45,46</sup>

Nous avons également tenté de développer une nouvelle famille de molécules capables de mimer la triple hélice de collagène. Ce travail repose sur l'incorporation d'acides aminés trifluorométhylés et sera présenté en chapitre 3. Avant cela, l'objet de cette première partie est de regarder en détail les équilibres qui existent entre un monomère de CMP et sa triple hélice. Les analyses effectuées sur ces peptides nous serviront de références pour des études ultérieures, qui seront réalisées sur des molécules originales.

## 1.2. Synthèse et caractérisation de CMP

Nous avons entrepris une étude systématique de peptides mimes de collagène (Pro-Hyp-Gly)<sub>n</sub> avec  $3 < n < 10$  afin de pouvoir établir leur domaine de stabilité. L'ensemble de ces peptides a été caractérisé par dichroïsme circulaire et RMN.

### 1.2.1 Synthèse

Dans les années 60, R.B. Merrifield pose les bases de la synthèse supportée sur phase solide (SPPS).<sup>47</sup> Le principe repose sur l'utilisation d'une matrice solide à laquelle on vient lier des acides aminés par des liaisons covalentes de manière itérative. L'avantage de cette technique est la facilité de purification contrairement à une synthèse en phase liquide. En effet, la matrice, plus souvent appelée résine, est constituée d'un polymère insoluble dans les solvants organiques choisis pour la synthèse, la purification se résume alors en une succession de lavage, filtration jusqu'à décrochage du peptide de la résine (Figure 4). Deux stratégies sont utilisées, la stratégie Fmoc (9-



fluorénylméthoxycarbonyl) et la stratégie Boc (*tertio*-butyloxycarbonyl). Ces appellations proviennent de la fonction utilisée pour la protection de l'amine. De ce fait, la méthode de déprotection est différente, l'une se fait en milieu basique (stratégie Fmoc) et l'autre en milieu acide (stratégie Boc), le protocole de couplage des acides aminés restant identique. Dans la synthèse supportée, la présence de groupements protecteurs sensibles sur les chaînes latérales des acides aminés, que ce soit en milieu acide ou en milieu basique, guidera le choix de la stratégie à utiliser. La stratégie Fmoc est cependant privilégiée car le clivage du peptide sur le support est réalisé en présence d'acide trifluoroacétique (TFA) concentré alors que la stratégie Boc nécessite l'usage de l'acide fluorhydrique (HF) concentré, plus difficilement manipulable. Nous avons donc choisi de travailler en stratégie Fmoc car le seul acide aminé qui nécessite la protection de sa chaîne latérale est l'hydroxyproline de par son groupement hydroxyle en position 4. Cette fonction est protégée par un groupe tertibutyle qui n'est pas sensible en milieu basique et qui sera déprotégée lors du clivage final en milieu acide. Les acides aminés (AA) utilisés sont alors protégés par le groupement chlorure de fluorénylméthoxycarbonyl (Fmoc).

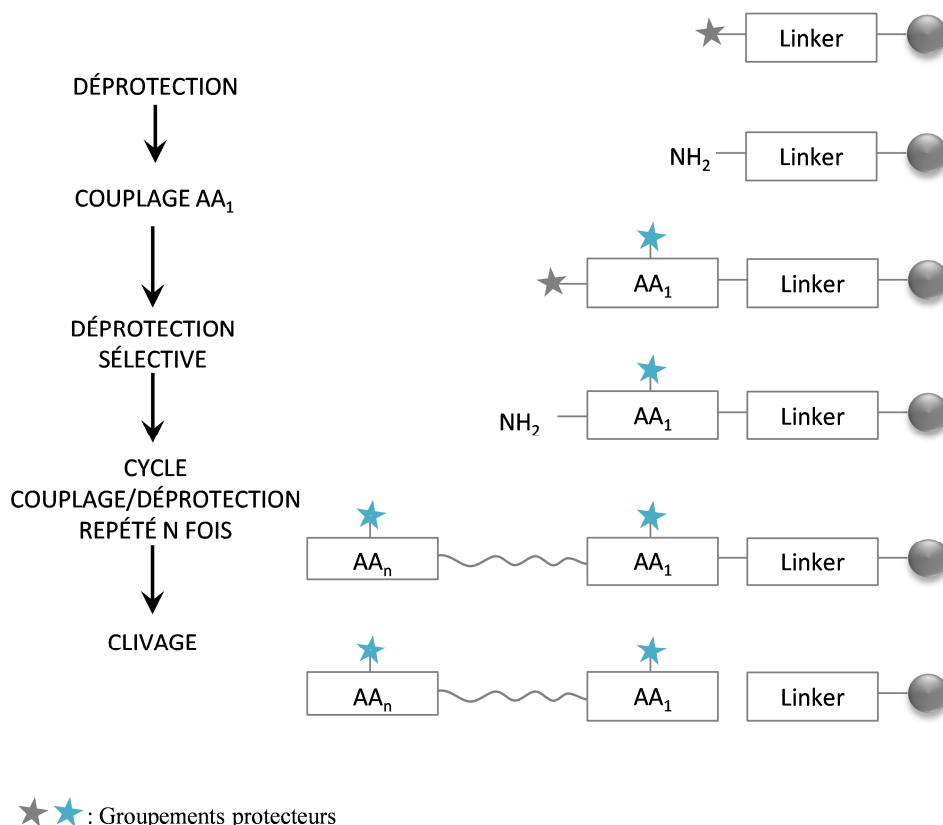


Figure 4: Schéma du principe de la synthèse supportée

La première stratégie que nous avons envisagée pour la synthèse de peptides mimes de collagène était une synthèse par blocs, le triplet Pro-Hyp-Gly revenant périodiquement dans la séquence primaire. Le couplage des blocs entre eux peut ensuite être effectué dans un temps raisonnable sans avoir recours au synthétiseur automatique.

*Synthèse par blocs* – Nous avons choisi de synthétiser en grande quantité le bloc (Pro-Hyp-Gly) afin de pouvoir synthétiser des CMP (Pro-Hyp-Gly)<sub>n</sub> avec n=3-10. Nous avons tenté d'optimiser le couplage du résidu Hyp qui est relativement coûteux. L'optimisation du temps de couplage (1h ou 2h) et du nombre d'équivalents (2 ou 4) a été effectuée, en simple ou en double couplages. Les rendements se sont avérés excellents (> 90%) même dans les conditions les plus défavorables. Nous avons donc retenu les conditions les plus économes, à savoir un couplage d'une heure avec 2 équivalents de résidus Hyp.

Procédure générale de synthèse : La résine utilisée est de type Fmoc 2-chlorotrityle (loading : 0.61mmol/g ou 0.51mmol/g) préchargée avec une glycine. Les synthèses ont été réalisées sur 0.1mmol. Tout d'abord, la résine est mise à gonfler avec du CH<sub>2</sub>Cl<sub>2</sub> pendant 5 min puis dans la NMP pendant 5 min également dans une seringue de 5 ml. Ensuite, un test Kaiser<sup>i</sup> est réalisé<sup>48</sup> afin de vérifier que la résine était bien protégée. On déprotège alors la résine avec une solution de pipéridine/NMP (20 :80) pendant 3 min, puis 5 min et enfin 10 min. Puis la résine est lavée avec de la NMP. On vérifie que la résine est bien déprotégée avec le test de Kaiser. Puis la Fmoc-O-tert-butyl-L-trans-4-hydroxyproline (2éq) est solubilisée dans 1.8éq HATU et 1.5 ml de NMP. Ensuite, 4éq de DIEA sont ajoutés à la solution pour activer la fonction acide. Après 3 min, cette solution est ajoutée à la résine dans la seringue. Le couplage est laissé 1h. Le test Kaiser permet seulement de mettre en évidence les amines primaires libres, un autre test est alors réalisé pour détecter la présence d'amines secondaires : le test chloranil.<sup>49</sup> Après couplage la solution est alors lavée avec de la NMP (2 fois minimum) puis on réalise la déprotection. Ensuite, une solution de proline (5éq) dans du HBTU (4,5éq)/DIEA (10éq) est ajoutée. La résine est lavée au CH<sub>2</sub>Cl<sub>2</sub> puis dans le MeOH et enfin séchée au dessiccateur pendant une nuit. Enfin, le peptide est clivé de la résine avec 3 ml d'une solution de HFIP/CH<sub>2</sub>Cl<sub>2</sub><sup>ii</sup> (20 :80), 5 fois pendant 15 min. Ensuite, le solvant est évaporé sous pression afin d'obtenir une poudre blanche (tr=20min, 5:95 à 0:100 A/B éluant A (H<sub>2</sub>O/TFA 0.1%) et éluant B (ACN/TFA 0.1%).

<sup>i</sup> Le test de Kaiser est un test qui permet de mettre en évidence la présence d'amine primaire libre.

<sup>ii</sup> L'HFIP est un réactif permettant de cliver le peptide de la résine sans enlever le groupement tertbutyle qui protège la fonction alcool.

Le bloc (Pro-Hyp-Gly) que l'on nommera par la suite (POG) est obtenu à un très haut degré de pureté ce qui permet de l'engager directement dans les étapes de couplages sans purification préalable. La synthèse a été réalisée à la main selon le schéma présenté ci-dessous (Figure 5).

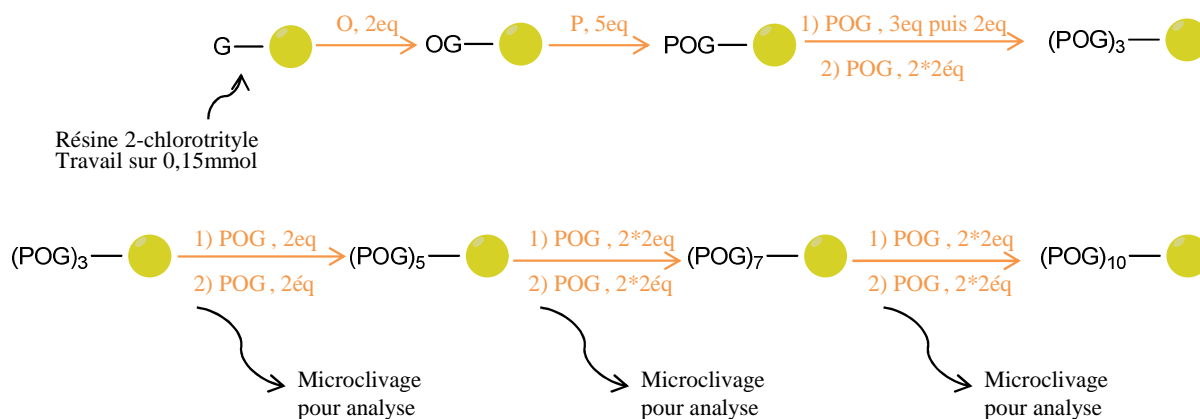


Figure 5: Schéma de synthèse par blocs à la main

Nous avons essayé d'optimiser les premières étapes car, dès le premier clivage, nous avons des profils HPLC révélant la présence de plusieurs composés. Malgré plusieurs essais d'optimisation (changement des temps de couplages, augmentation des nombres d'équivalents, diminution du loading de la résine), nous avons eu de très faibles rendements. Au vu de ces résultats décevants, nous nous sommes tournés vers une incorporation acide aminé par acide aminé pour synthétiser les plus longs peptides (POG)<sub>7</sub> et (POG)<sub>10</sub>. Des purifications extensives des bruts obtenus par cette synthèse par blocs nous ont permis d'isoler suffisamment de peptides (POG)<sub>3</sub> et (POG)<sub>5</sub> pour mener à bien leurs études structurales.

*Synthèse acide aminé par acide aminé* – Pour cette synthèse nous avons utilisé un synthétiseur automatique de peptides.

Procédure : La résine utilisée est de type Fmoc 2-chlorotrityle (loading : 0.51mmol/g) préchargée avec une glycine. La synthèse a été réalisée sur 0.25mmol. Les Fmoc-acides aminés (2,5eq) ont été ajoutés un à un pendant 1h20. Entre chaque ajout une déprotection de 40 min a eu lieu. Si le détecteur UV relié au synthétiseur ne détecte pas de signal à 280 nm alors il réitère la déprotection pendant 40 min, et si celle-ci est toujours infructueuse, il réalise une acétylation après le couplage suivant. Ensuite, nous avons réalisé un micro-clivage du peptide (POG)<sub>7</sub> avec 1ml d'une solution de TFA:H<sub>2</sub>O:TIS

(95:2,5:2,5) pendant 1h. Puis, le TFA a été évaporé et le peptide a été précipité dans l'éther diéthylique froid (3\*10ml). La solution est centrifugée à 4000 rpm pendant 6 min. Le culot est repris dans de l'eau puis lyophilisé. Le peptide a été analysé par HPLC puis par spectrométrie de masse MALDI-TOF, le peptide attendu (POG)<sub>7</sub> était présent (calculé : 1887,86g.mol<sup>-1</sup>, trouvé : 1888,96 g.mol<sup>-1</sup>(M+H<sup>+</sup>)). La synthèse a été poursuivie pour faire le peptide (POG)<sub>10</sub>. Le peptide attendu a été obtenu avec un rendement très faible de 16 % (calculé : 2689,23g.mol<sup>-1</sup>, trouvé : 2690,24 (M+H<sup>+</sup>)).

Plusieurs tentatives de synthèse ont été effectuées, aucune n'ayant permis d'améliorer ce mauvais rendement. Ce problème de synthèse n'a pu être résolu que bien plus tard, lorsqu'un troisième lot d'hydroxyproline a été utilisé. Malgré des vérifications par RMN qui n'avaient pas révélé d'anomalies, les deux premiers lots utilisés présentaient probablement des défauts préjudiciables pour la synthèse peptidique. La synthèse a donc pu être relancée dans les conditions exposées ci-dessus pour obtenir le peptide (POG)<sub>7</sub> avec un rendement de 60%. Le (POG)<sub>10</sub> a quant à lui été acheté auprès de PeptaNova© ou Synpeptide©.

### 1.2.2 Caractérisation par dichroïsme circulaire

L'analyse par dichroïsme circulaire (CD) est une technique précieuse pour l'analyse de la structure des peptides en solution.<sup>50</sup> Elle repose sur la capacité qu'ont les composés optiquement actifs à absorber de façon inégale la lumière polarisée circulairement à droite de la lumière polarisée circulairement à gauche (Figure 6). L'absorption préférentielle de l'une de ces deux polarisations résultera en une déviation de la résultante qui forme une ellipse (Figure 6B) au lieu d'une lumière polarisée plane (Figure 6A). Pour les protéines, l'absorption de la lumière polarisée entre 160 nm et 260 nm est principalement due à la présence de la liaison peptidique, pour laquelle on observe les transitions n→π\* vers 220 nm et π→π\* vers 190 nm. La mesure de la déviation θ à plusieurs longueurs d'onde permet des signatures caractéristiques pour chaque type d'élément de structure secondaire. Un exemple est donné en Figure 6 pour les cinq structures secondaires trouvées le plus souvent dans les protéines.

Cette mesure est généralement reportée en ellipticité molaire par résidu  $[\theta]$  ( $\text{deg.cm}^2.\text{dmol}^{-1}.\text{res}^{-1}$ ):

$$[\theta]_{\lambda} = \frac{100 * \theta_{\lambda}}{c * l * n}$$

$\theta$ : ellipticité (degrés)  
 $c$ : concentration ( $\text{mol.l}^{-1}$ )  
 $l$ : longueur de la cuve (cm)  
 $n$ : nombre de liaisons peptidiques

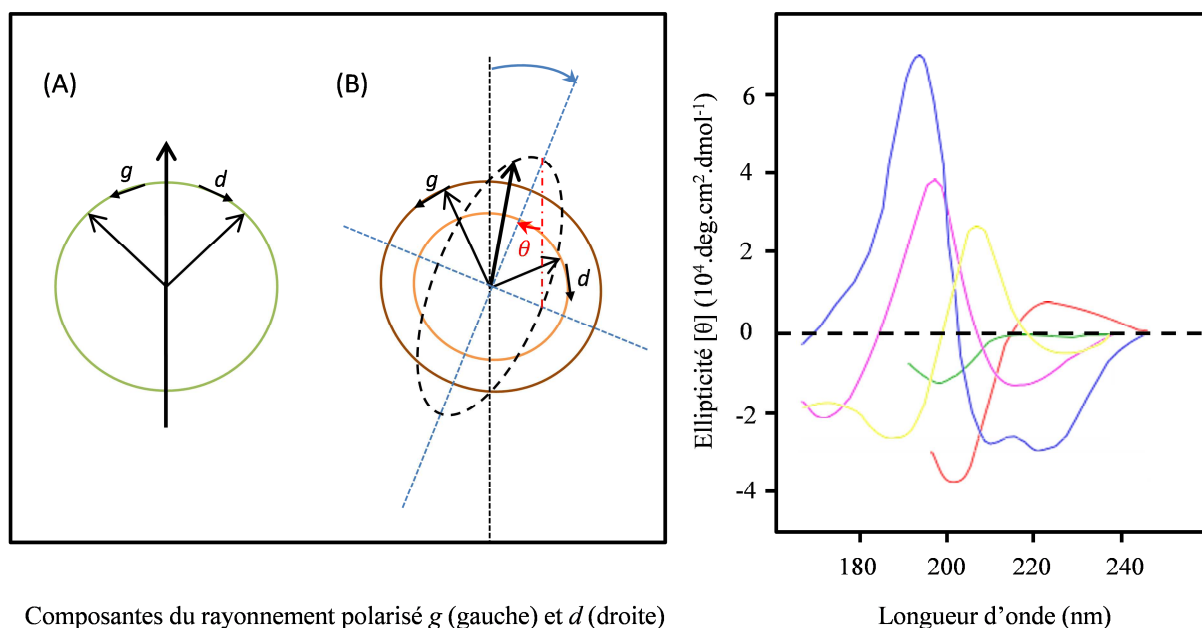


Figure 6: A gauche-Illustration de l'effet du dichroïsme, (A) absorption de la lumière égale (B) absorption de la lumière préférentielle. A droite-Spectres dichroïques caractéristiques de l'hélice  $\alpha$  (bleu), du feuillet  $\beta$  (rose), du coude  $\beta$  (jaune), de l'hélice polyproline II (rouge) et en absence de structure privilégiée (random coil en vert)

Dans le cadre des peptides CMP, nous nous sommes intéressés au suivi de la structure polyproline II (PPII) en fonction de différents paramètres. Celle-ci est caractérisée en CD par un maximum à 225 nm ainsi qu'un minimum à 200 nm (spectre tracé en rouge sur la Figure 6). Cette signature spectrale sera particulièrement visible au sein de la triple hélice. De plus, il a été montré que le rapport  $R_{pn} = |[\theta]_{\text{max}}/[\theta]_{\text{min}}|$  pouvait donner une estimation de la « triple hélicité », représentant la fraction de peptide présente sous forme de triple hélice.<sup>51-53</sup> Il a été montré que la longueur du peptide influençait la stabilité de la triple hélice, ce que nous avons voulu observer dans une première étude.<sup>54</sup>

*Analyse des peptides en fonction de la taille* – Nous avons tenté de mettre en évidence l'influence directe de la taille des peptides sur la structuration en triple hélice ; comme

évoqué précédemment, il avait été publié dans la littérature que 5 répétitions du triplet POG étaient suffisantes pour former une triple hélice.<sup>25</sup> Les peptides (POG)<sub>3</sub>, (POG)<sub>5</sub>, (POG)<sub>7</sub> et (POG)<sub>10</sub> ont ainsi été analysés par dichroïsme circulaire, les spectres sont présentés dans la Figure 7.

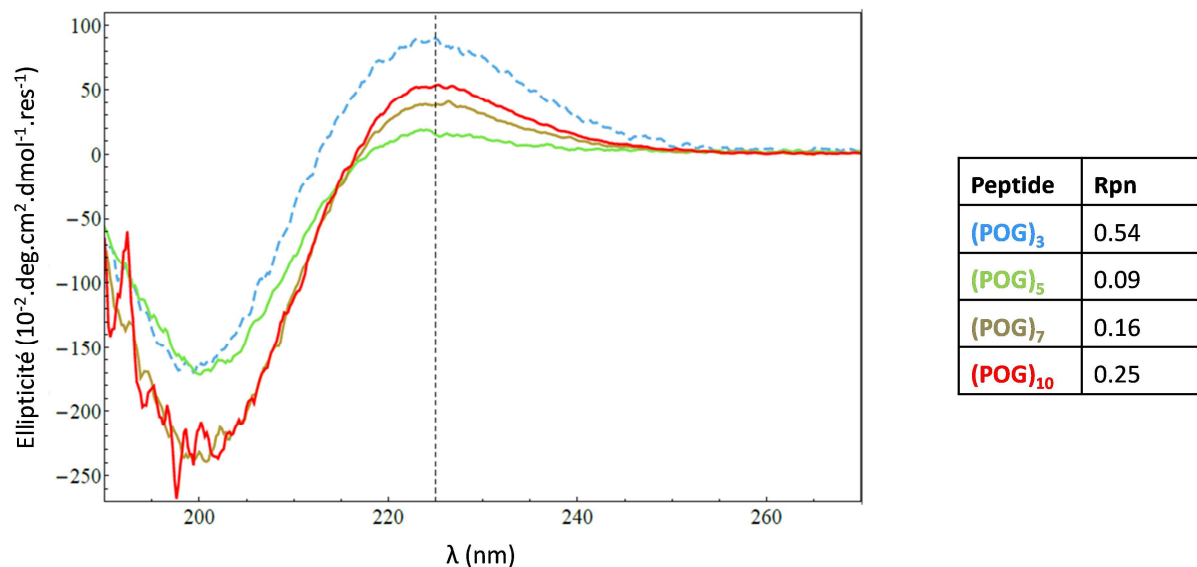


Figure 7: Spectres dichroïques de peptides POG<sub>n</sub>, 200μM, pH=7, 20°C

Le peptide (POG)<sub>3</sub> se comporte de manière particulière par rapport aux trois autres peptides. Les spectres de (POG)<sub>5</sub>, (POG)<sub>7</sub> et (POG)<sub>10</sub> montrent que le signal correspondant au PPII augmente avec la longueur du peptide. Le rapport Rpn  $[\theta]_{225}/[\theta]_{200}$  a été calculé pour ces trois peptides et indique une proportion croissante de triple hélice de 0.09 pour (POG)<sub>3</sub> à 0.25 pour (POG)<sub>10</sub>. A titre de comparaison, le collagène naturel de type I possède une valeur Rpn=0.11,<sup>52</sup> ce qui indique que (POG)<sub>10</sub> possède un taux important de triple hélice. Le fait que les spectres de (POG)<sub>5</sub>, (POG)<sub>7</sub> et (POG)<sub>10</sub> passent par un point isobestique indique qu'ils résultent essentiellement de la somme de deux conformations en proportions variables. La signature de la conformation PPII a été identifiée, celle-ci étant vraisemblablement convoluée avec la signature spectrale « random coil » (Figure 6 droite). Il est intéressant de noter que le spectre du composé (POG)<sub>3</sub> ne passe pas par ce point isobestique. Cela implique nécessairement qu'il subit un échange conformationnel avec un troisième type de structure secondaire. En particulier, on pourrait penser à une proportion non négligeable de conformation PPI. Celle-ci correspond à une hélice droite où toutes les liaisons peptidiques sont en conformation *cis* et est caractérisée par un signal dichroïque très intense entre 200 et 230 ppm. La liaison amide Xaa-Pro adoptant à près de 20% la

conformation *cis*<sup>55</sup> au sein de peptides courts, le spectre de (POG)<sub>3</sub> pourrait être influencé par une faible proportion de structure PPI comme cela a été mis en évidence sur les spectres CD de peptides plus longs solubilisés dans le n-propanol.<sup>56</sup>

Dans la littérature, les études CD sont principalement réalisées dans le tampon PBS (Phosphate Buffer Saline : NaCl 150 mM, Na<sub>2</sub>HPO<sub>4</sub> 10mM, NaH<sub>2</sub>PO<sub>4</sub> 10mM, pH=7) car il se rapproche des conditions physiologiques et n'absorbe pas aux faibles longueurs d'ondes. Les peptides étudiés ne comportent pas d'acide aminé sensible au pH, en revanche, leurs terminaisons ne sont pas protégées et possèdent un pKb d'environ 10.6 pour l'ammonium de la proline en N-terminal et de 3.3 pour la fonction acide du résidu glycine en C terminal.<sup>57</sup>

*Analyse des peptides en fonction du pH* – Afin de mettre en évidence l'effet de la protonation des peptides, nous avons fait des analyses dans un tampon phosphate à pH neutre (pH= 7) et dans un tampon acétate proche du pKa des peptides POG (pH=3) (Figure 8).

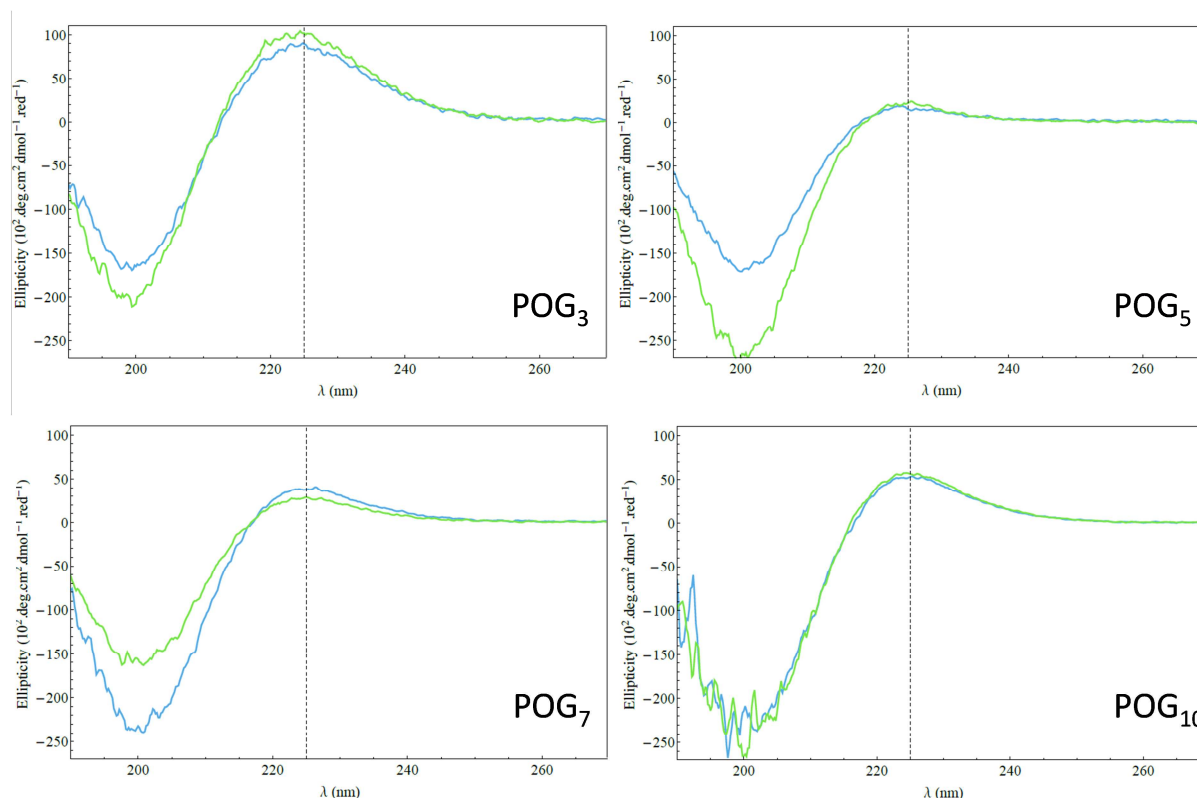


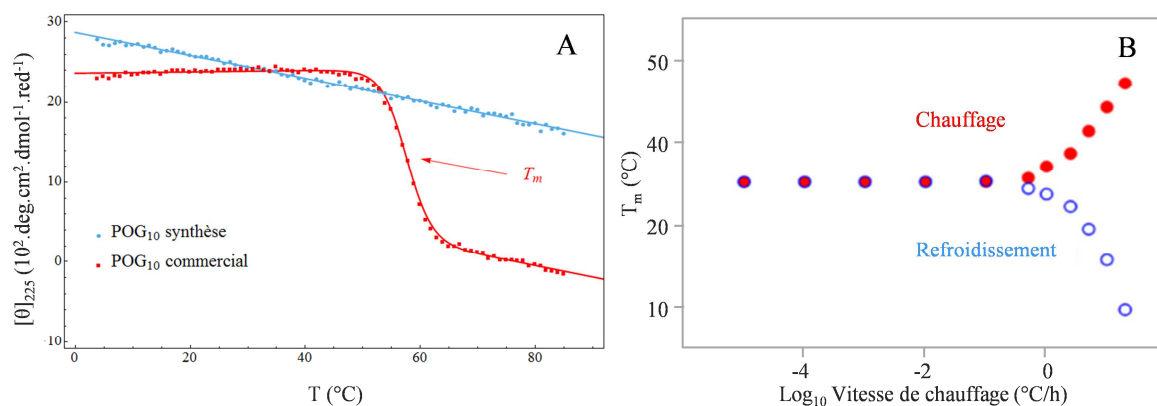
Figure 8: Spectre dichroïque de peptide (POG)<sub>n</sub>, 200μM, 20°C, pH=7(bleu) et pH=3 (vert)

On observe des différences subtiles entre les spectres enregistrés à pH=7 et à pH=3 qui dépendent du CMP étudié. Les deux spectres de (POG)<sub>3</sub> ont des allures similaires, celui

obtenu à pH 3 étant légèrement plus intense qu'à pH 7. Cette observation serait compatible avec une proportion « random coil » plus importante à pH 7, la présence de la charge négative sur la fonction carboxylate du résidu Gly terminal déstabilisant sensiblement les conformations PPII et PPI en équilibre. En revanche, les spectres du peptide (POG)<sub>10</sub> sont pratiquement superposables ce qui semble indiquer que la déprotonation de l'acide carboxylique C-terminal n'interfère pas avec la formation de la triple hélice. On pourrait aussi imaginer que les pKa des glycines sont influencés par le rapprochement des trois chaînes au sein de la triple hélice de (POG)<sub>10</sub> et que leur état de protonation reste le même entre pH=7 et pH=3. Les comportements des peptides (POG)<sub>5</sub> et (POG)<sub>7</sub> sont difficilement rationalisables dans la mesure où le tampon acétate stabilise la conformation PPII pour (POG)<sub>5</sub> mais la déstabilise pour (POG)<sub>7</sub>. On remarque cependant que la valeur Rpn est sensiblement la même aux deux pH pour (POG)<sub>7</sub>, ce qui peut indiquer que l'équilibre monomère/triple hélice n'est pas perturbé par le changement de charge de la glycine terminale. En revanche, la faible valeur Rpn = 0.09 obtenue à pH=7 pour (POG)<sub>5</sub> est encore abaissée à pH= 3, ce qui semble indiquer que le faible taux de triple hélice présent disparaît totalement au profit de la conformation PPII.

*Caractérisation de la triple hélice* – La formation de l'hélice PPII est liée à la présence des prolines qui fixent l'angle dièdre  $\phi$  et elle est guidée par des interactions stériques entre chaque cycle pyrrolidine et le résidu qui le précède. De cette façon, la formation d'hélices PPII est considérée comme un événement de pliage très local, donc non coopératif. Ce phénomène se traduit en dichroïsme circulaire par une décroissance linéaire du signal caractéristique du PPII (à 225nm) lorsque la température augmente.<sup>58</sup> Le dichroïsme circulaire est très utile pour suivre la fusion des assemblages en triples hélices car en présence d'une triple hélice, le signal PPII est stabilisé et un plateau est observé sur une certaine gamme de température (dépendante de chaque peptide). Au-delà, nous observons une chute brutale du signal PPII qui définit la température de fusion de la triple hélice ( $T_m$ ), caractéristique de sa stabilité, suivi d'une décroissance linéaire du signal, caractéristique du PPII seul (Figure 9). Afin de quantifier ce phénomène, il est indispensable de bien choisir la rampe de chauffage pour l'étude : si la rampe de chauffage est trop rapide, la chute du signal est moins brutale et donc la  $T_m$  sera différente comme le montre la Figure 9B.<sup>59</sup>





**Figure 9:** Caractérisation de la triple hélice par dichroïsme circulaire  $[\theta]_{225} = f(T)$  A.  $(\text{POG})_{10}$ ,  $100\mu\text{M}$ ,  $\text{pH}=7$ . B. Influence de la vitesse de chauffage sur la valeur de la  $T_m$  faite sur  $(\text{Gly-Pro-Pro})_{10}$ , issue de la publication de Mizuno *et al.*<sup>59</sup>

Cette première analyse en température a permis de montrer que le degré de pureté du peptide  $(\text{POG})_{10}$  synthétisé au laboratoire n'était pas suffisant. En effet, pour le peptide  $(\text{POG})_{10}$  commercial, nous observons tout d'abord un plateau reflétant la stabilité de la conformation PPII au sein d'une triple hélice, puis une chute rapide du signal suivie d'une décroissance linéaire caractéristique de l'hélice PPII ( $\Delta[\theta]/\Delta T = -30 \text{ deg.cm}^2.\text{dmol}^{-1}.\text{ }^\circ\text{C}^{-1}$ ).<sup>60,61</sup> Pour le peptide  $(\text{POG})_{10}$  synthétisé acide aminé par acide aminé, nous n'observons qu'une décroissance linéaire du signal dans des conditions opératoires identiques. Ce comportement montre que le peptide synthétisé n'est pas structuré en triple hélice, y compris à basse température. Il est tout de même important de constater que la pente  $\Delta[\theta]/\Delta T$  associée correspond à celles obtenues pour les peptides structurés en hélice PPII. Cependant, les courbes ne se superposent pas, ceci peut s'expliquer par l'impureté du peptide. En effet, cette impureté, que nous n'avons pas pu caractériser, possédant une masse de  $720.33\text{g.mol}^{-1}$  pourrait avoir un signal constant à 225 nm, indépendant de la température et donc surestimer la proportion d'hélice PPII.

Il est également possible de suivre la formation de la triple hélice en fonction du temps. Pour ce faire, il est impératif de dissocier toute la triple hélice en préchauffant l'échantillon à une température supérieure à sa température de fusion, puis on mesure le signal du PPII à 225 nm en fonction du temps et à une température donnée (Figure 10A).

Les données expérimentales peuvent ensuite être modélisées avec une fonction exponentielle croissante tel que :

$$[\theta]_{225,t} = a * \left(1 - e^{-\frac{(t-t_0)}{t_1}}\right)$$

Où  $[\theta]_{225,t}$  est l'ellipticité molaire par résidu à 225 nm en fonction du temps  $t$ ,  $t_0$  est le temps mort entre la fin du préchauffage et le début de l'enregistrement des données et  $t_1$  est le temps caractéristique de la vitesse de formation de la triple hélice.

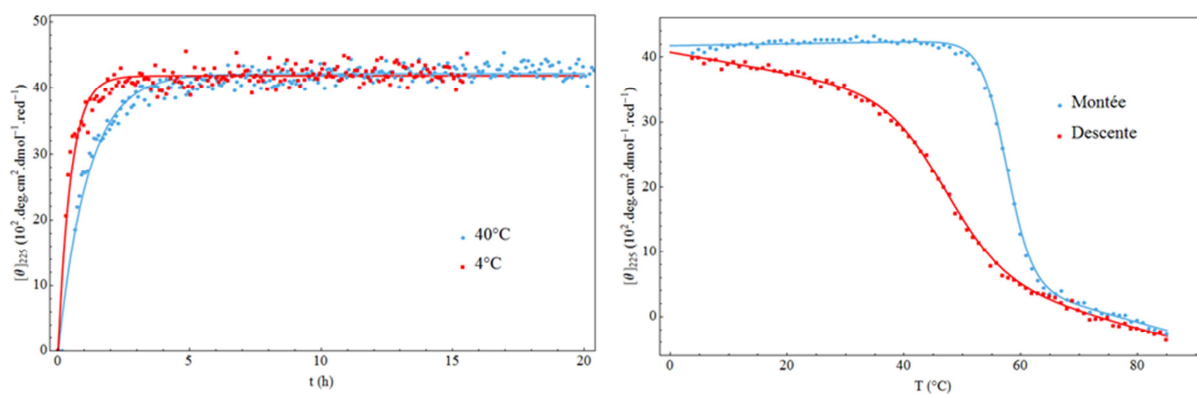


Figure 10: A. Suivi de la formation de la triple hélice en fonction du temps, (POG)<sub>10</sub>, 100µM, pH=7, B. Suivi de la déstabilisation et formation de la triple hélice en fonction de la température, (POG)<sub>10</sub>, 100µM, pH=7.

On voit alors que la triple hélice ne se forme pas instantanément, les temps  $t_1$  caractéristiques de formation sont de 3 h et 5h à 4°C et à 40°C, respectivement. Au-delà de la dépendance en température, ce résultat montre que la formation de la triple hélice dépend étroitement du temps. Ce temps nécessaire pour former la triple hélice est également visible en enregistrant un cycle de chauffage puis refroidissement (Figure 10B). On observe que les deux courbes ne se superposent pas, ce qui est dû à un phénomène d'hystérésis.<sup>56,59</sup> Le terme *hystérésis* vient du mot grec « *de hysteros* » (pour plus tard) et signifie que l'effet persiste même après sa cause abolie. Ce phénomène peut être atténué par une diminution de la vitesse de refroidissement. La triple hélice ne se reformant pas instantanément, il faut laisser un certain temps au système pour revenir dans le même état. C'est pourquoi par la suite, chacun de nos échantillons a été incubé pendant 24h à 4°C avant chaque analyse. Ce phénomène d'hystérésis est également observable dans les systèmes d'échanges PPI/PPII.<sup>56</sup> On note toutefois que le système PPII/ Triple hélice est réversible, ce qui n'est pas le cas pour le collagène naturel.<sup>62</sup>

La stabilité de la triple hélice est un paramètre que l'on peut évaluer par dichroïsme circulaire, notamment par la détermination de la température de fusion de la triple hélice.

La courbe correspondant à l'ellipticité molaire en fonction de la température peut être modélisée suivant l'équation de Richards :

$$[\theta]_T = a * T + \theta_f + \frac{b * T + \theta_i}{1 + e^{\frac{(T-T_m)}{c}}}$$

Où  $T_m$ , température de fusion, correspond à la température pour laquelle la moitié de la triple hélice est dissociée,  $[\theta]_T$  est l'ellipticité molaire par résidu à une température donnée  $T$ ,  $[\theta]_f$  est l'ellipticité molaire par résidu correspondant à l'extrapolation à 0 de la droite après la fusion de la triple hélice (Figure 10) avec  $[\theta]_i + [\theta]_f$  correspondant l'ellipticité molaire par résidu initiale,  $a$ ,  $b$  et  $c$  sont des paramètres ajustables qui correspondent aux pentes avant et après la fusion de la triple hélice ( $a+b$  et  $a$ , respectivement) et  $c$  caractérise la coopérativité du phénomène de fusion.

Les températures de fusion obtenues pour les peptides (POG)<sub>7</sub> et (POG)<sub>10</sub> en fonction de la concentration sont présentées dans le Tableau 1 (cf. §B p.72 pour les spectres correspondants)

Peptide	Concentration	$T_m$ (°C)
(POG) <sub>7</sub>	100 $\mu$ M	26
	10 mM	39,5
(POG) <sub>10</sub>	100 $\mu$ M	57
	10 mM	Précipitation

**Tableau 1: Températures de fusion obtenues par dichroïsme circulaire pour les peptides (POG)<sub>7</sub> et (POG)<sub>10</sub>, pH=7**

Ces mesures confirment que la stabilité de la triple hélice est dépendante de la taille du peptide, ce que nous avons déjà pu observer par la simple étude des spectres à une température donnée. Par ailleurs, elles mettent en évidence que la concentration joue également un rôle dans la stabilité de la triple hélice.

L'ensemble de ces analyses CD constituent des données de référence. Elles nous ont permis d'établir les domaines de stabilité des triples hélices, en fonction de la taille des peptides CMP, de leurs concentrations, de la température et du temps. Ces données précieuses seront utiles pour toutes les comparaisons que nous serons amenés à faire avec les résultats de Résonance Magnétique Nucléaire (RMN).

### 1.3. Etude des populations visibles en RMN liquide

La RMN, contrairement au dichroïsme circulaire, permet d'obtenir des informations locales, à l'échelle atomique. Les mesures effectuées précédemment par CD nous sont cependant très utiles pour nous placer dans des gammes de concentration ou de température compatibles avec l'observation par RMN du peptide monomérique, ou du peptide associé sous forme de triple hélice.

#### 1.3.1 Attribution des résonances

Nous nous sommes essentiellement consacrés à l'étude RMN du peptide (POG)<sub>7</sub> pour des raisons techniques. En effet, la température de fusion élevée du peptide (POG)<sub>10</sub> était incompatible avec l'utilisation de la cryosonde équipant notre spectromètre RMN. Des spectres <sup>1</sup>H et <sup>1</sup>H-<sup>13</sup>C HSQC ont été enregistrés et ont permis une attribution globale des différents types d'acides aminés, sur la base des déplacements chimiques connus des différents protons et carbones des trois résidus proline, hydroxyproline et glycine (Figure 11).

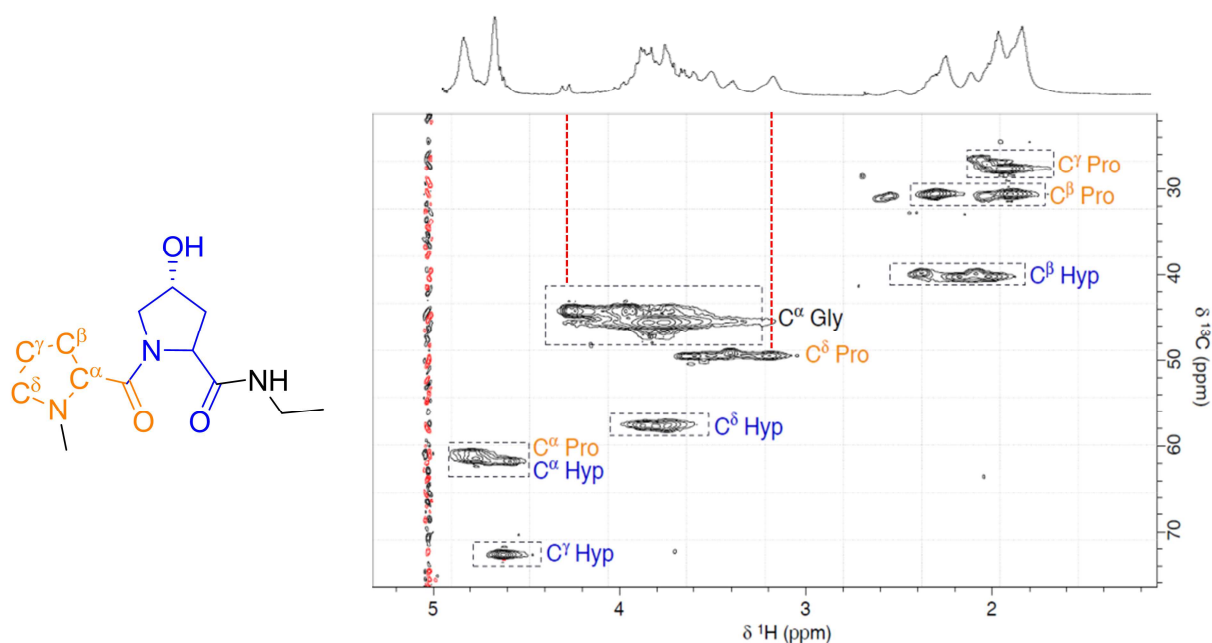
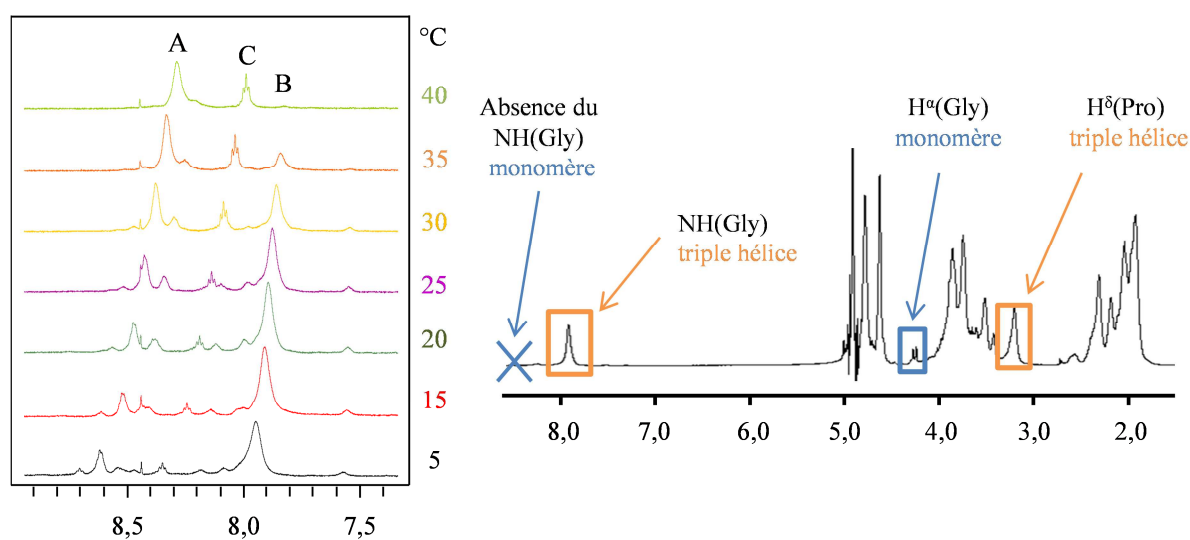


Figure 11: Attribution des signaux caractéristiques du peptide (POG)<sub>7</sub>, 1mM, 500 MHz, 5°C, 100% D<sub>2</sub>O, pH 7. Les traits pointillés mettent en évidence l'absence de recouvrement des H<sup>α</sup>(Gly) et H<sup>δ</sup>(Pro)

L'expérience HSQC est une expérience hétéronucléaire qui permet de relier chaque carbone à son ou ses protons *via* le couplage scalaire <sup>1</sup>J<sub>CH</sub>. Elle permet de mettre en évidence les protons diastéréotopes, mais aussi de différencier les protons dégénérés en

fréquence grâce à la grande dispersion des signaux dans la dimension carbone. Cependant, le spectre HSQC ne permet pas d'obtenir ici une résolution suffisante pour isoler les signaux de chaque acide aminé. La répétition de la séquence primaire des peptides  $(POG)_n$  implique en effet une dégénérescence des environnements magnétiques qui se traduit par une superposition des signaux. Nous verrons dans la partie 1.3.2 comment nous avons pu nous affranchir de ces problèmes de recouvrement et mesurer des informations structurales locales pour cette famille de peptide. L'analyse des signaux obtenus sur ce type de spectre est tout de même informative en complément des données obtenues par l'enregistrement de spectres 1D à différentes températures (Figure 12).



**Figure 12: Spectres  $^1\text{H}$ ,  $(POG)_7$ , 1mM, pH=7, 500MHz. A gauche : Agrandissement sur la zone des protons amides, 90:10  $\text{H}_2\text{O}/\text{D}_2\text{O}$ . A droite :  $\text{D}_2\text{O}$ , 20°C**

Sur les agrandissements présentés en Figure 12, correspondant à la région des protons amide des résidus glycine, nous observons trois populations majoritaires dont les proportions varient en fonction de la température. On remarque que les populations A et C augmentent en même temps que la température alors que la population B diminue. Pour chaque population, nous avons donc calculé le coefficient de température  $\frac{\Delta\delta_{NH}}{\Delta T}$ , représentant la variation du déplacement chimique des -NH en fonction de la température, afin de savoir si les NH étaient engagés dans une liaison hydrogène. La présence de liaisons hydrogène est associée à un coefficient de température élevé,  $\frac{\Delta\delta_{NH}}{\Delta T} > -6\text{ppb/K}$ . Seule la population B présente un coefficient de température supérieur à  $-6\text{ppb/K}$  (§D p.75). Nous avons donc conclu que la population B correspondait à la triple hélice, en accord avec sa disparition à température élevée et avec la persistance du signal NH dans

D<sub>2</sub>O à 20°C, signifiant un engagement dans une liaison hydrogène qui l'empêche de s'échanger avec le deutérium. Les populations A et C ont pour leur part été attribuées à deux formes monomériques. Les spectres enregistrés montrent bien que plusieurs conformères de (POG)<sub>7</sub> monomériques sont en échange lent à basse température. Certaines résonances coalescent entre 10°C et 40°C pour n'apparaître que sous la population A ou la population C. Ces deux populations restent en revanche bien distinctes à température élevée. Une barrière d'énergie connue au niveau des résidus proline correspond à l'isomérisation *cis-trans* de la liaison peptidique précédente. Cette barrière est élevée, de l'ordre de 20 kcal.mol<sup>-1</sup>, ce qui nécessiterait d'enregistrer des spectres à plus de 100°C pour observer des phénomènes de coalescence. Par intégration, le rapport des populations A et C a été évalué à 20:80, correspondant à l'équilibre *cis:trans*. Nous avons donc attribué la population A à l'ensemble des glycines suivies d'une liaison amide *trans*, et la population C aux glycines qui forment une liaison peptidique *cis* avec la proline suivante. Le petit épaulement situé au pied de la résonance A pourrait correspondre à la glycine terminale du peptide (POG)<sub>7</sub>, cette dernière ayant un environnement magnétique particulier. De la même manière, à basse température, plusieurs résonances larges sont visibles à proximité du signal correspondant aux résidus glycines de la triple hélice (population B). L'assemblage des trois monomères de CMP se faisant de manière décalée (cf. 1.5.2), les 3 protons amides des glycines terminales de la triple hélice ne seront pas chimiquement équivalents ce qui pourrait conduire au maximum à trois résonances satellites.

Une expérience complémentaire <sup>1</sup>H-<sup>1</sup>H NOESY nous a permis de relier les protons amides des glycines à leurs protons H $\alpha$  (Figure 13). On remarque que le proton amide de la population A est relié à un couple de protons H $\alpha$  diastéréotopes avec deux corrélations d'intensité différente. Cette observation nous permet de faire une attribution stéréospécifique des H <sup>$\alpha$ 2</sup> et H <sup>$\alpha$ 3</sup>. En effet, l'expérience NOESY permet de déduire des informations de distance entre protons. L'intensité de la corrélation NOE est proportionnelle à r<sup>-6</sup> où r représente la distance entre les deux atomes. Le proton le plus déblindé est alors le H <sup>$\alpha$ 2</sup>, car dans une structure canonique d'hélice PPII, il est le plus éloigné de son NH (dH<sup>N</sup>-H <sup>$\alpha$ 2</sup> = 2,63Å et dH<sup>N</sup>-H <sup>$\alpha$ 3</sup> = 2,13Å, d'après la structure PDB n°3B0S). La population C n'est que faiblement représentée à basse température et aucune corrélation n'est observée sur le spectre 2D. Des études similaires pourraient être faites à plus haute température pour caractériser les formes monomériques mais ces expériences n'ont pas été enregistrées.

Pour la population B correspondant à la triple hélice, les corrélations NOE sont très intenses. Ceci est directement lié à la taille importante de l'assemblage moléculaire. Le transfert d'aimantation NOE est relié à des phénomènes de relaxation dipolaire qui sont d'autant plus efficaces que l'assemblage moléculaire est de haut poids moléculaire. D'autre part, les transferts d'aimantations sont complexes au sein de la triple hélice en raison d'un réseau important de protons au niveau du squelette peptidique. Pour cette raison, la quantification précise et l'obtention de l'attribution stéréospécifique des protons de la glycine n'ont pas pu être effectuées. D'après la structure de triple hélice obtenue par cristallographie (PDB 1CAG), il est prédit que 9 atomes peuvent être visibles dans l'expérience NOESY à partir de la glycine ( $dH^N-X \leq 5 \text{ \AA}$ ),<sup>63</sup> ces observations résultant de corrélations inter-brins (rose) et intra-brins (vert). Il s'avère que nous observons 10 corrélations, celle non prédite correspondant à la corrélation  $H^N-H^{\beta 3}(\text{Pro})$  ( $dH^N-H^{\beta 3}(\text{Pro}) = 5,82 \text{ \AA}$ ), probablement observée en raison de phénomènes de diffusion de spins (Figure 13). Il est alors possible d'attribuer stéréospécifiquement les protons  $\text{ProH}^{\beta 2}/\text{ProH}^{\beta 3}$  et les protons  $\text{HypH}^{\beta 2}/\text{HypH}^{\beta 3}$  grâce l'intensité des corrélations NOE ( $dH^N-H^{\beta 2}(\text{Pro})=4,59 \text{ \AA}$  et  $dH^N-H^{\beta 3}(\text{Pro}) = 5,82 \text{ \AA}$  et  $dH^N-H^{\beta 2}(\text{Hyp})=3,30 \text{ \AA}$  et  $dH^N-H^{\beta 3}(\text{Hyp}) = 3,59 \text{ \AA}$ , PDB 3B0S). Les trois populations minoritaires (1, 2, 3 sur la Figure 13) ont des corrélations intenses avec les  $\text{HypH}^{\beta}$ . Cela signifie qu'elles correspondent à des glycines au sein d'une triple hélice, et donc à celles situées en bout de chaîne.

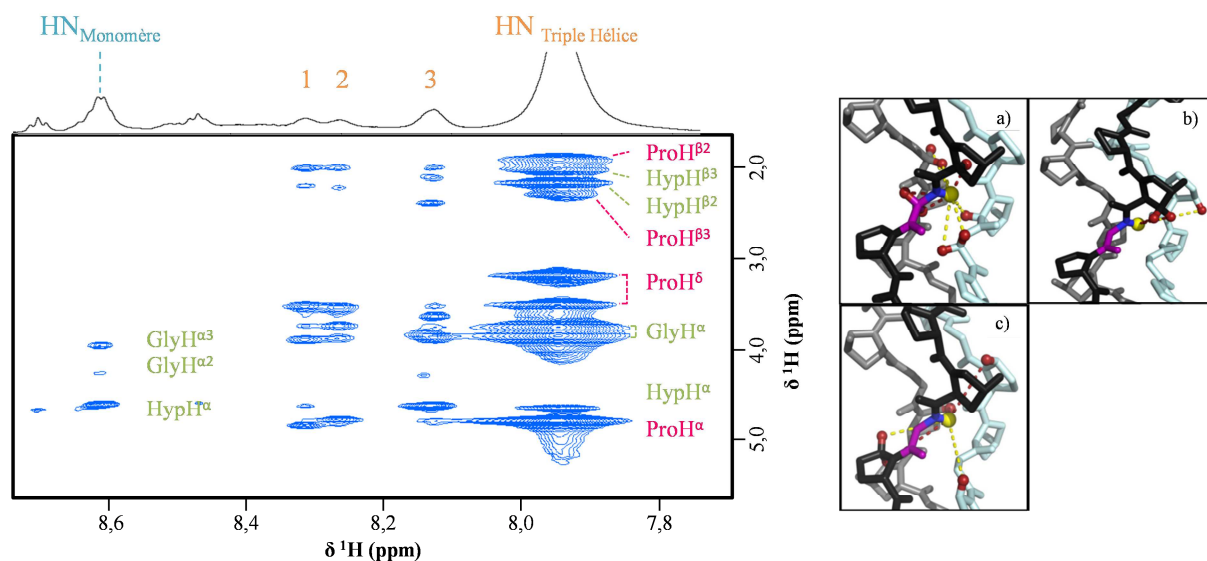


Figure 13: A gauche: Spectre  $^1\text{H}$ - $^1\text{H}$  NOESY sur le peptide  $^4(\text{POG})_7$ , 10mM, 500 MHz,  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (90:10),  $5^\circ\text{C}$ . Les corrélations inter-brins sont représentées en rose et les corrélations intra-brins sont représentées en vert. A droite : Prédiction des corrélations NOE observables dans la triple hélice grâce au marquage  $^{15}\text{N}$ -glycine. (a-d) la chaîne noire représente un site marqué: (a) représente les  $\text{H}\alpha$  de la glycine et de l'hydroxyproline; en jaune sont représentées les distances interchaines et en rouge les distances intrachaines observables; (b) représente tous les  $\text{H}\beta$  de l'hydroxyproline et de la proline; (c) représente les distances interchaines entre les glycines observables; (d) représente des atomes  $\text{H}\delta$  de la proline; 1CAG PDB. Figure tirée de la publication de *Hartgerink et al*<sup>63</sup>

L'analyse combinée des expériences  $^1\text{H}$ ,  $^{13}\text{C}$ - HSQC,  $^1\text{H}$ ,  $^1\text{H}$ - NOESY et  $^1\text{H}$ , nous a également permis de mettre en évidence la présence de protons isolés permettant la quantification des populations monomériques et triples hélices (Figure 11). Le proton à  $\delta=4.3\text{ppm}$  correspond au  $\text{H}^\alpha$  des glycines dans le monomère et le proton à  $\delta=3.2\text{ppm}$  correspond au  $\text{H}^\delta$  des prolines dans la triple hélice.

### 1.3.2 Etudes des populations monomériques et triples hélices

La RMN nous a permis d'étudier l'influence de la concentration et de la température sur les populations monomériques et triples hélices, en présentant l'avantage de pouvoir les quantifier contrairement au CD.

*Dépendance de la concentration* – Sur la Figure 14, dans la région des protons aliphatiques, nous pouvons observer l'apparition de la triple hélice lorsque la concentration augmente. Dans la région des NH, on voit également l'apparition de triple hélice, caractérisée par un déplacement chimique amide inférieur à 8 ppm à  $40^\circ\text{C}$ . Des signaux supplémentaires apparaissent, ceux-ci correspondant à des sous-populations qui seront étudiées plus en détail dans la partie suivante.



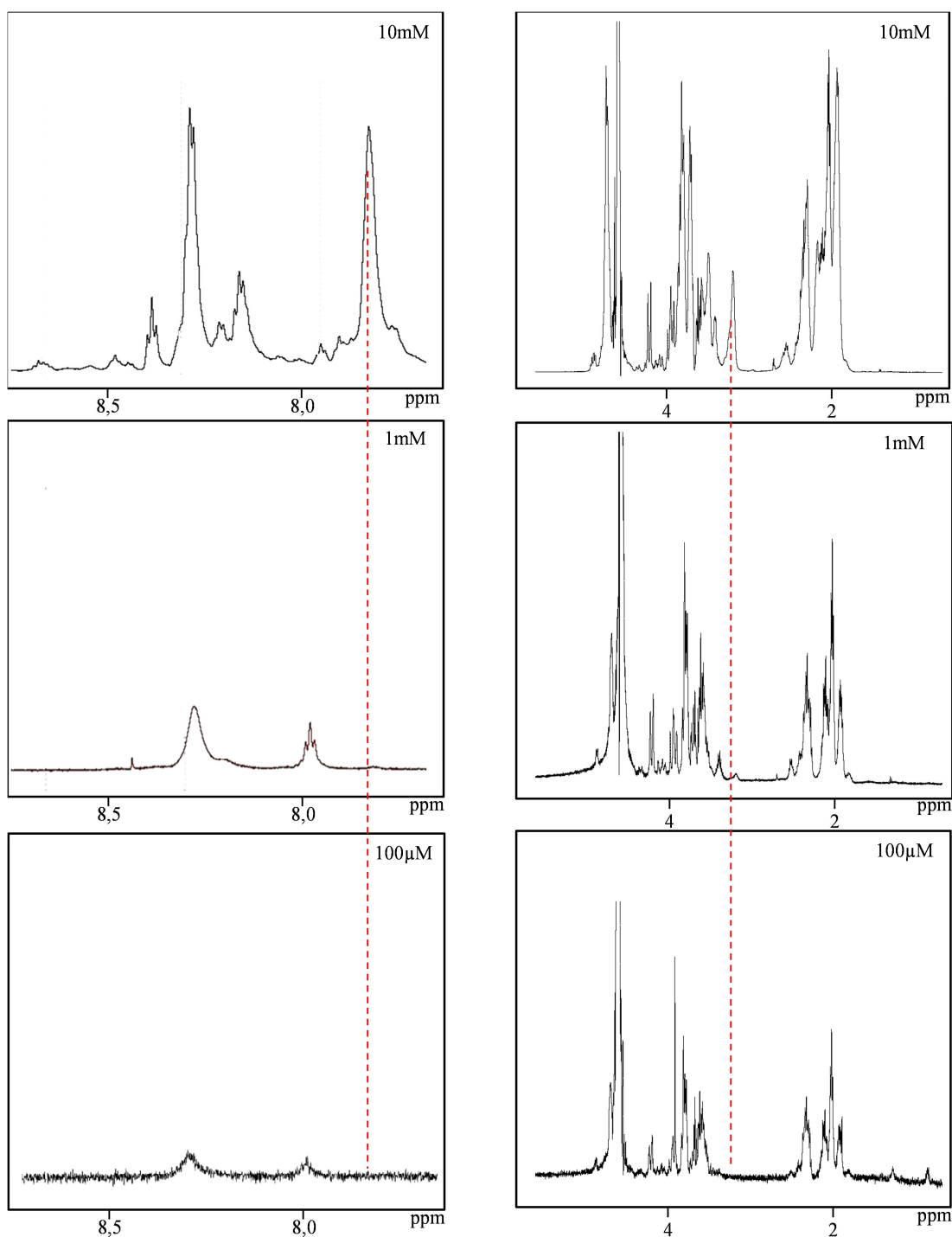


Figure 14: Influence de la concentration sur la stabilité de la triple hélice,  $(POG)_7$ ,  $40^\circ\text{C}$ ,  $\text{pH}=7$ ,  $500\text{MHz}$ .  
A gauche, zone des protons amides. A droite, zone des protons aliphatiques

L'étude en concentration nous montre donc que le rapport de concentration monomère/triple hélice dépend étroitement de la concentration totale en CMP, de manière analogue avec les molécules amphiphiles qui s'assemblent en micelles au-delà d'une CMC (concentration micellaire critique). Ici, nous observons qu'à partir d'une concentration de

30mM, le système conserve un ratio 80:20 (TH:M) (Figure 15). Ces deux populations sont en équilibre, la forme monomérique ne disparaît pas au profit de la triple hélice et inversement. Cependant, il faut nuancer cette observation en prenant en compte la limite de détection en RMN. Sur les spectres 1D  $^1\text{H}$ , nous n'observerons que les espèces de taille inférieure à  $\approx 100\text{kDa}$ . L'étude de la présence d'assemblages supramoléculaires, favorisés à concentration élevée, sera discutée dans le chapitre suivant.

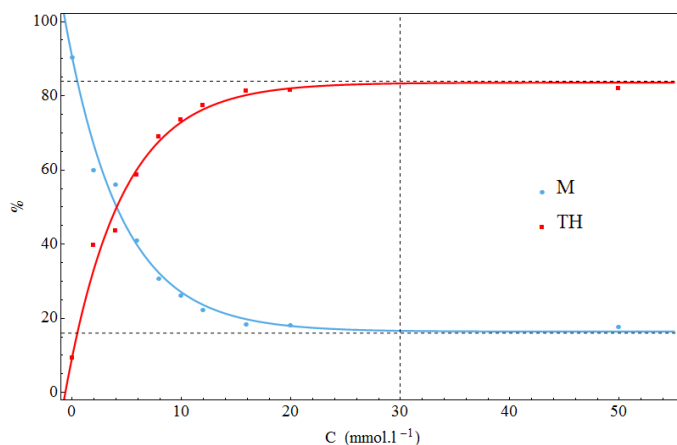


Figure 15: Evolution des populations TH et M à 40°C en fonction de la concentration

*Influence de la température* – Afin de compléter cette étude comparative, nous avons également préparé un échantillon à  $100\mu\text{M}$  afin de déterminer la  $T_m$  par RMN (Figure 16).

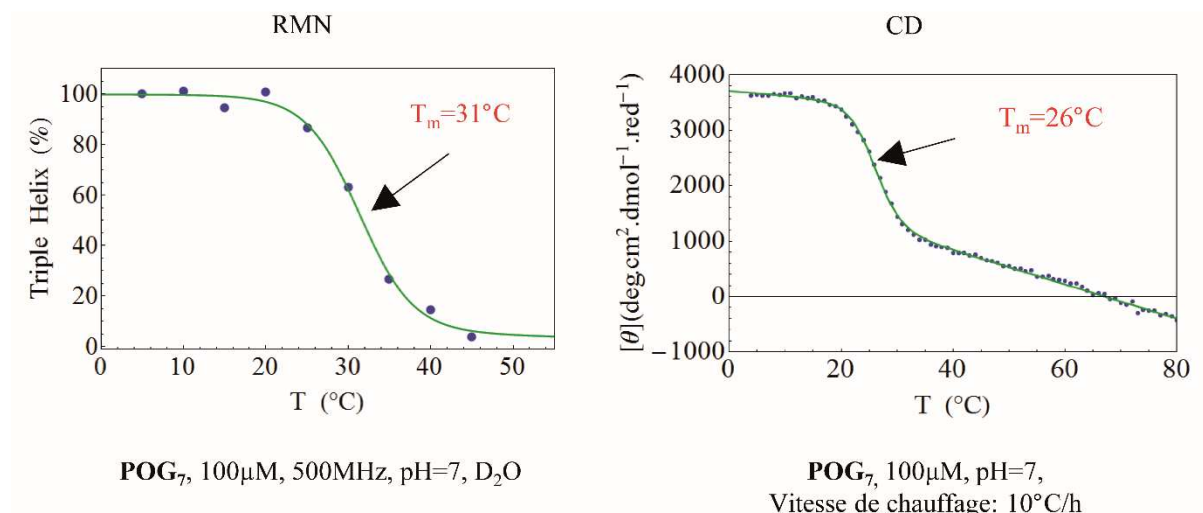


Figure 16: Comparaison des températures de fusion de la triple hélice obtenues par RMN et CD

Nous observons un décalage de  $5\text{ }^\circ\text{C}$  entre la température de fusion de la triple hélice trouvée par RMN et CD. Cette légère différence n'est pas expliquée mais pourrait provenir d'une différence de calibration en température des deux appareils ou être liée à des vitesses de chauffage différentes.

## 1.4. Marquage spécifique des résidus Glycine

Afin de palier le problème de recouvrement des signaux lié à la répétition des triplets POG, nous avons synthétisé sept peptides marqués  $^{13}\text{C}_2\text{-Gly}$  ou  $^{15}\text{N-}^{13}\text{C}_2\text{-Gly}$ . Pour chaque peptide, une seule glycine est marquée, ce qui permet leur attribution spécifique et l'extraction de paramètres structuraux relatifs à chaque position de la chaîne polypeptidique (Figure 17). La synthèse de ces peptides a été réalisée par SPPS et est présentée en partie expérimentale (cf.§C p.74). La convention utilisée pour nommer ces peptides est  $^n(\text{POG})_7$ , avec n symbolisant la position marquée, (n=1 correspond à la première position dans le sens de la SPPS Cter→Nter).

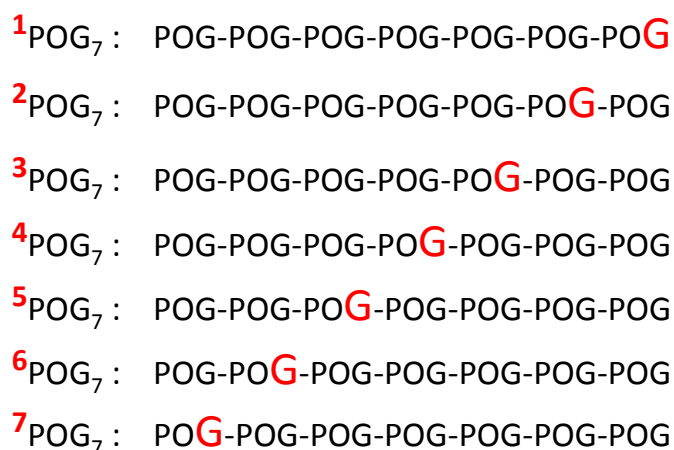
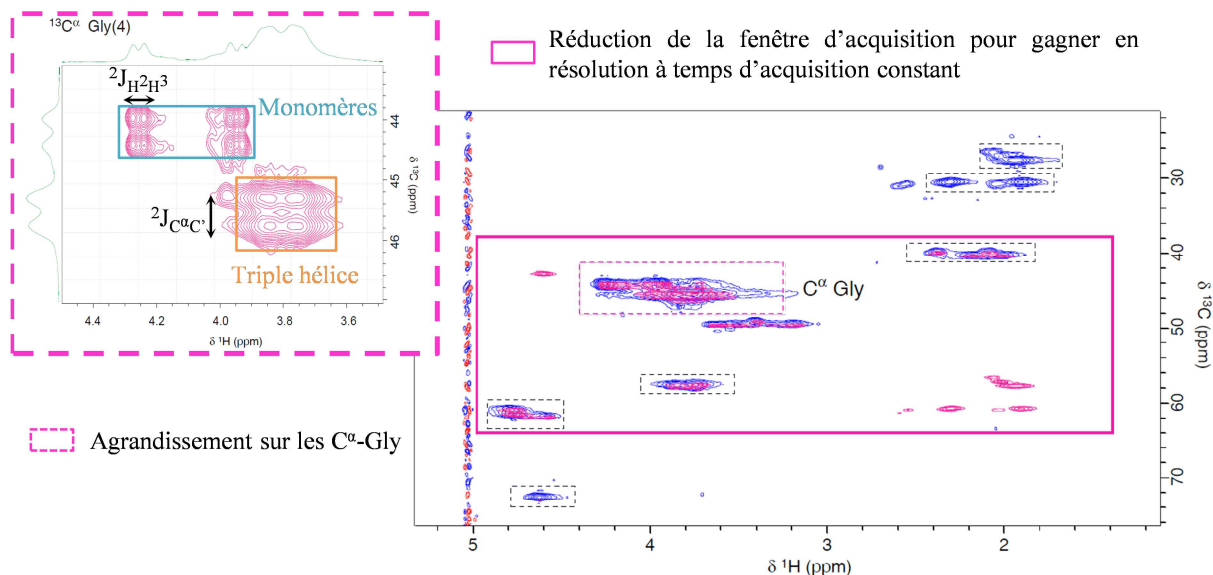


Figure 17 : Identification des glycines marquées

La glycine marquée est facilement identifiable sur le spectre  $^1\text{H}, ^{13}\text{C}$ -HSQC car elle présente d'une part des corrélations beaucoup plus intenses par rapport aux autres signaux qui sont enregistrés en abondance naturelle du  $^{13}\text{C}$ , et d'autre part, ses corrélations sont dédoublées dans la dimension carbone en raison du couplage entre le carbone  $\text{C}^\alpha$  et le  $\text{C}'$  du carbonyle ( $^1\text{J}_{\text{C}^\alpha\text{C}'}$ ) (Figure 18). La possibilité d'observer une glycine unique parmi les 7 présentes dans  $(\text{POG})_7$  permet d'obtenir des corrélations plus fines ; il n'y a pas de superposition de plusieurs signaux voisins. En outre, l'optimisation des paramètres d'acquisition nous permet d'augmenter la résolution dans la dimension carbone. Ainsi, pour le CMP marqué  ${}^4(\text{POG})_7$ , on a pu mettre en évidence deux paires de protons diastéréotopes correspondant au méthylène  $\text{C}^\alpha\text{H}_2$  de la glycine 12 marquée.

La population correspondant au monomère possède des pics de corrélations fins ( $\delta=4,24$  et  $\delta=3,95$ ), bien résolus permettant de visualiser la constante de couplage  $^2J_{H\alpha 2H\alpha 3}$  et la population correspondant à la triple hélice a des corrélations plus larges ( $\delta=3,87$  et  $\delta=3,74$ ), incompatibles avec l'analyse de leur structure fine.



**Figure 18:** Spectres  $^1\text{H}$ - $^{13}\text{C}$  HSQC enregistrés sur le peptide (POG)<sub>7</sub>, non marqué (spectre bleu) et sur le peptide  $^4$ (POG)<sub>7</sub>, marqué  $^{13}\text{C}_2$  (spectre rose). L'optimisation de la résolution permet de mettre en évidence des doublets sur les corrélations les plus intenses dans la dimension  $^{13}\text{C}$ . 1mM, 500 MHz, 100% D<sub>2</sub>O, 5°C

L'enregistrement d'expériences  $^1\text{H}$ ,  $^{15}\text{N}$ -HSQC nous permet également de mettre en évidence la présence de sous-populations (Figure 19). A nouveau, l'intégration des corrélations en fonction de la température nous informe sur l'évolution des populations même si les protons amides subissent des processus d'échange chimique importants à température élevée ce qui ne permet pas d'obtenir des quantifications précises. Nous avons en revanche calculé les coefficients de température pour chacune des populations visibles et nous pouvons voir que seule une population (repérée en orange sur la figure 18) présente un NH engagé dans une liaison H. Ceci confirme l'attribution de la triple hélice établie sur la base des spectres 1D et nous informe par ailleurs que la glycine 12 n'est pas engagée dans une liaison H pour les sous-populations visibles.

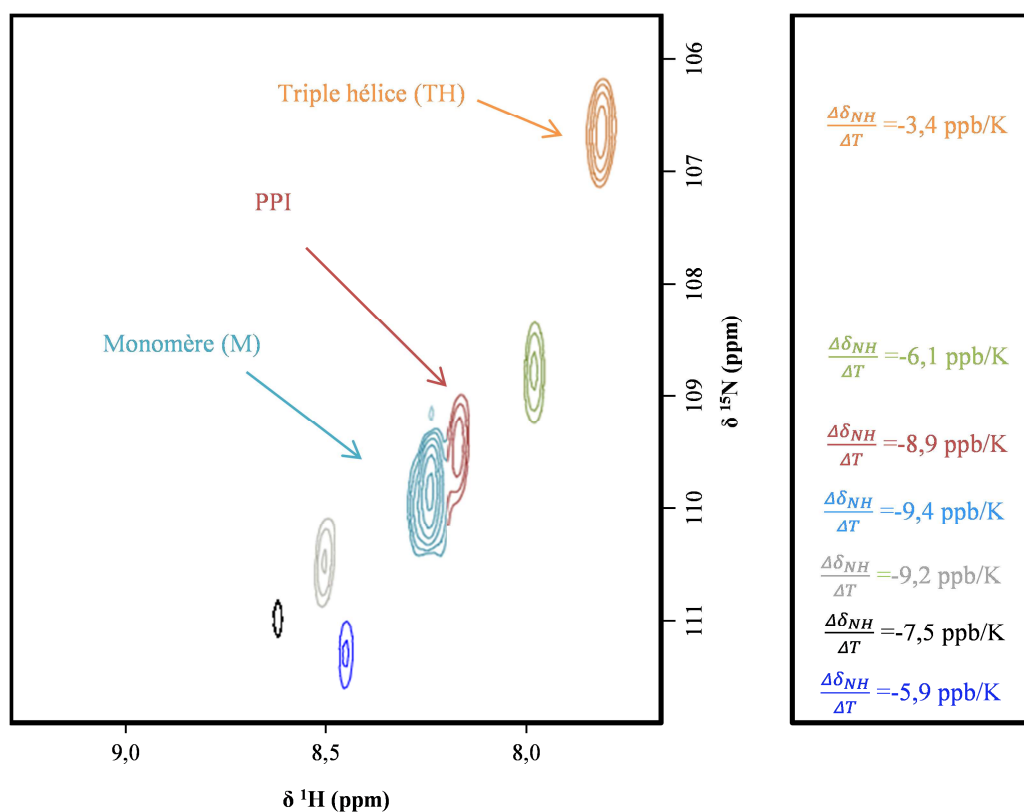


Figure 19: Spectre  $^1\text{H}, ^{15}\text{N}$ -HSQC, 500MHz,  $^4(\text{POG})_7$ , 10mM, pH=7,  $\text{H}_2\text{O}/\text{D}_2\text{O}(90:10)$ , 45°C. Trois corrélations ont été attribuées. Elles correspondent à la triple hélice, au peptide monomérique majoritaire (noté PPII), et à une troisième population notée PPI, pour laquelle la liaison peptidique Gly-Pro est en conformation *cis* (cf texte pour davantage d'explications)

*Influence de la température* – De manière similaire aux analyses 1D présentées précédemment, l'intensité des corrélations 2D renseigne sur les populations de peptides monomériques ou assemblés en triple hélice. Le gain de résolution associé aux marquages spécifiques et l'édition d'une deuxième dimension ( $^{13}\text{C}$  ou  $^{15}\text{N}$ ), nous permet de surcroît d'affiner ces analyses car les populations minoritaires, non visibles en RMN 1D, peuvent à présent être suivies en fonction de la température. Nos premières études ont été réalisées avec un échantillon  $^4(\text{POG})_7$ ,  $c=1\text{mM}$ . Puis, afin d'améliorer la sensibilité tout en limitant le temps d'analyse nous avons réalisé des études avec un échantillon à  $c=10\text{mM}$ .

Sur les spectres  $^1\text{H}-^{13}\text{C}$  HSQC, nous pouvons voir l'évolution des populations en fonction de la température ainsi que la présence d'une troisième population, notée PPI, impossible à détecter sur la seule base du spectre  $^1\text{H}$  (Figure 20). Celle-ci croît de manière identique à la population PPII en fonction de la température, dans un ratio constant de ~20% ce qui est compatible avec une forme monomérique où la glycine éditée est en relation *cis* vis-à-vis

de la proline suivante. Nous retrouverons ce ratio lors de nos études sur le triplet modèle (Pro-Hyp-Gly) qui est présenté dans le chapitre 3.

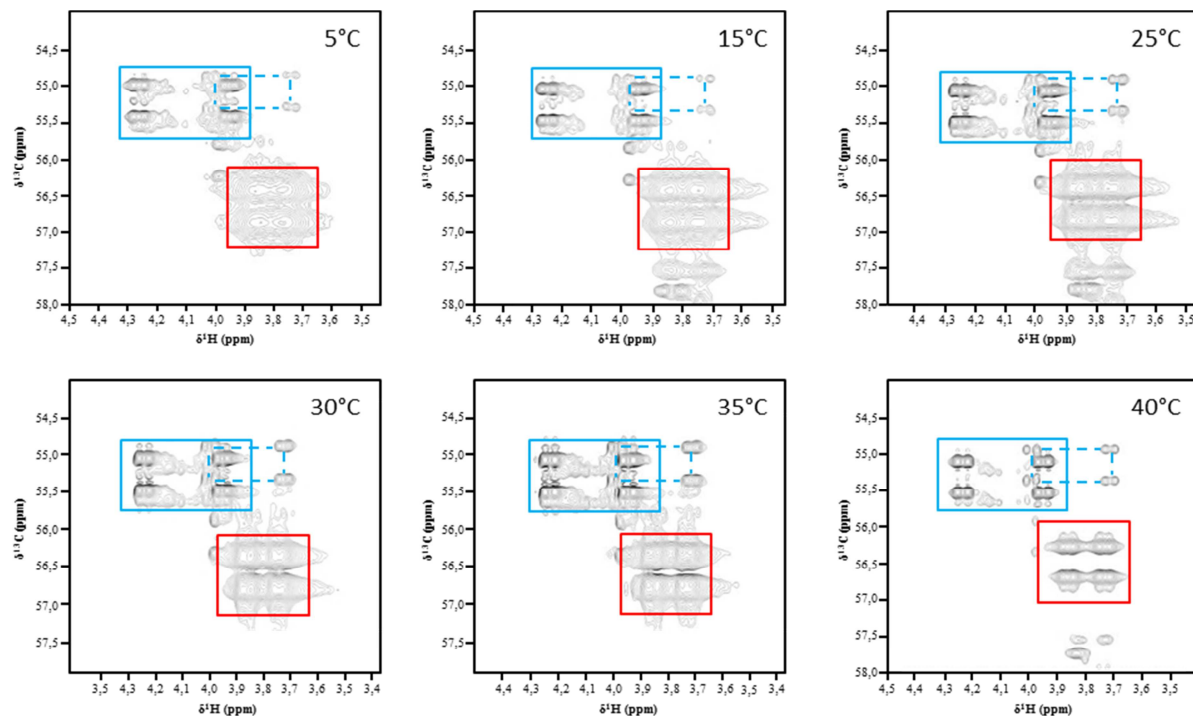


Figure 20: Evolution des populations M (bleu)/TH (rouge) en fonction de la température pour  $^4(\text{POG})_7$ . Les corrélations repérées par le cadre en pointillé correspondent à une liaison peptidique Gly<sub>12</sub>-Pro<sub>13</sub> en *cis*. Conditions : 10mM, pH=7, 500MHz, D<sub>2</sub>O

Chacune des populations a été quantifiée en fonction de la température en vue de comparer ces données aux résultats de dichroïsme circulaire et à l'analyse faite en RMN 1D (Figure 21).

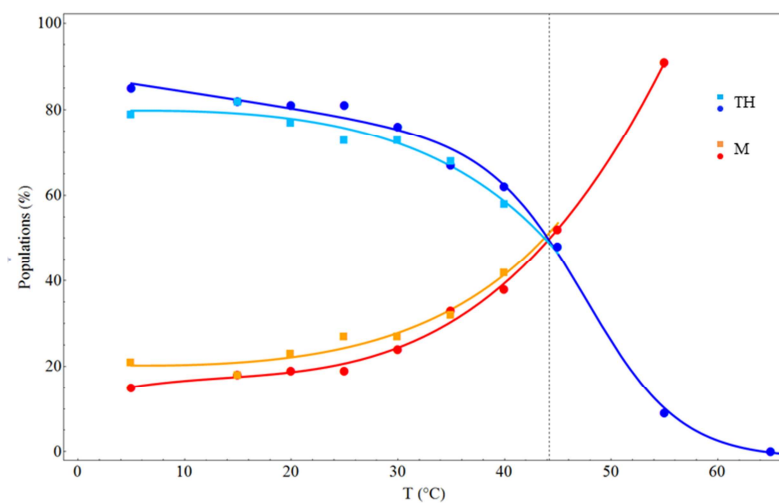


Figure 21: Quantification des populations TH et M sur le spectre 1D (•) et 2D (■) pour le peptide  $^4(\text{POG})_7$ , 10mM, pH=7, 500MHz, D<sub>2</sub>O.

Nous observons que les deux approches RMN donnent des résultats similaires en ce qui concerne la quantification. A cette concentration, la triple hélice est toujours présente à 40°C, ce qui coïncide avec la  $T_m = 39.5^\circ\text{C}$  trouvée par CD (cf. Tableau 1). Cependant, la détermination précise de la  $T_m$  par RMN est de  $44^\circ\text{C}$  ce qui est sensiblement supérieur à celle trouvée par CD. On note que nous avons également trouvé un décalage de  $\sim 5^\circ\text{C}$  sur les études RMN 1D réalisées à  $100\mu\text{M}$ .

Les quantifications obtenues sur les spectres 2D et 1D sont comparables, mais l'avantage de l'expérience 2D est la possibilité de voir les sous-populations.

*Influence de la position sur la chaîne* – L'enregistrement de spectres similaires a été effectué pour les autres peptides marqués et nous a permis d'établir les attributions précises  $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{15}\text{N}$  de chaque glycine en fonction de sa position dans la chaîne CMP et de sa forme libre (PPII) ou triple hélice. Il faut souligner que les spectres correspondant à  $^1(\text{POG})_7$ ,  $^2(\text{POG})_7$  et  $^7(\text{POG})_7$  sont bien plus complexes car au sein de la triple hélice, les trois chaînes PPII sont assemblées de manière décalée le long de l'axe de la triple hélice (décalage d'un résidu entre chaque brin), ce qui conduit à une inéquivalence chimique des trois glycines à chaque extrémité et à une multiplication des corrélations. Les spectres des peptides  $^3(\text{POG})_7$ ,  $^4(\text{POG})_7$ ,  $^5(\text{POG})_7$ , et  $^6(\text{POG})_7$  sont en revanche parfaitement superposables même en utilisant l'expérience  $^1\text{H}$ - $^{13}\text{C}$   $\text{CH}_2$ -TROSY<sup>64</sup> (cf 1.5.1) qui permet une augmentation considérable de la résolution dans les deux dimensions  $^{13}\text{C}$  et  $^1\text{H}$  (Figure 21).

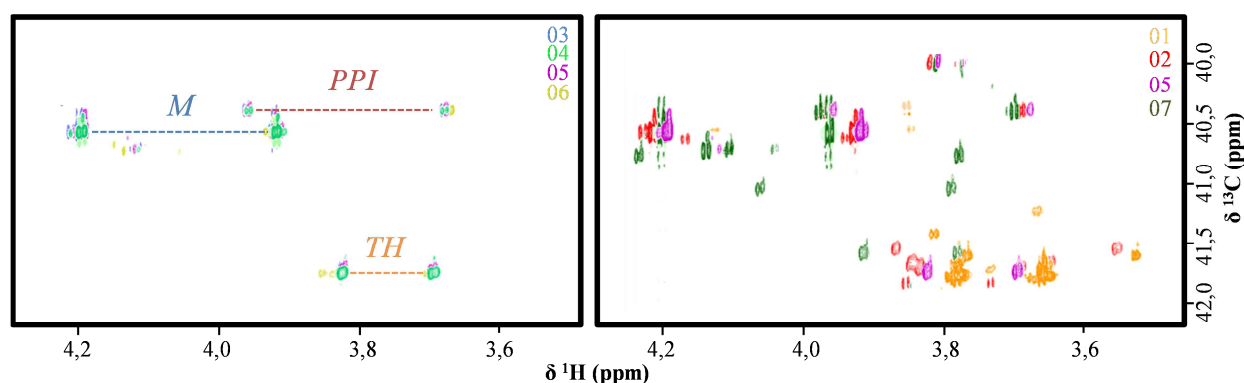


Figure 22: Spectres  $\text{CH}_2$ -TROSY (cf 1.5.1)<sup>64</sup> centrés sur la zone de déplacement chimique  $\text{C}^\alpha\text{H}_2$  des glycines, 800MHz, 10mM,  $\text{D}_2\text{O}$ , pH=7,  $40^\circ\text{C}$

Pour le peptide  $^1(\text{POG})_7$ , nous observons plusieurs populations rassemblées dans une région proche des corrélations obtenues pour les glycines centrales faisant partie de la triple hélice. On distingue principalement quatre populations distinctes dans des

proportions (1 ;0.22 ;0.17 ;0.24), ce qui pourrait correspondre au monomère PPII pour la première population et ensuite aux trois brins où chaque glycine marquée se trouve dans un environnement unique en raison de l'inéquivalence inhérente aux décalages des chaînes peptidiques au niveau des terminaisons (cf. Figure 31 qui illustre ce décalage).<sup>65</sup> De manière attendue, aucune corrélation ne correspond à une fraction PPI car cette glycine n'est pas suivie d'une liaison peptidique.

## 1.5. Etude structurale des CMP modèles

L'utilisation de peptides marqués nous a permis d'observer et de quantifier les différentes conformations adoptées par les peptides CMP mais également d'avoir accès à des informations structurales locales. L'étude structurale de CMP modèles a été entreprise par RMN en déterminant des constantes de couplage spécifiques. Ces mesures ont été réalisées grâce au développement d'expériences RMN originales, à 2 et 3 dimensions, basées sur un effet TROSY.<sup>64,66,67</sup> Des études de dynamique moléculaire ont également été menées à partir d'une structure aux rayons X en vue de comparer nos données expérimentales aux données simulées.

### 1.5.1 Utilisation d'une expérience TROSY pour accéder à des informations locales

L'étude de macromolécules de haut poids moléculaire par des méthodes de RMN liquide classique pose deux problèmes majeurs. Le premier est celui du grand nombre de résonances qui provoque un chevauchement des pics et donc une analyse très fastidieuse. Le second problème est lié à la relaxation rapide des molécules de masse importante, ce qui provoque un élargissement des raies et une diminution de la sensibilité conduisant à une perte du signal. Nous avons résolu ce premier problème grâce à un marquage sélectif de nos glycines <sup>13</sup>C et <sup>15</sup>N. La solution que nous avons envisagée pour limiter les problèmes de relaxation nucléaire a été de développer des séquences RMN basées sur un effet TROSY<sup>67</sup> pour *Transverse-Relaxation-optimized NMR spectroscopy*. Cet effet, qui provient d'interférences entre différents phénomènes de relaxation permet en effet d'obtenir des largeurs de raies satisfaisantes pour l'étude de molécules de très grandes tailles (jusqu'à 1 MDa pour l'expérience Methyl-TROSY)<sup>68</sup> (Figure 23).



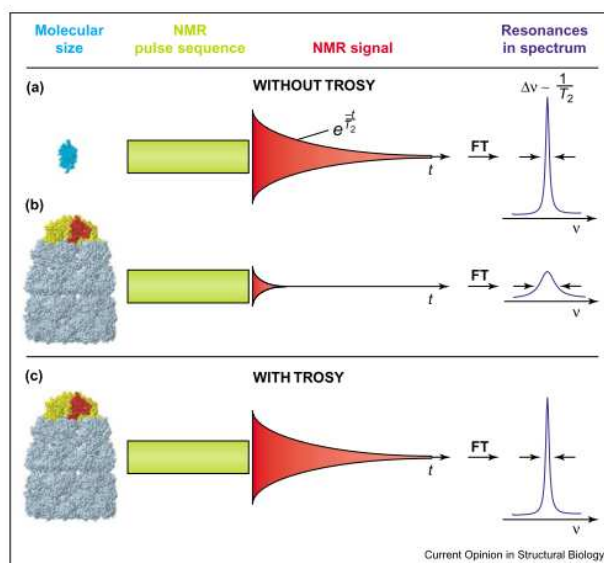


Figure 23: Effet de l'expérience TROSY. Schéma extrait de la publication de Wider et Fernandez<sup>68</sup>

L'expérience TROSY – Cette expérience permet un affinement des raies par sélection d'une corrélation spécifique. En effet, pour un spectre de corrélation  $^1\text{H}-^{15}\text{N}$ , chaque noyau peut être couplé à l'autre dans un état de spin  $\alpha$  ou  $\beta$ . Ainsi, sans découplage dans un spectre  $^1\text{H}-^{15}\text{N}$  HSQC, nous observons quatre corrélations (Figure 24). Dans le cas des petites molécules, les largeurs de raies sont identiques pour ces composantes et quatre corrélations de même intensité sont observées. Pour une grosse molécule, les largeurs de raies sont en moyenne plus larges en raison d'un temps de relaxation T2 plus court. Cependant la largeur de raie d'une des composantes est diminuée en raison de phénomènes de relaxation croisée décrit par Guéron en 1983.<sup>69</sup> En effet, la relaxation T2 est principalement due à un mécanisme de relaxation spin-spin qui résulte en grande partie de la modulation de l'interaction dipôle-dipôle par les mouvements moléculaires mais également de l'anisotropie de déplacement chimique. Ces deux phénomènes interfèrent entre eux et donnent lieu à des relaxations différentielles sur chacune des composantes des multiplets obtenues pour deux spins couplés. Lorsqu'un découplage  $^{15}\text{N}$  est réalisé, il n'est alors observé qu'une corrélation dans la dimension proton qui moyenne les propriétés de relaxation du noyau  $^{15}\text{N}$  dans ses états  $\alpha$  et  $\beta$ . Dans le cas d'une grosse molécule, il y aura alors une largeur de raie importante qui masquera une éventuelle multiplicité. L'expérience  $^1\text{H}-^{15}\text{N}$  TROSY, qui est en définitive un spectre de corrélation  $^1\text{H}-^{15}\text{N}$  dont la corrélation la plus fine a été sélectionnée, pourrait permettre en principe d'accéder aux constantes de couplages et donc d'obtenir des données structurales.

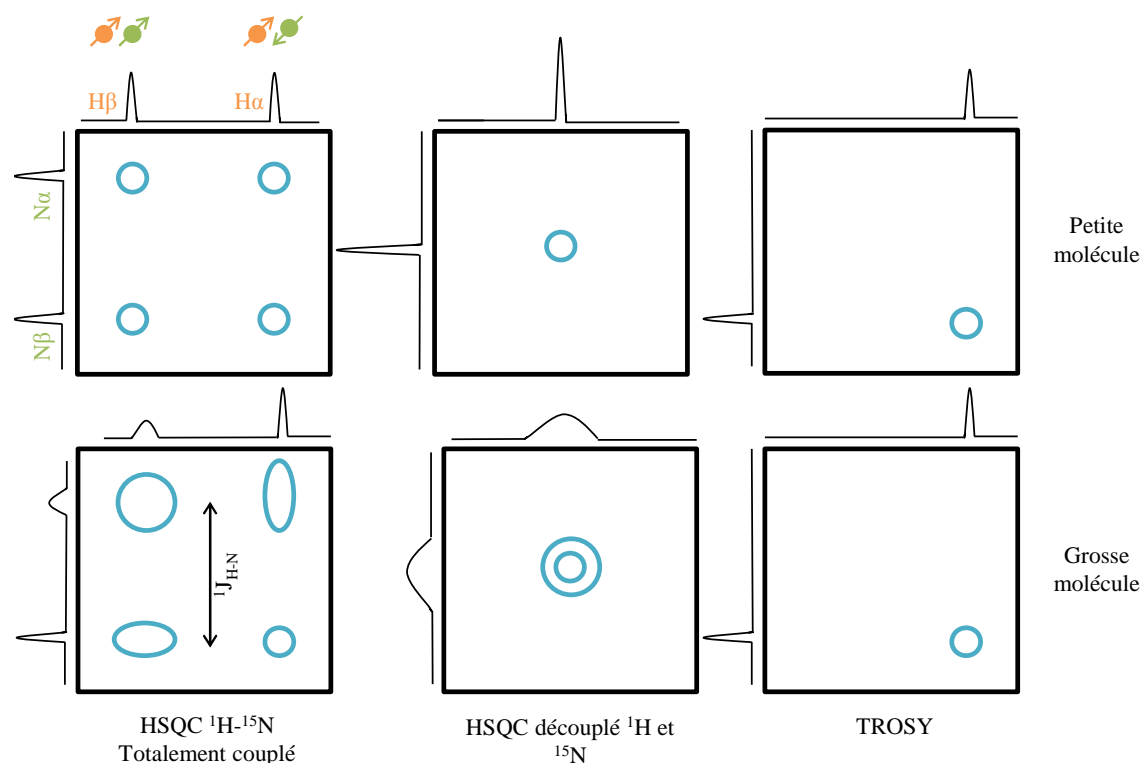


Figure 24: Représentation schématique de l'expérience  $^1\text{H}$ - $^{15}\text{N}$  TROSY

Différentes expériences TROSY ont été développées afin de répondre aux attentes multiples des expérimentateurs : l'optimisation pour l'observation des groupements amides  $^{15}\text{N}$ - $^1\text{H}$ ,<sup>66,67</sup> des sites aromatiques  $^{13}\text{C}$ - $^1\text{H}$ ,<sup>70-73</sup> des groupements méthyles  $^{13}\text{C}$ - $^1\text{H}_3$ ,<sup>74</sup> mais également des groupements méthylènes  $^{13}\text{C}$ - $^1\text{H}_2$ .<sup>64</sup> L'utilisation de cette dernière expérience paraît particulièrement intéressante dans le cadre de nos études, les peptides  $(\text{POG})_n$  présentant un nombre important de résidus glycine. Il faut noter par ailleurs qu'au-delà de l'effet TROSY qui sera modeste en raison de la taille modérée de la triple hélice (équivalent à 63 acides aminés dans le cas de  $(\text{POG})_7$ ), l'expérience  $\text{CH}_2$ -TROSY permet de s'affranchir systématiquement des couplages géminaux  $^2\text{J}_{\text{HH}}$  de  $\sim 15$  Hz. Ce gain en résolution sera très précieux pour des analyses détaillées. Il pourrait permettre la détermination des constantes de couplage  $^3\text{J}_{\text{NH-H}\alpha}$  des résidus glycine, reliées à l'angle  $\Phi$  par l'équation de Karplus, nous donnant ainsi des informations structurales locales. Dans un second temps, nous avons développé des expériences à 3 dimensions afin de pouvoir obtenir le maximum de constantes de couplage dans le but de comparer ces données à des analyses de dynamique moléculaire.

Détermination de constantes de couplage  $^3J_{H-NH}$  – Afin d’apprécier le gain en résolution apporté par l’expérience CH<sub>2</sub>-TROSY, nous avons tout d’abord enregistré une expérience  $^1H$ - $^{13}C$ -HSQC dans D<sub>2</sub>O (Figure 25 A). La mise en place d’un découplage spécifique des carbonyles permet de supprimer le couplage  $^1J_{C\alpha C'}$  dans la dimension indirecte, simplifiant le spectre et augmentant le rapport signal sur bruit (Figure 24 B). Nous avons enfin utilisé une séquence CH<sub>2</sub>-TROSY qui supprime la constante de couplage géminale entre les H $\alpha$  de la glycine et améliore les résolutions  $^{13}C$  et  $^1H$  par interférence de relaxation (Figure 24C). En solubilisant les peptides dans H<sub>2</sub>O, la réintroduction du proton HN nous permet d’observer sur chaque glycine la constante de couplage  $^3J_{H-NH}$  qui est reliée aux valeurs de l’angle  $\Phi$  (Figure 24D).

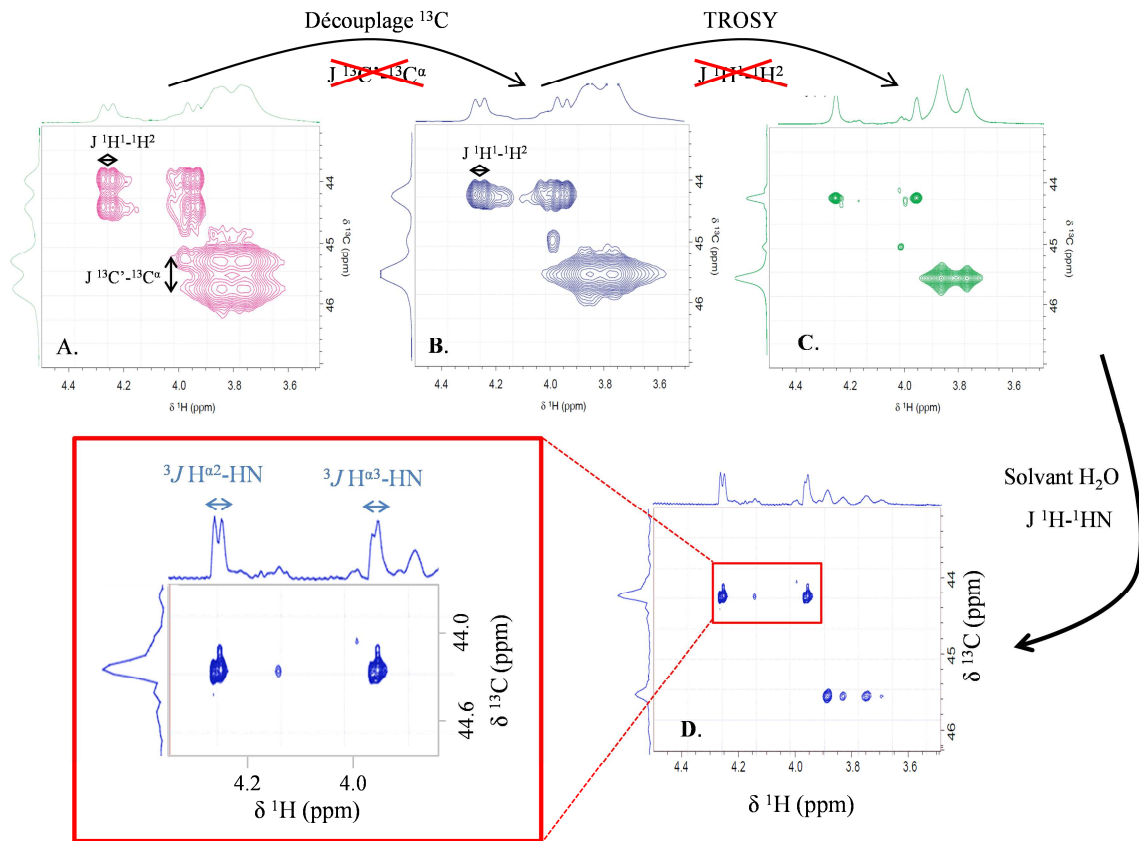


Figure 25: Simplification du spectre de corrélation  $^{13}C$ - $^1H$ .  $^4(POG)_7$ , 1mM, pH=7, 500 MHz, 5°C

L’angle dièdre  $\Phi$ , défini l’angle  $C'_{i-1}-N_i-C_i^\alpha-C'_i$ , est relié aux constantes de couplage  $^3J$  par la relation de Karplus :

$$J=A\cos^2(\Phi+\Delta\Phi)+B\cos(\Phi+\Delta\Phi)+C$$

où A, B et C sont des coefficients empiriques ou calculés par des méthodes DFT (Density Functional Theory) dépendant des substituants et  $\Delta\Phi$  est un déphasage à prendre en

compte pour relier  $H^N-N-C^\alpha-H^{\alpha 2}$  et  $H^N-N-C^\alpha-H^{\alpha 3}$  à l'angle  $\phi$  (Figure 26). *Courtier-Murias et al.* ont proposé des valeurs A, B, C théoriques pour les glycines appartenant au peptide Gly-Pro-Gly-Gly<sup>75</sup> et comparé leurs valeurs de couplages théoriques à des mesures expérimentales faites sur différentes glycines dans la flavodoxine. Leurs résultats étant convainquants, nous avons choisi d'utiliser ces paramètres pour nos études.

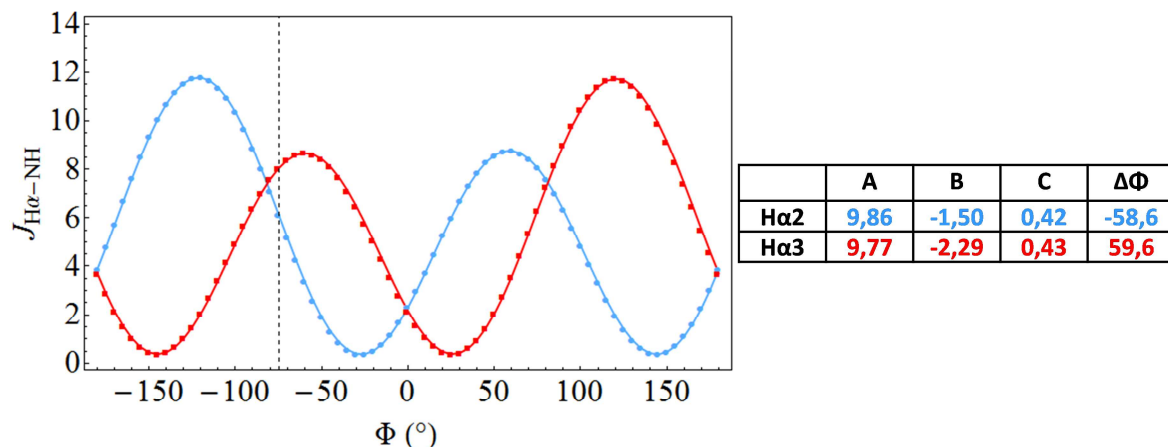


Figure 26: Représentation des courbes de Karplus avec les coefficients prédits par un modèle B3LYP EPR- III. Les valeurs de  $J$  sont indiquées en Hz

Les spectres CH<sub>2</sub>-TROSY ont été enregistrés sur un spectromètre à 800MHz afin de gagner en sensibilité et en résolution. Les échantillons ont été préparés à 10mM dans un tampon phosphate, pH=7 dans un mélange H<sub>2</sub>O/D<sub>2</sub>O (90:10) ou 100% D<sub>2</sub>O. Pour chacun des peptides, nous avons pu extraire les constantes de couplage des deux populations principales (M et TH). Les spectres ont été traités sur NMRpipe et les données brutes sont présentées en partie expérimentale (E p.75). Les résultats pour les populations majoritaires sont présentés sur un diagramme en Figure 27. Nous pouvons remarquer que les constantes de couplage  $^3J_{H-NH}$  mesurées pour le monomère comme pour la triple hélice sont en mauvais accord avec les valeurs théoriques. En particulier, pour l'hélice PPII canonique ( $\phi = -70^\circ$  et  $\psi = 150^\circ$ ), les valeurs attendues avec ces paramètres de Karplus sont de  $^3J_{NH-Ha2} = 5.19$  Hz et  $^3J_{NH-Ha3} = 8.38$  Hz. Une première hypothèse pour expliquer ces résultats décevants serait la présence de conformères en échange rapide entraînant une moyenne des constantes de couplage. Cette hypothèse paraît plausible dans le cas du monomère qui peut explorer un large espace conformationnel, en revanche, cela semble peu probable pour le cas de la triple hélice qui maintient les angles du squelette peptidique relativement contraints. Pour approfondir ces possibilités de dynamique, nous avons entamé des études de dynamique moléculaire qui seront exposées plus loin dans ce chapitre. Par ailleurs, nous

avons tenté d'affiner encore les mesures RMN en développant des expériences 3D qui permettent d'accéder à de nombreuses constantes de couplage dont les couplages vicinaux  $^1\text{H}-^1\text{H}$  qui nous intéressent ici.

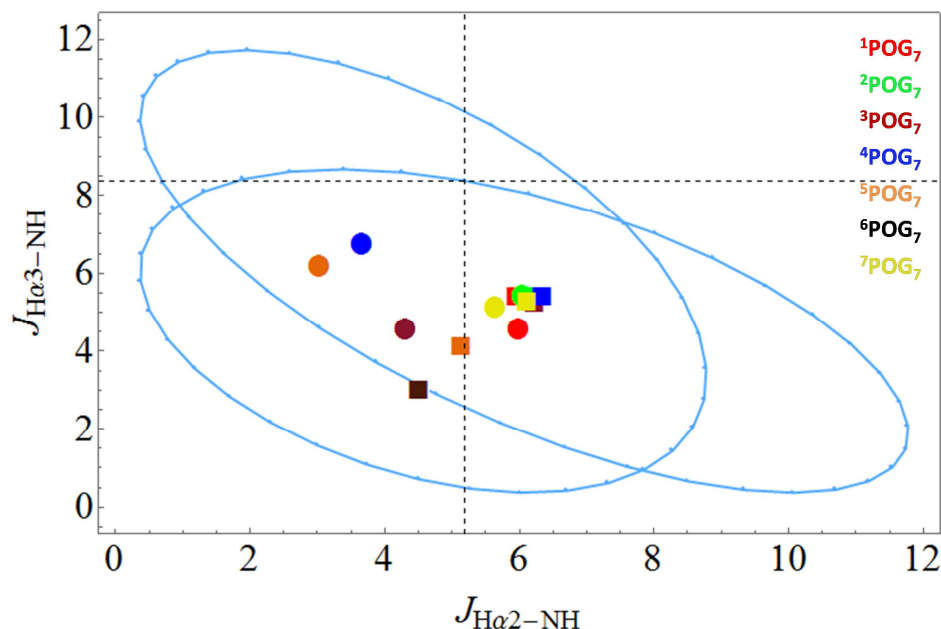


Figure 27: Représentation à deux dimensions des couples de valeurs  $^3J_{\text{HN-H}\alpha 3}$  et  $^3J_{\text{HN-H}\alpha 2}$ . Un code couleur a été utilisé pour distinguer chaque glycine, les deux populations majoritaires sont représentées par des marqueurs différents (M■ +TH●). La représentation en ligne continue symbolise les couples de constantes de couplage calculés pour chaque valeur de  $\phi$  à partir des paramètres de Karplus théoriques. Les lignes pointillées représentent les valeurs des couplages  $^3J_{\text{HN-H}\alpha 3}$  et  $^3J_{\text{HN-H}\alpha 2}$  correspondant à la structure PPII ( $\phi = -70^\circ$ )

*Mise au point d'expérience 3D sélectives d'état de spin* – La RMN à une dimension est fondée sur le schéma général : P-M- $t$  où la préparation (P) consiste à amener les spins hors de leur position d'équilibre (par une impulsion  $B_1$ ), le mélange (M) correspond aux optionnels transferts de polarisation et enfin, la détection de la fréquence de précession des spins ( $t$ ) permet d'obtenir un FID qui, après transformée de Fourier fournit un spectre 1D. La RMN multidimensionnelle introduit un ou plusieurs temps supplémentaires où les spins évoluent après plusieurs temps de mélange à leur fréquence de résonance. Dans le cas d'une 3D, le schéma est le suivant : P-M<sub>1</sub>- $t_1$ -M<sub>2</sub>- $t_2$ -M<sub>3</sub>- $t_3$ . Les temps  $t_1$  et  $t_2$  sont ajustés à 0 pour la première acquisition puis sont incrémentés graduellement au cours de l'expérience. Ainsi le signal,  $S^{\text{FID}}$  est une fonction  $f(t_1, t_2, t_3)$ . De la même manière que le FID enregistré pendant  $t_3$  peut être transformé par transformée de Fourier pour obtenir un spectre dans le domaine de fréquence  $F_3$  (Hz), cette opération peut être réalisée sur les temps  $t_2$  et  $t_1$ . Nous avons développé des séquences s3-HACACO et s3-HACAN (s3 pour « spin states selective ») basées sur des développements antérieurs dédiés aux systèmes

méthyne-méthylène. Ainsi que cela est détaillé sur la Figure 28, chacune de ces séquences peut être décomposée en six périodes :

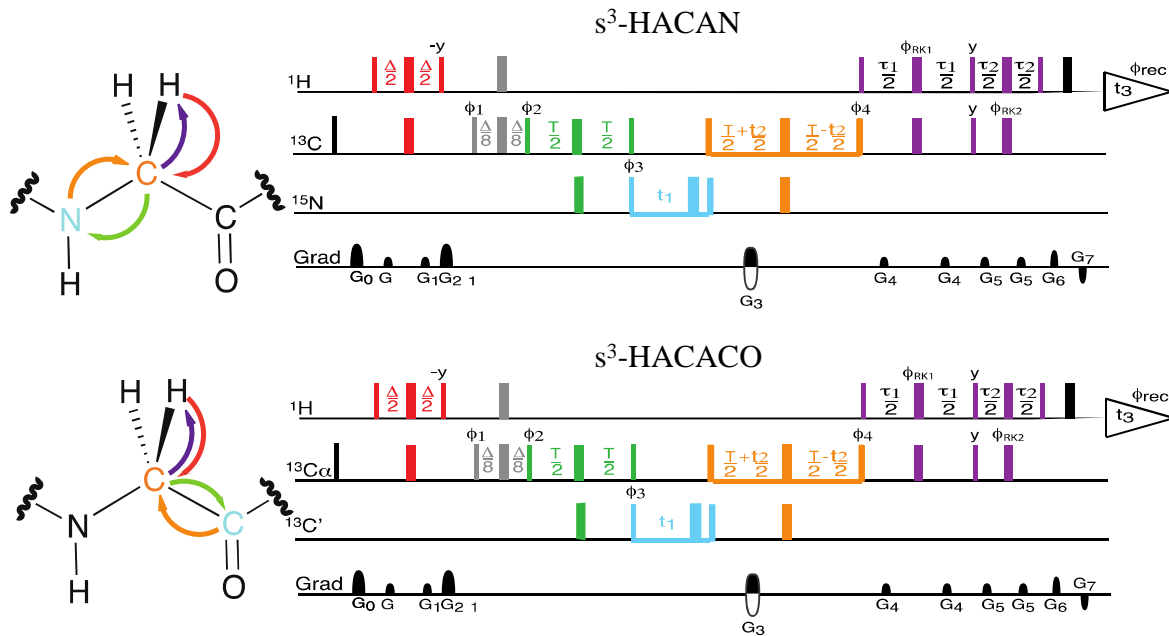


Figure 28: Séquences d'impulsions « spin states selective » (s3). Un code couleur est utilisé : chaque période de la séquence est associée à un transfert d'aimantation sur le résidu glycine. Les barres fines et épaisses indiquent respectivement des impulsions de 90° et 180°. Lorsque non spécifiée, la phase de l'impulsion est x. délais :  $\Delta = 1/(2J_{CaHa})$  ; pour la séquence s3-HACACO,  $T = 1/(2J_{CaC'})$  ; pour la séquence s3-HACAN,  $T = 1/(2J_{CaN})$  ;  $\tau_1 = 0.34/J_{CaHa}$  et  $\tau_2 = 0.23/J_{CaHa}$ .  $\Phi_1 = 45^\circ, 225^\circ$  ;  $\Phi_2 = y, -y$  ;  $\Phi_3 = x, x, -x, -x$  ;  $\Phi_4 = x, y, -x, -y$ . La sélection de la raie la plus fine est obtenue pour  $\Phi_{RK1} = x$  ;  $\Phi_{RK2} = x$ .

Il y a dans un premier temps un **transfert de polarisation  $H^a \rightarrow C^a$** . Si  $\Delta$  est fixé tel que  $\Delta = \frac{1}{2J_{CaHa}}$ , la séquence d'impulsion INEPT<sup>77</sup> transfère l'aimantation du proton vers le carbone créant ainsi le terme antiphase  $2C_zH_z$ . Puis après une impulsion 90° carbane ( $\phi_2$ ), l'aimantation peut évoluer en carbone par création du terme  $-2H_zC_x$ . La partie de la séquence  $S^3E$  (spin states selective excitation) est optionnelle. Elle permet de simplifier le multiplet dans la dimension carbone en ne retenant qu'une seule composante du triplet  $C^a$ . Elle est ajoutée pour obtenir le spectre « TROSY », où seule la corrélation  $^1H-^{13}C$  relaxant le moins vite est retenue. Cette partie  $S^3E$  ne doit pas être incorporée lorsque la séquence est utilisée pour mesurer la totalité des constantes de couplage. Il y a ensuite un **transfert de polarisation  $C^a \rightarrow N$  (ou  $C'$ )** par création du terme  $4C_zH_zN_y$  avec  $T = \frac{1}{2J_{CaN}}$  (ou  $T = \frac{1}{2J_{CaC'}}$  dans la séquence s3-HACACO. **L'aimantation sur l'azote (ou le carbonyle) évolue durant un temps  $t_1$** . Le signal sera modulé en fonction de  $t_1$ , donnant une première dimension indirecte correspondant au déplacement chimique azote (ou carbonyle). Aucun découplage n'est effectué durant cette période de marquage ce qui permet de conserver l'état de spin

des noyaux environnants. L'aimantation est ensuite transférée sur le carbone et évolue dans un temps constant d'une durée  $T = \frac{1}{2J_{C\alpha N}}$  permettant de refocaliser le couplage  $^1J_{C\alpha N}$  (ou  $T = \frac{3}{2J_{C\alpha C'}}$  pour  $^1J_{C\alpha C'}$ ). A chaque acquisition, le  $180^\circ_x$  est décalé avec un incrément (t2) ce qui introduit une modulation en fonction du déplacement chimique, c'est la deuxième dimension indirecte. Le gradient  $G_3$  encode spatialement l'aimantation de l'échantillon. Cela permettra la détection en quadrature de cette dimension (acquisition de type Echo-Antiecho). Le bloc Rance Kay  $S^3CT$  (spin state selective coherence transfert) ramène les aimantations carbone  $C_x$  et en  $C_y$  sur les protons  $H^\alpha$  diastéréotopes. L'utilisation des délais particuliers  $\tau_1$  et  $\tau_2$  permet en outre une sélection des états de spins ce qui augmente à la fois la sensibilité et la résolution de l'expérience.<sup>76</sup> Le décodage spatial est obtenu grâce aux gradients de champs  $G_6$  et  $G_7$ . Ceux-ci restaurent l'aimantation proton et permettent l'acquisition t3 qui est la dimension directe proton.

La Figure 29 détaille les spectres théoriques que l'on obtiendrait pour un proton  $H^{\alpha 2}$ -Gly dans différentes variantes de l'expérience 3D  $H^\alpha C^\alpha N$ . Le premier schéma représente le spectre théorique d'une expérience HACAN totalement découplée. Dans ce cas, seuls les déplacements chimiques  $\delta_{C\alpha}$ ,  $\delta_N$ ,  $\delta_{H\alpha 2}$  et le couplage  $^2J_{H\alpha 2 H\alpha 3}$  peuvent être déterminés. Le second schéma représente le spectre théorique d'une  $H^\alpha C^\alpha N$  sans découplage proton pendant le marquage carbone et sans découplage carbone à l'acquisition, les couplages  $^1J_{C\alpha H\alpha}$  sont ainsi réintroduits. On voit apparaître alors un doublet de doublet dans la dimension proton, et un triplet dans la dimension carbone ( $^1J_{C\alpha H\alpha 2} \approx ^1J_{C\alpha H\alpha 3}$ ). Cependant, le transfert  $S^3CT$  permet une sélection des états de spins et seules les corrélations marquées par des taches pleines sont observées. Dans une telle expérience, en plus des mesures de  $\delta_{C\alpha}$ ,  $\delta_N$ ,  $\delta_{H\alpha 2}$  et  $^2J_{H\alpha 2 H\alpha 3}$ , les constantes de couplage  $^1J_{C\alpha H\alpha 2}$  et  $^1J_{C\alpha H\alpha 3}$  peuvent être déterminées grâce aux décalages en fréquence entre les corrélations I/1 et III/4. Enfin, si on laisse s'exprimer les couplages  $^1H$  et  $^{13}C$  pendant le temps de marquage  $^{15}N$ , le proton apparaît sous forme d'un hexadécuplet, le carbone un dodécuplet et l'azote un octuplet, soit  $16 \cdot 12 \cdot 8 = 1536$  corrélations. A nouveau l'élément  $S^3CT$  permet une simplification de ce multiplet, ainsi nous observons seulement 16 corrélations. Un ensemble de douze constantes de couplage pourra être mesuré sur ce spectre :  $^1J_{C\alpha H\alpha 2}$  ;  $^1J_{C\alpha H\alpha 3}$  ;  $^2J_{H\alpha 2 H\alpha 3}$  ;  $^1J_{C\alpha C'}$  ;  $^2J_{C'H\alpha 2}$  ;  $^2J_{C'H\alpha 3}$  ;  $^1J_{NH\alpha}$  ;  $^2J_{NH\alpha 2}$  ;  $^2J_{NH\alpha 3}$  ;  $^3J_{H\alpha H\alpha 2}$  ;  $^3J_{H\alpha H\alpha 3}$  ;  $^2J_{C\alpha H\alpha}$ .

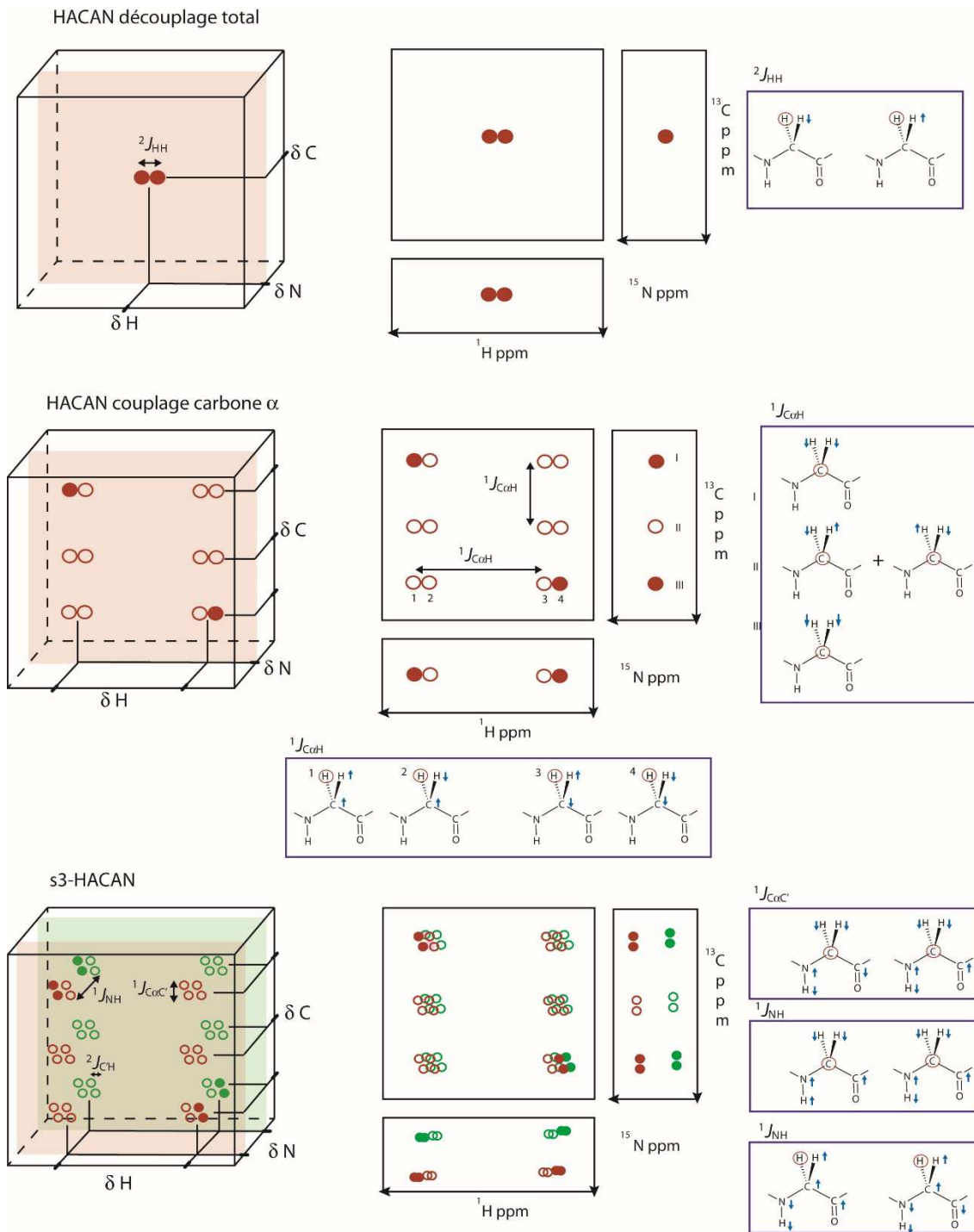


Figure 29: Spectres théoriques de corrélation  $^1\text{H}$ - $^{13}\text{C}$  obtenus pour un  $\text{H}^\alpha$  de résidu Glycine. En haut : expérience entièrement découplée, au milieu : expérience avec découplage N et en bas : expérience entièrement couplée. Le spectre est représenté dans un cube et en 2D sur 3 projections. Les états de spins associés aux différentes composantes observées sont représentés. Les corrélations pleines correspondent aux états de spins sélectionnés

La sélection des états de spins implique que certaines constantes de couplage ne peuvent pas être mesurées directement entre deux corrélations. Cependant, on peut mesurer des sommes de plusieurs constantes de couplage (Figure 30) et par un jeu d'additions/soustractions, remonter aux constantes de couplage individuelles.



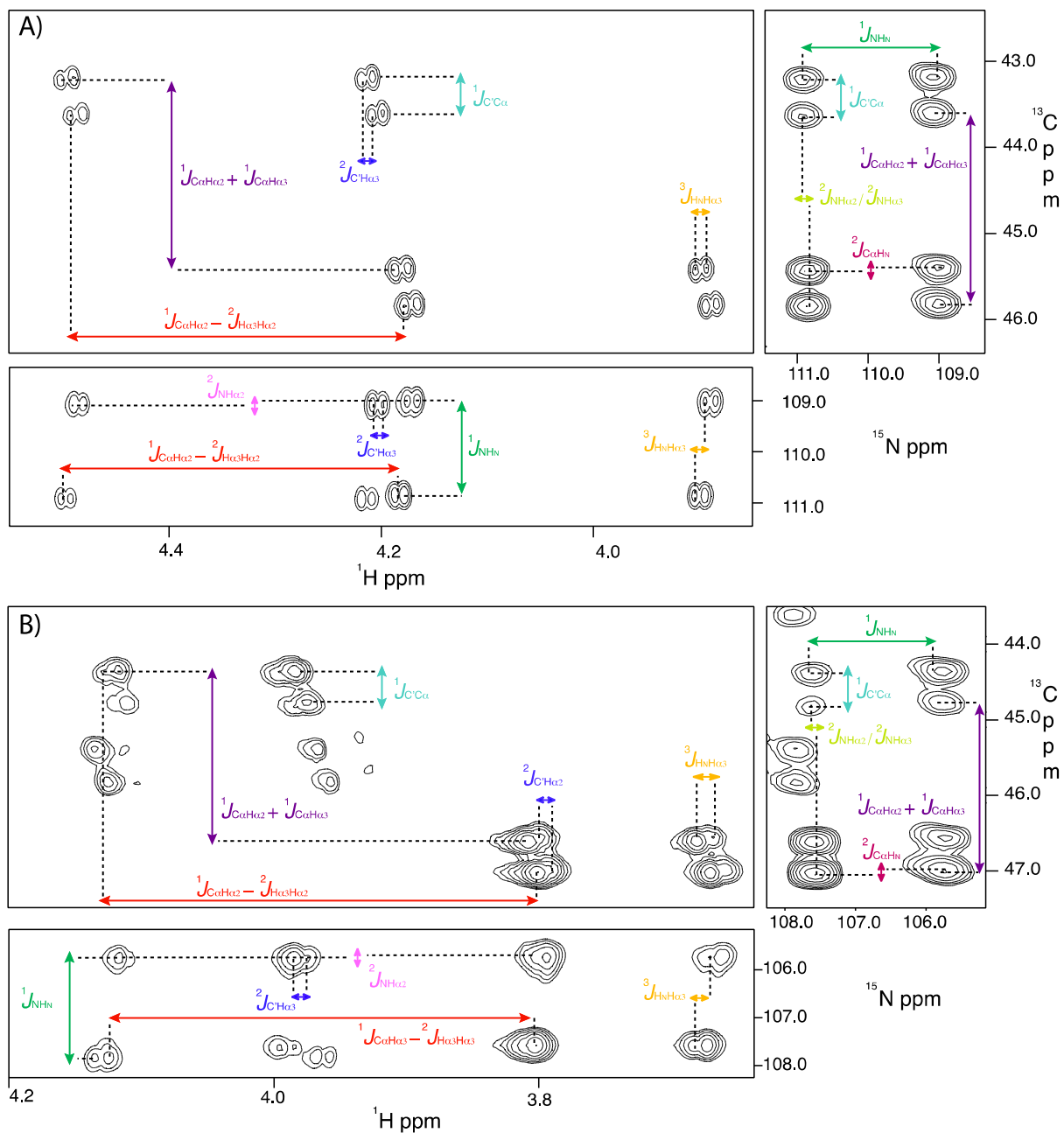


Figure 30: Projections dans les trois dimensions des expériences s3-HACAN sur  $^4(\text{POG})_7$ , 10mM, pH=7, 40°C, H<sub>2</sub>O/D<sub>2</sub>O (90:10), 500MHz. A) Projections restreintes au signal PPII, B) Projections restreintes au signal de la TH

La multiplication des corrélations pour une même fonction méthylène entraîne qu'un même couplage peut être mesuré ainsi jusqu'à 8 fois mais, dans un souci de clarté, chaque constante de couplage n'apparaît qu'une fois sur la figure 29. Les différentes mesures de couplages sont reportées dans le Tableau 2.

	$^1J_{CaHa2}$	$^1J_{CaHa3}$	$^2J_{Ha2Ha3}$	$^1J_{CaC'}$	$^2J_{C'Ha2}$	$^2J_{C'Ha3}$	$^1J_{NHN}$	$^2J_{NH\alpha 2}$	$^2J_{NH\alpha 3}$	$^3J_{HNHa2}$	$^3J_{HNHa3}$	$^2J_{CaHN}$
<b>Monomère (M)</b>												
Mesures	8	8	8	8	4	4	8	4	4	4	4	8
Moyenne (Hz)	139,44	139,35	17,78	52,66	4,45	4,85	93,41	3,45	3,68	6,00	5,07	3,72
Erreur	0,04	0,04	0,03	0,03	0,01	0,01	0,02	0,02	0,02	0,01	0,01	0,02
<b>Triple Hélice (TH)</b>												
Mesures	8	8	8	8	4	4	8	4	4	4	4	8
Moyenne (Hz)	140,38	138,32	18,26	52,53	3,40	5,00	93,55	2,87	2,24	4,30	5,63	5,04
Erreur	0,24	0,24	0,17	0,14	0,03	0,04	0,08	0,10	0,15	0,03	0,04	0,01

**Tableau 2: Mesures des constants de couplages sur le peptide  $^4(POG)_7$  (10mM, pH=7, 40°C) à partir de l'expérience s3-HACAN.**

Les couplages sont donc mesurés à partir de décalages en fréquence de deux composantes x et y dans les 3 dimensions  $^1H$ ,  $^{13}C$  et  $^{15}N$ . L'erreur de mesure  $\epsilon_{jx-y}$  correspond à :

$$\epsilon_{jx-y} = \sqrt{\frac{\delta v_x^2 + \delta v_y^2}{2 * \left(\frac{h_x + h_y}{2 * N}\right)^2}}$$

$\delta v$  : largeur du pic  
 $h$  : hauteur du pic  
 $N$  : bruit.

Une moyenne pondérée des constantes de couplage a ensuite été calculée ainsi que l'erreur sur la moyenne :

$$\langle J_{x-y} \rangle = \frac{\sum_{n=1}^N \frac{J_{x-y}}{\epsilon_{jx-y}^2}}{\sum_{n=1}^N \frac{1}{\epsilon_{jx-y}^2}} \quad \text{et} \quad \langle \epsilon_{jx-y} \rangle = \frac{1}{\sqrt{\sum_{n=1}^N \frac{1}{\epsilon_{jx-y}^2}}}$$

Les données brutes sont présentées en partie expérimentale (§F, p.77). Les erreurs sur les constantes de couplage sont relativement faibles bien que plus élevées sur la triple hélice. Cela s'explique par des largeurs de raies plus grandes associées à cette population et certains recouvrements qui faussent les mesures. Nous observons des différences significatives sur les valeurs des constantes de couplage  $^2J_{C'Ha}$ ,  $^2J_{NH\alpha}$ ,  $^2J_{CaHN}$  et  $^3J_{HNHa}$  entre les populations M et TH, notamment une proximité des constantes de couplage  $^nJ_{Ha2-Y}$  et  $^nJ_{Ha3-Y}$  dans le monomère, ce qui traduit une plus grande dynamique du squelette par rapport à l'assemblage en triple hélice. Les valeurs  $^3J_{HNHa}$  obtenues sont légèrement plus faibles que celles obtenues dans l'expérience 2D pour le monomère PPII, jusqu'à observer une différence de près de 1Hz dans le cas de la triple hélice. Bien entendu, les valeurs obtenues sur la base des spectres 3D sont plus précises qu'en 2D car chaque composante

apparaît sous forme d'un singulet dans le spectre 3D ; aucune convolution des multiplets ne peut fausser la mesure des couplages. Cela n'est pas le cas sur le spectre 2D CH<sub>2</sub>-TROSY car les deux composantes du doublet dans la dimension <sup>1</sup>H ne sont pas parfaitement résolues et se chevauchent en partie. Pour autant, ces nouvelles valeurs ne coïncident pas encore avec le modèle de Karplus présenté précédemment. Nous n'avons pas à ce jour d'explications satisfaisantes permettant de justifier cette dichotomie, à moins de remettre en cause les paramètres de Karplus déterminés par DFT pour les résidus glycine. On pourrait également invoquer l'échange des protons amide, le remplacement partiel H/D dans H<sub>2</sub>O:D<sub>2</sub>O 90:10, ou l'influence des liaisons H sur ces couplages. Un travail récent de Salvador et al. a par ailleurs montré pour les résidus alanine l'influence très importante (modulation de près de 6 Hz !) de l'angle Psi sur les valeurs de couplages au niveau de l'angle Phi.<sup>78</sup> Nous envisageons de mener une étude théorique plus complète, intégrant l'ensemble des valeurs de couplage mesurées sur nos expériences 3D pour sonder si de telles interdépendances existent entre les angles dièdres et constantes de couplage au niveau des résidus glycine. Au-delà de l'interprétation des constantes <sup>3</sup>J et des développements théoriques envisagés pour analyser les autres types de couplages (<sup>1</sup>J et <sup>2</sup>J), la mesure des couplages en milieu anisotrope permettrait de déterminer les valeurs des couplages dipolaires résiduels (RDC). Ces RDC sont fonction de la distance séparant deux noyaux ainsi que de leur orientation par rapport au champ magnétique B<sub>0</sub>. Cela fournira donc une information précieuse sur la structure tridimensionnelle des CMP et de leurs assemblages supramoléculaires. Des analyses similaires ont été réalisées avec l'enregistrement des expériences s3-HACACO sur le peptide <sup>3</sup>(POG)<sub>7</sub>. Les constantes de couplage sont présentées dans le tableau 4. Ces six mesures avaient déjà été réalisées pour <sup>4</sup>(POG)<sub>7</sub> dans l'expérience s3-HACAN. Des valeurs extrêmement proches sont trouvées à ces deux positions ce qui traduit la grande similitude structurale dans la partie centrale des CMP.

	$^1J_{CaHa2}$	$^1J_{CaHa3}$	$^2J_{Ha2Ha3}$	$^1J_{CaC'}$	$^2J_{C'Ha2}$	$^2J_{C'Ha3}$
Monomère (M)						
Mesures	2	2	8	4	2	2
Moyenne (Hz)	139,59	139,37	17,44	52,19	4,84	4,95
Erreur	0,21	0,21	0,06	0,03	0,09	0,09
Triple Hélice (TH)						
Mesures	2	2	8	4	2	2
Moyenne (Hz)	141,44	139,64	17,86	51,67	3,48	5,88
Erreur	0,30	0,32	0,26	0,07	0,10	0,1

Tableau 3: Mesures des constantes de couplage sur le peptide  $^3(POG)_7$  (10mM, pH=7, 40°C) réalisées à partir de l'expérience s3-HACACO

Ces mesures de constantes de couplage, incompatibles pour le moment avec le modèle de structure PPII figé, vont être par la suite comparées à des modèles obtenus par dynamique moléculaire.

### 1.5.2 Dynamique moléculaire

La dynamique moléculaire permet de simuler, à partir des lois de la mécanique classique de NEWTON, les trajectoires des atomes en phase cristalline, en solution ou en phase gazeuse. Elle donne l'évolution d'un système au cours du temps. Une résolution numérique des équations classiques du mouvement est utilisée pour calculer l'énergie du système le long de son parcours sur la surface de potentiel du système. Ainsi, une molécule est considérée comme un ensemble d'atomes dont les interactions mutuelles peuvent être décrites par une fonction d'énergie potentielle. C'est le "Champ de Force".

Ces travaux de dynamique moléculaire ont été réalisés dans le cadre d'une collaboration avec Claire Loison, chargée de recherche à Lyon. L'ensemble des calculs ont été effectué sur GROMACS (version 5.0.5) avec le champ de force Amber99sb-ILDN-phi et un modèle pour l'eau TIP3P. Ce champ de force a prouvé qu'il décrivait de façon très satisfaisante les couplages RMN pour divers peptides et de petites protéines<sup>79,80</sup> et particulièrement pour les peptides incorporant des prolines et hydroxyprolines.<sup>81</sup>

Deux productions ont été effectuées en parallèle sur la triple hélice (pose initiale correspondant à la structure pdb 1CAG)<sup>82</sup> d'une part, et sur le peptide monomérique (chaîne isolée de la structure 1CAG) d'autre part. Les différents paramètres de simulation sont décrits en partie expérimentale (§G p.80). Plusieurs paramètres ont été inspectés pour caractériser la dynamique de ces deux modèles : calcul du rmsd (Root-Mean-Square Deviation), vérification de la présence de liaisons hydrogène stabilisant la triple hélice, inspection des angles  $\phi$ ,  $\psi$  et des plissements des cycles pyrrolydine au cours de la dynamique.

*Calcul du rmsd dans la triple hélice* – Ce calcul permet d'évaluer la déviation des structures par rapport à la structure RX. Les écarts de coordonnées cartésiennes sont calculés entre chaque structure générée par dynamique moléculaire et la structure initiale selon :

$$rmsd = \sqrt{\frac{\sum_{i=1}^n (x_i - y_i)^2}{n}}$$

où  $x_i$  représente les coordonnées au cours de la dynamique et  $y_i$  les coordonnées de la structure initiale.

Il s'avère que le rmsd=2Å, ce qui montre que la structure fluctue légèrement, mais qu'elle reste stable au cours du temps, et proche de la structure initiale. La dynamique réalisée ici correspond à 635ns avec une structure extraite toute les 100ps, soit n=6350.

*Présence de liaisons hydrogène dans la triple hélice* – Le critère que nous avons retenu pour détecter la présence de liaisons hydrogène est défini par une limite de distance entre l'hétéroatome N qui porte l'hydrogène et l'accepteur de cette liaison hydrogène ( $d < 2,5 \text{Å}$ ) et par une contrainte sur l'angle  $\theta = \text{N-H-O}$  ( $\theta < 25^\circ$ ). La présence de liaisons H tout au long de la dynamique est schématisé sur la Figure 31 et les calculs sont présentés en partie expérimentale (cf.G2 p.81). Pour la chaîne C, nous observons que les glycines en N- et C-terminal sont les moins souvent engagées dans des liaisons hydrogène. Ceci est une conséquence directe de la flexibilité des chaînes aux extrémités. Nous n'observons pas de liaisons hydrogène pour l'extrémité C-terminale de la troisième chaîne (notée C) en raison du décalage des trois chaînes. Pour la chaîne A, la présence de liaisons hydrogène est plus importante aux extrémités qu'au milieu de la chaîne, ce qui est moins intuitif.

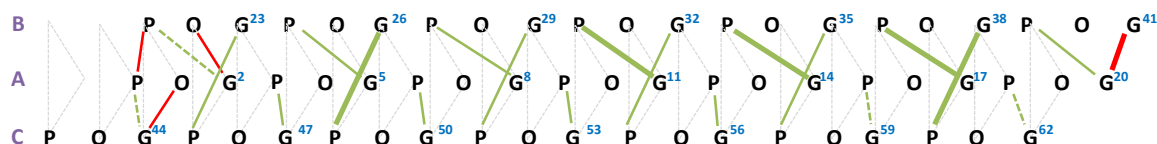


Figure 31: Illustration de la présence de liaisons hydrogène tout au long de la dynamique. Présence de liaisons H > 50% (trait plein épais), Présence de liaisons H < 50% (trait plein), Liaisons H qui devrait être présente (trait pointillé). En vert, les liaisons H habituelles et en rouge les liaisons H inhabituelles

Calcul des angles  $\phi$  et  $\psi$  dans la triple hélice – Les angles  $\phi$  et  $\psi$  trouvés au cours de la dynamique sont en accord avec une structure PPII. Ils se répartissent pour les résidus proline et hydroxyproline suivant une courbe gaussienne étroite, centrée autour des valeurs canoniques  $\phi = -70^\circ$  et  $\psi = 150^\circ$  (Figure 32). En ce qui concerne la glycine, les valeurs sont un peu plus diffuses, ce qui s’explique par le fait de leur flexibilité à l’extrémité des chaînes et à leur nature moins contrainte.

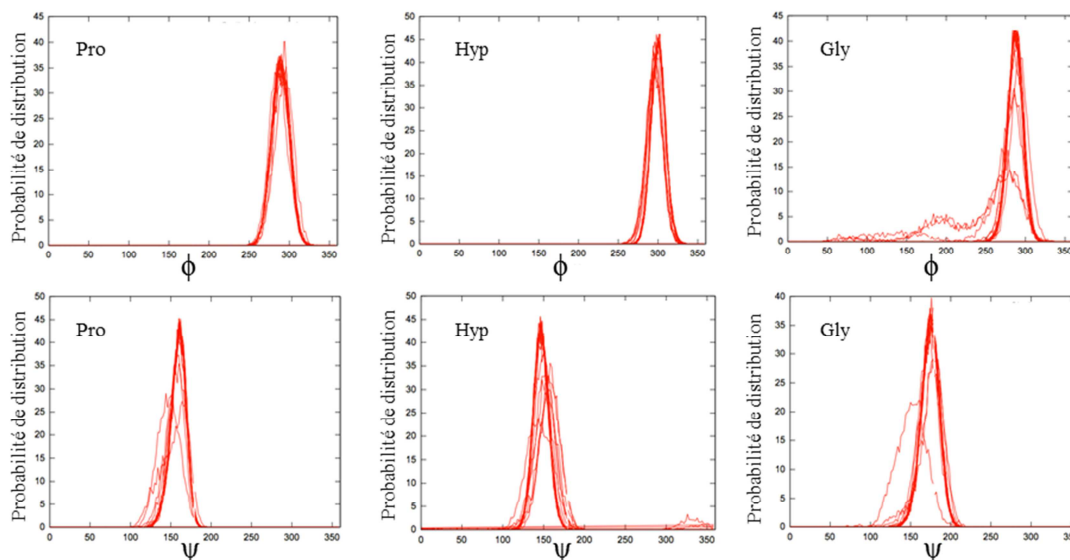
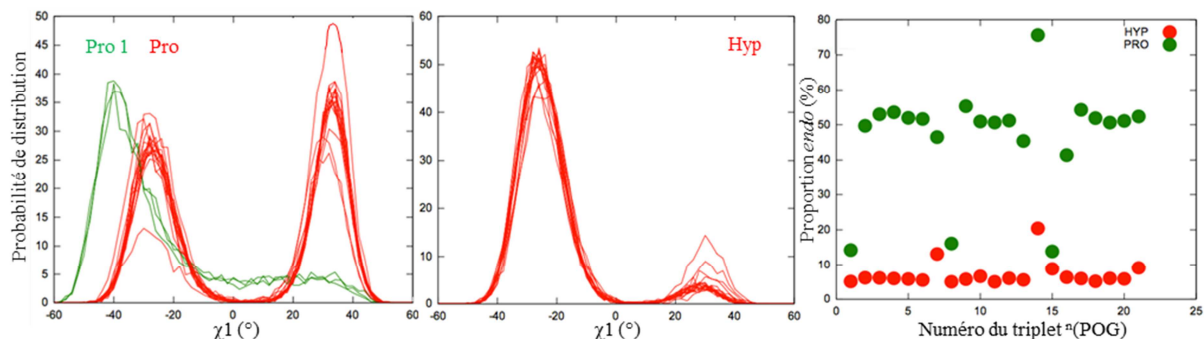


Figure 32: Distribution des angles  $\phi$  et  $\psi$  en fonction du type d’acide aminé (la première proline et la dernière glycine ont été exclues de cette représentation)

Plissement de cycle pyrrolidine dans la triple hélice – De précédentes études ont été réalisées pour adapter le champ de force afin qu’il respecte au mieux les préférences de plissement de cycle de la proline (*endo* :  $\chi_1 \approx 25^\circ$ ) et de l’hydroxyproline (*exo*  $\chi_1 \approx -25^\circ$ ).<sup>83</sup> Pour chaque proline et hydroxyproline, le pourcentage d’*endo* ( $\chi_1 > 6^\circ$ ) a été calculé et est représenté sur la Figure 33. On peut voir que les prolines positionnées en début de chaîne ont un  $\chi_1 \approx -40^\circ$ , ce qui peut s’expliquer par la nature particulière de l’azote protoné, non engagé dans une liaison peptidique. Nous pouvons également voir que les prolines adoptent aussi bien un plissement de cycle *endo* que *exo* avec une très légère préférence

pour la forme *endo*, en accord avec ce qui est décrit dans la littérature. Le plissement de cycle *exo* des hydroxyprolines est largement favorisé, sauf en bout de chaînes ( $C_{ter}$ ), ce qui semble être corrélé à la flexibilité de la chaîne.



**Figure 33:** Distribution de l'angle  $\chi_1$  pour la proline et l'hydroxyproline suivant sa place dans la chaîne. Les résidus sont numérotés par rapport au numéro du triplet auxquels ils appartiennent ( $n = 1$  à  $7$  pour la chaîne A,  $n = 8$  à  $14$  pour la chaîne B et  $n = 15$  à  $21$  pour la chaîne C)

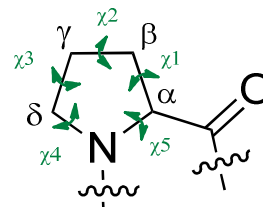
Dans les simulations réalisées par *Aliev et al.*<sup>81</sup> sur des petits peptides (Ace-Hyp-NMe et GPGG), les pourcentages *endo* sont d'environ 10% pour l'hydroxyproline et 60% pour la proline. Il semblerait qu'au sein de la triple hélice le plissement *exo* pour l'hydroxyproline soit encore plus favorisé.

*Calcul des angles  $\chi_1$  dans le monomère* – De la même façon, les angles  $\chi_1$  ont été calculés dans le monomère pour les prolines et les hydroxyprolines (Figure 34). Nous observons bien la préférence pour le plissement de cycle *exo* pour les hydroxyprolines et une légère préférence pour le plissement de cycle *endo* pour les prolines. Afin de caractériser la conformation des cycles, il est également possible de calculer l'amplitude de la pseudorotation ( $\chi_m$ ) et l'angle de phase de pseudorotation (P).<sup>81,84</sup>

$$P = \tan^{-1}\left(\frac{A}{B}\right) \quad \text{et} \quad \chi_m = (A^2 + B^2)^{1/2}$$

Avec 
$$A = \frac{2}{5} \sum_{i=1}^5 \chi_i \cos\left(\frac{4\pi}{5}(i-2)\right)$$

et 
$$B = -\frac{2}{5} \sum_{i=1}^5 \chi_i \cos\left(\frac{4\pi}{5}(i-2)\right)$$



Les valeurs trouvées pour ces deux paramètres suivant la position de la glycine marquée sont similaires à celles calculées pour *Aliev et al.* et sont présentées de la tableau 8 de la partie expérimentale (§G2 p.81).

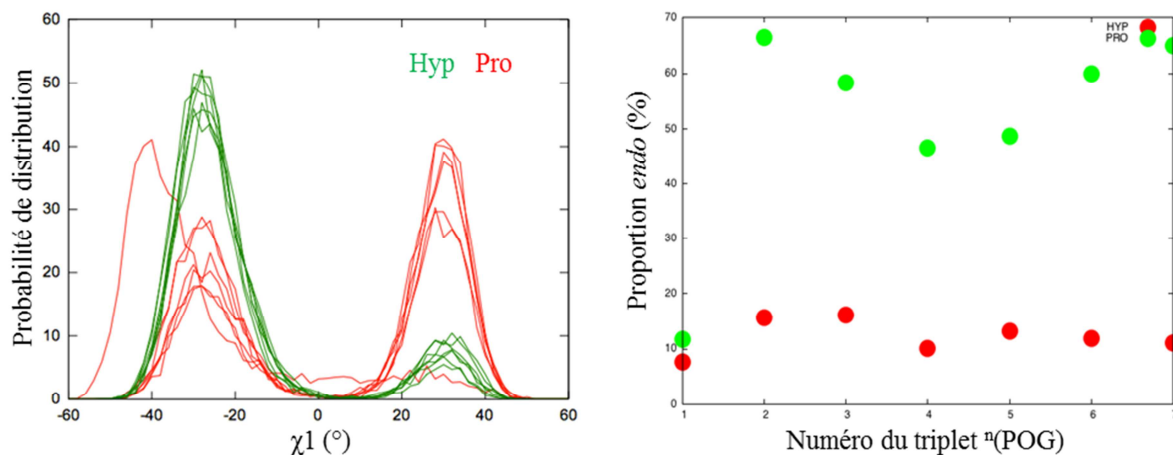


Figure 34: Distribution de l'angle  $\chi_1$  pour la proline et l'hydroxyproline suivant sa place dans la chaîne. Les résidus sont numérotés selon leur appartenance au triplet  $n = 1$  à 7

*Conformation du monomère* – Tout au long de la simulation, le polymère est en mouvement et n'adopte pas de conformation fixe. Il a été principalement observé des formes compactes telles que des hélices, des cycles, et des formes zippées (Figure 35).

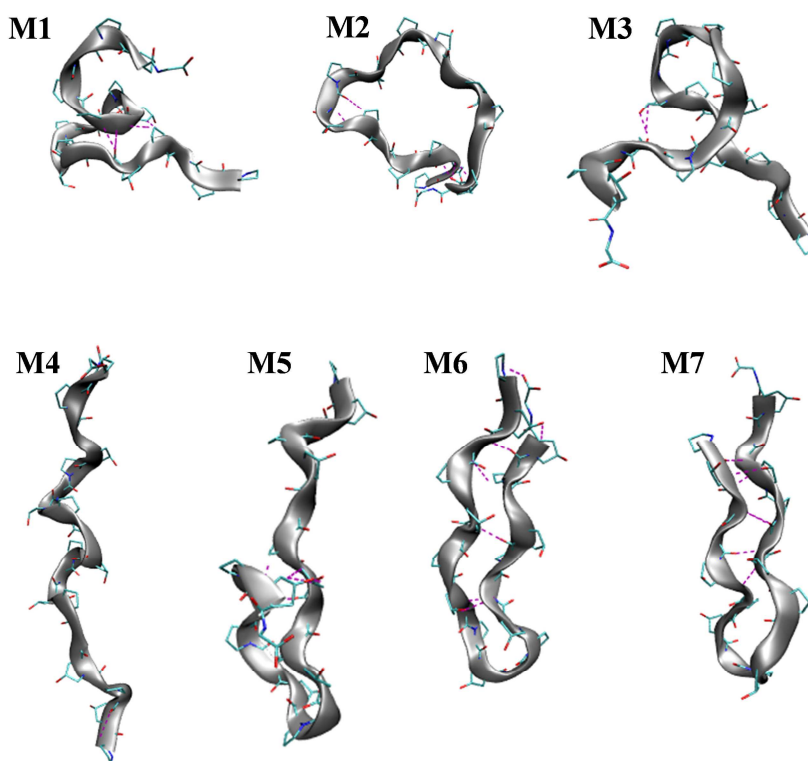


Figure 35: Structures représentatives adoptées au cours d'une simulation de 500ns.

Des paramètres tels que le rayon de giration décrivant la compacité de la molécule, la distance entre le premier et le dernier acide aminé, l'asphéricité ( $\delta^*$ , compris entre 0 et 1, 0 pour une sphère et 1 pour une tige) et l'acylindricité ( $S$ ) permettent de décrire la nature de



la conformation. A l'issue de la simulation nous obtenons les valeurs moyennes suivantes :  $\delta^*=0.42$  et  $S=0.53$  ce qui, de prime abord, se rapproche beaucoup des valeurs caractéristiques des structures aléatoires ( $\delta^*=0.45$  et  $S=0.57$ ).<sup>85</sup> Le monomère, sous sa conformation étendue M4, a un rayon de giration égal à 17Å et dans notre cas, après une simulation de 500 ns, le rayon de giration moyen est égal à 8,76Å, ce qui signifie que notre monomère est relativement compact et que la structure PPII n'est pas très représentée au cours de la dynamique. Si l'on s'intéresse à la distance entre le premier et le dernier acide aminé, la valeur moyenne calculée (10,1Å) est relativement faible et ne reflète pas une structure aléatoire. La forte proportion de liaisons H entre la Pro-2 et Gly-22, mais également entre la Gly-7 et l'Hyp-18, nous montre que le monomère reste en conformation « zippée » pendant plus de 20% de la dynamique structure M7 (Tableau 4).

Donneur	Accepteur	Présence de liaisons H (%)
GLY-7	PRO-17	8.08
GLY-4	HYP-21	9.91
GLY-4	PRO-20	13.64
GLY-19	PRO-5	17.70
GLY-22	PRO-2	18.00
GLY-7	HYP-18	21.20
PRO-2	GLY-22	21.50

Tableau 4: Présence de liaisons hydrogène dans le monomère

Une autre manière de visualiser la flexibilité conformationnelle du peptide monomérique est de reporter dans un diagramme de Ramachandran les angles dièdres  $\phi$  et  $\psi$  adoptés au cours de la dynamique. Les prolines ainsi que les hydroxyprolines restent exclusivement dans la région PPII (§G2 p.81). Les glycines quant à elles possèdent un angle  $\psi$  relativement contraint dans la région étendue mais l'angle  $\phi$  ne reste pas dans une région spécifique, expliquant à lui seul la flexibilité conformationnelle du monomère (POG)<sub>7</sub>. Il est intéressant d'observer que la présence de ces différents conformères est en accord avec l'observation par RMN de plusieurs formes monomériques dont les signaux coalescent lorsque la température est augmentée. La présence de liaisons hydrogène pour certaines structure « zippées » expliquerait leur relative stabilité ce qui serait compatible avec un échange lent à basse température (Figure 36).

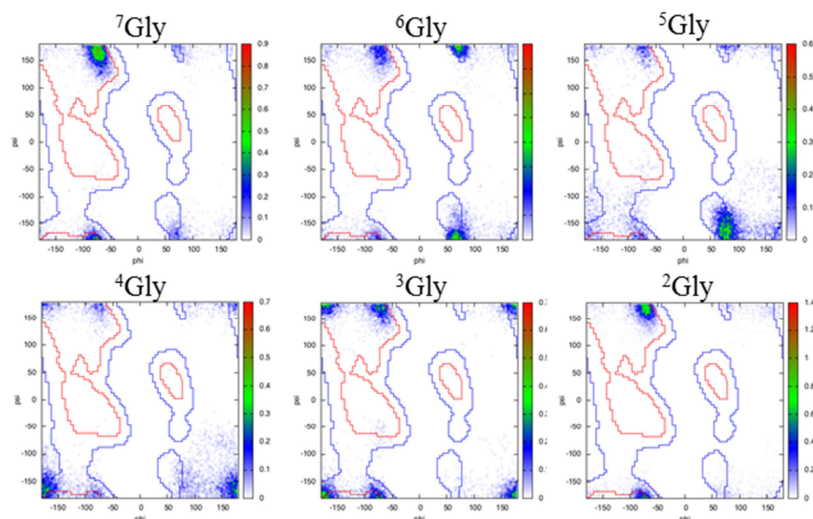


Figure 36: Diagrammes de Ramachandran pour les glycines sauf la <sup>1</sup>Gly

Des simulations plus longues ont été programmées afin de s'assurer que nos observations sont statistiquement valables. De plus des simulations de type « replica exchange » vont être menées pour prendre en considération l'influence des isoméries *cis-trans*.

*Calcul des constantes de couplage*  $^3J_{H\alpha-NH}$  – Nous avons pu extraire les constantes de couplage  $^3J_{H\alpha-NH}$  pour la triple hélice et le monomère sur la base des différentes structures générées au cours de ces deux dynamiques moléculaires. Le calcul des couplages moyens obtenus lors de la dynamique pourra être directement comparé à nos mesures RMN (Tableau 5).

			<sup>7</sup> Gly	<sup>6</sup> Gly	<sup>5</sup> Gly	<sup>4</sup> Gly	<sup>3</sup> Gly	<sup>2</sup> Gly	<sup>1</sup> Gly
<b><sup>3</sup>J<sub>H<math>\alpha</math>-H<math>\alpha</math>2</sub></b>									
M	Dynamique		6,03	6,78	6,30	5,01	5,29	5,75	6,12
	RMN		6,13	4,53	5,14	6,35	6,24	6,30	5,96
TH	Dynamique	A	6,85	5,60	5,60	5,57	5,64	5,29	7,21
		B	5,98	5,62	5,60	5,61	5,65	4,62	7,24
		C	5,50	5,50	5,66	5,58	5,60	5,60	5,45
	RMN		5,66	2,32	3,03	3,67	4,33	6,05	6,00
	PDB		6,08	5,55	5,64	4,41	6,48	7,88	4,85
<b><sup>3</sup>J<sub>H<math>\alpha</math>-H<math>\alpha</math>3</sub></b>									
M	Dynamique		7,21	6,33	6,23	5,47	5,37	7,43	4,60
	RMN		5,29	3,02	4,15	5,43	5,26	5,21	5,43
TH	Dynamique	A	7,14	8,01	8,01	8,03	8,00	8,11	5,44
		B	7,81	8,01	8,01	8,00	8,00	8,21	5,65
		C	8,85	7,85	7,99	8,02	8,01	8,01	7,78
	RMN		5,12	?	6,19	6,74	4,57	5,42	4,57
	PDB		8,05	8,25	8,22	8,56	7,88	4,05	8,46

Tableau 5: Constantes de couplage  $^3J_{H\alpha-H\alpha}$  des glycines calculées par dynamique moléculaire (A, B, C : trois brins de la TH) et mesurées par RMN. Les valeurs dérivées de la structure pdb 1CAG sont également indiquées

Les valeurs obtenues par dynamique moléculaire pour la triple hélice sont représentées sur la Figure 37 (trois valeurs par Glycine pour chaque chaîne A, B et C). Seules deux corrélations  ${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha 2}}/{}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha 3}}$  s'écartent des valeurs théoriques. Celles-ci correspondent à la glycine C-terminale sur les deux chaînes A et B ce qui traduit simplement la grande dynamique de ces extrémités, la chaîne C étant davantage stabilisée dans l'hélice en raison du décalage entre les trois brins.

Pour le monomère en revanche, l'ensemble des valeurs obtenues par dynamique moléculaire s'écartent de la courbe de corrélation  ${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha 2}}/{}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha 3}}$  théorique. Ceci traduit à nouveau la dynamique importante du peptide qui, d'après les simulations de modélisation moléculaire, n'est pas stabilisé dans une conformation PPII à température ambiante. La comparaison avec les valeurs des couplages scalaires obtenues par RMN est relativement satisfaisante avec un écart de 1 Hz en moyenne. En revanche, la dynamique moléculaire confirme la grande stabilité de la triple hélice. Sa flexibilité ne peut donc pas être invoquée pour expliquer la divergence entre les couplages mesurés par RMN et les couplages théoriques. L'influence des liaisons hydrogène ou au contraire, de l'échange chimique quand les liaisons sont faiblement stabilisées pourrait être évaluée pour expliquer ces différences.

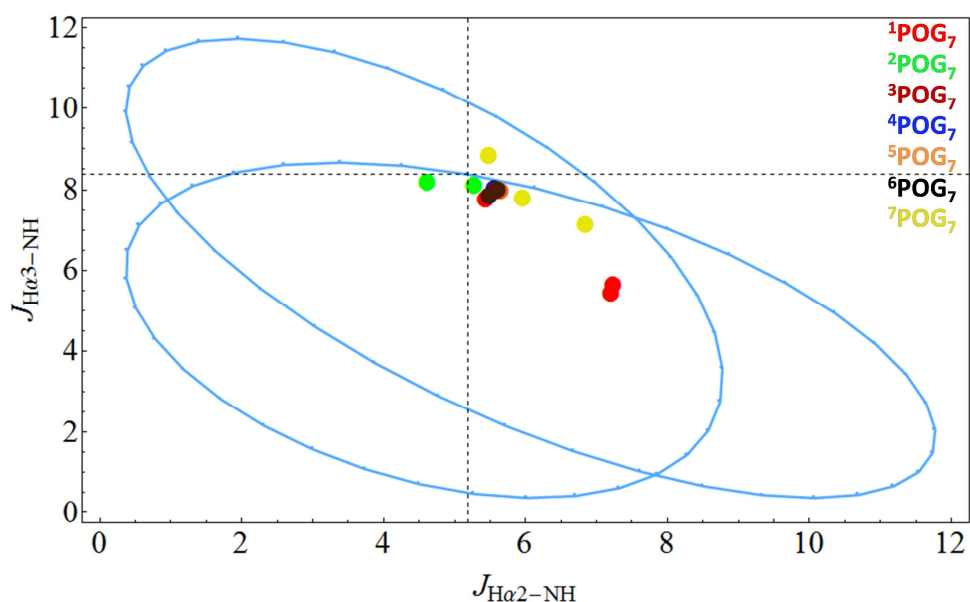


Figure 37: Représentation à deux dimensions des couples de valeurs  ${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha 3}}$  et  ${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha 2}}$ . Un code couleur a été utilisé pour distinguer chaque glycine. La représentation en ligne continue symbolise les couples de constantes de couplage calculés pour chaque valeur de  $\phi$  à partir des paramètres de Karplus théoriques. Les lignes pointillées représentent les valeurs des couplages  ${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha 3}}$  et  ${}^3J_{\text{H}^{\text{N}}\text{H}^{\alpha 2}}$  correspondant à la structure PPII ( $\phi = -70^\circ$ ).

## Conclusions et perspectives

En conclusion, nous avons pu synthétiser et étudier la structure de peptides mimes de collagène (POG)<sub>n</sub>. Des analyses par CD et par RMN en solution ont été menées et ont mis en évidence la présence de deux populations majoritaires en échange. Celles-ci ont été attribuées à une forme monomérique et à une forme structurée en triple hélice. Des résonances <sup>1</sup>H caractéristiques ont pu être attribuées et nous ont servis pour des analyses quantitatives. L'influence de la longueur de chaîne (POG)<sub>n</sub>, du pH, de la température, et de la concentration a été évaluée par CD et RMN, nous donnant des informations concordantes. Des expériences RMN complémentaires ont été enregistrées grâce à un marquage <sup>13</sup>C et <sup>15</sup>N des glycines. Celles-ci ont permis de mettre en évidence des populations monomériques minoritaires. Nous avons également mesuré localement les constantes de couplage <sup>3</sup>J<sub>HN-H $\alpha$</sub>  caractéristiques de l'angle  $\phi$  grâce à l'enregistrement d'expérience CH<sub>2</sub>-TROSY à deux et trois dimensions. Les valeurs obtenues n'étant pas en accord un modèle statique en conformation étendue de type PPII pour le monomère de (POG)<sub>7</sub>, ou avec une conformation figée pour la triple hélice, nous avons entrepris des études de dynamique moléculaire. Alors que les valeurs prédites par simulation numérique pour la forme monomérique se rapprochent des valeurs mesurées, un écart important demeure pour la triple hélice. Des développements théoriques supplémentaires pourront être effectués, intégrant la grande quantité de couplages mesurés dans les expériences 3D que nous avons mises au point ou en prenant en considération les échanges chimiques au niveau des protons amides. Des mesures RDC pourront également être réalisées afin d'obtenir davantage de contraintes structurales sur ces oligomères.

Un phénomène important que nous avons pu mettre en évidence lors de ces études en solution est la formation d'assemblages supramoléculaires de taille variable, au-delà de la triple hélice. Ces aspects seront traités en détail dans le chapitre suivant.

## Partie expérimentale

### A. Tests colorés :

#### A1. Kaiser Test

*Réactifs :* Ninhydrine

Phénol

Cyanure de potassium

Ethanol

Pyridine

*Solutions pour les tests :* A. Ninhydrine (5%, w/v) dans de l'éthanol

B. Phénol (4:1, w/v) dans de l'éthanol

C. Cyanure de potassium (2%, v/v, d'une solution à 1 mmol.l<sup>-1</sup>) dans la pyridine

*Procédure :* Le test est réalisé par addition de 4 gouttes de A, 2 gouttes de B, et 2 gouttes de C à l'échantillon (environ 4-5 mg de peptide-résine) contenue dans un tube à essai. Le mélange est ensuite chauffé à 100°C pendant 5 min. Une coloration bleue de la solution indique qu'il y a une amine primaire libre en solution. Il est important de noter que certains acides aminés peuvent donner des colorations inhabituelles allant du rouge au bleu (notamment Asn, Cys, Ser, Thr).

#### A2. Test chloranil

*Réactifs:* Acétaldéhyde

Chloranil

DMF

*Solutions pour les tests :* A. Acétaldéhyde (2%, v/v) dans du DMF

B. Chloranil (2%, v/v) dans du DMF

*Procédure :* Le test est réalisé par addition de 2 gouttes de A et 2 gouttes de B à l'échantillon (environ 4-5 mg de peptide-résine) contenu dans un tube à essai. La réaction est instantanée. Une coloration verte de la solution indique qu'il y a une amine primaire/secondaire libre en solution.

### B. Préparation des échantillons de CD et traitement de données

Les mesures ont été réalisées sur un spectrophotomètre JASCO. Les échantillons à 100µM et 200µM sont préparés par dilution d'une solution à 10mM de peptide afin de limiter les erreurs de mesures dues aux faibles pesées. Le tampon acétate (pH=3.13) est préparé avec CH<sub>3</sub>COONa et CH<sub>3</sub>COOH et le tampon phosphate de potassium (pH=7) avec du K<sub>2</sub>HPO<sub>4</sub> et KH<sub>2</sub>PO<sub>4</sub>. Les mesures sont réalisées dans des cuves en quartz de trajet optique 1mm (Hellma GmbH) pour les échantillons à 100µM et 200µM et dans des cuves de 0,01mm (Hellma GmbH) pour les échantillons à 10mM. Tous les échantillons ont été incubés pendant 24h avant les mesures.

Paramètres d'acquisition :

*Spectre à température constante :*

Bandwith : 2nm  
 Balayage : 190-280nm  
 Vitesse de balayage : 50nm/min  
 Pas de mesure : 0,2nm  
 Nombre d'accumulations : 3  
 Température: 20°C ou 4°C

*Etude en température:*

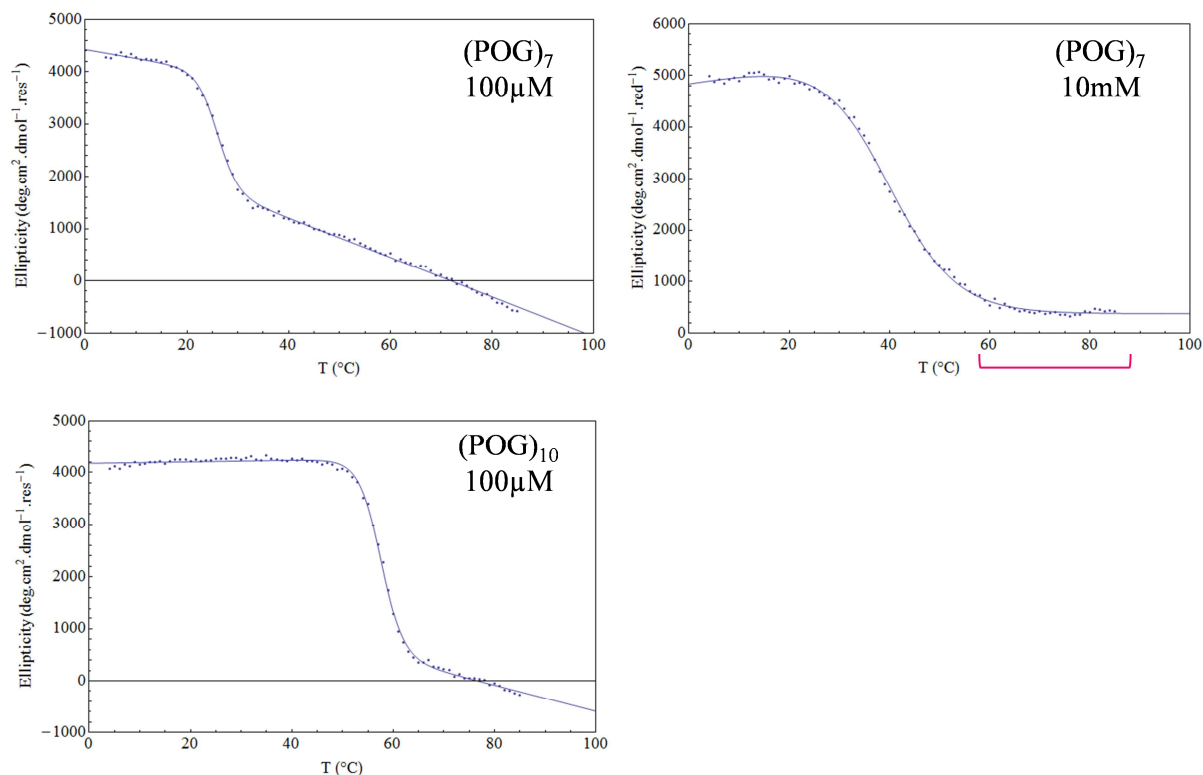
Bandwith : 2nm  
 Balayage : 210-270nm  
 Vitesse de balayage : 50nm/min  
 Pas de mesure : 0,2nm  
 Rampe de température: 10C/h  
 Pas de température : 1°C  
 Intervalle de température : 4 à 85°C  
 Nombre d'accumulations : 2

Traitement des données :

Le logiciel permettant l'acquisition et le traitement de données est SpectraManager®. Les données ont été exploitées sur Mathematica®. La correction de la ligne de base a été réalisée de la façon suivante : Correction de ligne base =  $[\theta]_{190-270}$  - Moyenne  $[\theta]_{260-270}$ . Puis les courbes ont été lissées sur 25 points.

Pour l'étude en température, chaque courbe a été corrigée et lissée puis les valeurs à 225 nm ont été prises afin de tracer la courbe  $f(T) = [\theta]_{225}$  et de déduire la température de fusion par ajustement avec l'équation de Richards.

*Spectre dichroïque des peptides (POG)<sub>7</sub> et (POG)<sub>10</sub> à pH=7 permettant la détermination de la température de fusion de la triple hélice :*



Nous remarquons que la température de fusion déterminée pour le peptide (POG)<sub>7</sub> à 10mM est approximative, en raison d'une évaporation de la solution à haute température. Ceci est observable par l'absence de décroissance linéaire du signal caractéristique du PPII (repérée ici en rouge) après la décroissance brutale.

### C. SPPS pour les peptides marqués <sup>13</sup>C et <sup>15</sup>N

Les synthèses comportant la glycine marquée <sup>13</sup>C ont été effectuées sur Syro2 sans micro-ondes et celles comportant la glycine marquée <sup>13</sup>C et <sup>15</sup>N ont été effectuées sur un Liberty Blue avec micro-ondes.

La résine utilisée pour les synthèses de <sup>n</sup>(POG)<sub>7</sub>, n=2-7, est une résine Wang préchargée avec une glycine, loading 0,36 mmol/g. Les synthèses ont été réalisées en parallèle, en travaillant sur 0.054mmol (150mg). La résine est alors incorporée dans le réacteur.

Pour les peptides marqués en position 1, la résine utilisée est une résine Wang, loading 0.22mmol/g (230mg). Le premier acide aminé à être incorporé est la glycine marquée. Ce couplage se fait à la main. A 150 mg de Fmoc-Gly-OH (0.5mmol, 10eq) sont ajoutés 0,5 ml de DCM avec du DIC (113.6mg, 0.22mmol, 5eq) au goutte à goutte à 0°C. La réaction est laissée pendant 20min avant d'évaporer le DCM, puis reprise dans le DMF. Pendant ce temps, la résine est mise à imbiber dans 3ml de DMF, puis la glycine dans le DMF est ajoutée à la résine avec du DMAP (0.1éq) pendant 1h. La résine est alors incorporée au réacteur. Les acides aminés (4eq) ont été ajoutés uns à uns pendant 40min, excepté la glycine marquée (2eq). Un double couplage a été réalisé avec les prolines car nous avons remarqué que ce couplage était critique. Chaque acide aminé est activé par un mélange HATU/DIEA (sans micro-ondes) ou DIC/oxyma pure (avec micro-ondes) avant couplage. Entre chaque ajout une déprotection de 3 min avec une solution à 40% pipéridine dans le DMF a eu lieu, suivie d'une déprotection de 15 min avec une solution à 20% pipéridine dans le DMF. Ensuite, le clivage des peptides de la résine est réalisé avec 3 ml d'une solution de TFA:H<sub>2</sub>O:TIS (95:2,5 :2,5) pendant 1h. Ensuite, le TFA a été évaporé puis les peptides sont précipités dans l'éther diéthylique froid (3\*10ml). La solution est ensuite centrifugée à 4000rpm pendant 6 min. Le culot est repris dans de l'eau puis lyophilisé. Les peptides ont été analysés par UPLC/MS (Acquity UPLC BEH C18 1.7µm, 2.1\*50mm). Les peptides ont été purifiés sur une colonne C18 (Xbridge 5µM, 19\*50mm).

Peptide	Rendements <sup>13</sup> C (%)	Rendements <sup>13</sup> C- <sup>15</sup> N (%)
<sup>1</sup> (POG) <sub>7</sub>	43	*
<sup>2</sup> (POG) <sub>7</sub>	36	19
<sup>3</sup> (POG) <sub>7</sub>	47	
<sup>4</sup> (POG) <sub>7</sub>	48	42
<sup>5</sup> (POG) <sub>7</sub>	45	
<sup>6</sup> (POG) <sub>7</sub>	43	26
<sup>7</sup> (POG) <sub>7</sub>	55	9

\* La synthèse n'a pas fonctionné

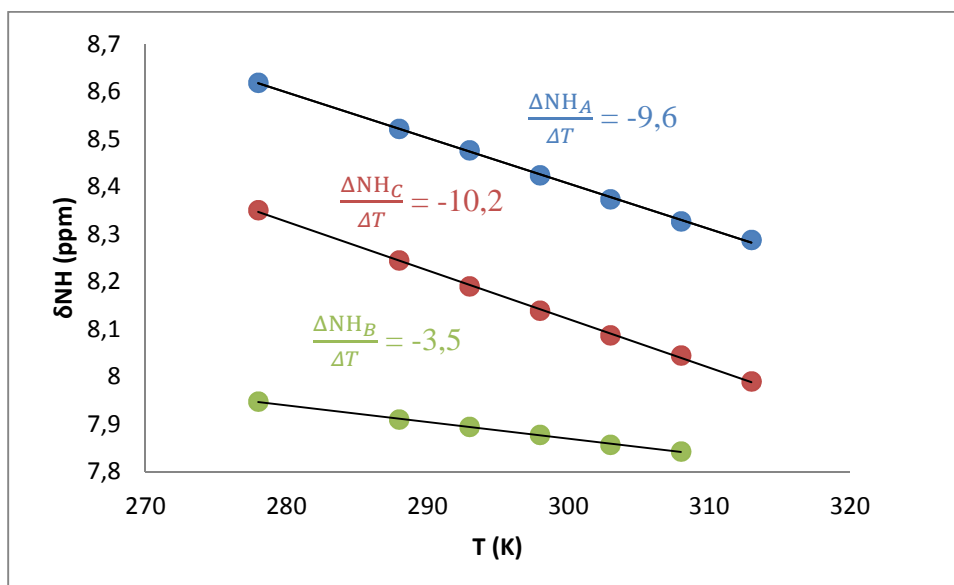
Tableau 6: Rendements des synthèses avec les peptides marqués <sup>13</sup>C et <sup>15</sup>N

### D. Calculs des coefficients de température des populations majoritaires de (POG)<sub>7</sub>

T (K)	δ (ppm)		
	A	C	B
278	8,618	8,35	7,948
288	8,521	8,245	7,91
293	8,476	8,19	7,895
298	8,424	8,139	7,878
303	8,373	8,087	7,857
308	8,327	8,045	7,843
313	8,288	7,99	disparition

Tableau 7: Déplacement chimiques (en ppm) des populations A, B et C en fonction de la température

Le graphique suivant représente les déplacements chimiques des NH en fonction de la température pour les trois populations majoritaires. Les données ont été ajustées par des fonctions affines  $ax+b$  où  $a$  : coefficient de température ( $\Delta NH/\Delta T$  en ppb/K).





E. Calculs des constantes de couplage  $^3J_{\text{H}\alpha\text{-NH}}$ 

$^1(\text{POG})_7$				
Populations	JHN-H $\alpha$ 2	Erreur	JHN-H $\alpha$ 3	Erreur
TH	4.93	0.07	5.24	0.13
	5.94	0.05	4.74	0.05
	<b>6.00</b>	<b>0.02</b>	<b>4.57</b>	<b>0.02</b>
PPII	<b>5.96</b>	<b>0.002</b>	<b>5.43</b>	<b>0.003</b>

$^2(\text{POG})_7$				
Populations	JHN-H $\alpha$ 2	Erreur	JHN-H $\alpha$ 3	Erreur
A	6.10	0.07	5.37	0.07
PPII	<b>6.30</b>	<b>0.01</b>	<b>5.25</b>	<b>0.01</b>
C	6.54	0.04	?	?
PPI	<b>5.47</b>	<b>0.04</b>	<b>5.70</b>	<b>0.07</b>
TH	<b>6.05</b>	<b>0.05</b>	<b>5.42</b>	<b>0.05</b>
	3.82	0.09	5.93	0.15
	3.71	0.16	5.84	0.12
	2.45	0.28	5.26	0.18

$^3(\text{POG})_7$				
Populations	JHN-H $\alpha$ 2	Erreur	JHN-H $\alpha$ 3	Erreur
PPI	5.08	0.07	5.50	0.10
PPII	6.24	0.01	5.26	0.01
TH	4.33	0.14	4.57	0.18

$^4(\text{POG})_7$				
Populations	JHN-H $\alpha$ 2	Erreur	JHN-H $\alpha$ 3	Erreur
PPI	5.48	0.07	6.04	0.10
PPII	6.35	0.01	5.43	0.01
TH	3.67	0.10	6.74	0.09

$^5(\text{POG})_7$				
Populations	JHN-H $\alpha$ 2	Erreur	JHN-H $\alpha$ 3	Erreur
PPI	1.88	0.20	3.02	0.28
PPII	5.14	0.07	4.15	0.05
TH	3.03	0.19	6.19	0.17

$^6(\text{POG})_7$				
Populations	JHN-H $\alpha$ 2	Erreur	JHN-H $\alpha$ 3	Erreur
PPI	3.66	0.23	2.24	0.30
PPII	4.56	0.03	3.01	0.04
TH	2.32	0.20	?	?

$^7(\text{POG})_7$				
Populations	JHN-H $\alpha$ 2	Erreur	JHN-H $\alpha$ 3	Erreur
PPI	5.37	0.04	5.82	0.03
PPII	6.13	0.01	5.29	0.01
TH	<b>5.66</b>	<b>0.09</b>	<b>5.12</b>	<b>0.08</b>
	4.65	0.14	4.95	0.14
	3.52	0.15	5.79	0.14

## F. Valeurs des constantes de couplage obtenues sur les expériences 3D

Constante de couplage pour le monomère M									Moyenne	
${}^2J_{C'H\alpha 2}$										
Valeurs (Hz)	4,48	4,36	4,35	4,65						<b>4,45</b>
Erreur (Hz)	0,01	0,01	0,01	0,01						<b>0,01</b>
${}^2J_{C'H\alpha 3}$										
Valeurs (Hz)	4,86	4,85	4,81	4,88						<b>4,85</b>
Erreur (Hz)	0,01	0,01	0,01	0,01						<b>0,01</b>
${}^1J_{C\alpha C'}$										
Valeurs (Hz)	52,31	52,03	53,22	53,34	53,22	52,03	53,04	52,66	<b>52,66</b>	
Erreur (Hz)	0,07	0,05	0,07	0,05	0,07	0,05	0,06	0,06	<b>0,03</b>	
${}^1J_{C\alpha H\alpha 2} + {}^1J_{C\alpha H\alpha 3}$										
Valeurs (Hz)	278,11	277,84	279,88	278,69	278,52	278,64	279,52	279,14	<b>278,78</b>	
Erreur (Hz)	0,06	0,05	0,06	0,06	0,06	0,06	0,06	0,06	<b>0,02</b>	
${}^1J_{C\alpha H\alpha 2} + {}^2J_{H\alpha 2 H\alpha 3}$										
Valeurs (Hz)	157,47	157,35	157,13	157,43						<b>157,35</b>
Erreur (Hz)	0,01	0,01	0,01	0,01						<b>0,01</b>
${}^1J_{C\alpha H\alpha 3} + {}^2J_{H\alpha 2 H\alpha 3}$										
Valeurs (Hz)	157,06	157,05	156,94	157,01						<b>157,01</b>
Erreur (Hz)	0,01	0,01	0,01	0,01						<b>0,01</b>
${}^1J_{NHN}$										
Valeurs (Hz)	93,72	93,46	92,95	92,50	93,99	93,25	94,04	93,41	<b>93,41</b>	
Erreur (Hz)	0,04	0,04	0,04	0,05	0,05	0,05	0,05	0,05	<b>0,02</b>	
${}^2J_{NH\alpha 2}$										
Valeurs (Hz)	6,10	5,80	6,01	6,15						<b>6,00</b>
Erreur (Hz)	0,01	0,01	0,01	0,01						<b>0,01</b>
${}^2J_{NH\alpha 3}$										
Valeurs (Hz)	5,05	5,01	5,09	5,13						<b>5,06</b>
Erreur (Hz)	0,01	0,01	0,01	0,01						<b>0,00</b>
${}^2J_{C\alpha HN}$										
Valeurs (Hz)	2,94	3,35	4,24	3,97	3,74	3,80	4,65	3,62	<b>3,72</b>	
Erreur (Hz)	0,05	0,05	0,05	0,06	0,07	0,07	0,07	0,06	<b>0,02</b>	
${}^3J_{HNH\alpha 2}$										
Valeurs (Hz)	4,12	3,47	3,04	3,20						<b>3,45</b>
Erreur (Hz)	0,05	0,04	0,05	0,04						<b>0,02</b>
${}^3J_{HNH\alpha 3}$										
Valeurs (Hz)	4,12	3,61	3,21	3,82						<b>3,68</b>
Erreur (Hz)	0,05	0,04	0,05	0,05						<b>0,02</b>

Constante de couplage pour la triple hélice TH									Moyenne
${}^2J_{C'H\alpha 2}$									
Valeurs (Hz)	3,66	3,49	1,57	3,45					<b>3,40</b>
Erreur (Hz)	0,21	0,04	0,15	0,06					<b>0,03</b>
${}^2J_{C'H\alpha 3}$									
Valeurs (Hz)	5,69	5,11	4,92	4,50					<b>5,00</b>
Erreur (Hz)	0,16	0,06	0,09	0,10					<b>0,04</b>
${}^1J_{CaC'}$									
Valeurs (Hz)	53,24	51,81	56,10	53,13	55,89	52,16	52,38	52,87	<b>52,53</b>
Erreur (Hz)	0,82	0,14	0,51	0,26	0,50	0,23	0,33	0,42	<b>0,14</b>
${}^1J_{CaH\alpha 2} + {}^1J_{CaH\alpha 3}$									
Valeurs (Hz)	281,70	280,28	279,82	276,09	278,25	275,27	278,18	278,67	<b>278,81</b>
Erreur (Hz)	0,26	0,24	0,33	0,29	0,35	0,33	0,33	0,42	<b>0,11</b>
${}^1J_{CaH\alpha 2} + {}^2J_{Ha2Ha3}$									
Valeurs (Hz)	158,57	158,39	157,79	159,68					<b>158,60</b>
Erreur (Hz)	0,08	0,94	0,08	0,08					<b>0,04</b>
${}^1J_{CaH\alpha 3} + {}^2J_{Ha2Ha3}$									
Valeurs (Hz)	156,84	156,26	157,10	156,69					<b>156,71</b>
Erreur (Hz)	0,09	0,08	0,08	0,11					<b>0,04</b>
${}^1J_{NH'N}$									
Valeurs (Hz)	92,83	93,83	93,61	93,66	98,86	93,82	94,73	94,06	<b>93,54</b>
Erreur (Hz)	0,12	0,20	0,18	0,25	0,53	0,39	0,39	0,25	<b>0,08</b>
${}^2J_{NH\alpha 2}$									
Valeurs (Hz)	4,26	4,29	2,97	5,07					<b>4,30</b>
Erreur (Hz)	0,04	0,06	0,26	0,12					<b>0,03</b>
${}^2J_{NH\alpha 3}$									
Valeurs (Hz)	5,43	6,04	5,00	5,78					<b>5,63</b>
Erreur (Hz)	0,06	0,08	0,15	0,09					<b>0,04</b>
${}^2J_{CaHN}$									
Valeurs (Hz)	5,13	4,32	6,44	5,02	0,12	6,89	2,98	3,38	<b>5,04</b>
Erreur (Hz)	0,16	0,27	0,21	0,34	0,86	0,51	0,49	0,32	<b>0,10</b>
${}^3J_{HNH\alpha 2}$									
Valeurs (Hz)	2,14	4,13	1,03	1,90					<b>2,87</b>
Erreur (Hz)	0,20	0,15	0,30	0,23					<b>0,10</b>
${}^3J_{HNH\alpha 3}$									
Valeurs (Hz)	1,86	2,51	1,46	2,51					<b>2,24</b>
Erreur (Hz)	0,30	0,21	0,43	0,34					<b>0,15</b>

Constante de couplage pour la triple hélice TH					Moyenne	Constante de couplage pour le monomère M					Moyenne
${}^2J_{C\alpha 2}$											
Valeurs (Hz)	3,18	3,67	<b>3,48</b>			4,97	4,66	<b>4,84</b>			
Erreur (Hz)	0,16	0,13	<b>0,10</b>			0,11	0,14	<b>0,09</b>			
${}^2J_{C\alpha 3}$											
Valeurs (Hz)	5,70	6,15	<b>5,88</b>			5,06	4,81	<b>4,95</b>			
Erreur (Hz)	0,12	0,15	<b>0,10</b>			0,14	0,12	<b>0,09</b>			
${}^1J_{C\alpha C'}$											
Valeurs (Hz)	50,09	48,28	54,16	54,23	<b>51,67</b>	52,02	51,73	52,63	52,23	<b>52,19</b>	
Erreur (Hz)	0,10	0,20	0,16	0,13	<b>0,07</b>	0,12	0,15	0,12	0,12	<b>0,06</b>	
${}^1J_{C\alpha H\alpha 2} + {}^1J_{C\alpha H\alpha 3}$											
Valeurs (Hz)	281,00	280,33	281,40	281,46	<b>281,09</b>	279,32	279,74	278,94	278,75	<b>278,96</b>	
Erreur (Hz)	0,26	0,33	0,26	0,31	<b>0,14</b>	0,24	0,28	0,28	0,27	<b>0,13</b>	
${}^1J_{C\alpha H\alpha 2} + {}^2J_{H\alpha 2 H\alpha 3}$											
Valeurs (Hz)	158,91	159,74	<b>159,30</b>			157,24	156,89	<b>157,03</b>			
Erreur (Hz)	0,06	0,07	<b>0,05</b>			0,04	0,03	<b>0,03</b>			
${}^1J_{C\alpha H\alpha 3} + {}^2J_{H\alpha 2 H\alpha 3}$											
Valeurs (Hz)	157,16	157,93	<b>157,50</b>			156,81	156,79	<b>156,81</b>			
Erreur (Hz)	0,09	0,10	<b>0,07</b>			0,04	0,04	<b>0,03</b>			

La constante de couplage  ${}^2J_{H\alpha 2 H\alpha 3} = [({}^1J_{C\alpha H\alpha 3} + {}^2J_{H\alpha 2 H\alpha 3}) + ({}^1J_{C\alpha H\alpha 3} + {}^2J_{H\alpha 2 H\alpha 3}) - ({}^1J_{C\alpha H\alpha 2} + {}^1J_{C\alpha H\alpha 3})]/2$  avec une erreur égale à la somme des erreurs moyennes.

Les constantes de couplages  ${}^1J_{C\alpha H\alpha n} = ({}^1J_{C\alpha H\alpha n} + {}^2J_{H\alpha 2 H\alpha 3}) - {}^2J_{H\alpha 2 H\alpha 3}$  où les erreurs sont égales à la somme des erreurs sur chaque constante de couplage.

## G. La dynamique moléculaire sur [(POG)<sub>10</sub>]<sub>3</sub>

### G1. Les paramètres de simulations

*Généralités* – Toutes les simulations ont été effectuées en utilisant GROMACS (version 5.0.5) avec le champ de force Amber99sb-ILDN-phi (inspirés par *Aliev et al.*)<sup>81</sup> et le modèle de l'eau TIP3P. Ce champ de force est basé sur Amber99, avec des paramètres améliorés pour le squelette, les paramètres ont été améliorés pour la proline et l'hydroxyproline afin de simuler les bons plissements de cycle. La triple hélice initiale a été construite en utilisant la structure cristalline du peptide [(GPO)<sub>10</sub>]<sub>3</sub> (structure PBD:1CAG), les acides aminés non désirés ont été supprimés afin de construire un [(POG)<sub>7</sub>]<sub>3</sub>. Ces modifications ainsi que l'ajout des hydrogènes manquant ont été faits avec l'outil de pdb2gmx.

*Paramètres de la boîte de production* – Pour chacune des simulations des conditions aux limites périodiques ont été utilisées. La triple hélice n'est pas protégée dans les extrémités (COO<sup>-</sup> et NH<sub>2</sub><sup>+</sup>), trois cations Na<sup>+</sup> ainsi que trois anions Cl<sup>-</sup> ont donc été ajoutés pour la neutralité. Elle a été solvatée avec 3714 molécules d'eau dans une boîte parallélépipède de dimensions 3,67, 3,64 et 8,99 nm, où le dernier axe est aligné avec l'axe de l'hélice. Le monomère (POG)<sub>7</sub> a été construit en extrayant une chaîne de la triple hélice. De la même manière, un cation de Na<sup>+</sup> et un anion de Cl<sup>-</sup> ont été ajoutés pour la neutralité. Un premier équilibrage de 10 ns dans un ensemble NVT (N : nombres de particule, V : volume, T : température) a été réalisé dans une boîte de simulation de grande taille (cube de longueur L = 8,45 nm) remplie de 19802 molécules d'eau. Après plusieurs nanosecondes, la conformation étendue obtenue à partir de la triple hélice initiale tend vers une forme plus compacte. Pour le cycle de production, ce monomère plus compact a été utilisé comme point de départ, et solvaté avec 5244 molécules d'eau dans une boîte cubique plus petite (L = 5,49 nm).

Toutes les 2 fs, une trajectoire est calculée et toutes les 1 ou 10 ps une structure est sélectionnée. La liste des voisins, qui est une approximation des atomes en interaction, est mise à jour toutes les 10 étapes. La méthode de maillage d'Ewald (PME) a été employée pour le calcul des énergies électrostatiques avec une interpolation du 4<sup>e</sup> ordre. Dans la dernière optimisation du champ de force par *Aliev et al.*, deux cut-off sont utilisés pour définir les liaisons de Van des Waals afin de rendre l'algorithme plus stable, le premier de 10Å et le deuxième de 1Å.

La température a été contrôlée à 313K en utilisant la vitesse de rééchelonnement avec un terme stochastique (V-rescale) égal à 42 et une constante de temps de 0,1 ps, pour la protéine et le reste du système.

Un schéma de Parrinello-Rahman a été employé pour le contrôle de la pression à 1 bar en utilisant une constante de couplage de 2 ps et une compressibilité isotherme de 4.5x10<sup>-5</sup> bar.

Avant le lancement de la simulation, le système a été minimisé en utilisant la plus grande pente et des algorithmes de gradient conjugués. Les étapes de minimisation ont été suivies par trois étapes d'équilibration. Les deux premières étapes ont été réalisées avec la molécule de soluté fixe, pour permettre aux molécules d'eau d'équilibrer le système, la première dynamique moléculaire a été faite dans un ensemble NVT de 100 ps, et la seconde dans un ensemble TNP (P : Pression) de 100 ps. Ensuite, une simulation TNP est

lancée pour vérifier la détente du volume sans aucune contrainte sur la protéine. Enfin, pour la production, une dynamique NVT de 2000 ns a été effectuée en utilisant le dernier volume obtenu. Les configurations ont été référencées chaque tranche de 100 ps.

Pour la triple hélice, la direction de la triple hélice a été imposée afin qu'elle reste approximativement le long de l'axe le plus long de la boîte de simulation, appelé axe z. Ceci a été réalisé en imposant des contraintes sur deux atomes (les C $\alpha$  de la GLY6 et GLY12 du premier monomère), nommé ici A et B, afin que la liaison factice A-B soit alignée avec l'axe z. La première contrainte a été de fixer les trois coordonnées de l'atome A ( $x_a$ ,  $y_a$ ,  $z_a$ ), la deuxième contrainte a été d'imposer à l'atome B de se déplacer le long de l'axe z seulement, ce qui équivaut à fixer ses deux coordonnées ( $x_b$ ,  $y_b$ ). Pour éviter les artefacts apparaissant lors de l'utilisation à la fois de l'ensemble TNP et de la contrainte sur les positions des atomes, les simulations de production ont été réalisées dans un ensemble NVT.

*Analyse de la production* – La trajectoire a été analysée par VMD. Les liaisons hydrogène ont été analysées en imposant que celles-ci soient formées avec les atomes polaires (X=N, O), distants de maximum 3,5Å et un angle X-H-X' inférieur à 25 °. Les couplages RMN  $^3J_{\text{HN-H}\alpha}$  pour les glycines ont été calculés en utilisant la relation de Karplus obtenu par *Aliev et al.*

## G2. Données issues de la simulation

### *La triple hélice*

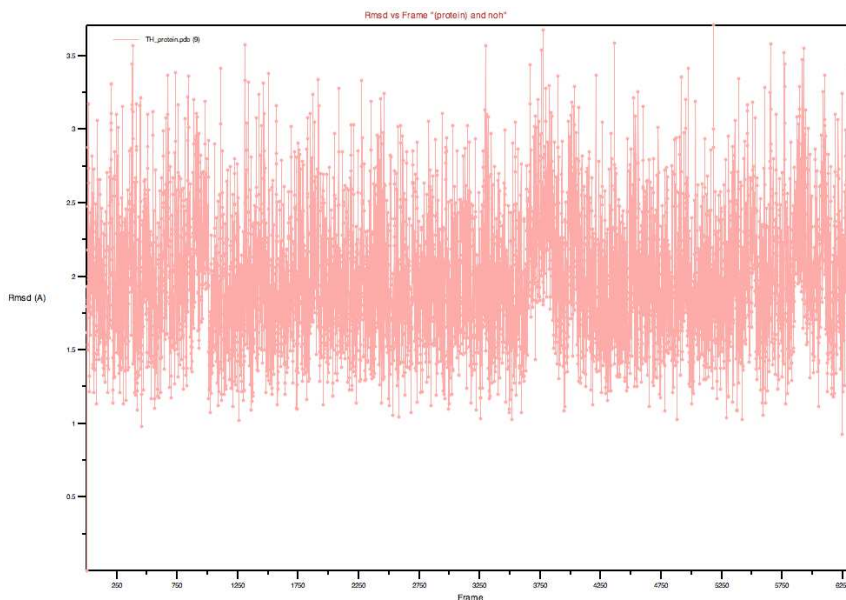


Figure 38: Graphique représentant le *rmsd* en fonction du temps de la simulation

Donneur	Accepteur	Présence de liaisons H (%)	A	B	C
PRO-21	PRO-0	4.98			
GLY-2	HYP-22	9.86	7		
GLY-29	PRO-51	10.39		5	
GLY-56	PRO-12	11.28			3
GLY-32	PRO-54	12.13		4	
GLY-5	PRO-24	13.01	6		
GLY-20	PRO-39	22.43	1		
GLY-44	HYP-1	23.5			7
GLY-8	PRO-27	26.80	5		
GLY-23	PRO-45	33.83		7	
GLY-47	PRO-3	40.47			6
GLY-53	PRO-9	42.58			4
GLY-35	PRO-57	46.52		3	
GLY-26	PRO-48	50.05		6	
GLY-14	PRO-33	50.80	3		
GLY-41	GLY20	52.82		1	
GLY-38	PRO-60	54.55		2	
GLY-50	PRO-6	60.88			5
GLY-11	PRO-30	68.06	4		
GLY-17	PRO-36	72.09	2		

Tableau 8: Calcul de la présence de liaisons hydrogène suivant la chaîne (A, B, C) et sa place dans la chaîne (Cter→Nter)

### Le monomère

		<sup>7</sup> X	<sup>6</sup> X	<sup>5</sup> X	<sup>4</sup> X	<sup>3</sup> X	<sup>2</sup> X	<sup>1</sup> X
<b>X</b>								
Pro	$\chi_m$	41,4	37,6	37,9	37,8	38,0	37,7	37,8
	P	24	122	110	90	93	112	120
Hyp	$\chi_m$	39,0	39,0	38,8	38,8	39,0	39,4	39,0
	P	26	41	41	29	36	32	32

Tableau 9: Calcul des paramètres  $\chi_m$  et P en fonction de la position de la glycine marquée

	Distance entre <sup>1</sup> AA et <sup>21</sup> AA (Å)	Rayon de giration	R <sub>g</sub> <sup>3</sup>	Asphéricité	Acyllindricité	Forme anisotropie
Valeurs	10,1	8,76	7,73	0,42	0,53	0,64
Erreur	0,1	0,1	0,1	0,01	0,02	0,01

Tableau 10: Paramètres décrivant la conformation moyenne du monomère au cours de la dynamique moléculaire

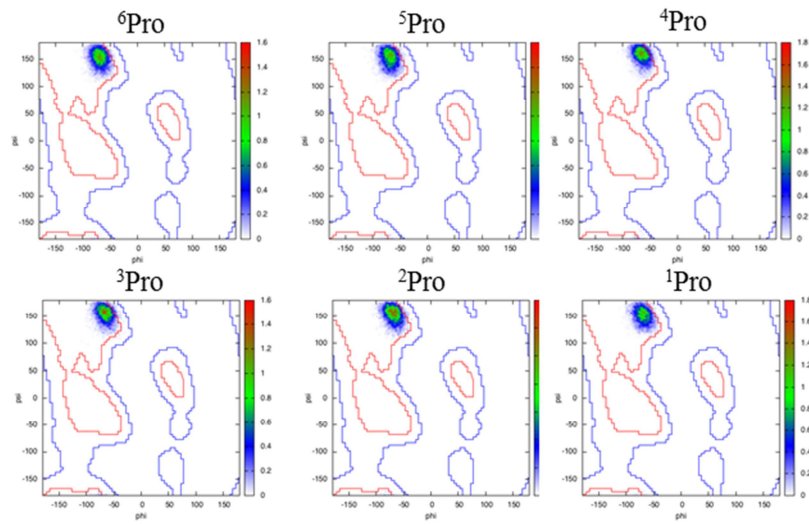


Figure 39: Diagrammes de Ramachandran pour les prolines du monomère, excepté celle en position  $N_{ter}$

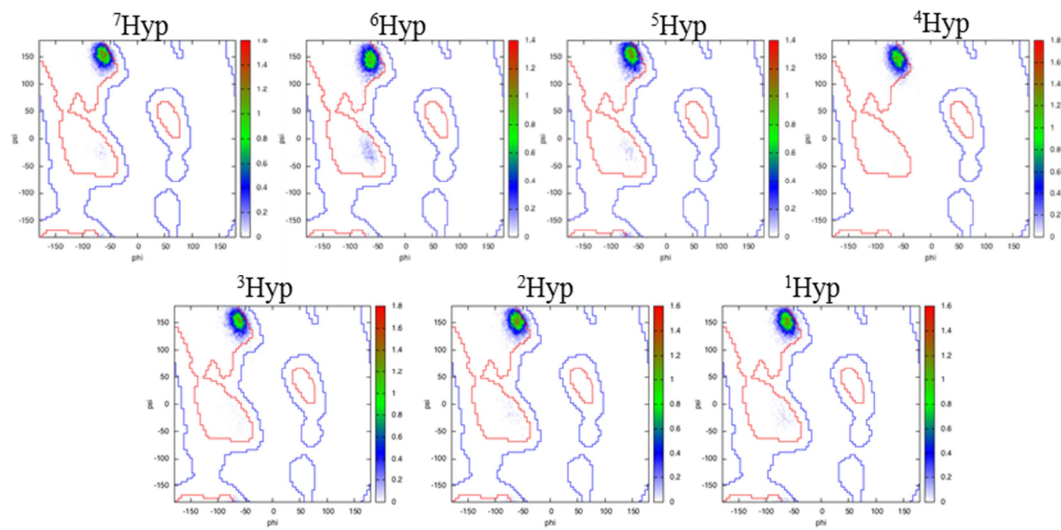


Figure 40: Diagrammes de Ramachandran pour les hydroxyprolines du monomère



## Références du Chapitre 1

- 1 Van Der Rest, M. & Garrone, R. Collagen family of proteins. *FASEB* **5**, 2814-2823 (1991).
- 2 Ricard-Blum, S. The collagen family. *Cold Spring Harbor perspectives in biology* **3**, a004978, doi:10.1101/cshperspect.a004978 (2011).
- 3 Engel, J. & Bachinger, H. P. Structure, Stability and Folding of the Collagen Triple Helix. *Top. Curr. Chem.* **247**, 7-33, doi:10.1007/b103818 (2005).
- 4 Engel, J. THE ZIPPER-LIKE FOLDING OF COLLAGEN TRIPLE HELICES AND THE EFFECTS OF MUTATIONS THAT DISRUPT THE ZIPPER. *Annu Rev Biophys. Biophys. Chem.* **20**, 137-152 (1991).
- 5 Myllyharju, J. & Kivirikko, K. I. Collagens, modifying enzymes and their mutations in humans, flies and worms. *Trends Genet* **20**, 33-43, doi:10.1016/j.tig.2003.11.004 (2004).
- 6 Leikina, E., Merts, M. V., Kuznetsova, N. & Leikin, S. Type I collagen is thermally unstable at body temperature. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 1314-1318, doi:10.1073/pnas.032307099 (2002).
- 7 Persikov, A. V. & Brodsky, B. Unstable molecules form stable tissues. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 1101-1103, doi:10.1073/pnas.042707899 (2002).
- 8 Boedtker, H. & Doty, P. A study of gelatin molecules, aggregates and gels. *J Phys Chem* **58**, 968-983 (1954).
- 9 Team, M. G. C. P. Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proceedings of the National Academy of Sciences* **99**, 16899-16903, doi:10.1073/pnas.242603899 (2002).
- 10 Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annual review of biochemistry* **78**, 929-958, doi:10.1146/annurev.biochem.77.032207.120833 (2009).
- 11 Brodsky, B. & Ramshaw, J. A. M. The collagen triple helix structure. *Matrix Biology* **15**, 545-554 (1997).
- 12 Adzhubei, A. A., Sternberg, M. J. & Makarov, A. A. Polyproline-II helix in proteins: structure and function. *Journal of molecular biology* **425**, 2100-2132, doi:10.1016/j.jmb.2013.03.018 (2013).
- 13 Holmgren, S. K., Taylor, K. M., Bretscher, L. E. & Raines, R. T. Code for collagen's stability deciphered. *Nature* **392** (1998).
- 14 Yang, W., Chan, V. C., Kirkpatrick, A., Ramshaw, J. A. M. & Brodsky, B. Gly-Pro-Arg Confers Stability Similar to Gly-Pro-Hyp in the Collagen Triple-helix of Host-Guest Peptides. *The Journal of biological chemistry* **272**, 28837-28840 (1997).
- 15 Ackerman, M. S. *et al.* Sequence Dependence of the Folding of Collagen-like Peptides. *J Bio Chem* **274**, 7668-7673 (1999).
- 16 Horng, J. C., Kotch, F. W. & Raines, R. T. Is glycine a surrogate for a D-amino acid in the collagen triple helix? *Protein science : a publication of the Protein Society* **16**, 208-215, doi:10.1110/ps.062560107 (2007).
- 17 Weiss, M. S., Jabs, A. & Hilgenfeld, R. Peptide bonds revisited. *Nat Struct Biol* **5**, 676 (1998).
- 18 Aszodi, A., Legate, K. R., Nakchbandi, I. & Fassler, R. What mouse mutants teach us about extracellular matrix function. *Annual review of cell and developmental biology* **22**, 591-621, doi:10.1146/annurev.cellbio.22.010305.104258 (2006).

- 19 Taylor, D. Fracture and repair of bone: a multiscale problem. *Journal of Materials Science* **42**, 8911-8918, doi:10.1007/s10853-007-1698-3 (2007).
- 20 Athanasiou, K. A. & Johns, D. E. Design characteristics for temporomandibular joint disc tissue engineering: learning from tendon and articular cartilage. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine* **221**, 509-526, doi:10.1243/09544119jeim158 (2007).
- 21 Torbet, J. *et al.* Orthogonal scaffold of magnetically aligned collagen lamellae for corneal stroma reconstruction. *Biomaterials* **28**, 4268-4276, doi:10.1016/j.biomaterials.2007.05.024 (2007).
- 22 Wei, W. *et al.* A multicenter, double-blind, randomized, controlled phase III clinical trial of chicken type II collagen in rheumatoid arthritis. *Arthritis Res Ther* **11**, R180, doi:10.1186/ar2870 (2009).
- 23 Graumlich, J. F. *et al.* Healing Pressure Ulcers with Collagen or Hydrocolloid: A Randomized, Controlled Trial. *J Am Geriatr Soc* **51**, 147-154 (2003).
- 24 Lee, S. K., Posthauer, M. E., Dorner, B., Redovian, V. & Maloney, M. J. Pressure Ulcer Healing with a Concentrated, Fortified, Collagen Protein Hydrolysate Supplement: A Randomized Controlled Trial. *Advances in Skin & Wound Care* **19**, 92-96 (2006).
- 25 Sakakibara, S. *et al.* Synthesis of (Pro-Hyp-Gly)<sub>n</sub> of defined molecular weights Evidence for the stabilization of collagen triple helix by hydroxyproline. *Biochimica et biophysica acta* **303**, 198-202 (1973).
- 26 Shah, N. K., Ramshaw, J. A. M., Kirkpatrick, A., Shah, C. & Brodsky, B. A Host-Guest Set of Triple-Helical Peptides: Stability of Gly-X-Y Triplets Containing Common Nonpolar Residues†. *Biochemistry* **35**, 10262-10268 (1996).
- 27 Ramshaw, J. A. M., Shah, N. K. & Brodsky, B. Gly-X-Y Tripeptide Frequencies in Collagen: A Context for Host-Guest Triple-Helical Peptides. *Journal of structural biology* **122**, 86-91 (1998).
- 28 Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A. & Brodsky, B. Peptide investigations of pairwise interactions in the collagen triple-helix. *Journal of molecular biology* **316**, 385-394 (2002).
- 29 Persikov, A. V., Ramshaw, J. A. M., Kirkpatrick, A. & Brodsky, B. Electrostatic Interactions Involving Lysine Make Major Contributions to Collagen Triple-Helix Stability. *Biochemistry* **44**, 1414-1422 (2005).
- 30 Rainey, J. K. & Goh, M. C. An interactive triple-helical collagen builder. *Bioinformatics* **20**, 2458-2459, doi:10.1093/bioinformatics/bth24710.102/ci034177z (2004).
- 31 Melacini, G., Feng, Y. & Goodman, M. Acetyl-Terminated and Template-Assembled Collagen-Based Polypeptides Composed of Gly-Pro-Hyp Sequences. 3. Conformational Analysis by <sup>1</sup>H-NMR and Molecular Modeling Studies. *Journal of the American Chemical Society* **118**, 10359-10364 (1996).
- 32 Feng, Y., Melacini, G., Taulane, J. P. & Goodman, M. Acetyl-Terminated and Template-Assembled Collagen-Based Polypeptides Composed of Gly-Pro-Hyp Sequences. 2. Synthesis and Conformational Analysis by Circular Dichroism, Ultraviolet Absorbance, and Optical Rotation. *Journal of the American Chemical Society* **118**, 10315-10358 (1996).
- 33 Kwak, J., Da Capua, A., Locardi, E. & Goodman, M. TREN (Tris(2-aminoethyl)amine): An Effective Scaffold for the Assembly of Triple Helical Collagen Mimetic Structures. *Journal of the American Chemical Society* (2002).
- 34 Kinberger, G. A., Cai, W. & Goodman, M. Collagen Mimetic Dendrimers. *Journal of the American Chemical Society* **124**, 15162-15163 (2002).

- 35 Cai, W. *et al.* Facile and efficient assembly of collagen-like triple helices on a TRIS scaffold. *Bioorg Chem* **35**, 327-337, doi:10.1016/j.bioorg.2007.01.003 (2007).
- 36 Cai, W. *et al.* Scaffold Assembly of Collagen-Like Triple Helices at the C-Terminus. *Letters in Organic Chemistry* **4**, 96 (2007).
- 37 Goodman, M., Melacini, G. & Feng, Y. Collagen-like triple helices incorporating peptoid residues. *Journal of the American Chemical Society* **118**, 10928-10929 (1996).
- 38 Feng, Y., Melacini, G. & Goodman, M. Collagen-Based Structures Containing the Peptoid Residue N-Isobutylglycine (Nleu): Synthesis and Biophysical Studies of Gly-Nleu-Pro Sequences by Circular Dichroism and Optical Rotation†. *Biochemistry* **36**, 8716-8724 (1997).
- 39 Melacini, G., Feng, Y. & Goodman, M. Collagen-Based Structures Containing the Peptoid Residue N-Isobutylglycine (Nleu): Conformational Analysis of Gly-Nleu-Pro Sequences by 1H-NMR and Molecular Modeling. *Biochemistry* **36**, 8725-8732 (1997).
- 40 Barth, D., Kyrieleis, O., Frank, S., Renner, C. & Moroder, L. The role of cystine knots in collagen folding and stability, part II. Conformational properties of (Pro-Hyp-Gly)<sub>n</sub> model trimers with N- and C-terminal collagen type III cystine knots. *Chemistry* **9**, 3703-3714, doi:10.1002/chem.200304918 (2003).
- 41 Henkel, W. *et al.* Synthesis and folding of native collagen III model peptides. *Biochemistry* **38**, 13610-13622 (1999).
- 42 Kotch, F. W. & Raines, R. T. Self-assembly of synthetic collagen triple helices. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3028-3033, doi:10.1073/pnas.0508783103 (2006).
- 43 Cai, W., Kwok, S. W., Taulane, J. P. & Goodman, M. Metal-assisted Assembly and Stabilization of Collagen-like Triple Helices. *Journal of the American Chemical Society* **126**, 15030-15031 (2004).
- 44 Kinberger, G. A., Taulane, J. P. & Goodman, M. FeIII binding collagen mimetics. *Inorg. Chem.* **45**, 961-963 (2006).
- 45 Yu, Y.-C., Berndt, P., Tirrell, M. & Fields, G. B. Self-Assembling Amphiphiles for Construction of Protein Molecular Architecture. *Journal of the American Chemical Society* **118**, 12515-12520 (1996).
- 46 Yu, S. M., Tirrell, M. & Fields, G. B. Minimal Lipidation Stabilizes Protein-Like Molecular Architecture. *Journal of the American Chemical Society* **120**, 9979-9987 (1998).
- 47 Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* **85**, 2149-2154, doi:10.1021/ja00897a025 (1963).
- 48 Wellings, D. A. Standard Fmoc Protocols. *Methods in enzymology* **289**, 44-67 (1997).
- 49 Boas, U. & Mirsharghi, S. Color test for selective detection of secondary amines on resin and in solution. *Organic letters* **16**, 5918-5921, doi:10.1021/ol502936d (2014).
- 50 Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochimica et biophysica acta* **1751**, 119-139, doi:10.1016/j.bbapap.2005.06.005 (2005).
- 51 Suehiro, T. *et al.* Temperature-dependent higher order structures of the (Pro-Pro-Gly)(1)(0)-modified dendrimer. *Biopolymers* **95**, 270-277, doi:10.1002/bip.21576 (2011).

- 52 Suehiro, T., Kojima, C., Tsumura, S., Harada, A. & Kono, K. Higher order structure of short collagen model peptides attached to dendrimers and linear polymers. *Biopolymers* **93**, 640-648, doi:10.1002/bip.21413 (2010).
- 53 Feng, Y., Melacini, G., Taulane, J. P. & Goodman, M. Collagen-Based Structures Containing the Peptoid Residue N-Isobutylglycine (Nleu): Synthesis and Biophysical Studies of Gly-Pro-Nleu Sequences by Circular Dichroism, Ultraviolet Absorbance, and Optical Rotation. *Biochemistry* **39**, 859-872 (1996).
- 54 Goodman, M., Bhumralkar, M., Jefferson, E. A., Kwak, J. & Locardi, E. Collagen mimetics. *Biopolymers* **47**, 127-142 (1998).
- 55 Miclet, E., Jacquot, Y., Goasdoué, N. & Lavielle, S. Solution structural study of a proline-rich decapeptide. *Comptes Rendus Chimie* **11**, 486-492, doi:10.1016/j.crci.2007.06.016 (2008).
- 56 Kuemin, M., Engel, J. & Wennemers, H. Temperature-induced transition between polyproline I and II helices: quantitative fitting of hysteresis effects. *Journal of peptide science : an official publication of the European Peptide Society* **16**, 596-600, doi:10.1002/psc.1245 (2010).
- 57 Sanz-Nebot, V., Toro, I., Benavente, F. & Barbosa, J. pKa values of peptides in aqueous and aqueous-organic media Prediction of chromatographic and electrophoretic behaviour. *J Chromatogr A*. **942**, 145-156 (2002).
- 58 Creamer, T. P. Left-Handed Polyproline II Helix Formation Is (Very) Locally Driven. *Proteins* **33**, 218-226 (1998).
- 59 Mizuno, K., Boudko, S. P., Engel, J. & Bachinger, H. P. Kinetic hysteresis in collagen folding. *Biophysical journal* **98**, 3004-3014, doi:10.1016/j.bpj.2010.03.019 (2010).
- 60 Makarov, A. A., Adzhubei, A. A., Protasevich, I. I., Lobachov, V. M. & Fasman, G. D. Melting of the left-handed helical conformation of charge poly-L-lysine. *Biopolymers* **34**, 1123-1124 (1994).
- 61 Lakshminarayanan, R. *et al.* Analysis of secondary structure and self-assembly of amelogenin by variable temperature circular dichroism and isothermal titration calorimetry. *Proteins* **76**, 560-569, doi:10.1002/prot.22369 (2009).
- 62 Mu, C., Li, D., Lin, W., Ding, Y. & Zhang, G. Temperature induced denaturation of collagen in acidic solution. *Biopolymers* **86**, 282-287, doi:10.1002/bip.20742 (2007).
- 63 Acevedo-Jake, A. M., Jalan, A. A. & Hartgerink, J. D. Comparative NMR analysis of collagen triple helix organization from N- to C-termini. *Biomacromolecules* **16**, 145-155, doi:10.1021/bm501281a (2015).
- 64 Miclet, E. *et al.* Relaxation-Optimized NMR Spectroscopy of Methylene Groups in Proteins and Nucleic Acids. *Journal of the American Chemical Society* **126**, 10560-10570, doi:10.1021/ja047904v (2004).
- 65 Melacini, G., Bonvin, A. M. J. J., M., G., Boelens, R. & Kaptein, R. Hydration dynamics of the collagen triple helix by NMR. *Journal of molecular biology* **300**, 1041-1048 (2000).
- 66 Riek, R., Wider, G., Pervushin, K. & Wüthrich, K. Polarization transfer by cross-correlated relaxation in solution NMR with very large molecules. *Proceedings of the National Academy of Sciences of the United States of America* **96**, 4918-4923 (1999).
- 67 Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in

- solution. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 12366-12371 (1997).
- 68 Fernandez, C. & Wider, G. TROSY in NMR studies of the structure and function of large biological macromolecules. *Curr Opin Struct Biol* **13**, 570-580, doi:10.1016/j.sbi.2003.09.009 (2003).
- 69 Gueron, M., Leroy, J. L. & Griffey, R. H. Proton nuclear magnetic relaxation of <sup>15</sup>N-labeled nucleic acids via dipolar coupling and chemical shift anisotropy. *Journal of the American Chemical Society* **105**, 7262-7266 (1983).
- 70 Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in <sup>13</sup>C-Labeled Proteins. *Journal of the American Chemical Society* **120**, 6394-6400, doi:10.1021/ja980742g (1998).
- 71 Brutscher, B., Boisbouvier, J., Pardi, A., Marion, D. & Simorre, J.-P. Improved Sensitivity and Resolution in <sup>1</sup>H-<sup>13</sup>C NMR Experiments of RNA. *Journal of the American Chemical Society* **120**, 11845-11851, doi:10.1021/ja982853l (1998).
- 72 Fiala, R., Czernek, J. & Sklenar, V. Transverse relaxation optimized triple-resonance NMR experiments for nucleic acids. *J. Biomol. NMR* **16**, 291-302 (2000).
- 73 Weininger, U., Diehl, C. & Akke, M. <sup>13</sup>C relaxation experiments for aromatic side chains employing longitudinal- and transverse-relaxation optimized NMR spectroscopy. *Journal of Biomolecular NMR* **53**, 181-190, doi:10.1007/s10858-012-9650-5 (2012).
- 74 Tugarinov, V., Hwang, P. M., Ollerenshaw, J. E. & Kay, L. E. Cross-Correlated Relaxation Enhanced <sup>1</sup>H-<sup>13</sup>C NMR Spectroscopy of Methyl Groups in Very High Molecular Weight Proteins and Protein Complexes. *Journal of the American Chemical Society* **125**, 10420-10428, doi:10.1021/ja030153x (2003).
- 75 Aliev, A. E. & Courtier-Murias, D. Experimental Verification of Force Fields for Molecular Dynamics Simulations Using Gly-Pro-Gly-Gly. *The journal of physical chemistry. B* **114**, 12358-12375 (2010).
- 76 Miclet, E., Boisbouvier, J. & Bax, A. Measurement of eight scalar and dipolar couplings for methine-methylene pairs in proteins and nucleic acids. *J Biomol NMR* **31**, 201-216, doi:10.1007/s10858-005-0175-z (2005).
- 77 Bodenhausen, G. & Ruben, D. J. Natural abundance nitrogen-15 by enhanced heteronuclear spectroscopy. *Chem. Phys. Lett.* **69** (1980).
- 78 Salvador, P., Tsai, I. H. & Dannenberg, J. J. J-coupling constants for a trialanine peptide as a function of dihedral angles calculated by density functional theory over the full Ramachandran space. *Phys Chem Chem Phys* **13**, 17484-17493, doi:10.1039/c1cp20520j (2011).
- 79 Beauchamp, K. A., Lin, Y. S., Das, R. & Pande, V. S. Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements. *J Chem Theory Comput* **8**, 1409-1414, doi:10.1021/ct2007814 (2012).
- 80 Piana, S., Lindorff-Larsen, K. & Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization? *Biophysical journal* **100**, L47-49, doi:10.1016/j.bpj.2011.03.051 (2011).
- 81 Aliev, A. E. *et al.* Motional timescale predictions by molecular dynamics simulations: case study using proline and hydroxyproline sidechain dynamics. *Proteins* **82**, 195-215, doi:10.1002/prot.24350 (2014).
- 82 Bella, J., Eaton, M., Brodsky, B. & Berman, H. M. Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science (New York, N.Y.)* **266**, 75-81 (1994).

- 83 Park, S., Radmer, R. J., Klein, T. E. & Pande, V. S. A new set of molecular mechanics parameters for hydroxyproline and its use in molecular dynamics simulations of collagen-like peptides. *J Comput Chem* **26**, 1612-1616, doi:10.1002/jcc.20301 (2005).
- 84 Markley, J. L. *et al.* Recommendations for the presentation of NMR structures of proteins and nucleic acids. *Journal of molecular biology* **280**, 933-952 (1998).
- 85 Cannon, J. W., Aronovitz, J. A. & Goldbart, P. Equilibrium distribution of shapes or linear and star molecules. *J Phys I* **1**, 629-645 (1991).

## ANNEXE 1. Classement des collagènes suivant leur appartenance, leur localisation et leur composition<sup>10</sup>

Type	Classe	Composition	Localisation	Pathologie
I	Fibrillaire	$\alpha 1[\text{I}]_2\alpha 2[\text{I}]$	Abondante et répandue: derme, os, tendons, ligaments	Ostéogénèse imparfaite, syndrome d'Ehlers-Danlos, ostéoporose
II	Fibrillaire	$\alpha 1[\text{II}]_3$	Cartilage, vitreux	Arthrite, chondrodysplasies
III	Fibrillaire	$\alpha 1[\text{III}]_3$	La peau, les vaisseaux sanguins, l'intestin	Syndrome d'Ehlers-Danlos, anévrismes artériels
IV	Réseau	$\alpha 1[\text{IV}]_2\alpha 2[\text{IV}]$ $\alpha 3[\text{IV}]\alpha 4[\text{IV}]\alpha 4[\text{IV}]$ $\alpha 5[\text{IV}]\alpha 6[\text{IV}]$	Les membranes basales	Syndrome d'Alport
V	Fibrillaire	$\alpha 1[\text{V}]_3$ $\alpha 1[\text{V}]_2\alpha 2[\text{V}]$ $\alpha 1[\text{V}]\alpha 2[\text{V}]\alpha 3[\text{V}]$	Répandue: derme, os, cornée, placenta	Syndrome d'Ehlers-Danlos
VI	Réseau	$\alpha 1[\text{VI}]\alpha 2[\text{VI}]\alpha 3[\text{VI}]$ $\alpha 1[\text{VI}]\alpha 2[\text{VI}]\alpha 4[\text{VI}]$	Répandue: derme, os, cornée, cartilage	La myopathie de Bethlem
VII	Fibrilles d'ancrage	$\alpha 1[\text{VII}]_2\alpha 2[\text{VII}]$	Derme, vessie	Epidermolyse bulleuse acquise
VIII	Réseau	$\alpha 1[\text{VIII}]_3$ $\alpha 2[\text{VIII}]_3$ $\alpha 1[\text{VIII}]_2\alpha 2[\text{VIII}]$	Répandue: derme, cerveau, cœur, rein	Dystrophie cornéenne endothéliale de fuchs
IX	FACIT	$\alpha 1[\text{IX}]\alpha 2[\text{IX}]\alpha 3[\text{IX}]$	Cartilage, vitreux, cornée	Arthrite, dysplasie épiphysaire multiple
X	Réseau	$\alpha 1[\text{X}]_3$	Cartilage	Chondrodysplasies
XI	Fibrillaire	$\alpha 1[\text{XI}]\alpha 2[\text{XI}]\alpha 3[\text{XI}]$	Cartilage, disques intervertébraux	Chondrodysplasies, arthrite
XII	FACIT	$\alpha 1[\text{XII}]_3$	Derme, tendons	-
XIII	MACIT	-	Cellules endothéliales, derme, yeux, cœur	-
XIV	FACIT	$\alpha 1[\text{XIV}]_3$	Répandue: os, derme, cartilage	-
XV	MULTIPLEXIN	-	Capillaires, testicules, rein, cœur	-
XVI	FACIT	-	Derme, reins	-
XVII	MACIT	$\alpha 1[\text{XVII}]_3$	Hémidesmosomes dans l'épithélium	Epidermolyse bulleuse jonctionnelle généralisée atrophique
XVIII	MULTIPLEXIN	-	Membranes basales, foie	Syndrome de knobloch
XIX	FACIT	-	Membranes basales	-
XX	FACIT	-	Cornée (poussin)	-
XXI	FACIT	-	Estomac, reins	-
XXII	FACIT	-	Jonctions des tissus	-
XXIII	MACIT	-	Cœur, rétine	-
XXIV	Fibrillaire	-	Os, cornée	-
XXV	MACIT	-	Cerveau, cœur, testicules	Maladies amyloïdes ?
XXVI	FACIT	-	Testicules, ovaires	-
XXVI I	Fibrillaire	-	Cartilage	-
XXVI II	-	-	Derme, nerf sciatique	Maladies neurodégénératives ?

## ANNEXE 2. Séquence des brins $\alpha 1$ et $\alpha 2$ du collagène de type I<sup>9</sup>

### Séquence du brin $\alpha 1$ du collagène de type I avant les modifications post-traductionnelles

1	MFSFVDLRLL	LLLAATALLT	HGQEEGQVEG	QDEDI <b>PP</b> ITC	VQNGLRYHDR	DVWK <b>PE</b> PCRI
61	CVCNDGKVLG	DDVICDETKN	CPGA <b>EV</b> PEGE	CC <b>PVC</b> PDGSE	SPTDQ <b>ETT</b> GV	EG <b>PK</b> GDTGPR
121	GRGPAGPPG	RDGIPGQ <b>PGL</b>	PGPPGPPGPP	GPPGLG <b>GN</b> FA	PQLSYGYDEK	STGGISV <b>PPG</b>
181	MGPSGPRGLP	GPPGAPG <b>PQG</b>	FQGGPPG <b>EPGE</b>	PGASG <b>PMGP</b>	GPPGPPG <b>KNG</b>	DDGEAG <b>KPGR</b>
241	PGERGPPG <b>PQ</b>	GARGL <b>PGTAG</b>	LP <b>GMK</b> GHRGF	SGLD <b>GAK</b> GDA	SGLD <b>GAK</b> GDA	SGLD <b>GAK</b> GDA
301	MGPRGL <b>PGER</b>	GR <b>PGAP</b> GPAG	GR <b>PGAP</b> GPAG	GR <b>PGAP</b> GPAG	GPPGF <b>PGAVG</b>	AKGEAG <b>PPG</b>
361	RGSEGPQ <b>GVR</b>	GE <b>PGPP</b> GPAG	AAG <b>PAGN</b> PGA	DGQ <b>PGAK</b> GAN	GAP <b>GIAG</b> APG	F <b>PGARG</b> PSGP
421	QGGPGPPG <b>PK</b>	GNSG <b>EPGA</b> PG	SKGDT <b>GAK</b> GE	PGPV <b>GVQ</b> GGP	GPAGE <b>E</b> GKRG	ARG <b>EPG</b> PTGL
481	PGPPGERG <b>GP</b>	GSRGF <b>PGAD</b> G	VAG <b>PKG</b> PAGE	RGS <b>PGP</b> AGPK	GSP <b>GEAG</b> RPG	EAGL <b>PGA</b> KGL
541	TGSPGS <b>PGPD</b>	GKT <b>GPPG</b> PAG	QDGR <b>PGPP</b> GP	PGAR <b>GQAG</b> V	GF <b>PPK</b> GAA <b>G</b>	EP <b>GKAG</b> ERGV
601	PGPPGAV <b>GA</b>	GKDGEAG <b>AQ</b> G	PPG <b>PAG</b> PAGE	RGEQ <b>GPA</b> GS <b>P</b>	GFQGL <b>PG</b> PAG	PPGEAG <b>PPG</b> E
661	QGVPGDL <b>GAP</b>	GPSGARG <b>ER</b> G	FPGER <b>GVQ</b> GP	PG <b>PAG</b> PRGAN	GAPGND <b>GAK</b> G	DAGAP <b>GA</b> PGS
721	QGAPGLQ <b>GMP</b>	GERGA <b>AGL</b> PG	PKGDR <b>GDA</b> GP	KGAD <b>GS</b> PGKD	GVRGLT <b>GP</b> IG	PPG <b>PAG</b> APGD
781	KGESG <b>PS</b> GA	GPTGARG <b>AP</b> G	DRGE <b>GPP</b> GP	AGFAG <b>PP</b> GAD	GQ <b>PGA</b> KGE <b>PG</b>	DAGAK <b>GDA</b> GP
841	PGPAG <b>PAG</b> PP	GPIGN <b>V</b> GAPG	AKGARG <b>SAG</b> P	PGAT <b>G</b> FP <b>GA</b>	GRV <b>PPG</b> PSG	NAG <b>PPG</b> PPG <b>P</b>
901	AGKEG <b>GK</b> GPR	GETG <b>PAG</b> RG	EVG <b>PPG</b> PPG <b>P</b>	AGEK <b>GS</b> PGAD	GPAG <b>A</b> GT <b>PG</b>	FQ <b>GIA</b> GQRGV
961	VGL <b>PG</b> QRGER	GF <b>PGL</b> PGPSG	EPG <b>KQ</b> GPS <b>GA</b>	SGER <b>GPP</b> GP <b>M</b>	GPPGL <b>AG</b> PPG	ESGREG <b>A</b> PGA
1021	EGSPGRD <b>GSP</b>	GAKGDR <b>GET</b> G	PAG <b>PPG</b> AP <b>GA</b>	PGA <b>PGP</b> V <b>GA</b>	GKSGDR <b>GET</b> G	PAG <b>PAG</b> PV <b>GA</b>
1081	VGARG <b>PAG</b> PQ	GPRGD <b>K</b> GETG	EQGDR <b>G</b> IKGH	RFS <b>GLQ</b> GGP	GPPGS <b>PG</b> EQG	PSGAS <b>GP</b> AG <b>P</b>
1141	RGPPGSAG <b>AP</b>	GKDGL <b>NGL</b> PG	PIG <b>PPG</b> PRGR	TGDAG <b>PV</b> GGP	GPPGPPG <b>PPG</b>	PPSAG <b>FD</b> FSF
1201	L <b>PQ</b> PPQEKAH	DGGRYR <b>RA</b> DD	ANVVRDRD <b>LE</b>	VDTTLK <b>SLS</b> Q	QIENIR <b>S</b> PEG	SRKN <b>P</b> ARTCR
1261	DLKMCHSD <b>WK</b>	SGEY <b>WID</b> PNQ	GCNLD <b>AIK</b> VF	CNMET <b>G</b> ETCV	YPTQ <b>P</b> SVAQK	NWYISK <b>N</b> PKD
1321	KRHVWF <b>G</b> ESM	TDG <b>FQ</b> FEYGG	QGS <b>D</b> PADVAI	QLTFLR <b>LM</b> ST	EASQN <b>I</b> TYHC	KNSVAY <b>MD</b> QQ
1381	TGNLKK <b>ALL</b> L	QGSNEI <b>E</b> IRA	EGNSR <b>F</b> TYSV	TVD <b>G</b> CT <b>SHT</b> G	AWG <b>K</b> T <b>VI</b> EYK	TTK <b>T</b> SRLR <b>II</b>
1441	DVA <b>PLD</b> V <b>GAP</b>	DQEF <b>G</b> FDVGH	VCFL			

### Séquence du brin $\alpha 2$ du collagène de type I avant les modifications post-traductionnelles

1	MLSFVDTRTL	LLLA <b>V</b> TL <b>C</b> LA	TCQSLQE <b>ET</b> V	RK <b>G</b> PAGDR <b>GP</b>	RGER <b>GPP</b> GP <b>P</b>	GRDGED <b>GPT</b> G
61	PPGPPGPPG <b>P</b>	PGLG <b>GN</b> FAAQ	YD <b>GK</b> GV <b>GL</b> P	G <b>PM</b> GLM <b>GP</b> RG	PPGA <b>AG</b> AP <b>GP</b>	QGFQ <b>GP</b> AG <b>EP</b>
121	GE <b>PG</b> QT <b>GP</b> AG	ARG <b>PAG</b> PPG <b>K</b>	AGED <b>GH</b> PG <b>KP</b>	GR <b>P</b> GERGV <b>G</b>	PQ <b>GARG</b> FP <b>GT</b>	PGL <b>PG</b> FK <b>GIR</b>
181	GHNGLD <b>GL</b> KG	Q <b>PGA</b> PGV <b>K</b> GE	PG <b>AP</b> GEN <b>T</b> P	G <b>QT</b> GAR <b>GL</b> PG	ERGR <b>VGA</b> PG <b>P</b>	AGARGSD <b>GS</b> V
241	GPV <b>GP</b> AG <b>PI</b> G	SAG <b>PPG</b> FP <b>GA</b>	PG <b>PK</b> GEI <b>GAV</b>	G <b>NAG</b> PAG <b>PAG</b>	PRGE <b>V</b> L <b>PGL</b>	SG <b>PV</b> GG <b>PN</b> P
301	GANGLT <b>GAK</b> G	AAGL <b>PG</b> VAG <b>A</b>	PGL <b>PG</b> PR <b>GI</b> P	GPVGA <b>AG</b> AT <b>G</b>	ARGL <b>V</b> GE <b>PGP</b>	AGSKG <b>ES</b> GN <b>K</b>
361	GE <b>PG</b> SAG <b>PQ</b> G	PPG <b>PS</b> GEE <b>GK</b>	RGN <b>GE</b> AG <b>SA</b>	GPPG <b>PP</b> L <b>R</b> G	SPGS <b>R</b> L <b>PGA</b>	DGRAG <b>VM</b> GP <b>P</b>
421	GS <b>R</b> AS <b>GP</b> AG	VRG <b>P</b> NGD <b>AG</b> R	PG <b>E</b> PLM <b>GP</b> R	GL <b>PG</b> SP <b>GN</b> IG	PAG <b>KE</b> GP <b>V</b> GL	PGID <b>GR</b> PG <b>PI</b>
481	GPAG <b>ARG</b> EP <b>G</b>	NIG <b>F</b> PG <b>PK</b> GP	TGD <b>P</b> GK <b>NG</b> DK	G <b>HAG</b> LAG <b>ARG</b>	AP <b>GP</b> D <b>GN</b> GA	Q <b>GPP</b> GP <b>Q</b> GV <b>Q</b>
541	GGK <b>GE</b> Q <b>GP</b> AG	PPG <b>FQ</b> GL <b>PG</b> P	SG <b>PAG</b> EV <b>GK</b> P	GER <b>GLH</b> GE <b>FG</b>	L <b>PG</b> PAG <b>PR</b> GE	R <b>GPP</b> GES <b>GAA</b>
601	GPT <b>GP</b> IG <b>S</b> RG	PSG <b>PPG</b> PD <b>GN</b>	KGE <b>P</b> GV <b>V</b> GA <b>V</b>	G <b>TAG</b> PS <b>GP</b> SG	L <b>PG</b> ERGA <b>AGI</b>	PGG <b>K</b> GE <b>K</b> EP
661	GLR <b>GE</b> IG <b>N</b> PG	RDGARG <b>AP</b> GA	VGAP <b>GP</b> AG <b>AT</b>	GDR <b>GE</b> AG <b>AA</b> G	PAG <b>PAG</b> PR <b>GS</b>	PGER <b>GE</b> V <b>GA</b>
721	PN <b>G</b> FAG <b>PAG</b>	AAGQ <b>PG</b> AK <b>GE</b>	RGA <b>K</b> GP <b>K</b> GEN	G <b>VV</b> GP <b>T</b> GP <b>V</b> G	AAG <b>PAG</b> NP <b>G</b> P	PG <b>PAG</b> SR <b>G</b> D <b>G</b>
781	GPP <b>GM</b> T <b>G</b> FP <b>G</b>	AAGRT <b>GPP</b> GP	SG <b>I</b> SG <b>PP</b> GP <b>P</b>	GPAG <b>KE</b> GL <b>R</b> G	PRD <b>Q</b> GP <b>V</b> GR	T <b>GE</b> V <b>GA</b> V <b>GP</b> P
841	GFAG <b>E</b> K <b>GP</b> SG	EAG <b>TAG</b> PP <b>GT</b>	PG <b>PQ</b> LL <b>G</b> AP	G <b>IL</b> GL <b>PG</b> SR <b>G</b>	ERGL <b>PG</b> VAG <b>A</b>	V <b>GE</b> PG <b>PL</b> G <b>IA</b>
901	GPP <b>G</b> ARG <b>PP</b> G	AVGS <b>PG</b> V <b>NG</b> A	PG <b>E</b> AGR <b>D</b> GN <b>P</b>	GND <b>GP</b> PR <b>D</b> G	Q <b>P</b> G <b>H</b> K <b>GER</b> GY	PN <b>IG</b> PV <b>GAA</b>
961	GAP <b>GP</b> H <b>GP</b> V <b>G</b>	PAG <b>K</b> H <b>GN</b> RG <b>E</b>	TG <b>PS</b> GP <b>V</b> GA <b>P</b>	GAV <b>GP</b> RP <b>PS</b> G	PQ <b>GIR</b> GD <b>K</b> GE	P <b>G</b> E <b>K</b> PR <b>GL</b> P
1021	GL <b>K</b> HN <b>GL</b> Q <b>G</b>	L <b>PG</b> IAG <b>HH</b> GD	QG <b>AP</b> SV <b>GA</b> P	G <b>PR</b> GP <b>AG</b> PS <b>G</b>	PAG <b>KD</b> GR <b>T</b> GH	PG <b>T</b> V <b>GP</b> AG <b>IR</b>
1081	SLNNQ <b>I</b> ET <b>LL</b>	TP <b>E</b> GS <b>R</b> KN <b>PA</b>	PG <b>V</b> SGGG <b>Y</b> DF	G <b>YD</b> DFY <b>RAD</b>	Q <b>PR</b> SAP <b>SL</b> RP	KDY <b>EV</b> DAT <b>LK</b>
1141	ETC <b>I</b> RA <b>Q</b> PEN	ETC <b>I</b> RA <b>Q</b> PEN	RTCRD <b>L</b> R <b>L</b> SH	PE <b>W</b> SS <b>G</b> Y <b>Y</b> WI	D <b>PN</b> Q <b>G</b> CT <b>MDA</b>	IKV <b>Y</b> CD <b>F</b> ST <b>G</b>
1201	ETC <b>I</b> RA <b>Q</b> PEN	ETC <b>I</b> RA <b>Q</b> PEN	KDK <b>K</b> H <b>V</b> W <b>L</b> G	ET <b>I</b> NA <b>G</b> S <b>Q</b> FE <b>Y</b>	N <b>VE</b> G <b>V</b> T <b>S</b> K <b>EM</b>	AT <b>Q</b> LAF <b>M</b> R <b>LL</b>
1261	ANYAS <b>Q</b> N <b>I</b> TY	HCK <b>NS</b> I <b>A</b> Y <b>MD</b>	EET <b>G</b> N <b>L</b> K <b>K</b> AV	IL <b>Q</b> GS <b>ND</b> VEL	VA <b>E</b> GN <b>S</b> R <b>F</b> TY	TV <b>L</b> VD <b>G</b> CS <b>KK</b>
1321	TNE <b>W</b> G <b>K</b> T <b>II</b> E	YK <b>T</b> N <b>K</b> PS <b>R</b> LP	FLD <b>IA</b> PL <b>D</b> IG	GAD <b>Q</b> EFF <b>V</b> DI	GP <b>V</b> CF <b>K</b>	





# Chapitre 2

## Caractérisation d'assemblages supramoléculaires de CMP

*“Au-delà de la chimie moléculaire, fondée sur la liaison covalente, s'étend ainsi un domaine qu'on peut nommer supramoléculaire : la chimie des interactions moléculaires, des associations de deux ou plusieurs espèces chimiques, les complexes, et de la liaison intermoléculaire” – Jean Marie Lehn, 1980*

Dans ce chapitre, nous présenterons dans un premier temps les assemblages supramoléculaires solubles que nous avons observés pour les peptides (POG)<sub>n</sub>, avec n = 7 ou 10. Nous décrirons les techniques de diffusion de la lumière et de microscopie qui nous ont permis de les mettre en évidence. Dans un second temps, nous étudierons la présence d'assemblages supramoléculaires, formant des particules insolubles en solution aqueuse. Ensuite, nous proposerons un modèle cinétique, basé sur des études par RMN, permettant d'identifier les processus et de quantifier les vitesses de formation de ces différentes populations en solution. Et enfin, nous présenterons nos tentatives de dopage du collagène par ces peptides.

## 2.1. Mise en évidence de la présence d'assemblages supramoléculaires en solution

Les techniques de diffusion de la lumière ainsi que la microscopie nous ont permis de mettre en évidence la présence d'assemblages supramoléculaires dépendante des contre-ions en solution et dans différentes conditions de concentration et de température.

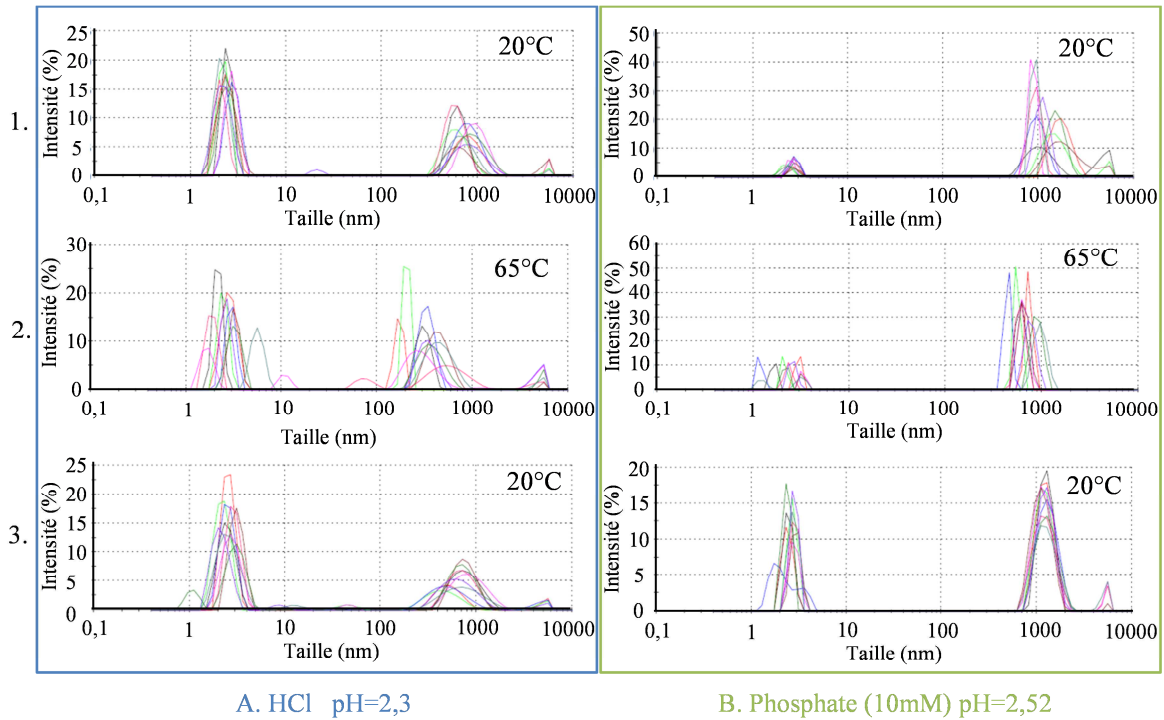
### 2.1.1 La diffusion dynamique de la lumière

*Principe* – La DLS (Dynamic Light Scattering) est une technique permettant d'accéder à la taille des particules en solution, plus précisément à leur rayon hydrodynamique. Cette technique est basée sur le principe de diffusion de la lumière par des particules soumises à des mouvements browniens. La diffusion de lumière peut être en effet directement reliée au coefficient de diffusion des particules dans un milieu donné et, par la suite, à un diamètre de particule par la relation de Stokes-Einstein :

$$R_h = \frac{k_B * T}{6\pi * \eta * D_T}$$

Où  $R_h$  est le rayon hydrodynamique,  $k_B$  est la constante de Boltzmann,  $T$  la température en K,  $D_T$  le coefficient de diffusion translationnel directement mesuré par l'appareil et  $\eta$  la viscosité du solvant. La mesure de rayons hydrodynamiques implique donc que toutes les particules en solution soient assimilées à des sphères ce qui est une approximation assez grossière dans le cas de morphologies allongées. C'est notamment le cas des triples hélices de CMP, toutefois, des coefficients tenant compte d'un facteur de forme peuvent être appliqués pour corriger ces mesures.<sup>1</sup>

*Résultats* – Des études ont été réalisées en parallèle sur les peptides (POG)<sub>7</sub> et (POG)<sub>10</sub> à une concentration de 10mM et en présence de différents contre-ions (Cl<sup>-</sup> et H<sub>2</sub>PO<sub>4</sub><sup>-</sup>) (Figure 1). Les conditions expérimentales ainsi que les estimations de la taille des assemblages sont présentées en partie expérimentale (cf. §A p.133).



**Figure 1: Etude par DLS du peptide (POG)<sub>7</sub>, 10 mM. (1.) 20°C, (2.) 65°C après chauffage à 90°C pendant 30min, (3.) 20°C après chauffage à 90°C puis laissé à température ambiante (20°C) pendant 2h. Chaque mesure a été réalisée 10 fois pour vérifier la répétabilité.**

On attend un rayon hydrodynamique de l'ordre du nanomètre pour le monomère et la triple hélice<sup>2</sup> (voir ANNEXE 3), le rapport de coefficient de diffusion entre une espèce trimérique et monomérique allant de 1,2 à 1,5 selon que l'on considère une absence de structuration ou des modèles globulaires.<sup>3</sup> Dans toutes les conditions testées, immédiatement après la dissociation des triples hélices (chauffage à 90°C) ou après un long temps d'incubation, une population caractérisée par un rayon dynamique moyen de ~2.5 nm est présente, en proportion variable. Cette valeur est compatible avec la présence de monomère et/ou de triples hélices. On observe de surcroît de manière systématique une population correspondant à un haut poids moléculaire (rayon hydrodynamique compris entre 500 nm et 1µm) qui indique la présence d'assemblages supramoléculaires (on parlera de nano-objets O par la suite). L'ordre de grandeur du taux d'agrégation peut être calculé sur la base du rapport des rayons hydrodynamiques :  $R_O/R_M = \sqrt[3]{(V_O/V_M)}$ . Ceci permet d'estimer un assemblage moyen de  $\sim 10^8$  peptides par agrégats. Nous observons une plus grande polydispersité de ces assemblages dans la solution HCl (A) que dans la solution phosphate (B). De plus, les nano-objets sont plus petits dans A (700 nm à 20°C et 350 nm à 65°C) que dans B (1200 nm à 20°C et 700 nm à 65°C). Ceci s'explique par la présence de l'anion Cl<sup>-</sup> chaotrope.<sup>4,5</sup> En effet, *Hofmeister* l'a inclus dans une série de sels qui ont

des effets sur la solubilité des protéines mais également sur la stabilité de leur structure secondaire et tertiaire. Les ions  $\text{Cl}^-$  interfèreraient avec les réseaux de molécules d'eau qui stabilisent la structure des assemblages supramoléculaires. Il a déjà été démontré qu'en présence d'une forte concentration en ions  $\text{Cl}^-$ , l'assemblage du collagène sur des feuilles de mica était différent de celui adopté dans une solution composée d'ions  $\text{SO}_4^{2-}$ .<sup>4</sup> Nous observons également une plus forte intensité du signal monomère-triple hélice pour l'échantillon A. En effet, après un long temps d'incubation dans le tampon phosphate, la proportion de monomère (M) et/ou triple hélice (TH) est très minoritaire par rapport aux assemblages supramoléculaires (Figure 1B 1.), ce qui indique que ceux-ci représentent les espèces stables thermodynamiquement dans le tampon phosphate.

Ceci est cohérent avec nos observations précédentes : les nano-objets (O) étant déstabilisés par les anions  $\text{Cl}^-$ , une plus grande quantité de monomères et de TH est attendue en solution. A partir de ces premières observations, nous pouvons déjà proposer un modèle simple (Figure 2). Nous avons déjà pu mettre en évidence l'équilibre entre le M et la TH par des études de dichroïsme circulaire et de RMN. Les études de DLS nous amènent à imaginer soit un échange entre les M et les O, soit entre la TH et les O.

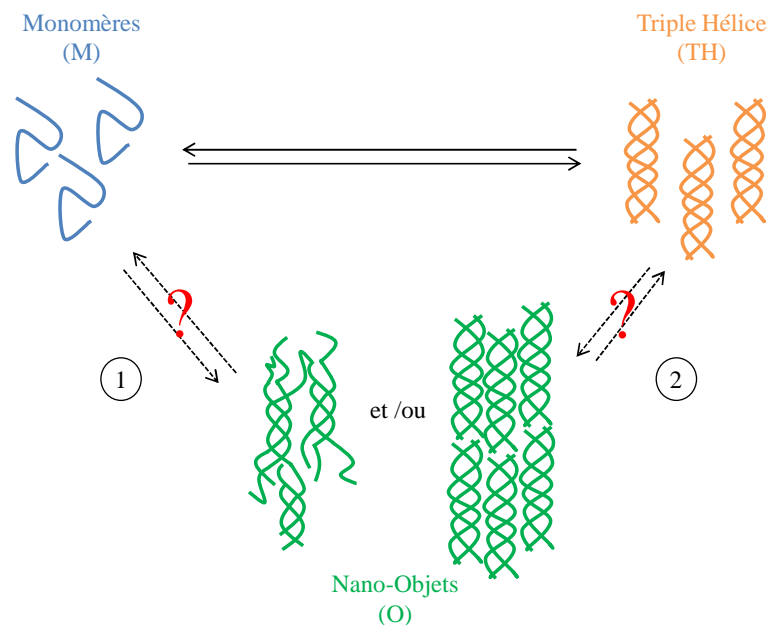
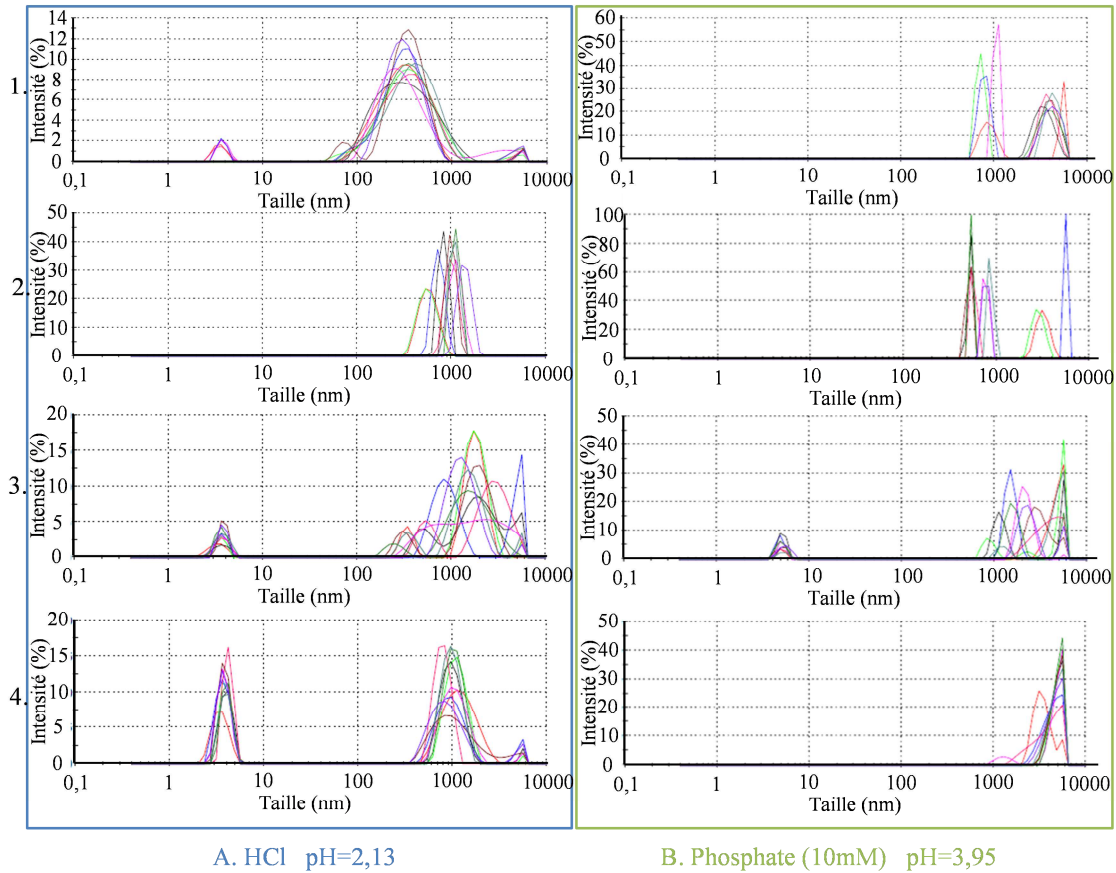


Figure 2: Hypothèses de formation des nano-objets

Nous avons également observé qu'après chauffage des échantillons à 90°C (température nécessaire pour dissocier les assemblages selon les analyses de *Brodsky*,<sup>6</sup> les O étaient toujours présents, ce qui signifie qu'ils ne se sont pas dissociés totalement ou qu'ils se reforment très rapidement (< 5 min).

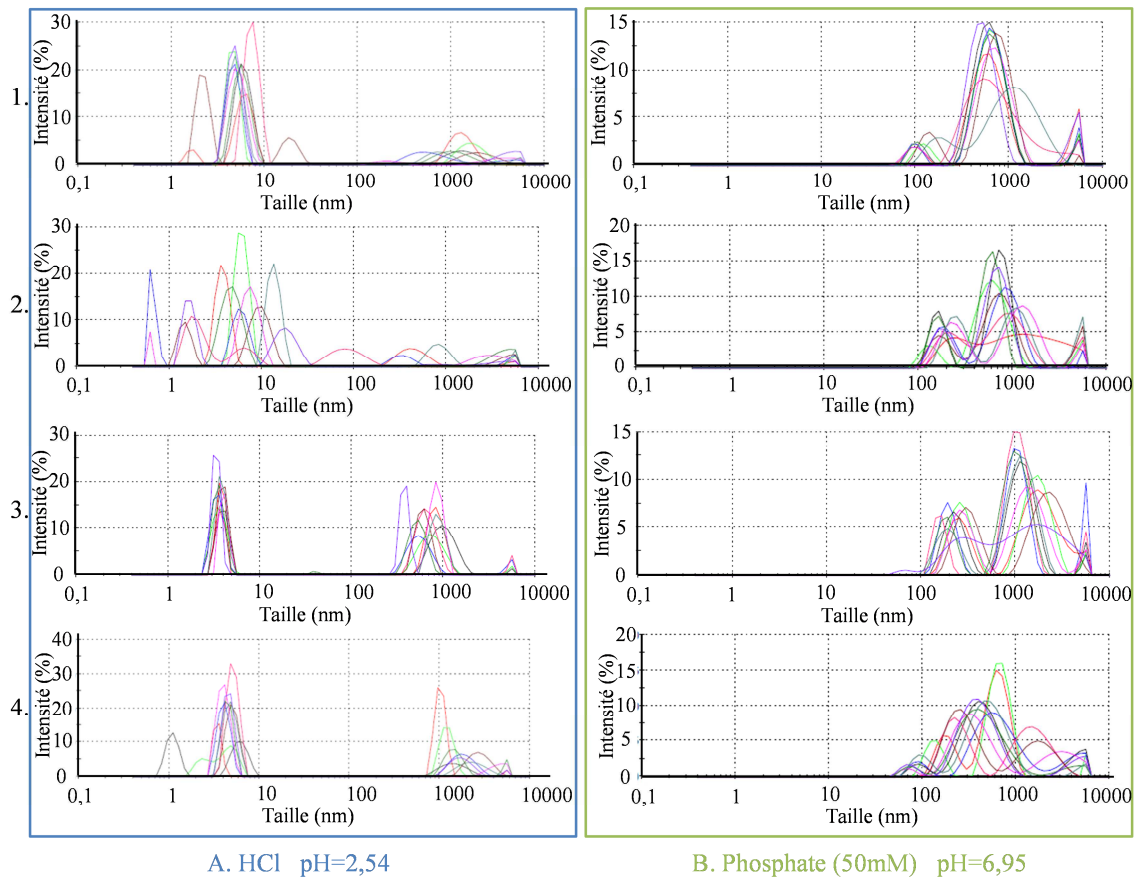
Les études comparatives menées sur le peptide (POG)<sub>10</sub> nous ont apportées des éléments de réponse (Figure 3). On observe en premier lieu un rayon hydrodynamique légèrement plus élevé pour les espèces de faible poids moléculaire (monomère et triples hélices), en accord avec l'allongement de la chaîne peptidique. Ces espèces semblent moins favorisées, étant même absentes à 20°C dans le tampon phosphate après un long temps d'incubation. La polydispersité est également plus grande dans la solution HCl (A) que dans le tampon phosphate (B). Les agrégats ont toujours en moyenne une plus petite taille. Dans les deux cas, à 65°C, on remarque la présence persistante de nano-objets et, à la différence du (POG)<sub>7</sub>, aucune espèce de faible poids moléculaire. Ceci montre la formation rapide des nano-objets de (POG)<sub>10</sub> à partir des espèces dissociées sans passage par la forme TH. Nous avons en effet montré par dichroïsme circulaire, à plus basse concentration (C=100µM), que la formation de la triple hélice de (POG)<sub>10</sub> était beaucoup plus lente que la formation de la triple hélice (POG)<sub>7</sub>. La séquestration de peptides (POG)<sub>10</sub> au sein des nano-objets pourrait expliquer cette observation. Dans la solution A, on s'aperçoit que les triples hélices se forment progressivement au cours du temps à 20°C (Figure 3A.3-4), tandis qu'elles n'apparaissent que transitoirement dans la solution B, moins chaotrope, laissant la place à des espèces de très haut poids moléculaire (Figure 3B.3-4).



**Figure 3: Etude par DLS du peptide (POG)<sub>10</sub>, 10 mM. (1.) 20°C, (2.) 65°C après chauffage à 90°C pendant 30min, (3.) 20°C après chauffage à 90°C puis laissé à température ambiante (20°C) pendant 1h30, (4.) 20°C après incubation à 20°C pendant 1 jour. Chaque mesure a été réalisée 10 fois pour vérifier la répétabilité.**

Des études complémentaires ont été réalisées à une concentration plus faible ( $c = 2.6\text{mM}$ ) dans HCl pH 2,5. Les allures des profils DLS obtenus sont sensiblement différents (Figure 4A). Les espèces de faible poids moléculaire sont à présent omniprésentes, et dans des proportions importantes. On s'aperçoit qu'après un long temps d'incubation, (Figure 4A 1) à 20°C, les espèces monomériques et triple hélices sont largement majoritaires, représentant donc les espèces thermodynamiquement les plus stables à cette concentration. A 65°C et après chauffage à 90°C, les mesures effectuées sont relativement imprécises mais révèlent que les nano-objets sont pratiquement absents. L'ensemble de ces mesures indique donc que la stabilité des O est directement fonction de la concentration des échantillons. *Brodsky et al.*<sup>6</sup> ont réalisé des études similaires sur le peptide (POG)<sub>10</sub> solubilisés dans un tampon PBS à pH=7. Leurs données suggèrent également que la présence d'assemblages supramoléculaires est dépendante de la température d'étude. En particulier, à faible température (4°C) et faible concentration (2.6mM), aucun agrégat n'a été détecté, ce qui est en désaccord avec nos observations (Figure 4B 1-4). Cependant, à plus haute température, les assemblages mis en évidence dans leurs études présentent des

rayons hydrodynamiques près de 10 fois inférieurs aux nôtres (~ 100 nm). Ces différences suggèrent que les assemblages de peptides CMP et leur cinétique de formation sont très dépendants des conditions expérimentales. Dans des conditions similaires à celles de *Brodsky* (Figure 4B 1-4), nos mesures révèlent uniquement la présence de nano-objets qui restent stables même après un temps d'incubation à 90°C (Figure 4B 2). Le changement de charge de la glycine terminale à pH 6,95 pourrait expliquer la stabilisation de ces agrégats par interaction électrostatique.



**Figure 4: Etude par DLS du peptide (POG)<sub>10</sub>, 2.6mM. (1.) 20°C, (2.) 65°C après chauffage à 90°C pendant 30min, (3.) 20°C après incubation à 20°C pendant 1h30, (4.) 20°C après incubation à 20°C pendant 1 jour. Chaque mesure a été réalisée 10 fois par assurer la répétabilité.**

De ces études DLS, nous avons pu mettre en évidence la présence de nano-objets de haut poids moléculaire. Deux types d'organisations peuvent être proposés pour expliquer ces observations : un assemblage ordonné de plusieurs triples hélices ou un assemblage davantage désordonné, constitué de triples hélices imparfaites pontées par les monomères (POG)<sub>n</sub> eux-mêmes. Le fait d'observer (i) la formation rapide des nano-objets après dissociation des monomères (chauffage à 90°C), (ii) une augmentation lente des populations de faible poids moléculaire à 20°C dans HCl (monomère et triple hélice), et



(iii) sachant que la formation de la triple hélice est cinétiquement limitée dans le cas du (POG)<sub>10</sub> (observation par CD), nous amène à privilégier un modèle d'agrégats constitués de l'assemblage de monomères reliés par de courts fragments triples hélices (modèle 1 Figure 2).

### 2.1.2 Observations des assemblages supramoléculaires par microscopie électronique

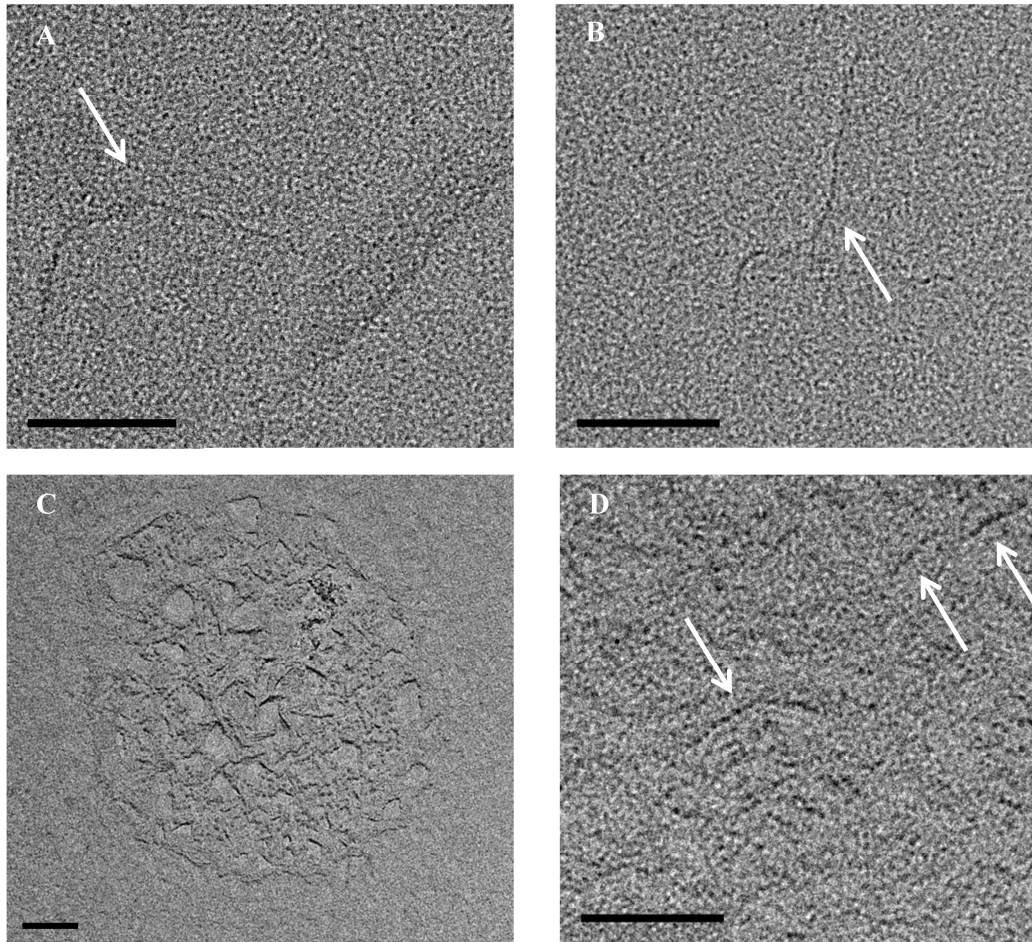
*Principe* – La microscopie électronique utilise des électrons dont la longueur d'onde est beaucoup plus petite que celle des photons visibles et donc permet de s'affranchir de la limitation de résolution optique. Les meilleurs microscopes électroniques atteignent aujourd'hui des résolutions inférieures à l'Angström dans le domaine de la nanocaractérisation des matériaux. Il existe plusieurs types de microscopie électronique, nous nous focaliserons ici sur la microscopie électronique à transmission (TEM). Le microscope électronique à transmission est composé d'une source d'électrons ; de lentilles, appelées condenseurs, qui dirigent ces électrons vers l'échantillon ; d'un porte échantillon ; d'une première lentille agrandisseur, appelé objectif ; de plusieurs lentilles intermédiaires et d'une lentille de projection. Le rôle de la source est de fournir un faisceau d'électrons accéléré par une différence de potentiel adéquate comprise entre quelques dizaines et quelques centaines de kilovolts (20kV ou 400kV) dans les cas les plus courants.

Les échantillons biologiques complexes doivent passer par de nombreuses étapes de préparation pour pouvoir être observés sans modifier leur état morphologique. Pour cela, une des méthodes de préparation consiste en une fixation chimique permettant de conserver la structure et l'ultrastructure du matériel biologique. Puis, ils sont déshydratés par imprégnation dans des bains successifs d'alcools. Ils sont ensuite imprégnés dans une résine, qui sous l'effet de la chaleur va polymériser. On obtient alors un matériau suffisamment résistant qui peut être découpé en lames minces sous ultramicrotome (outil permettant de réaliser des coupes de l'ordre de la centaine de nanomètre). Enfin pour améliorer le contraste, l'échantillon est exposé à un produit contenant des atomes lourds (généralement de l'acétate d'uranyle) (cf. partie expérimentale §B p.135). Dans le cas de l'étude des collagènes, la microscopie électronique sert par exemple à étudier la structure

des fibrilles, ainsi qu'à analyser leur formation.<sup>7-9</sup> Nous nous sommes intéressés à l'étude des assemblages supramoléculaires formés par les CMP.

Après les études menées en parallèle par RMN et présentées dans la partie suivante (cf.§2.1), nous avons pu déterminer la présence de ces assemblages à des concentrations bien plus faibles ( $C=200\mu\text{M}$ ) que celles utilisées pour la DLS. Des études en coloration par ombrage sont en cours d'investigation ainsi que des études en cryo-microscopie.

*L'ombrage*<sup>10</sup> – L'ombrage est une technique qui permet de mettre en évidence les surfaces suivant un angle défini. On fabrique une réplique de l'échantillon par vaporisation de métaux lourds puis de carbone sur la surface pour stabiliser la réplique. On observe alors la réplique de l'échantillon qui permet de visualiser les rugosités de surface. Des études sur les peptides (POG)<sub>7</sub> et (POG)<sub>10</sub> ont été réalisées à une concentration de 200 $\mu\text{M}$ . La préparation des échantillons est détaillée en partie expérimentale (cf. § B3 p.137). Nous avons pu observer des morphologies caractéristiques de chaque peptide (Figure 5). Le peptide (POG)<sub>7</sub> possède des fibres flexibles d'environ 300 nm alors que les fibres du peptide (POG)<sub>10</sub> s'apparentent à des bâtonnets d'une longueur de 100 nm. Afin d'améliorer la caractérisation de ces particules, des études en cryo-microscopie sont en cours.



**Figure 5: Fibres observées en ombrage sur le peptide (POG)<sub>7</sub> (A-B) et (POG)<sub>10</sub> (C-D) à 200μM, HCl 3mM. La barre d'échelle représente 200nm.**

*La cryo-microscopie* – La cryo-microscopie est une méthode de préparation d'échantillon qui a l'avantage de le garder hydraté (méthode détaillée en partie expérimentale §B2 p.136). La congélation de l'échantillon dans l'éthane liquide préalablement refroidi à l'azote liquide permet d'obtenir de la glace vitreuse qui, contrairement à la glace cristalline, ne possède pas de structure ordonnée qui nuirait à l'observation. Nos premiers essais ne nous ont pas permis d'observer nos espèces car nous avons été gênés par la présence de contaminants et des problèmes de congélation.

La caractérisation structurale d'espèces de haut poids moléculaires solubles n'a pas pu être obtenue de manière satisfaisante. Cependant, nous poursuivons nos efforts pour trouver des conditions expérimentales et une approche technique compatible avec leur analyse. Nous avons pu observer, pour le peptide (POG)<sub>10</sub> uniquement et dans certaines conditions, la formation d'agrégats insolubles et donc facilement isolables. Ceux-ci ont pu être étudiés par des méthodes de microscopies et par RMN (cf. § 2.3).

## 2.2. Analyse par RMN liquide

La RMN à l'état liquide permet d'observer des molécules atteignant quelques centaines de kDa, et même jusqu'au MDa en utilisant des techniques de marquages isotopiques particulières.<sup>11</sup> Une masse de 1 MDa, correspond à un objet moléculaire constitué d'environ 10000 acides aminés ce qui est considérable mais encore bien en deçà des assemblages moléculaires solubles de (POG)<sub>7</sub> ou (POG)<sub>10</sub> que nous avons observés par DLS. Cependant, il existe des expériences RMN capables de mettre en évidence de manière indirecte la présence d'assemblages de taille très importante, notamment l'expérience STD (Saturation-Transfer-Difference).<sup>12,13</sup> Lorsque les systèmes ne sont pas à l'équilibre thermodynamique, il est également possible de suivre leur évolution au cours du temps, la formation d'espèces de très haut poids moléculaire conduisant à la disparition du signal RMN.

### 2.2.1 Mise en évidence d'assemblages supramoléculaires par des expériences STD

L'expérience STD est largement employée pour caractériser les complexes ligand-récepteur, y compris quand ceux-ci possèdent une masse trop élevée pour être observés en RMN liquide. Cette approche RMN peut être utilisée pour cribler des banques de molécules et mettre en évidence de nouveaux ligands. Elle permet également de caractériser leur mode de liaison.

*Principe de l'expérience STD<sup>12,13</sup>* – Les molécules de haut poids moléculaire ne sont pas visibles directement par RMN en raison de leur relaxation rapide. Cependant, l'expérience STD permet de mettre en évidence l'interaction entre une molécule de haut poids moléculaire et une molécule plus petite qui, elle, est visible. La Figure 6 montre le principe de l'expérience STD. Il est tout d'abord nécessaire d'enregistrer un spectre <sup>1</sup>H qui servira de référence. En présence d'une molécule de haut poids moléculaire (protéine) représentée schématiquement en bleu, et de ligands potentiels représentés en rouge et vert, ce spectre de référence consiste en des raies fines pour les ligands, la protéine n'étant quant à elle pas visible (raies trop larges). Une irradiation sélective de la protéine, généralement à des valeurs de déplacement chimique inférieures à zéro, va permettre un transfert de saturation (par diffusion de spin) vers le ligand, uniquement si celui-ci est en interaction avec la protéine. Ce transfert de saturation se traduit par une diminution du signal des protons du

ligand en contact avec la surface protéique. Ainsi, en soustrayant le spectre irradié au spectre de référence, nous aurons le spectre STD où les résonances du ligand apparaissent.

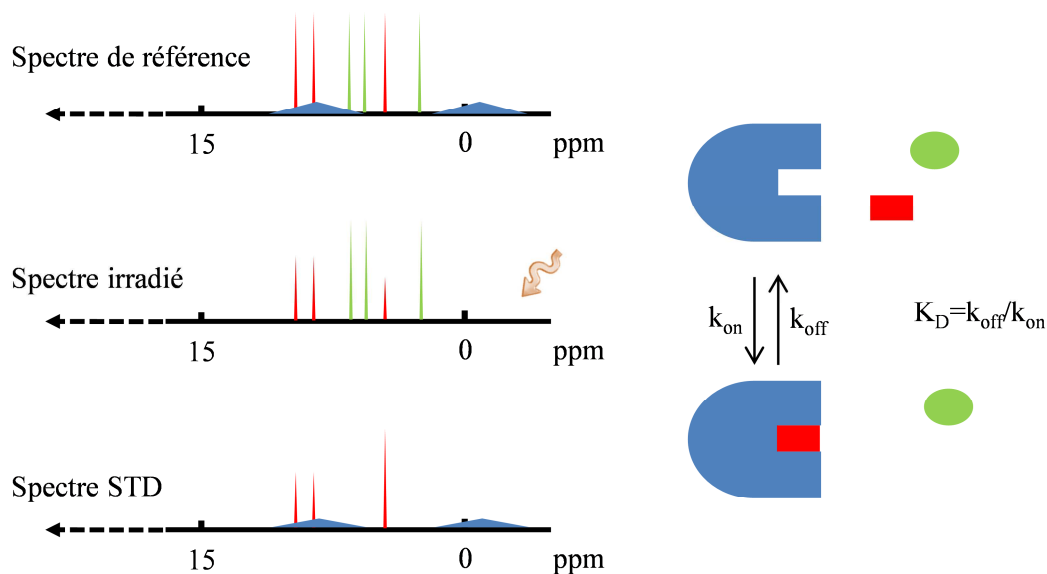
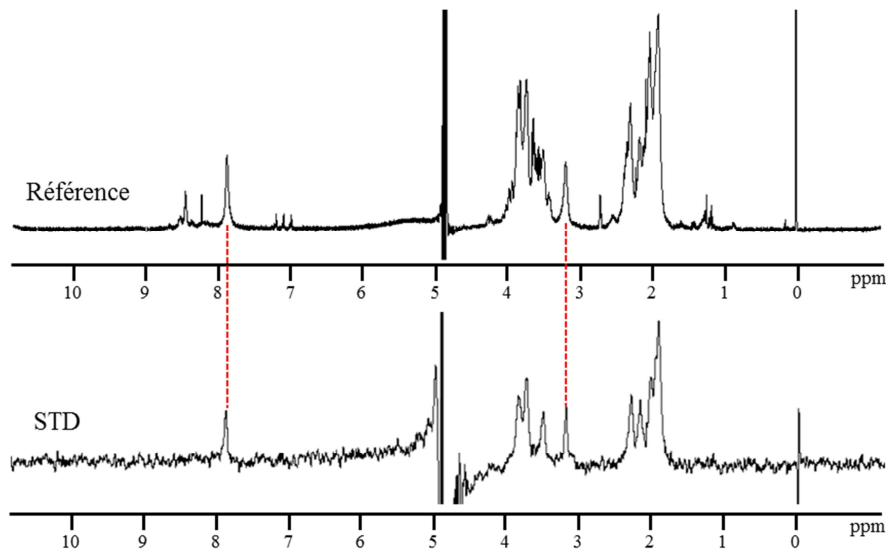


Figure 6: Schéma de principe de l'expérience STD.

Cette technique offre l'avantage de pouvoir tester des ligands peptidiques non marqués et de ne consommer que très peu de protéine. Cependant, il existe des limites à cette technique : en cas d'échange lent, lorsque  $k_{on}$  et  $k_{off}$  sont petits par rapport au temps d'irradiation (souvent le cas lorsque la valeur du  $K_D$  est inférieure à *ca.* 10 nM), le transfert de saturation ne touche que la faible proportion de ligands liés et ne se propage pas à la fraction de ligands libres. L'autre point critique de cette expérience est la nécessité de pouvoir irradier sélectivement la protéine sans toucher les résonances du ligand potentiel. Les biomolécules de haut poids moléculaire présentent cependant des résonances très blindées, liées à la présence de cœurs hydrophobes. Les irradiations sélectives peuvent être effectuées dans cette région spectrale sans affecter directement les protons des petites molécules. Dans tous les cas, des spectres de contrôle peuvent être enregistrés pour s'assurer que la perte de signal observée sur une petite molécule est bien liée à un transfert de saturation de la protéine.

*Résultats obtenus*– Dans le cadre des expériences que nous avons menées pour mettre en évidence des interactions entre CMP et collagène (cf § 2.4), nous avons enregistré des spectres STD de contrôle, où les peptides CMP sont présents mais en absence de collagène. Nous n'avons pas encore caractérisé la présence d'assemblages de CMP de haut poids moléculaire par les autres techniques et nous avons été surpris d'observer du signal STD pour ces spectres de contrôle (Figure 7). Le peptide  $(POG)_{10}$ , seul en solution à 200  $\mu$ M,

présente un signal STD à 20°C ( $T < T_m$ ), que l'impulsion sélective soit appliquée à 0.8 ppm, -0.85 ppm ou même à -2 ppm. Cette approche STD s'est donc révélée difficile à exploiter pour l'observation d'interactions entre le collagène et les CMP mais nous a permis de caractériser, une fois de plus, la présence de nano-objets. Ceux-ci sont responsables du transfert de polarisation vers les espèces CMP de faible poids moléculaire. Plus précisément, nous avons remarqué que le signal STD correspondait uniquement au signal de la triple hélice (protons aliphatiques caractéristiques et protons amides visibles en Figure 10). Cette observation intéressante est également faite pour le peptide (POG)<sub>7</sub> à plusieurs températures (5°C, 10°C, 20°C). En revanche, lorsque la température a été portée au-delà de la  $T_m$  de (POG)<sub>7</sub> ( $T_m = 30^\circ\text{C}$  à 200  $\mu\text{M}$ ) aucun signal n'a été obtenu (Figure 8), ce résultat faisant office de témoin négatif pour nos expériences de STD.



**Figure 7: Mise en évidence de signal STD sur le peptide (POG)<sub>10</sub>, 200 $\mu\text{M}$ , HCl 3mM, 20°C, 500MHz. Impulsion sélective à -2 ppm**

Ces expériences STD, répétées de nombreuses fois, nous ont donc permis (i) d'établir la présence d'espèces de haut poids moléculaire, non visibles directement en RMN liquide, et (ii) de caractériser un transfert de saturation entre ces espèces et les triple hélices, exclusivement. Nous pouvons alors supposer que ce transfert est dû soit à une interaction faible entre les formes triples hélices et les nano-objets (interactions de surface) ou soit à un échange rapide entre les triples hélices en solution et l'assemblage de haut poids moléculaire. Des études complémentaires présentées dans la partie suivante, nous ont permis de distinguer ces deux hypothèses.

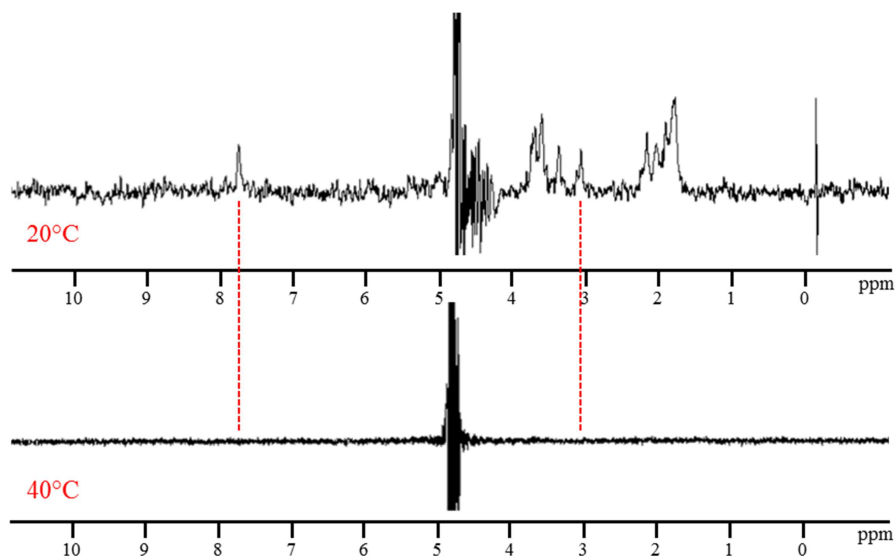


Figure 8: Expériences STD enregistrées avec le peptide (POG)<sub>7</sub>, 200μM, HCl 3mM, 500MHz. Impulsion sélective à -2ppm. Le signal STD est uniquement observé pour  $T < T_m$

A partir de ces observations, nous avons tenté d'obtenir des informations relatives à la cinétique d'agrégation de nos peptides par STD. En chauffant les échantillons à 90°C pendant 20 min, il est possible de dissocier les nano-objets, y compris les triples hélices. Un refroidissement rapide à une température inférieure à  $T_m$  pourrait permettre de suivre en temps réel par STD la formation des assemblages supramoléculaire. Malheureusement, la cinétique de formation de ces assemblages a été trouvée très rapide par rapport au temps nécessaire pour l'acquisition d'expériences STD exploitables (20 min). Une alternative a été de suivre les signaux du monomère et de la triple hélice en RMN <sup>1</sup>H à intervalle régulier (~1min 30s). Nous avons ainsi pu constater une disparition significative du signal au cours du temps que nous avons attribuée à la formation des assemblages de haut poids moléculaires, invisibles en RMN en phase liquide. Nous avons bien sûr vérifié qu'aucun précipité n'apparaissait.

### 2.2.2 Etudes cinétiques

L'ensemble des techniques employées jusqu'ici, que ce soit la RMN, la DLS ou les microscopies, nous a permis de caractériser les peptides (POG)<sub>n</sub> sous trois formes : monomériques (plusieurs monomères ont été identifiés), triples hélices ou sous des formes d'assemblages supramoléculaires. Nous avons par ailleurs observé, dans le cas de (POG)<sub>10</sub> uniquement, la formation d'espèces insolubles, hautement organisées. En nous restreignant

à l'étude des (POG)<sub>7</sub>, stables en solution, nous avons donc proposé le modèle cinétique suivant et tenté de quantifier les différents échanges en présence :

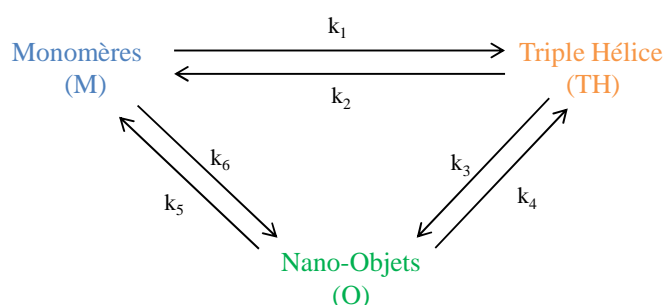
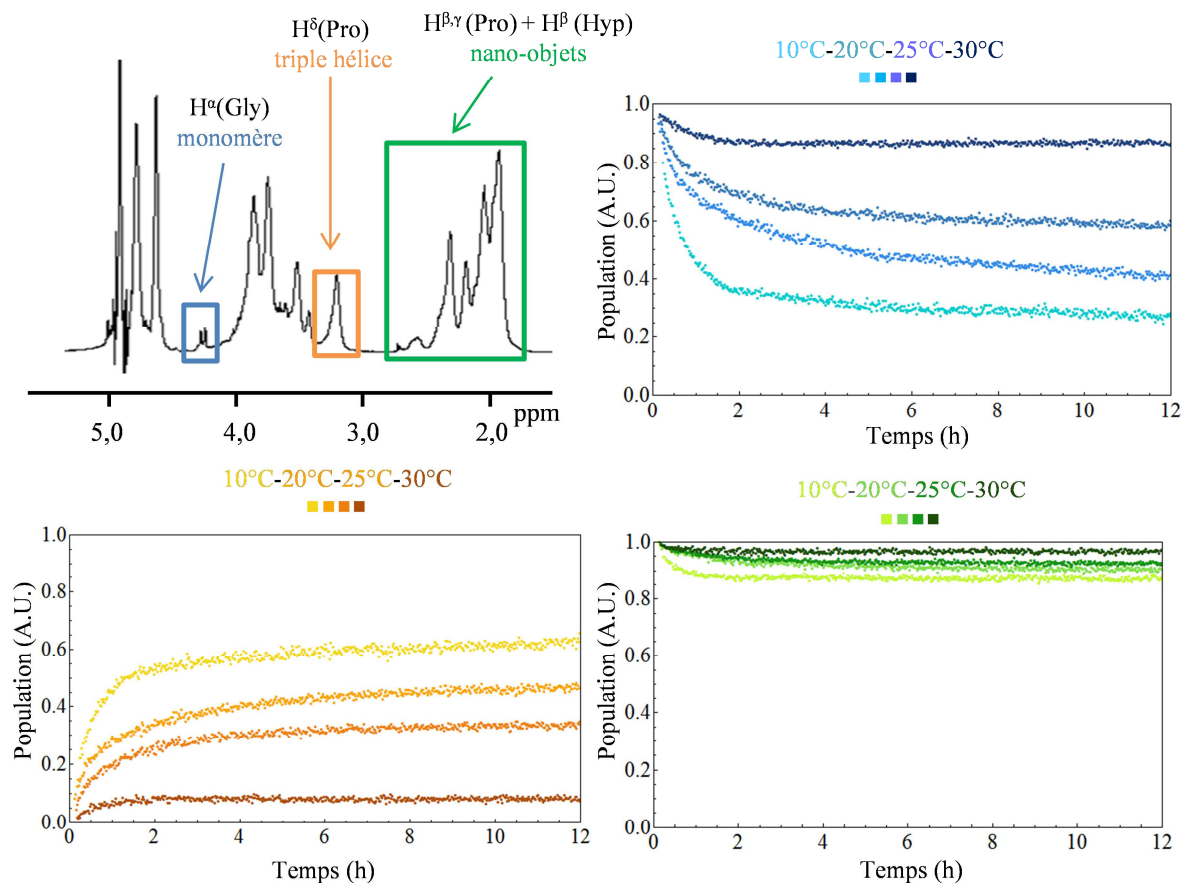


Figure 9: Modèle cinétique proposé pour les échanges  $M \rightleftharpoons TH \rightleftharpoons O$  de (POG)<sub>7</sub>

Pour cela, nous avons repéré deux régions du spectre <sup>1</sup>H correspondant spécifiquement au monomère majoritaire d'une part et à la triple hélice d'autre part (Figure 10, cf chapitre 1 pour l'attribution). Nous avons par ailleurs identifié une troisième région où les protons H<sup>β</sup>, H<sup>γ</sup> de la proline et H<sup>β</sup> de l'hydroxyproline résonnent, qu'ils soient au sein du monomère ou de la triple hélice. Ainsi, l'intégration de cette région devrait être insensible aux échanges M ↔ TH. Immédiatement après chauffage à 90°C (pendant 20 min), nous avons supposé que les nano-objets de (POG)<sub>7</sub> étaient entièrement dissociés ; sous forme monomérique à 100%. Nous nous sommes ensuite placés à différentes températures (10°C, 20°C, 25°C, 30°C) et avons suivi les intégrations de ces trois régions en fonction du temps (Figure 10). Nous avons ainsi pu mettre en évidence des cinétiques de formation pour les triples hélices, avec, selon la température, des temps caractéristiques de quelques dizaines de minutes à plus d'une heure. De manière tout à fait prévisible, la disparition simultanée de l'espèce monomérique a été observée. Le fait que le signal diminue également dans la troisième région, qui comprend pourtant à la fois les protons des espèces monomériques et de la triple hélice, nous indique la formation d'une population de haut poids moléculaire, invisible en RMN à l'état liquide. Le détail des conditions expérimentales est commenté en partie expérimentale (§.D2 p.139).





**Figure 10:** Suivi des populations à M/TH/O à différentes températures pendant 12h sur  $(\text{POG})_7$ ,  $200\mu\text{M}$ , HCl  $3\text{mM}$ . La population en nano-objets est déduite de la disparition des signaux aliphatiques entre 1,8 et 3 ppm.

Ces expériences résolues en temps rejoignent les observations faites précédemment par CD ou par RMN à l'équilibre où nous avons quantifié la diminution de la TH à mesure que la température augmentait. Les plateaux obtenus en Figure 10, qui caractérisent l'équilibre thermodynamique, correspondent bien à des proportions de TH allant de 65 % à basse température à 8% quand on se rapproche de la température de fusion de  $(\text{POG})_7$ . Nous observons également que la vitesse de formation de la triple hélice (ou de disparition du monomère) est plus lente lorsque la température augmente. La quantité de nano-objets présents à  $10^\circ\text{C}$  ( $T < T_m$ ) est plus importante qu'à  $30^\circ\text{C}$  ( $T = T_m$ ), ce qui implique que les nano-objets se forment à partir des triples hélices et/ou qu'ils sont eux-mêmes instables à une température avoisinant la température de fusion.

La détermination des constantes de vitesse a été entreprise en considérant le schéma cinétique présenté sur la Figure 9. Il s'agit d'un modèle rudimentaire ne tenant pas compte

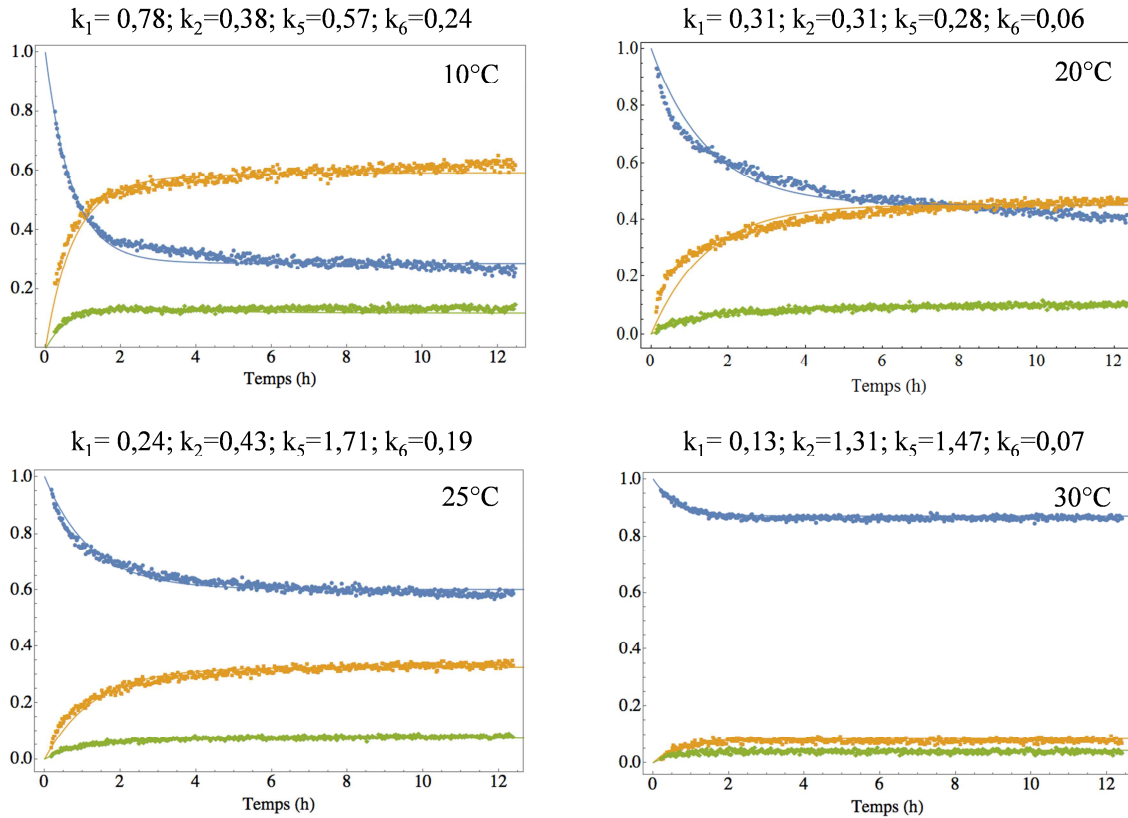
des stœchiométries que nous aurions du mal à établir en ce qui concerne les nano-objets. Nous avons donc tiré les trois équations différentielles suivantes :

$$\frac{d[M]}{dt} = k_2[TH] + k_5[O] - (k_1 + k_6)[M] \quad Eq. (1)$$

$$\frac{d[TH]}{dt} = k_1[M] + k_4[O] - (k_2 + k_3)[TH] \quad Eq. (2)$$

$$\frac{d[O]}{dt} = k_3[TH] + k_6[M] - (k_5 + k_4)[O] \quad Eq. (3)$$

La résolution numérique de ce système d'équations différentielles a été faite à l'aide du logiciel Mathematica. Nous avons constaté des problèmes de convergence lorsque les six constantes de vitesses étaient minimisées simultanément. Nous avons donc tenté de simplifier le système en annulant  $k_3$  et  $k_4$  (Figure 11). Ce modèle néglige donc l'équilibre qui pourrait avoir lieu entre les formes triples hélices et les nano-objets solubles. Le monomère se trouve ainsi en position centrale dans le modèle cinétique, pouvant servir à la fois de réservoir pour la formation de triples hélices et des espèces de haut poids moléculaire. Dans les deux cas, une hausse de la température déstabilise les assemblages et augmente la population des espèces monomériques. On observe une diminution de  $k_1$  en fonction de la température, ce qui est en accord avec la déstabilisation de la TH à l'approche de sa température de fusion ( $T_m=30^\circ\text{C}$ ). On remarque que la constante  $k_6$  suit cette même tendance (à  $20^\circ\text{C}$ , la valeur très basse obtenue peut être reliée à la mauvaise modélisation des données expérimentales), traduisant le comportement similaire des nano-objets.



**Figure 11:** Les données expérimentales ont été modélisées avec les solutions analytiques des 3 équations différentielles. Les unités des constantes de vitesses sont en  $\text{h}^{-1}$ . Ici,  $k_3$  et  $k_4$  sont égales à 0.

Si nous calculons les constantes d'équilibre ( $K_{\text{TH/M}}$  et  $K_{\text{O/M}}$ ) nous pouvons mettre en évidence les populations favorisées à chaque température (Tableau 1). Pour l'équilibre TH/M, on observe un décalage de la température de fusion prédite par cette méthode (correspondant à  $K_{\text{TH/M}} = 1$ ), qui serait attendue à 20°C alors que les mesures CD indiquent une valeur proche de 30°C à cette concentration de peptide. Cette différence est probablement liée à l'approximation faite au niveau de la stœchiométrie de réaction. En tout état de cause, la température de fusion associée aux nano-objets apparaît plus basse, traduisant une cohésion moindre de ces assemblages par rapport à la triple hélice. On peut noter à ce stade qu'ils diffèrent nettement avec les nano-objets insolubles de  $(\text{POG})_{10}$ , dont la formation nécessitait une température élevée.

Température (°C)	10	20	25	30
$K_{TH/M}=k1/k2$	2,05	1,00	0,56	0,10
Population favorisée	TH	TH~M	M	M
$K_{O/M}=k6/k5$	0,42	0,21	0,11	0,05
Population favorisée	M	M	M	M

**Tableau 1: Calculs des constantes d'équilibre**

Une dernière exploitation de ces valeurs de constantes de vitesse peut être faite en considérant,

d'une part la définition des constantes d'équilibre par rapport l'énergie libre de Gibbs :

$$\Delta G = -RT \ln K$$

où  $\Delta G$  représente l'enthalpie libre,  $R$  la constante des gaz parfaits ( $8,314 \text{ J.K}^{-1}.\text{mol}^{-1}$ ),  $T$  représente la température (en K) et  $K$  la constante d'équilibre,

et d'autre part la relation entre l'énergie libre de Gibbs et les grandeurs thermodynamiques  $\Delta H$  et  $\Delta S$  :

$$\Delta G = \Delta H - T\Delta S$$

où  $\Delta H$  représente la variation d'enthalpie et  $\Delta S$  représente la variation d'entropie.

Si on applique l'approximation d'Ellingham, considérant que, sur une gamme restreinte, l'enthalpie et l'entropie standards de réaction sont indépendantes de la température, on peut alors écrire :

$$\Delta G^\circ(T) = \Delta H^\circ - T\Delta S^\circ = -RT \ln K$$

En traçant les droites  $\ln(\frac{1}{T}) = \ln K$ , pour  $K_{TH/M} = k1/k2$  (équilibre monomère-triple hélice) et  $K_{O/M} = k6/k5$  (équilibre monomère-nano-objet), nous obtenons des régressions linéaires satisfaisantes (Figure 14).

Ces droites nous permettent de définir les paramètres suivants :

- $\Delta H_{M-TH}^\circ = -100 \text{ kJ.mol}^{-1}$  et  $\Delta S_{M-TH}^\circ = -347 \text{ J.mol}^{-1}.\text{K}^{-1}$
- $\Delta H_{M-O}^\circ = -75 \text{ kJ.mol}^{-1}$  et  $\Delta S_{M-TH}^\circ = -272 \text{ J.mol}^{-1}.\text{K}^{-1}$

Pour les deux équilibres, nous obtenons des réactions exothermiques ( $\Delta H^\circ < 0$ ) et  $\Delta S^\circ < 0$ , ce qui signifie bien que le désordre diminue lors de la formation de ces assemblages. Dans le cas des nano-objets, il résulte d'un regroupement d'un nombre très important de monomères, qui demeurent probablement partiellement désordonnés. Pour ce qui concerne la triple hélice, le nombre de monomère recruté est faible mais l'organisation est très grande. La présence des nano-objets solubles pourrait être due à des triples hélices

présentant des brins décalés permettant ainsi de les lier les unes aux autres par leurs extrémités.

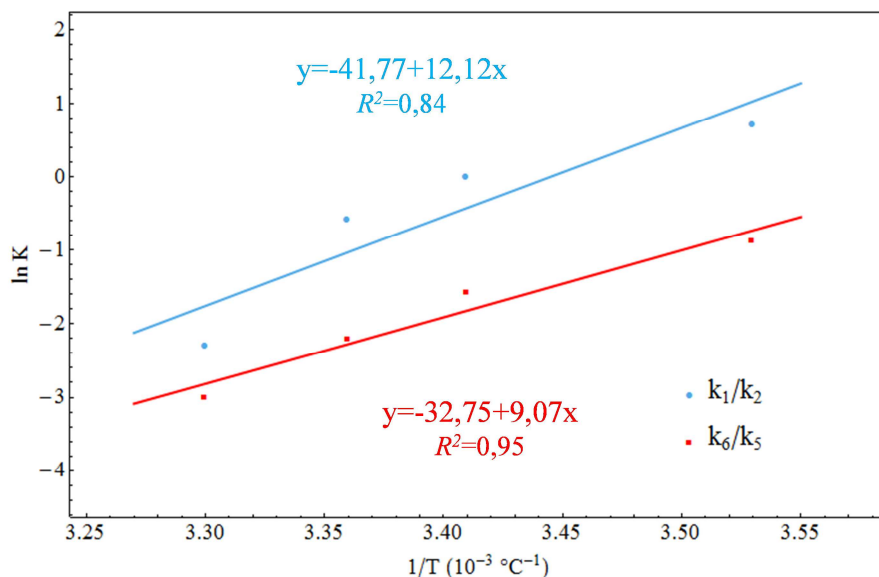
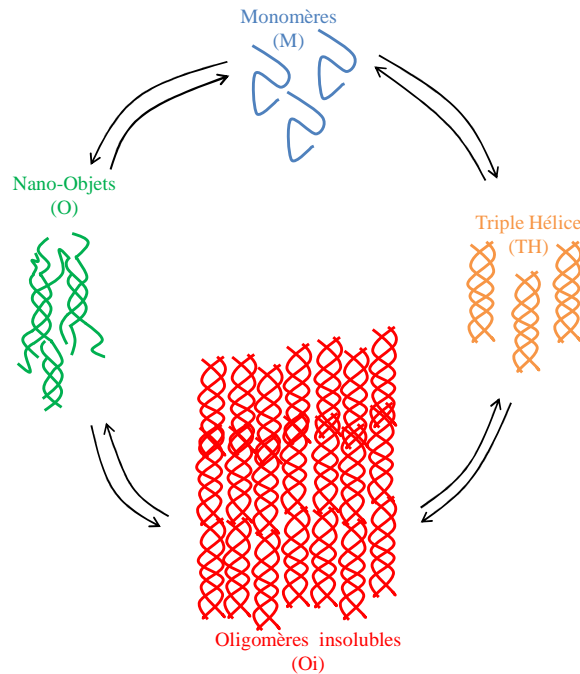


Figure 12: Représentation des courbes  $\ln K = f\left(\frac{1}{T}\right)$

Il convient de rappeler que, pour de courts peptides, la liaison peptidique Gly-Pro n'est qu'à 80% en conformation *trans* à l'équilibre thermodynamique. La probabilité que les peptides  $(\text{POG})_n$  soient à 100% en conformation *trans* est ainsi de l'ordre de  $0,8^{(n-1)}$  ce qui ne représente que ~25% pour  $(\text{POG})_7$ . On peut ainsi imaginer que des triples hélices, qui requièrent des liaisons peptidiques *trans*, ne puissent se former que partiellement. Celles-ci pourraient alors présenter des portions de peptides déviant de son axe, à l'origine d'un assemblage supramoléculaire supérieur. L'isomérisation de la liaison peptidique est connue pour être une étape limitante dans les processus de repliement des protéines, des enzymes particulières, les peptidyl-prolyl isomérases (PPIases), pouvant intervenir pour catalyser ce réarrangement conformationnel.<sup>14</sup> La vitesse d'isomérisation *cis-trans* est donc relativement longue (de quelques secondes à plusieurs minutes), et dans le cas des nano-objets de  $(\text{POG})_n$ , les réarrangements conformationnels permettant d'obtenir des triples hélices stables sont limités cinétiquement. On a remarqué dans le cas de  $(\text{POG})_{10}$  qu'il fallait avoisiner la température de fusion pour obtenir des assemblages supramoléculaires insolubles. Cela est bien en lien avec la nécessité de déstabiliser partiellement les triples hélices, les « effiloche », pour obtenir ces nano-objets. La Figure 13 décrit de manière schématique ces observations. Les assemblages supramoléculaires insolubles résulteraient de triples hélices stabilisées par des interactions non covalentes, en accord avec la très forte organisation observée par SHG et en RMN du solide (cf. 2.3.1). Deux types d'interactions

assureraient leur stabilité : des interactions « têtes-queues » entre les extrémités de chaînes le long de l'axe de l'hélice, et des interactions latérales pouvant être médiées par des molécules d'eau. Dans le 3<sup>e</sup> chapitre, nous avons envisagé d'utiliser des groupes fluor afin de favoriser de telles interactions inter-hélices.



**Figure 13: Mécanisme d'auto-association du peptide (POG)<sub>7</sub> en solution.**

Ce modèle est en accord avec les études de *Raines et al.* qui ont montrées qu'en synthétisant des triples hélices liées entre elles par des ponts disulfure, les extrémités peptidiques libres permettaient une association continue de triples hélices pour former des assemblages semblables aux fibrilles de collagène<sup>15</sup>.

*Perspectives* – Des études complémentaires vont être menées sur le peptide (POG)<sub>10</sub> afin de comparer les cinétiques d'agrégation et de formation de triple hélice. Par ailleurs, nous avons observé des effets du pH sur ces assemblages (expériences de DLS § 2.1.1). Son influence sur ces cinétiques de formation devra être explorée. En effet à pH < 3, le peptide possède seulement une charge positive à son extrémité N-terminale alors que pour un pH supérieur une charge négative apparaît à son extrémité C-terminale ce qui devrait promouvoir des interactions électrostatiques entre triple hélices. De plus, comme nous avons pu le voir par DLS, le contre ion présent en solution semble modifier le type d'assemblage. L'influence des sels que nous avons observée par DLS pourra également être contrôlée par cette méthode RMN résolue en temps.

## 2.3. Mise en évidence et étude d'assemblages supramoléculaires non-solubles

### 2.3.1 Mise en évidence d'assemblages supramoléculaires non-solubles

Nous avons mené des études par turbidimétrie qui ont montré que des agrégats non solubles de  $(\text{POG})_{10}$  se formaient à partir d'une concentration de 2,6mM, en accord avec la littérature (Figure 14).<sup>6</sup> La turbidimétrie est la mesure du degré de turbidité d'une suspension. Elle peut être déterminée grâce à un système optique, en général un spectrophotomètre qui mesure l'augmentation de l'absorbance, elle-même proportionnelle à la concentration des particules en suspension.

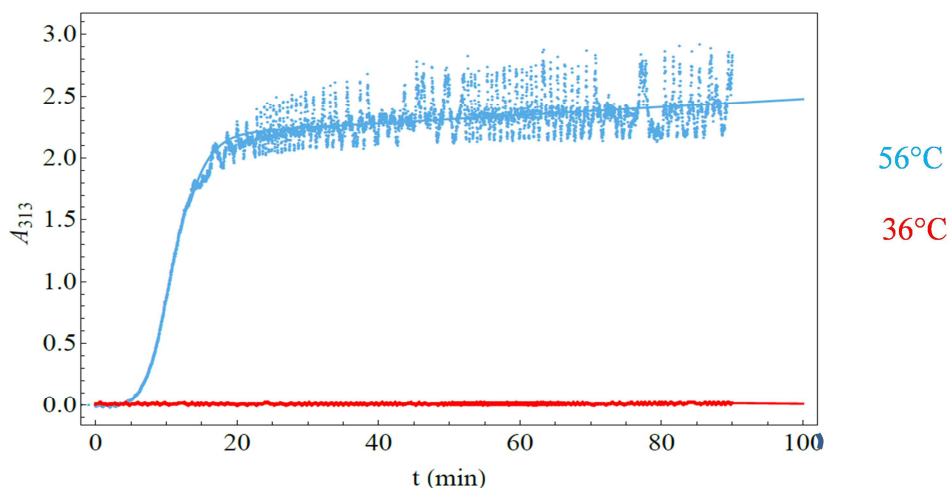


Figure 14: Suivi de l'absorbance à 313 nm en fonction du temps et de la température  $(\text{POG})_{10}$ , 2,6mM, pH=7, tampon  $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$  (50mM)

L'apparition d'assemblages supramoléculaires insolubles est caractérisée par une augmentation de l'absorbance à 313 nm. Un temps de latence de 5 minutes est observé à 56°C, celui-ci devenant plus important si la température diminue. A 36°C, nous n'observons pas d'augmentation de l'absorbance après plus d'1h30 d'incubation. Les agrégats insolubles obtenus se sont révélés stables à température ambiante (+1mois), leur solubilisation pouvant être rapidement obtenue en les plaçant à 4°C, ou au contraire, en les chauffant à 90°C, bien au-delà de la température de fusion de la triple hélice (~65°C pour  $(\text{POG})_{10}$  dans ces conditions).

Nous avons alors entrepris une caractérisation de ces assemblages, notamment pour savoir s'il s'agissait d'agrégats amorphes ou d'assemblages supramoléculaires organisés.

### 2.3.1 Caractérisation des assemblages supramoléculaires non solubles

Des techniques de microscopie (SHG : Second Harmonic Generation et à lumière polarisée) et la RMN à l'état solide nous ont permis de répondre à cette question.

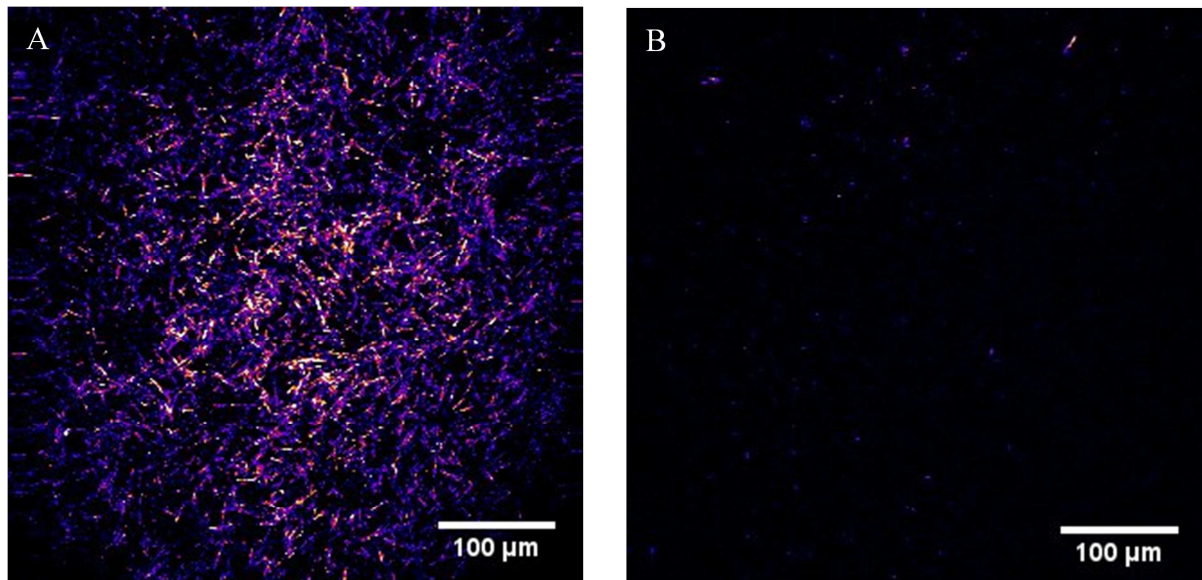
*La microscopie SHG* – C'est une technique biphotonique qui possède de nombreux avantages, notamment celui d'une résolution tri-dimensionnelle intrinsèque. Cette résolution permet une imagerie 3D en profondeur dans les tissus biologiques, en raison entre autre des longueurs d'ondes utilisées, qui se situent dans l'infra-rouge et limitent la diffusion et l'absorption par les tissus. De plus, cette technique se combine facilement à la fluorescence biphotonique qui permet d'exciter des marqueurs fluorescents mais aussi des molécules endogènes (NADH, FAD, riboflavines...) et peut ainsi rendre visible, sans marquage, le tissu ou les cellules. La microscopie de génération de seconde harmonique (SHG) consiste en la génération d'un signal à la fréquence double de la fréquence d'excitation. Ce phénomène est dit « cohérent » car la molécule n'absorbe pas l'énergie incidente, mais présente un effet de diffusion instantanée non linéaire. Ainsi, les champs électriques diffusés par chacun des harmonophores (élément polarisable par l'excitation lumineuse) vont interférer entre eux. La parité du processus est telle que la génération de SHG est interdite pour une répartition centro-symétrique des harmonophores. L'émission de deux harmonophores sera maximale lorsque ceux-ci seront orientés parallèlement (interférence constructive), alors qu'aucun signal ne sera généré si ces deux harmonophores sont orientés antiparallèlement (interférence destructive). En ce qui concerne les protéines, les signaux SHG sont donc spécifiques principalement du collagène, des microtubules et de la myosine, protéines organisées de manière dense et non centro-symétrique dans les tissus.

C'est en 1971 que le signal SHG du collagène est observé pour la première fois au sein d'un tissu.<sup>16,17</sup> Seuls les collagènes fibrillaires génèrent un signal SHG,<sup>18,19</sup> principalement le collagène de type I, le plus fréquent. Il a été proposé que le signal SHG provienne des liaisons peptidiques elles-mêmes.<sup>20</sup> En effet, celle-ci est non centro-symétrique et présente une délocalisation d'un électron  $\pi$  entre les groupes C=O et N-H, se



comportant comme de faibles donneurs et accepteurs d'électrons respectivement. Pour avoir une forte réponse SHG, les harmonophores doivent être alignés dans la même direction. En ce qui concerne les monomères de collagène ( $\alpha 1$  et  $\alpha 2$ ), toutes les liaisons peptidiques sont bien alignées, conduisant à des interactions constructives. Il en est de même au niveau macromoléculaire lorsque les molécules de collagènes s'organisent en fibrilles. A un niveau supérieur, les fibrilles peuvent s'organiser en fibres, souvent avec des directions antiparallèles, ce qui peut diminuer le signal SHG.

Par analogie avec les études faites sur les phases denses de collagène, nous avons tenté d'obtenir un signal SHG sur nos assemblages supramoléculaires de  $(\text{POG})_{10}$ , à température ambiante puis après chauffage à  $90^\circ\text{C}$  (Figure 15).



**Figure 15:** Mise en évidence de l'organisation au sein des agrégats insolubles du peptide  $(\text{POG})_{10}$  par microscopie SHG (à  $20^\circ\text{C}$ ). (A) chauffage préalable à  $56^\circ\text{C}$ , (B) chauffage préalable à  $90^\circ\text{C}$

Nous obtenons alors un signal SHG intense pour l'échantillon incubé à  $56^\circ\text{C}$ . La présence d'agrégats insolubles est visible à l'œil nu sur cet échantillon mais le fait d'observer un signal SHG traduit également un alignement dans les assemblages. Les particules mises en évidence sont en forme de bâtonnets aux contours mal définis, de quelques dizaines de  $\mu\text{m}$  de longueur. Après chauffage à  $90^\circ\text{C}$ , on observe visuellement la dissolution des agrégats, ce qui est en accord avec les images de microscopies SHG qui montrent une disparition quasi complète du signal, un très léger signal diffus est observé pouvant correspondre aux

monomères ou triple hélices en solution. Ces observations de microscopies sont en accord avec les données de RMN solide présentées dans la section suivante.

*La RMN à l'état solide* – Dans les liquides, les interactions anisotropes comme l'anisotropie de déplacement chimique et l'interaction dipolaire sont moyennés par le mouvement brownien ce qui conduit à l'observation de raies fines. Dans les solides, la présence de ces interactions anisotropes se traduit par un élargissement très important des raies qui masque les informations recherchées habituellement comme le déplacement chimique et le couplage scalaire. Dans les années 60, *Andrew et al.*<sup>21</sup> ont montré qu'il était possible d'éliminer en grande partie les interactions anisotropes et obtenir des spectres correspondant à la moyenne isotrope des interactions en faisant tourner l'échantillon à l'angle magique  $\theta=54.74^\circ$ . La technique de rotation à l'angle magique est en général désignée par son acronyme anglais MAS (Magic Angle Spinning).

Nous avons donc enregistré des expériences  $^{13}\text{C}$ -CPMAS (polarisation croisée en présence de rotation à l'angle magique) dans des conditions identiques, sur de la poudre lyophilisée de  $(\text{POG})_{10}$  ainsi que sur les assemblages supramoléculaires observés par SHG (Figure 16). L'expérience CPMAS va permettre d'augmenter la sensibilité du noyau  $^{13}\text{C}$  par transfert de polarisation à partir du proton. La poudre lyophilisée présente des raies larges caractéristiques des composés amorphes (largeurs à mi-hauteur de  $\sim 350$  Hz). La résolution obtenue est cependant suffisante pour effectuer l'attribution des résonances des différents carbones de  $(\text{POG})_{10}$ . Les assemblages supramoléculaires possèdent des raies bien plus fines, (largeurs à mi-hauteur de  $\sim 100$  Hz) une anisotropie de déplacement chimique moindre caractéristique d'une organisation tridimensionnelle.

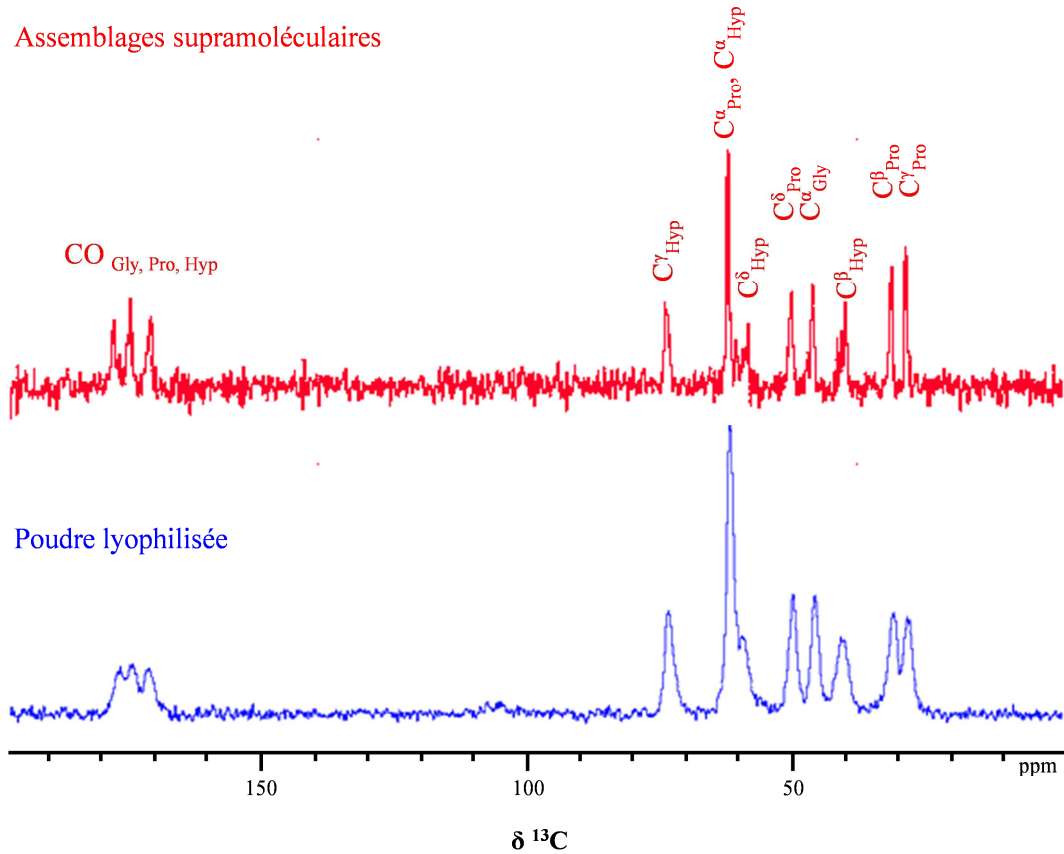


Figure 16: Spectres  $^{13}\text{C}$ -CPMAS de  $(\text{POG})_{10}$  en abondance naturelle enregistré à 9.4T, 7kHz : fréquence de rotation en utilisant un rotor 4mm HRMAS (12 $\mu$ l). Conditions pour les assemblages supramoléculaires : 7mg/ml, pH=7, tampon phosphate 50mM.

L'enregistrement de ces expériences simples nous permet donc de confirmer, par une technique totalement différente de la microscopie SHG, que les molécules au sein des assemblages supramoléculaires insolubles sont organisées. Nous avons ensuite envisagé d'enregistrer des spectres réintroduisant l'anisotropie de certaines interactions afin de mettre en évidence la proximité structurale (interaction dipolaire) ou des informations sur la géométrie (anisotropie de déplacement chimique ou interaction dipolaire). *Mosser et al.* ont d'ailleurs pu mettre en évidence des différences de déplacement chimique sur des échantillons de collagène soluble et fibrillaire, grâce à l'enregistrement de spectre  $\{H^1\}$ - $^{13}\text{C}$  INEPT MAS, qui serait relié à des différences de préférences de plissements des cycles pyrrolidines induit par la fibrillogénèse.<sup>22</sup> Des études à 2D sont également envisageables pour étudier nos échantillons. Nous avons cependant manqué de matériel pour approfondir cette étude structurale par RMN solide (limite de sensibilité). Nous avons alors eu recours à un nouveau lot de  $(\text{POG})_{10}$  pour générer davantage d'échantillons et remplir convenablement les rotors. Malheureusement ce nouveau lot s'est comporté de manière

très différente par rapport aux précédents et ne permettait plus d'obtenir des agrégats stables à température ambiante. Nous avons donc dû différer ces études.

*Microscopie optique à lumière polarisée* – Pour tenter de décrire la morphologie des assemblages que nous formions, leur diversité et leurs caractéristiques en fonction des conditions expérimentales, nous avons utilisé la microscopie optique à lumière polarisée. Celle-ci permet de mettre en évidence la biréfringence des milieux. La biréfringence<sup>23</sup> (ou double réfraction) d'un échantillon résulte de son anisotropie et est déterminée en le plaçant entre polariseurs croisés à 90°. Le premier polariseur sélectionne la lumière polarisée linéairement dans une direction donnée. En l'absence d'échantillon (ou en présence d'un échantillon isotrope), l'analyseur orienté à 90° empêche la transmission de cette lumière, provoquant une extinction. L'intensité de la biréfringence dépend principalement de l'alignement et de l'ordre dans l'agencement moléculaire, mais aussi de la nature des groupements chimiques rencontrés ainsi que de l'épaisseur de l'échantillon observé.

La microscopie optique nous a permis d'identifier la morphologie des assemblages mais également de voir leurs modifications en fonction du pH et de la température. Les premières observations réalisées sur l'échantillon (POG)<sub>10</sub> incubé à 56°C se sont avérées spectaculaires (Figure 17).

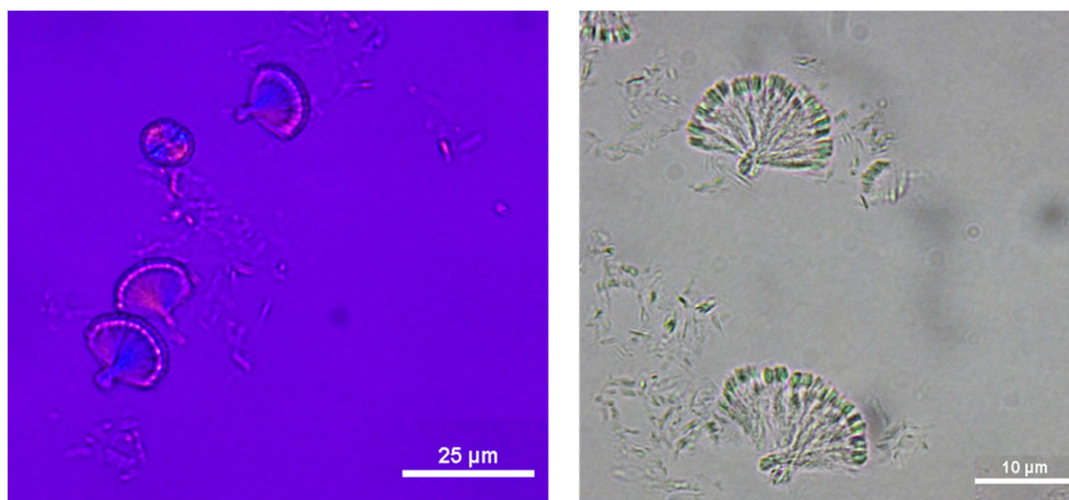
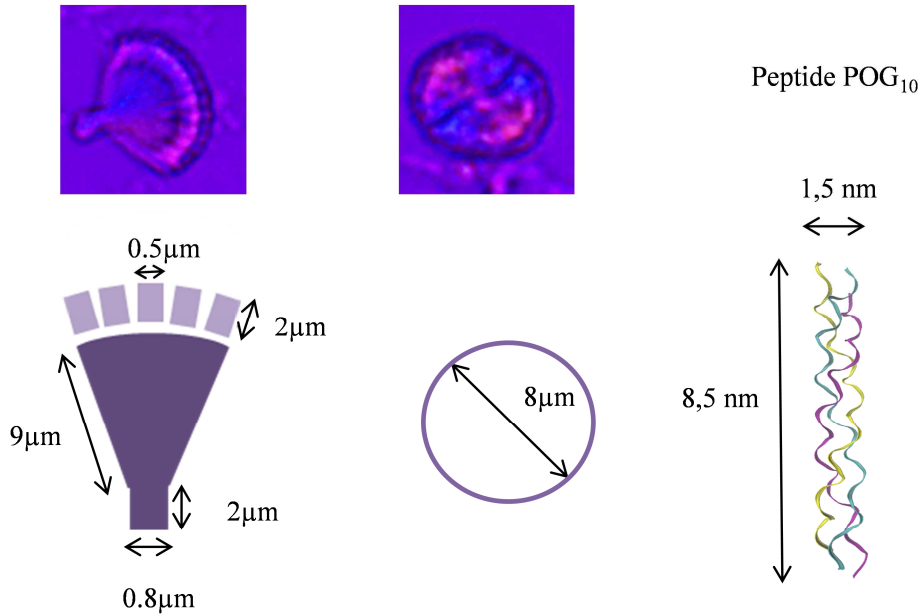


Figure 17: Image de microscopie à lumière polarisée sur l'échantillon (POG)<sub>10</sub> préalablement chauffé à 56°C, pH=7

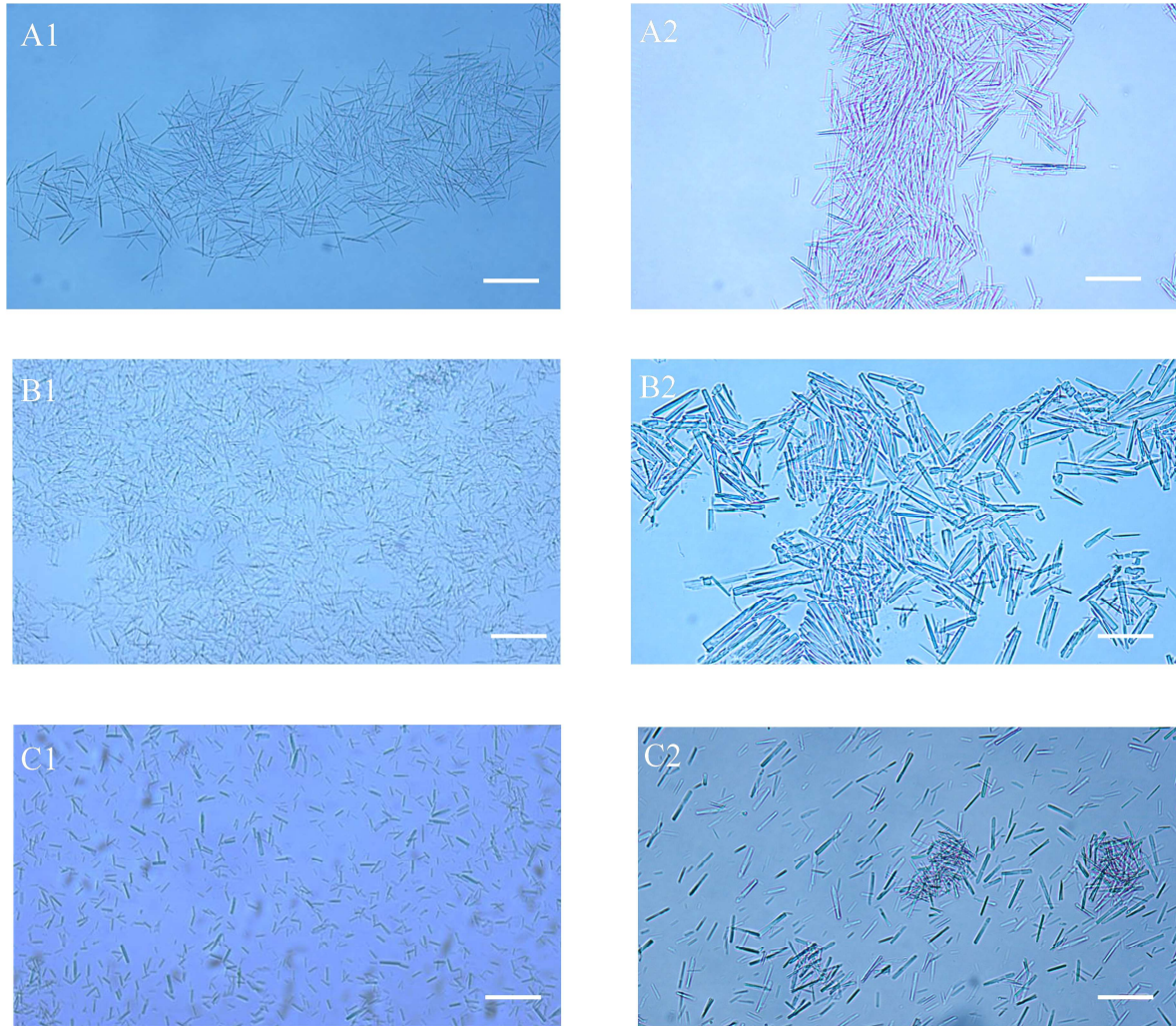
Nous avons observé des morphologies en « bouquets » ou « éventails » qui semblent posséder un arrangement différent sur la partie haute du bouquet. Nous avons pu mesurer approximativement la taille de chaque partie de ces assemblages particuliers, ce qui nous

donne une idée de la quantité de triples hélices pouvant être présentes (Figure 18). Pour l'extrémité inférieure des bouquets, si nous supposons que les triples hélices sont alignées verticalement, il y aurait plus de 60 millions de triples hélices. Cette ordre de grandeur nous rappelle la taille des agrégats observés par DLS (rayons hydrodynamiques de  $\sim 1\mu\text{m}$ ).



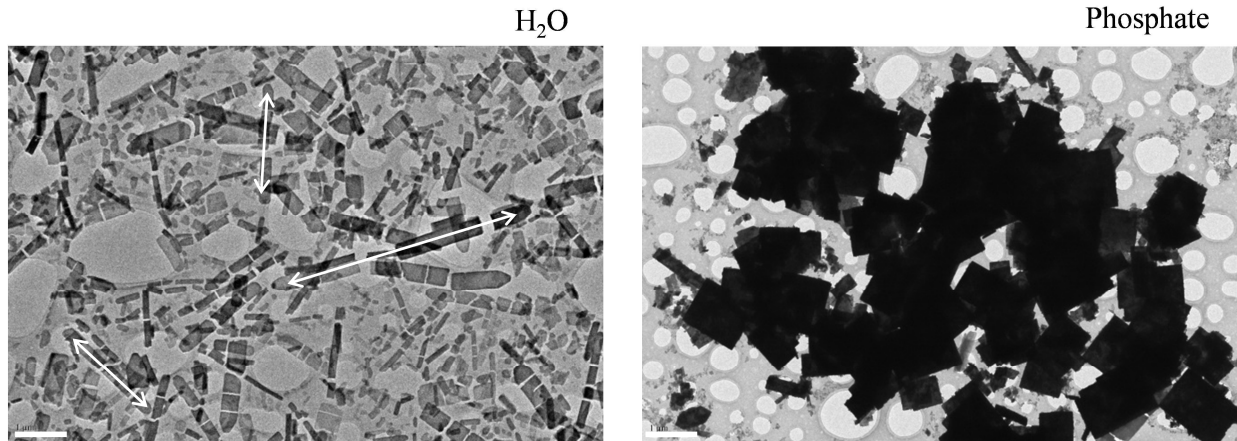
**Figure 18: Schéma représentatif des bouquets et des tailles moyennes des différents éléments du bouquet**

Nous avons tenté de voir quels étaient les rôles des tampons et de la température dans la formation de ces assemblages et s'ils pouvaient influencer leur morphologie. Nous avons donc préparé des échantillons de  $(\text{POG})_{10}$  à partir d'une solution à 2,6mM pré-incubée 24h à 4°C dans les conditions suivantes : (A) incubation à 35°C, pH=7 (tampon phosphate 10mM), (B) à 60°C, pH=7 (tampon phosphate 10mM), (C) à 60°C, pH=2.5 (acide acétique 5mM). Nous avons obtenu dans chacun des cas des aiguilles de 5µM à 10µM de long. Curieusement, aucune de ces conditions ne nous a permis de retrouver la morphologie en bouquet observée précédemment, y compris l'échantillon B qui correspondait aux mêmes conditions. Nous avons ensuite effectué des cycles de chauffage-refroidissement pour comparer les assemblages obtenus à partir des hautes températures (sans présence de triples hélices) et ceux formés après incubation à 4°C, où les triples hélices sont prépondérantes. Nous avons alors mis en évidence des augmentations significatives des diamètres des aiguilles, sans pour autant retrouver la structure en bouquet observée initialement (Figure 20).



**Figure 19:** Clichés de microscopie optique à lumière polarisée sur un échantillon de  $(POG)_{10}$ , 7mg/ml, (A1) Assemblages formées à 35°C, pH=7, (B1) Assemblages formées à 60°C, pH=7, (C1) Assemblages formées à 60°C, pH=4.1. Chaque échantillon a ensuite été chauffé à 90°C puis refroidit à 40°C (A2, B2, C2). Les barres d'échelle correspondent toutes à 20µm

Nous avons donc montré la présence d'assemblages supramoléculaires organisés non solubles présents dans une solution de  $(POG)_{10}$  concentrée ( $c=2,6$  mM). Ces assemblages, une fois formés, sont stables à température ambiante mais se dissocient rapidement à basse (4°C) ou haute température (90°C). Nous avons également pu remarquer que le peptide  $(POG)_7$ , lui, ne formait pas ces assemblages insolubles dans les mêmes gammes de concentration (cf. chapitre 1 1.3.2.). Il serait intéressant de voir si celui-ci s'agrège à plus forte concentration. Nous avons tenté d'inspecter par microscopie électronique certains des échantillons permettant de générer des aiguilles de  $(POG)_{10}$ . Il est évident que les assemblages vus par microscopie optique sont trop gros pour être analysés par cette technique mais nous espérons voir des amorces de ces assemblages de tailles inférieures (Figure 20).



**Figure 20:** Clichés de microscopie électronique à transmission sur  $(\text{POG})_{10}$ , 2,6mM, à gauche : pH=2.9 et à droite pH=7. Barre d'échelle 1 $\mu$ M. Les flèches blanches montrent les aiguilles.

Par microscopie optique, nous avons pu remarquer que les assemblages dans l'eau semblaient être de tailles significativement inférieures, ce qui explique pourquoi nous pouvons observer des aiguilles dans cet échantillon. Les aiguilles sont allongées et « plates », ce qui signifie qu'une direction de croissance est privilégiée contrairement à l'échantillon solubilisé dans le tampon phosphate où la croissance semble être équivalente dans trois directions, ce qui explique ces formes "cubiques". Les aiguilles apparaissent cassées sur toute la longueur de leur axe principal, ce qui semble être une conséquence directe de la méthode de dépôt sur la grille. Il est cependant possible d'estimer leurs dimensions, on trouve des longueurs d'aiguilles comprises entre 2 $\mu$ m et 5 $\mu$ m et des largeurs variant de 50nm à 300nm. Ces dimensions semblent coïncider avec les dimensions des constituants élémentaires des morphologies en bouquets observées précédemment. Une étude structurale par cristallographie pourrait être envisagée pour vérifier si les éléments cristallisés sont identiques ou non dans les deux cas ou si il s'agit uniquement d'une différence de maille cristalline.

## 2.4. Etude de la modification des phases denses de collagène par dopage avec des peptidomimétiques

### 2.4.1 Préparation des solutions de collagène

Il y a des décennies, des chercheurs ont montrés que le collagène acido-soluble pouvait former des fibrilles présentant des striations semblables à celles observées dans le collagène natif lorsque le pH était élevé.<sup>7</sup>

*Sources de collagène* – Différents protocoles d'extraction du collagène sont décrits dans la littérature.<sup>24,25</sup> Les sources de collagène de type I sont principalement les tendons et la peau animale tels que les tendons de queues de rats ou le derme de veau. Dans notre cas, nous avons extrait le collagène des tendons de queues de rat. Dans la mesure du possible nous n'avons utilisé que des queues de rats Long Evans jeunes (moins de 8 semaines) et n'ayant pas subi de traitement particulier (contrôles négatifs). En effet, en vieillissant, il se crée des liaisons covalentes entre les fibrilles de collagène, ce qui les rend moins solubles et nuit à la pureté des solutions. Un protocole d'extraction stricte est suivi car les expériences de physico-chimie requièrent des solutions très semblables d'une extraction à l'autre pour être reproductible.

*Protocole d'extraction* – Les tendons sont prélevés sur les queues de rats à l'aide de pinces, sous une hotte de culture cellulaire stérile, et immédiatement recueillis dans un bécher de PBS (Phosphate Buffer Saline : NaCl 137 mM, KCl 2,68mM, Na<sub>2</sub>HPO<sub>4</sub> 8,07mM, NaH<sub>2</sub>PO<sub>4</sub> 1,47mM). Les tendons sont lavés et centrifugés plusieurs fois dans du PBS afin d'éliminer les cellules et les traces de sang jusqu'à ce que le surnageant soit limpide. Les tendons sont ensuite placés dans une solution de NaCl 4M afin de lyser les cellules résiduelles par choc osmotique et de faire précipiter d'éventuelles protéines de haut poids moléculaire. Après un nouveau rinçage au PBS, les tendons propres sont solubilisés dans de l'acide acétique 500mM stérile. S'en suivent alors des cycles de précipitation/centrifugation avec des solutions de NaCl à des concentrations différentes pour éliminer les éventuels contaminants. Enfin, des dialyses sont réalisées afin d'éliminer le sel du collagène. La solution est conservée à 4°C et centrifugée à 41000 × g pendant 4 heures avant usage pour éliminer les particules insolubles restantes. (Le protocole détaillé est fourni en partie expérimentale §F p.141). La concentration des solutions mères ainsi préparées est comprise entre 2 et 5 mg/ml.



*Caractérisation de la pureté* – La pureté des solutions mères est caractérisée par électrophorèse en conditions dénaturantes sur gels de polyacrylamide. L'électrophorèse permet de séparer des molécules ou des fragments de molécules en fonction de leur charge électrique. Le milieu de migration est constitué d'un gel aqueux de polyacrylamide coulé entre deux plaques de verre maintenues verticalement. Une différence de potentiel est appliquée entre la base et le sommet du gel qui baigne dans un tampon conducteur. Les échantillons sont introduits au sommet du gel dans différents puits. Un marqueur contenant différentes protéines de poids connus est également déposé dans un puit du même gel, ce qui permet de rapporter la distance de migration à un poids moléculaire. Dans le système de tampons Laemmli<sup>26</sup> (SDS-PAGE) que nous avons utilisé, le tampon de migration est différent du tampon d'échantillon. Ce dernier contient un tensioactif, le Sodium Dodécyl Sulfate (SDS), et du  $\beta$ -mercapto-éthanol, un agent réducteur de groupements thiol, qui a pour but de réduire d'éventuels ponts disulfures entre chaînes polypeptidiques. Les protéines ainsi dénaturées prennent une conformation en pelote statistique (« random coil »). Elles ont alors un rapport charge/masse uniforme, proportionnel à leur masse moléculaire, puisque la surface de la molécule est intégralement recouverte de SDS chargé négativement. Les témoins des marqueurs de poids moléculaires sont eux aussi dénaturés et se comportent de la même manière. De plus, la vitesse de migration et la résolution dépendent de la taille des pores du gel. Cette taille est contrôlée par la concentration en monomères d'acrylamide utilisée pour fabriquer les gels ou par la proportion d'agent réticulant bis-acrylamide. En routine, nous avons utilisé des gels à 10% en acrylamide/bis-acrylamide (30%). Le protocole détaillé est fourni en partie expérimentale (§H p.144).

Un gel où différents lots de collagène ont été analysés est présenté en Figure 21. Les bandes caractéristiques du collagène sont les quatre bandes les plus intenses. Elles correspondent, par ordre croissant de masse moléculaire aux brins  $\alpha 2$  (~95 kD) et  $\alpha 1$  (~100 kD) et aux dimères  $\beta 12$  (~195 kD) et  $\beta 11$  (~200 kD). Des bandes beaucoup moins intenses sont visibles vers 52kDa et pourraient correspondre à de l'albumine, de l'élastine ou à des Heat-Shock Proteins (HSD47).

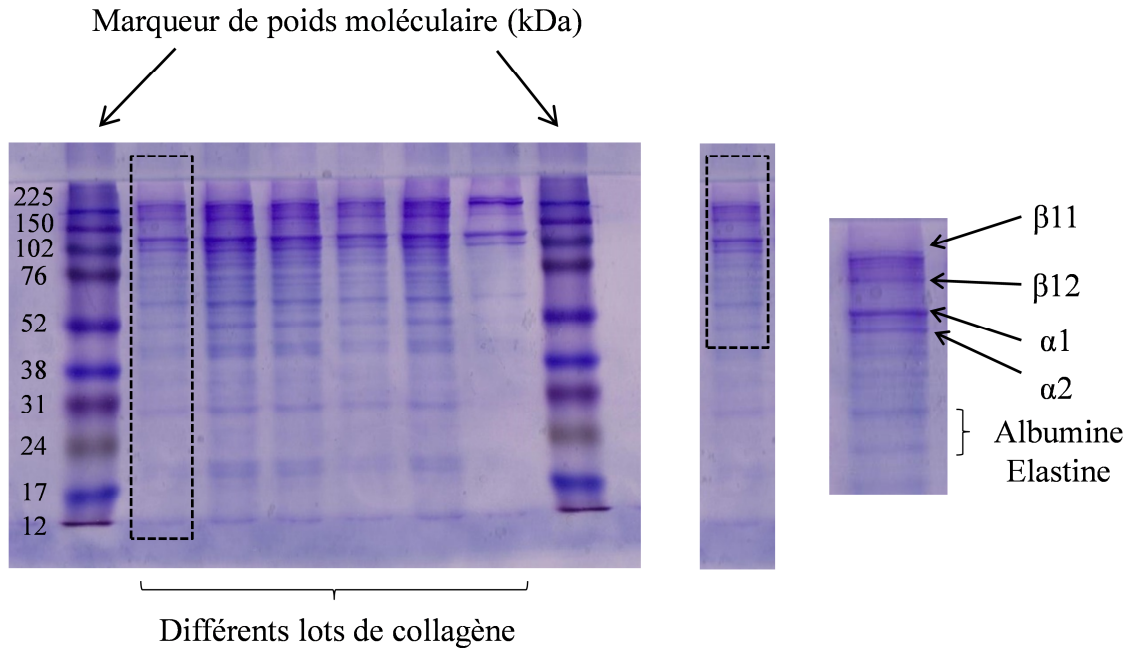


Figure 21: Gel d'électrophorèse en conditions dénaturantes

Toutefois, cette technique ne nous permet pas de détecter des contaminants non protéiques tels que des protéoglycanes dont la présence est connue.

#### 2.4.2 Les phases cristallines liquides

Le nom « cristal liquide » est né en 1900 après la découverte de cristaux de benzoate de cholestérol coulant comme des liquides par l'autrichien Friedrich Reinitzer et son collègue Lehmann. Les cristaux liquides sont caractérisés par un état intermédiaire entre la phase cristalline, où règne un ordre tridimensionnel, et la phase liquide où aucun ordre n'existe. Il existe trois grandes familles de cristaux liquides : les phases *nématiques* qui ne possèdent qu'un ordre d'orientation, les phases *cholestériques* où le vecteur directeur, caractéristique d'un ordre d'orientation, est hélicoïdale, et les phases *smectiques* où les molécules possèdent une orientation unique qui est conservée dans les 3 dimensions (Figure 22).

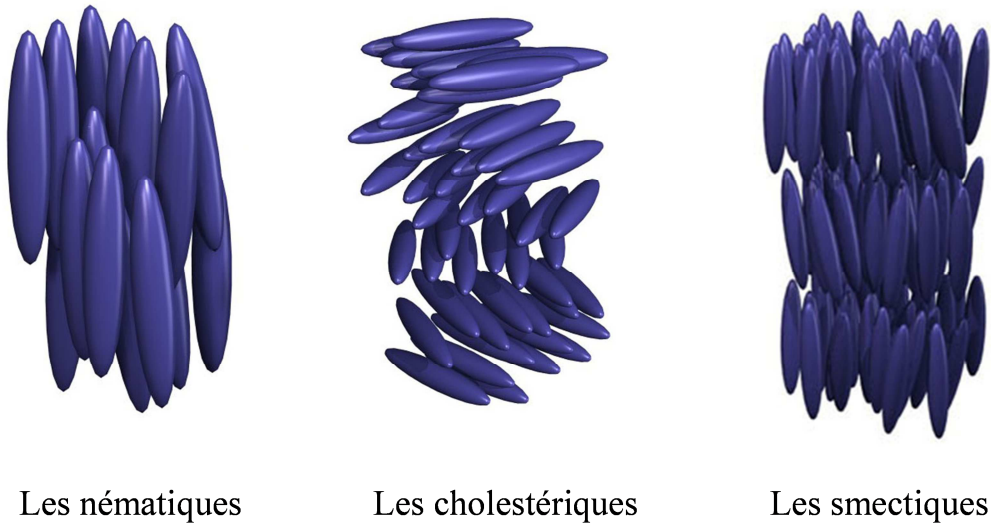


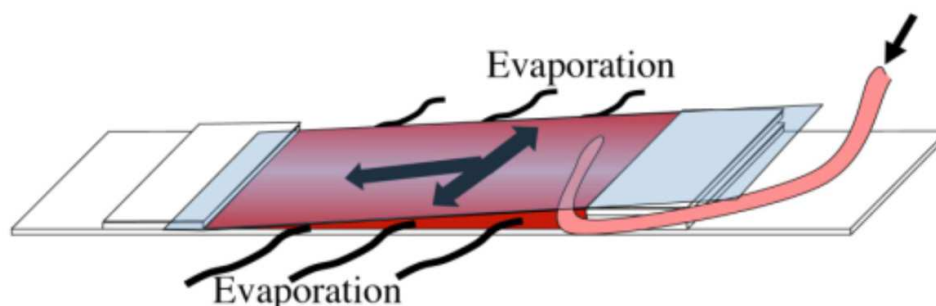
Figure 22: Les trois grandes familles de phases dans les cristaux liquides

Ces différentes phases se différencient par leurs propriétés optiques telles que la biréfringence (propriété d'un matériau dans lequel la lumière se propage de façon anisotrope). On pourra alors observer avec un microscope à lumière polarisée des zones de textures distinctes correspondant à des domaines où les molécules sont orientées dans des directions différentes.

Des études précédentes ont montrées que, dans des conditions acides, les molécules de collagène de type I pouvaient s'organiser à forte concentration pour former spontanément phases cristallines liquides.<sup>27</sup> Ces phases présentent donc des propriétés de biréfringence lorsqu'on les observe en lumière polarisée. Il est alors possible de distinguer différents arrangements des monomères de collagène en fonction de la concentration notamment des ondulations précholestériques et des phases cholestériques lorsqu'on atteint une concentration de 80mg/ml.<sup>28,29</sup> La torsion cholestérique est notamment décrite pour diverses matrices extracellulaires *in vivo*, cette torsion se produit entre les fibrilles de collagène alors que les torsions observées *in vitro* se produisent entre les molécules fragmentées. Cependant, la présence de phases cholestériques à plus petite échelle pourrait correspondre au premier état de la sécrétion du collagène par les cellules. Des études structurales utilisant des techniques classiques de fixation chimique sont inappropriées pour les phases cristallines liquides en raison de l'absence de liens stables conservant leur ordre en trois dimensions. Le processus de fibrillation est alors induit en plaçant les phases cristallines liquides sous vapeur d'ammoniac. Les vapeurs, en se solubilisant dans le solvant, neutralisent l'acide acétique et induisent une augmentation de pH. Cette méthode

permet de gélifier des solutions sans risque de dilution. Néanmoins, elle ne permet pas de contrôler les conditions physico-chimiques telles que le pH et la force ionique du système. Les fibres sont ensuite étudiées par microscopie électronique. De ces études, *Besseau et al.* ont pu montrer que l'ultrastructure du collagène liquide cristallin stabilisé imite fidèlement les modèles fibrillaires décrits *in vivo* dans des matrices extracellulaires.<sup>29</sup> Un dispositif expérimental a été développé pour l'étude de ces phases denses de collagène par *Gervaise Mosser* (Equipe Matériaux et Biologie, LCMCP, UPMC) : les microcellules.

*Les microcellules*<sup>30</sup> – Afin d'analyser les phases cristallines liquides, des solutions de collagène à 5 mg/ml sont injectées en continu dans des microchambres de verre pour augmenter la teneur en collagène tout en compensant l'évaporation (Figure 23). Ces microchambres de verre ou microcellules sont inclinées et ouvertes sur les deux côtés espacées de manière décroissante de 300  $\mu\text{m}$  à quelques micromètres. Cet ensemble nous a permis d'étudier l'effet du confinement en plus de la concentration en collagène.



**Figure 23: Schéma d'une microcellule permettant la formation de phases denses de collagène**

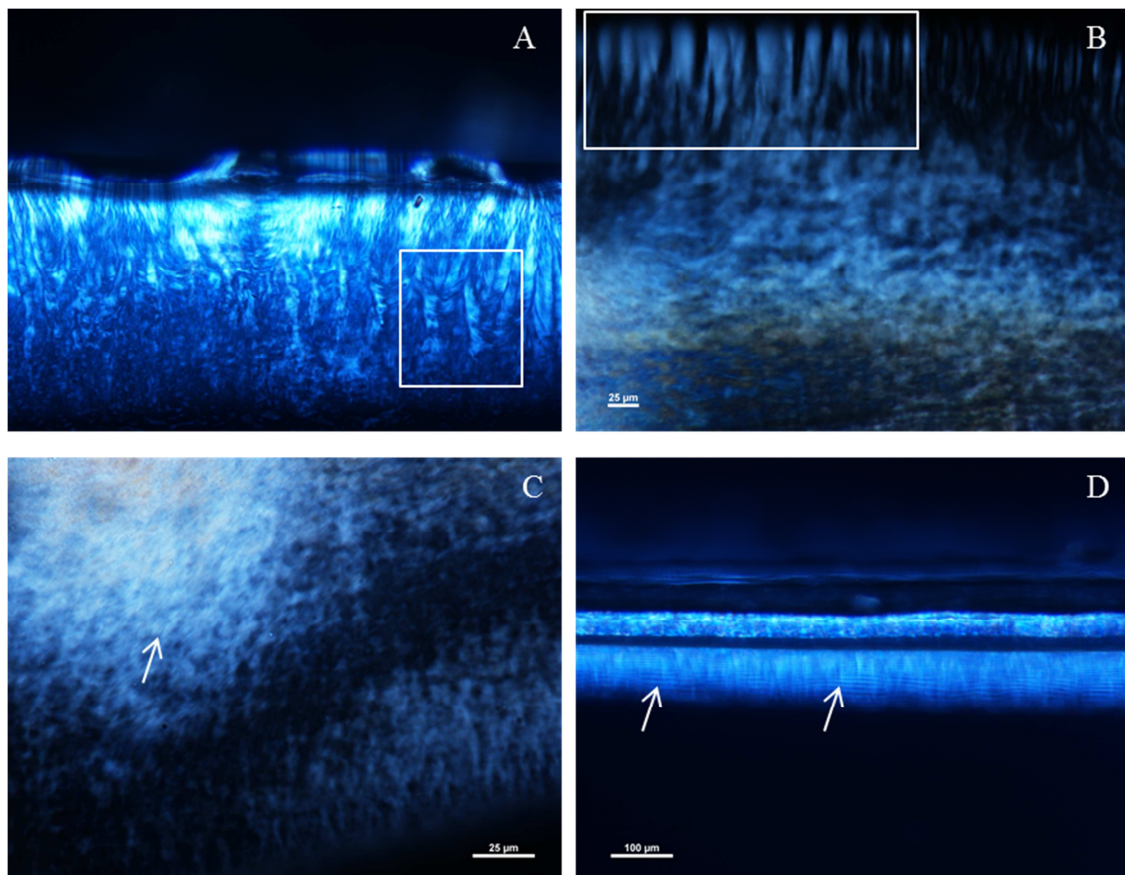
Le flux continu est maintenu plusieurs jours (5 ml sont injecté). Les microcellules sont observées une première fois après 4 ou 5 jours par microscopie à lumière polarisée et présentent des zones biréfringentes à l'interface air/eau. Après 7 jours, il est possible d'identifier la présence de phases cholestériques. Le processus de fibrillation peut être ensuite induit par vapeur d'ammoniac, fixant ces assemblages, en vue d'avoir des informations structurales plus précises par microscopie électronique (Protocole détaillé de la fixation pour la microscopie électronique est fourni en partie expérimentale § B1 p.135).

### 2.4.3 Dopage du collagène par des peptidomimétiques

De nombreux matériaux inorganiques (métaux, la silice, apatite)<sup>31-33</sup> ou polymères synthétiques (polyéthylène, téréphtalate, alcool polyvinyle, polypyrrole, polyéthylène glycol)<sup>34-36</sup> ont été utilisés pour modifier/stabiliser les phases de collagène. Il a notamment

été montré que ces matériaux/polymères mélangés ou liés de manière covalente au collagène naturel forment des matériaux hybrides ou composites. Ces matériaux trouvent d'ailleurs leurs applications dans le biomédical ou les biomatériaux. De manière analogue, nous nous sommes intéressés au dopage du collagène par les peptides mimes de collagène (POG)<sub>n</sub>, des interactions entre peptides et collagène de type I ayant été déjà mises en évidence.<sup>37,38</sup> Les similarités de structures primaire, secondaire et tertiaire entre les peptides CMP et le collagène naturel pourrait en effet permettre des interactions perturbant les phases denses qui sont obtenues en absence de peptides.

*Le collagène* – Des études réalisées précédemment au laboratoire de Chimie de la Matière Condensée ont montré qu'à pH=2,5, des phases cholestériques étaient observées.<sup>30</sup> Nous avons donc réalisé des microcellules en injectant des solutions de 1mg/ml à 0,2mg/ml en collagène solubilisé dans de l'acide acétique 500mM, HCl 3mM ou HCl 0.3mM (Figure 24).



**Figure 24:** Clichés de microscopie à lumière polarisée d'échantillons de collagène dans différentes conditions (A) 1mg/ml, Acide Acétique 500mM, (B) 0.2mg/ml, Acide Acétique 500mM, (C) 0.4mg/ml, HCl 3mM, (D) 0.4mg/ml, HCl 0.3mM. Les carrés entourent les structures en bandes et les pointes des flèches montrent les phases cholestériques.

A 1mg/ml dans de l'acide acétique, nous observons une texture dite « texture en bandes » qui est généralement rencontrée lorsque les solutions de collagène sont injectées en concentration initiale trop élevée. Nous avons donc diluée cette solution mais, même à 0.2mg/ml, une texture similaire est observée (cliché B). Selon le lot de collagène purifié, les domaines de stabilité des différentes phases peuvent varier. Nous avons alors tenté de changer la nature du tampon pour observer les phases cholestériques. Dans HCl 3mM et HCl 0.3mM (cliché C et D), nous avons pu observer du cholestérique caractérisé par des ondulations. Nous avons donc déterminé deux types de conditions menant à des organisations distinctes. Les arrangements obtenus en phase dense conditionnent les propriétés des biomatériaux, générés dans un second temps par passage de vapeur d'ammoniac. La possibilité de modifier les assemblages supramoléculaires en phase liquide à l'aide des molécules CMP a été explorée.

*Phases denses de peptidomimétiques* – Avant d'étudier l'interaction de nos peptides (POG)<sub>7</sub> et (POG)<sub>10</sub> avec les phases denses de collagène, nous avons inspecté le comportement de nos peptides seuls en microcellules. Les images de microscopies obtenues se sont révélées bien différentes de celles décrites précédemment pour le collagène. Les peptides forment des cristaux de morphologies différentes suivant les conditions de solubilisation (Figure 25). Le peptide (POG)<sub>10</sub> solubilisé dans des solutions acides (clichés A-B) forment des cristaux rectangulaires principalement à l'interface air-solution alors que quand le peptide est solubilisé dans l'eau, nous observons des cristaux triangulaires présents dans toute la microcellule. Le peptide (POG)<sub>7</sub> forme des cristaux en aiguilles (clichés E-F). Lorsqu'il est solubilisé dans l'acide acétique (cliché E), nous pouvons remarquer que l'organisation des aiguilles diffère suivant la zone d'observation, probablement en raison du gradient de concentration qui existe entre l'intérieur de la cellule et le bord de la cellule à l'interface air-solution : les aiguilles sont parallèles les unes aux autres à l'interface air-solution (apparaissant ici en bleu), puis on observe une bande intermédiaire de ~20µm dans laquelle elles sont arrangées perpendiculairement, et une troisième zone enfin, lorsque l'on se rapproche du centre de la microcellule, où elles sont désordonnées.

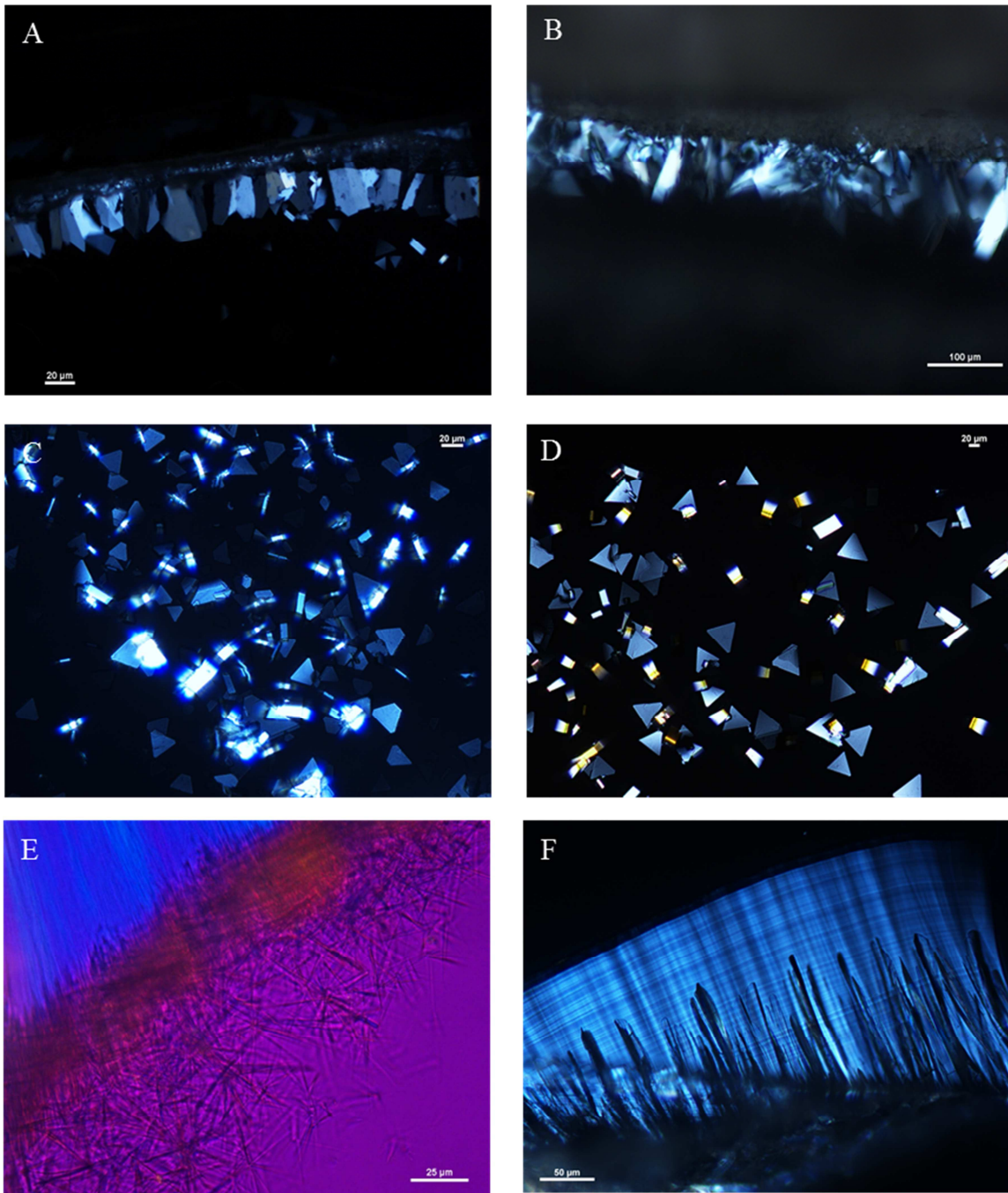


Figure 25: Clichés de microscopie à lumière polarisée des échantillons de  $(POG)_{10}$  (A-D) et  $(POG)_7$  (E-F) dans différentes conditions (A) 1mg/ml, Acide Acétique 500mM, (B) 1 mg/ml, Acide Acétique 5mM, (C) 1mg/ml, eau, (D) 0.2mg/ml, eau, (E) 1mg/ml, Acide Acétique 5mM, (F) 1 mg/ml, eau

Lorsque le peptide  $(POG)_7$  est solubilisé dans l'eau (cliché F), nous observons la coexistence de deux organisations distinctes. Il y a des aiguilles qui sont bien plus épaisses que celles présentes dans le cliché E ainsi qu'un bandeau. Dans ce bandeau, nous observons une organisation selon deux axes. Bien que cela ne soit pas le but initial de ces études, nous avons envisagé d'étudier ces formes cristallines plus en détail, que ce soit par diffraction des rayons X ou par RMN. Nous avons réalisé quelques essais pour produire des cristaux de taille plus importante (pour les RX) ou d'en produire davantage (pour la

RMN). Ces essais n'ayant pas été concluants, nous avons différé ces études pour nous concentrer sur l'étude des interactions CMP/collagène.

*Dopage par des peptidomimétiques* – Nous avons dopé du collagène (1mg/ml solubilisé dans l'acide acétique 500mM) par le peptide (POG)<sub>10</sub> (1mg/ml solubilisé dans l'acide acétique 500mM) à différents ratios (v/v) (Figure 26). Aucun changement majeur n'a été observé. Quelle que soit la concentration en peptide injectée (0%, 5%, 10% ou 15%), nous avons toujours obtenu la même signature optique.

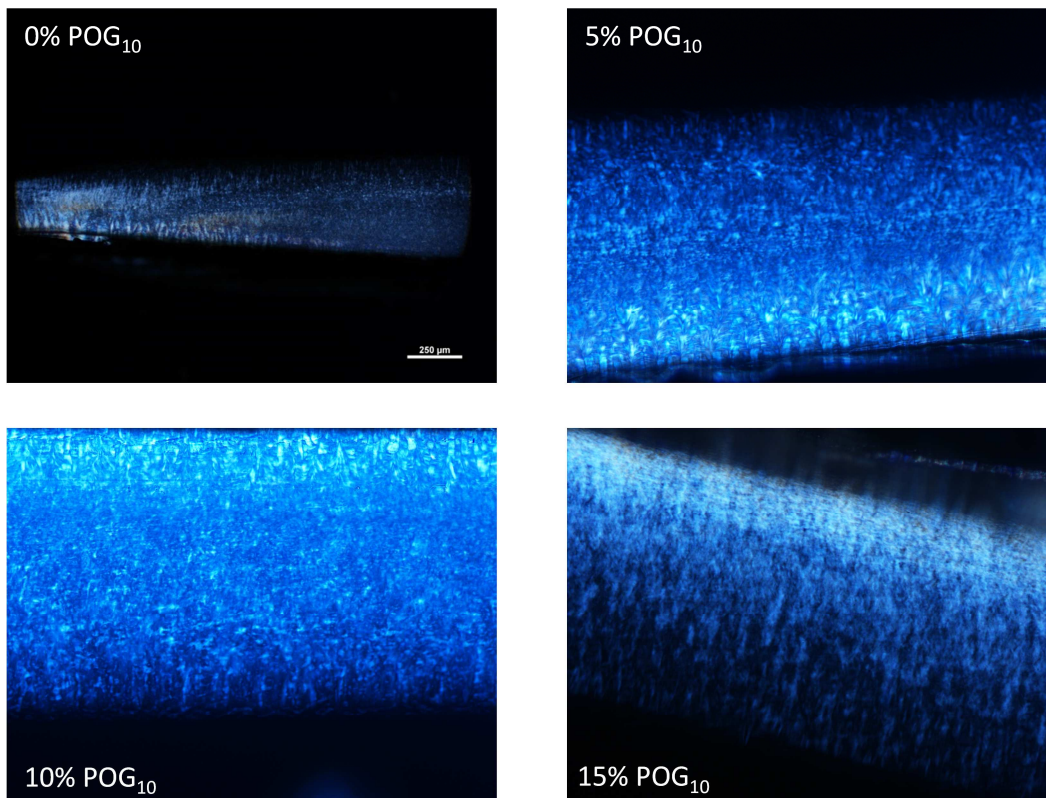


Figure 26: Clichés de microscopie à lumière polarisée des échantillons de collagène (1mg/ml dans AA 500mM) + (POG)<sub>10</sub> (1mg/ml dans AA 500mM) dans différents ratios

Ce résultat traduit une probable absence d'interaction dans les conditions testées, conduisant à une ségrégation de phases. En effet, le collagène conserve son organisation propre, et l'on peut imaginer qu'à une concentration de CMP supérieure, des cristaux pourraient apparaître indépendamment des structures en bande. Des résultats préliminaires indiquent que lorsque le collagène a été destabilisé par chauffage (gélatine), il se comporte différemment en présence (25%) ou en absence de CMP. Bien que ces résultats aient besoin d'être confirmés, ils rejoignent des observations faites par Wang *et al.* qui avaient montré que des peptides modifiées pouvait se lier au collagène seulement quand celui commençait à se dénaturer.<sup>38</sup>



## Conclusions et perspectives

Nous avons utilisé la DLS et la RMN liquide pour mettre en évidence la présence de différents assemblages supramoléculaires (triple hélices, particules solubles et insolubles) et pour étudier leurs cinétiques de formation. Les résultats de dynamique moléculaire obtenus précédemment illustrent parfaitement les raisons pour lesquelles la formation de la triple hélice n'est pas instantanée. Celle-ci requiert un rapprochement de 3 chaînes peptidiques en conformation PPII étendue, or les trajectoires obtenues indiquaient que les monomères étaient rarement dépliés.

Combinés à des approches de microscopies, nos observations ont permis de proposer deux modèles d'assemblages supramoléculaires. L'un d'eux, insoluble, présente un taux de triple hélice et un ordre très important mais n'est observé que pour (POG)<sub>10</sub>. Une oligomérisation des monomères inhomogènes d'un point de vue conformationnel pourrait expliquer la présence des assemblages solubles observés pour (POG)<sub>7</sub>, relativement désordonnés. Nos résultats montrent que la nature des contre-ions présents en solution et le pH influent directement sur leur stabilité. Un mécanisme de formation de ces deux types d'assemblages est présenté, en accord avec nos données cinétiques et RMN (STD). Afin de mieux caractériser les paramètres thermodynamiques guidant ces processus, des expériences complémentaires seront menées utilisant différentes conditions physicochimiques (concentrations, sels, pH, température...).

En ce qui concerne le dopage des phases denses de collagène, nous n'avons pas observé d'interaction directe entre CMP et collagène natif. Des études sur le dopage de la gélatine par nos CMP sont en cours et semblent plus prometteuses.

## Partie expérimentale

### A. Etude DLS

#### Matériels

- Spectromètre Brookhaven équipé d'un analyseur potentiel  $\zeta$  (BI-ZetaPlus)
- Cuve en plastique de capacité 500 $\mu$ l

#### Mode opératoire

Les études par DLS ont été réalisées sur 2 peptides (POG)<sub>7</sub> (C=10mM) et (POG)<sub>10</sub> (C=2,6mM et C=10mM). Les solutions de dilutions sont HCl 3mM pH 2,3 et un tampon K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub> pH 7. Nous nous sommes assuré que les échantillons avaient été incubés préalablement à 4°C pendant 24h avant chaque mesure afin de permettre aux triples hélices de s'assembler (sauf indication contraire dans le tableau). Pour une condition donnée, les mesures ont été réalisées 10 fois afin d'obtenir une valeur moyennée. Les conditions expérimentales ainsi que les rayons hydrodynamiques (R<sub>h</sub>) de chaque population sont décrites dans le Tableau 2 pour le peptide (POG)<sub>7</sub> et dans le Tableau 3 pour le peptide (POG)<sub>10</sub>.

#### Résultats

Le commentaire *Après chauffage à 90°C* signifie que les mesures à 65°C ont été réalisées après un chauffage dans un bain marie à 90°C afin de dissocier les assemblages supramoléculaires. Le commentaire *h+2* signifie que les mesures ont été réalisées 2h après la mesure précédente. L'échantillon est resté à température ambiante durant ces 2h. Le commentaire *j+1* signifie que les mesures ont été réalisées 1 jour après la mesure précédente. L'échantillon est resté à température ambiante durant cette journée.

Solvant	Température	Commentaires	R <sub>h</sub> (nm)
HCl pH=2,3 (20°C)	20°C		2,4
			754,6
	65°C	Après chauffage à 90°C	2,9
			362
	20°C	h+2	2,6
			711,2
Phosphate pH=2,5	20C		2,6
			1241,5
	65°C	Après chauffage à 90°C	2,4
			678
	20°C	h+2	2,6
			1238,4

Tableau 2 : (POG)<sub>7</sub>, 10mM

Solvant	Concentration et pH	Température	Commentaires	R <sub>h</sub> (nm)
HCl	2,6mM pH=2.54	20°C		5,6
				1753,6
		20°C	h+2	3,7
				719,8
		20°C	j+1	4,4
				1916,7
	10mM pH=2.13	20°C		3,6
				362,5
		65°C	Après chauffage à 90°C	936
		20°C	h+ 1,5	3,7
				1663,8
		20°C	j+1	3,9
1013,8				
Phosphate	2,6mM pH=6.95	20°C		122,8
				727,2
		55°C	15 min	100,3
				684,5
		65°C	15 min	123,7
				874,7
		20°C	h+2	167,4
				1467,7
	20°C	h+2	620,3	
	10mM pH=3.95	20°C		847,9
				3993,5
		65°C	Après chauffage à 90°C	521
				778,5
				2980,8
				5560
		20°C	h+ 1,5	4,7
1885,9				
		4940,5		
20°C	j+1	4500		

 Tableau 3 : (POG)<sub>10</sub>

## B. Préparation des échantillons pour la microscopie électronique

Les échantillons ont été préparés selon les trois techniques présentées ci-après. Les observations ont été réalisées sur un microscope électronique à transmission FEI Tecnai spirit G2 fonctionnant à 120 kV. Les images ont été enregistrées avec une caméra CCD (Orius Gatan 832 digital).

### B1. Protocole de fixation pour la microscopie électronique

Se fait dans des salières en verre :

#### 1ère fixation : glutaraldéhyde

- Préparer une solution de glutaraldéhyde :
  - glutaraldéhyde à 8% *1ml*
  - Tampon Caco 0,05M - pH 7,4 *2,2ml*
- Ajouter  $\approx$  1ml de solution de glutaraldéhyde dans chaque salière
- Laisser 1 heure à 4°C
- Effectuer 3 lavages successifs de 10 mn avec un mélange :
  - 1 vol. Tampon Caco 0,1M - pH 7, 4
  - 1 vol. saccharose 0,6M

#### 2ème fixation : osmium

- Préparer une solution d'Osmium :
  - Osmium tétraoxyde 4% *1vol.*
  - H<sub>2</sub>O distillée *1 vol.*
- Ajouter  $\approx$  1ml de solution d'Osmium dans les salières
- Laisser 1 heure à 4°C
- Effectuer 3 lavages successifs de 10 mn avec H<sub>2</sub>O distillée

#### Déshydratation

- Alcool 50° *5 mm*
- Alcool 70° *5 mm*
- (Il est possible d'arrêter à cette étape et mettre les salières à 4°C)
- Alcool 95° *5 mm*
- Alcool 100° *2\*10 mm*
- Alcool 100° 1 vol. + Oxyde de propylène 1 vol. *10 mn*
- Oxyde de propylène *10 mm*

#### Imprégnation

- Préparer une solution d'Araldite
  - Araldite CY212 *20 ml*
  - DDSA = durcisseur *22 ml*
  - BDMA = accélérateur *1,1 ml*
- Oxyde de propylène ¾ vol. + Araldite ¼ vol *60 mn*
- Oxyde de propylène ¼ vol. + Araldite ¾ vol *1 nuit à 4°C*
- Changer les échantillons de salières
- Araldite *1 h*

- Changer les échantillons de salières
- Remettre de l'Araldite

*¾ h à 1h heure en dégazant*

### **Inclusion**

- Se fait dans une plaque à inclusion
- Remplir presque complètement les alvéoles de solution d'Araldite (la même solution que celle préparée pour l'imprégnation)
- Placer l'échantillon au sommet de l'alvéole en le centrant bien
- Finir de remplir l'alvéole de solution d'Araldite à l'aide d'une petite seringue
- Mettre à l'étuve à 60°C pendant 3 jours

### **Echantillon préparé pour nos études**

- Echantillon de collagène pur solubilisé dans de l'acide acétique 500mM fibrillé à partir de la solution en microcellules
- Echantillon de collagène avec 5% (v/v) de peptide (POG)<sub>10</sub> fibrillé à partir de la solution en microcellules

## B2. Préparation des échantillons pour la cryomicroscopie

### **Préparation des grilles**

- Prévoir 3 grilles par échantillon
- Les grilles sont placées dans une effluveur pour les rendre hydrophiles et favoriser l'étalement des échantillons.

### **Congélation de l'échantillon**

- Remplir le puit d'azote liquide
- Remplir le puit d'éthane gazeux qui se liquéfie suite à la température imposée par l'azote liquide
- La grille à congeler est placée à l'extrémité d'un système de guillotine à l'aplomb du puit
- 2 à 4 µl d'échantillon sont déposés sur la grille (l'excès de liquide est éliminé avec du papier filtre whatman #4)
- Plonger la grille dans l'éthane liquide (l'excès d'éthane liquide est éliminé avec du papier filtre)
- Conserver l'échantillon dans l'azote liquide pour éviter le réchauffement et la formation de glace cristalline.

### **Préparation de l'analyse**

- La grille est ensuite installée dans le porte échantillon qui est sous atmosphère d'azote
- Le porte échantillon est alors introduit dans le cryomicroscope.

Le paramètre limitant pour cette technique est la formation de glaçons qui rendent l'observation impossible. Il est alors très important que la congélation se fasse très rapidement et qu'une fois l'échantillon congelé, il ne se réchauffe pas.

La difficulté majeure a été de préparer nos échantillons dans lesquels nous souhaitions voir les assemblages supramoléculaires solubles. Il a fallu déposer des échantillons « chauds » (préalablement chauffés à l'étuve à 90°C pendant 20 min) pour ensuite les congeler. Ces échantillons représentaient les contrôles négatifs, nous permettant ainsi de montrer que nous avons détruit tous les assemblages supramoléculaires. Un système de chauffage a été mis en place pour chauffer l'extrémité de la guillotine. Les échantillons congelés pour ces études n'ont rien donné, ce qui est dû à un mauvais dépôt sur les grilles.

### **Echantillons préparés pour nos études**

Chaque échantillon a été préalablement incubé à 4°C pendant 24h avant dépôt.

- (POG)<sub>10</sub>, 2,6 mM dans un tampon phosphate (50mM) pH=7
- (POG)<sub>10</sub>, 200µM dans HCl 3mM (la solution a été séparée en 2 afin de faire un témoin négatif qui consiste en un chauffage de l'échantillon à 90°C pendant 20 min pour détruire les assemblages supramoléculaires)
- (POG)<sub>7</sub>, 200µM dans HCl 3mM (la solution a été séparée en 2 pour les mêmes raisons que l'échantillon précédent)

### B3. Préparation des échantillons pour l'ombrage

#### **Dépôt des échantillons sur plaques de Mica**

- Découper les plaques de Mica en carré de 0,5cm de côté.
- Vaporisation de l'échantillon à l'aide d'air comprimé ou d'un spray (l'air comprimé semble être plus efficace) sur la plaque de Mica.

#### **Réalisation de l'ombrage**

- Dépôt d'un film de platine par évaporation sous vide, sous un angle d'incidence de 10°, 20° ou 30°.
- Dépôt de carbone à 90° pour faire un film continu.

#### **Réalisation des grilles**

- Les feuilles de Mica sont plongées dans l'eau afin de décoller la réplique de l'échantillon formée.
- A l'aide d'une pince, les grilles sont à leur tour plongées dans l'eau pour ensuite récupérer cette réplique qui flotte.
- Les grilles sont ensuite séchées sur papier filtre avant l'observation.

### **Echantillons préparés pour nos études**

Chaque échantillon a été préalablement incubé à 4°C pendant 24h avant dépôt.

- (POG)<sub>10</sub>, 200µM dans HCl 3mM (la solution a été séparée en 2 afin de faire un témoin négatif qui consiste en un chauffage de l'échantillon à 90°C pendant 20 min pour détruire les assemblages supramoléculaires)

- (POG)<sub>7</sub>, 200μM dans HCl 3mM (la solution a été séparée en 2 pour les mêmes raisons que l'échantillon précédent).

### C. Préparation des échantillons pour les études par microscopie optique

#### **Préparation des échantillons**

Quatre échantillons de (POG)<sub>10</sub> à 7mg/ml ont été réalisés :

- A et B : Solubilisation dans un tampon phosphate (K<sub>2</sub>HPO<sub>4</sub>/KH<sub>2</sub>PO<sub>4</sub>) pH=7
- C : Solubilisation dans de l'acide acétique 5mM pH=2,5
- D : Solubilisation dans de l'eau distillée

#### **Formation des agrégats**

- A : Formation des agrégats à 35°C (étuve) pendant 3 jours
- B, C et D : Formation des agrégats à 60°C (étuve) pendant 3 jours

#### **Observation des échantillons**

10 μl de chaque échantillon sont déposés sur une plaque de verre pour observer la morphologie des agrégats. Le microscope à lumière polarisée utilisé est un Nikon Eclipse E600 Pol, équipé de polariseurs croisés et d'une caméra DXM 1200CCD. Les agrégats formés sont de morphologies identiques.

#### **Evolution des échantillons**

Ensuite, les quatre échantillons ont subi les mêmes traitements. A savoir, un chauffage à 90°C pendant 20 min afin d'observer la dissolution des agrégats, puis 3 jours à température ambiante (clichés montrés dans le manuscrit).

### D. Etudes cinétiques

#### D1. Préparation des échantillons

##### **Matériel**

- Tubes shighemi 5mm

##### **Mode opératoire**

Les études ont été réalisées avec le peptide (POG)<sub>7</sub> et (POG)<sub>10</sub> à 200μM solubilisés dans l'acide chlorhydrique 3mM (solution dans laquelle est solubilisé le collagène). Chaque échantillon est chauffé au préalable à 90°C dans un bain marie pendant 20 minutes afin de détruire les assemblages formés puis directement introduit dans le spectromètre RMN stabilisé à la température souhaitée (étude réalisée à 10°C, 20°C, 25°C, 30°C et 40°C).

## D2. Expériences RMN et traitement des données

Les spectres ont été enregistrés sur un spectromètre Bruker Avance III 500MHz équipée d'une crysonde.

### Paramètres d'enregistrement des expériences

#### - Expériences $^1\text{H}$ STD :

Un spectre de référence  $^1\text{H}$  (impulsion sélective *off-resonance* à +30000Hz) est enregistré suivi de l'acquisition d'un spectre  $^1\text{H}$  auquel est appliqué une impulsion sélective comprise entre -1000Hz et 400 Hz (56,7dB, 0,1kHz) pour saturer les assemblages supramoléculaires (spectre *on-resonance*). Le spectre STD est obtenu par la soustraction du spectre *on-resonance* au spectre *off-resonance*. Des spectres STD sont alors enregistrés toutes les 20 min pendant 12h.

#### - Expériences protons :

L'expérience RMN consiste en une succession de spectres  $^1\text{H}$  (simple excitation) à la température souhaitée (10°C, 20°C, 25°C, 30°C et 40°C). L'intervalle de temps entre chaque expérience est de 86s (temps d'acquisition du spectre).

### Traitement des données

#### - Expérience STD :

Nous avons intégré les signaux entre 2,4-1,5 ppm à l'aide de TopSpin® ensuite nous les avons traité en utilisant un script Mathematica®. La cinétique d'agrégation étant rapide et le rapport signal sur bruit modeste, il est impossible d'observer une évolution dans le temps.

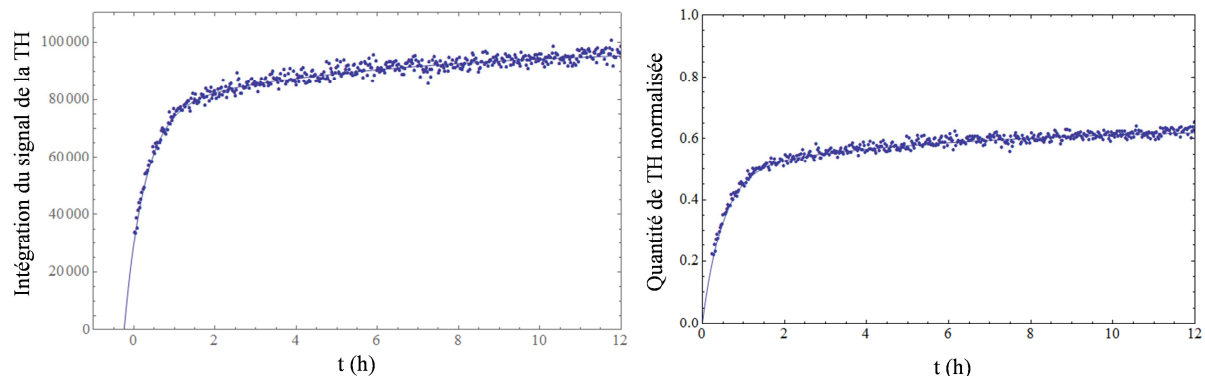
#### - Expérience protons :

Nous avons intégré les signaux entre 4,35-4,20 ppm pour le monomère, 3,30-3,10 ppm pour la triple hélice et 2,90-1,50 ppm pour les nano-objets à l'aide de NMRpipe®. Les données ont été ensuite traitées sur Mathematica®.

Le chauffage à 90°C des échantillons permet la dissociation des triples hélices (à  $t=0$ , triples hélices=0%). La modélisation par une équation à double exponentielle des données analytiques sur la triple hélice, nous permet alors d'avoir accès au temps mort (voir exemple). Le temps mort correspond au temps entre la fin du préchauffage à 90°C et le début de l'enregistrement des données. Après avoir déterminé ce temps mort, les données obtenues pour chaque population (M, TH et O) sont normalisées par un facteur correctif qui est obtenu par normalisation à 1 de la quantité de monomères à  $t=0$ .



Exemple sur les données expérimentales du suivi de la triple hélice à 10°C pendant 12h :



Données expérimentales brutes modélisées par une fonction double exponentielle

Données expérimentales normalisées par le temps mort et le facteur correctif

### Étude cinétique : extraction des constantes de vitesse

Les données expérimentales normalisées pour chaque population (monomère, triple hélice et monomère) à une température donnée sont utilisées pour extraire les constantes de vitesse. Ces données ont été traitées avec Mathematica®. Les solutions analytiques du système d'équations différentielles (cf. partie 2.2.2 p.106 Eq (1), (2) et (3)) ont été utilisées pour ajuster simultanément les courbes expérimentales des trois populations. Le protocole NonlinearModelFit avec la méthode Nelder-Mead a été implémenté dans NMinimize (voir <http://mathworld.wolfram.com/> pour plus de précisions) afin d'optimiser la recherche des constantes de vitesse. Un exemple du code Mathematica utilisé pour ajuster les courbes expérimentales des 3 états se trouve dans en ANNEXE 4.

## E. RMN Solide

### Conditions expérimentales CPMAS :

Fréquence de rotation : 10kHz

Temps de contact : 2 ms (poudre lyophilisée) / 10ms (assemblages supramoléculaires)

Impulsion sélective sur le proton : square.100 (impulsion carrée 100 points)

Condition de Hartmann-Hahn : 57 kHz ( $^1\text{H}$ ), 64kHz ( $^{13}\text{C}$ )

Découplage  $^1\text{H}$  : SPINAL64 à 90kHz

## F. Protocole d'extraction du collagène

### **Prélèvement des tendons**

Extraction à partir de tendons de queue de rat (30 à 40 queues de rats). L'extraction se fait sous la hotte à flux laminaire.

#### Matériel nécessaire :

- 2-3 litres de PBS
- 1 bécher contenant ½ litre de NaCl 4M dans H<sub>2</sub>O
- pince coupe-ongle
- ciseaux

#### Technique :

- À l'aide de la pince coupe-ongles, serrer légèrement la queue à 1cm de son extrémité fine.
- Casser la vertèbre par rotation du poignet (comme pour dénuder un fil électrique) et tirer pour sortir les tendons de la queue
- Couper les tendons dénudés au-dessus d'un bécher de 500 ml de PBS
- Recommencer l'opération de 1 cm en 1 cm. (Ne pas utiliser les 3 derniers centimètres de la base de la queue)

### **Lavage des tendons**

- Laver ABONDAMMENT et plusieurs fois avec PBS, centrifugations successives 5min/1500 g. Le surnageant doit être limpide et incolore (pas de sang résiduel : donc pratiquement pas d'albumine).
- Rinçage des tendons dans du NaCl (dans de l'eau distillée) à 4M (4M permet de retenir tout le collagène et d'éclater tout ce qui est cellulaire, notamment le sang, évacuation de l'albumine)
- Laver ABONDAMMENT et plusieurs fois les tendons, avec PBS, centrifugations successives / 5min / 1500 × g. (élimination au maximum du sel et de tous les débris cellulaires)

### **Mise en solution du collagène**

- Filtrer les tendons isolés sur une gaze stérile. Les peser. (NB : 10 queues de rats fournissent environ 1 gramme de tendon).
- Les mettre dans un bécher contenant du CH<sub>3</sub>COOH 0,5 M (pH 2,5 équilibré avec HCl 1 N) stérile. NB : Dissolution de 30 grammes de tendons dans 2,5 litres.
- Mettre à agiter (dans la pièce à 20°C).
- Laisser le collagène se dissoudre pendant 24 h.
- Filtrer la solution sur gaze. (Élimine les gros agrégats insolubles qui ne sédimentent pas à la centrifugation).

### **Purification du collagène**

- Centrifugation du filtrat à 19000 rpm pendant 2 heures. (Élimine les gros agrégats non dissous).
- Précipitation du surnageant avec NaCl 300 mM, centrifugation à 19000 rpm pendant 3 heures. Conserver le surnageant et éliminer le culot (Élimine les éventuels contaminants qui auraient précipité par le sel jusqu'à 300mM).

- Précipitation du collagène I avec NaCl 600 mM
- Centrifugation à 5500 rpm pendant 45 minutes (10°C) (permet la récupération du collagène précipité et l'élimination des molécules encore présente en solution). Centrifuger jusqu'à ce que le surnageant soit limpide. Rassembler et peser les culots.
- Resolubilisation des culots dans CH<sub>3</sub>COOH 0,5 M équilibré à pH 2.5 par HCl 1N. NB : 200 grammes de culots dans 1,8 litres qsp donnent une solution à 2-3 mg/ml.
- 4 dialyses successives contre du CH<sub>3</sub>COOH 0,5 M, 1V dans 10V. (Élimination du sel et de HCl).
- Transférer le contenu des boudins dans une bouteille stérile après la dialyse.
- Centrifugation à 40000-50000 × g pendant 4 h, 15°C. (Élimination des agrégats non dissous)

## G. Détermination de la concentration d'hydroxyproline

### Hydrolyse des échantillons

- Prélever 50 µl de solution à la propipette ; peser précisément le volume mis à hydrolyser dans l'eppendorf (le faire en triple).
- Hydrolyser avec 50 µl d'acide chlorhydrique 6 N. Vortexer, centrifuger et placer les eppendorfs hermétiquement fermés dans une étuve à 106°C pendant la nuit.
- Le lendemain laisser les eppendorfs à l'air jusqu'à ce que les tubes soient secs (40 min à 1h).

### Dosage de l'hydroxyproline

- Reprendre les échantillons secs avec 1 ml d'eau distillée. Vortexer.
- Dosage

Hydroxyproline ou échantillons (µl)	Gamme d'hydroxyproline à partir d'une solution mère à 20µg/ml					Echantillon collagène		
	0	40	80	100	140	20	40	80
H <sub>2</sub> O (µl)	400	360	320	300	260	380	360	320
Dilution de l'échantillon						1/40	1/20	1/10
Chloramine T	200µl à température ambiante pendant 20 min. Vortexer							
Acide perchlorique 3,15M	200µl à température ambiante pendant 5 min. Vortexer							
DMBA	200µl 20 min 60°C. Vortexer							

- Mettre 5 min à 4°C.
- Lecture de la densité optique à 557 nm sur un spectromètre UV-visible.
- Calcul de la concentration (C°) :

$$C^{\circ} \text{ de collagène (mg/ml)} = \frac{C^{\circ} \text{ Hydroxyproline } (\mu\text{g.ml}^{-1}) * 7.6}{m_{\text{échantillon collagène}}}$$

× 7,6 car il y a 13% d'hydroxyproline dans une molécule de collagène.

DOSAGE DE L'HYDROXYPROLINE – RÉACTIFS

Tampon de dosage

- Acide citrique monohydraté (0,25 M) 6,25g
- Acétate de sodium trihydraté (1 M) 17,5g
- Hydroxyde de sodium (0,85 M) 4,25g
- Acide acétique Purex 1,5ml
- Ajouter H<sub>2</sub>O distillé qsp 100ml 125ml

Ajuster le pH à 6 avec de l'acide acétique Purex

Conserver à 4°C

Hydroxyproline standard (20µg/ml)

- solution mère :
- Hydroxy-L-Proline 20mg
- H<sub>2</sub>O distillée 100ml

- solution de dosage :
- solution mère diluée au 1/10

Solution de DMBA

- DMBA (Diméthylamino-4 benzaldéhyde) 1g
- Méthyl Cellosolve 5ml

Acide Perchlorique 3,15M

- Acide Perchlorique 65% (10,39 M) 30ml
  - H<sub>2</sub>O distillée 70ml
- Conserver à 4°C

Solution de Chloramine T

- Chloramine T 400mg
- H<sub>2</sub>O distillée 5ml
- Méthyl Cellosolve (Éther monoéthylique de l'éthylène glycol) 7,5ml
- Tampon de dosage 1,25ml

## H. Préparation des gels d'électrophorèse

### Préparation des gels

Gel de migration 10% 2 gels=20ml

H <sub>2</sub> O distillée	8,1 ml
Tris HCl 1.5M, pH=8.8	5 ml
Acrylamide/Bis 30%	6,6 ml
SDS 10%	200 µl
Persulfate d'ammonium 10%	100 µl
Temed (à ajouter au moment de l'emploi du gel)	10 µl

Gel de concentration 5% 2 gels=10ml

H <sub>2</sub> O distillée	5,7 ml
Tris HCl 0.5M, pH=8.8	2,5 ml
Acrylamide/Bis 30%	1,7 ml
SDS 10%	100 µl
Persulfate d'ammonium 10%	50 µl
Temed (à ajouter au moment de l'emploi du gel)	10 µl

### Préparation des plaques - Montage

- Laver soigneusement à l'alcool une plaque support en verre et une plaque fine en verre par gel. Sécher.
- Installer les plaques sur le support de montage. Vérifier l'étanchéité avec H<sub>2</sub>O distillée.
- Remplir de gel de migration jusqu'à 4-5mm de la ligne verte.
- Ajouter doucement H<sub>2</sub>O distillée jusqu'en haut des plaques.
- Attendre la polymérisation du gel.
- Retourner le montage pour vider l'H<sub>2</sub>O au-dessus du gel.
- Ajouter le gel de concentration jusqu'à 2-3mm du bord supérieur.
- Mettre un peigne 10 puits. Laisser polymériser.

### Préparation de la cuve à électrophorèse

- Placer la cuve dans un plateau rempli de glace.
- Transférer les montages sur le support prévu pour l'électrophorèse, petite plaque vers l'intérieur (impératif). Mettre le support dans la cuve.
- Remplir de Tampon de migration 1X jusqu'au ras des plaques.

### Préparation des échantillons

- Diluer les échantillons à doser au ½ dans le tampon d'échantillon (généralement 10µl d'échantillon + 10µl de tampon).
- Pour les échantillons de collagène : évaporation à sec puis reprise dans le tampon d'échantillon.

### **Dépôt des échantillons**

- Déposer les échantillons dans les puits.
- Réserver un puits par plaque pour le marqueur de Poids Moléculaires (Rainbow<sup>TM</sup>mix, code RPN800E).

### **Migration**

- Recouvrir la cuve de son couvercle en respectant les polarités.
- Relier les électrodes à l'appareil.
- Migration en 2 temps :
  - 1- Régler l'appareil sur 80 volts. Mettre en marche (l'apparition de petites bulles dans le Tampon de migration indique que le courant passe).
    - Laisser 10 à 20mn, jusqu'à ce que les dépôts soient compactés en une seule ligne (les protéines ont alors atteint le front du gel de séparation).
  - 2 - Régler l'appareil sur 160 volts pour la migration proprement dite
    - Arrêter quand toutes les protéines ont complètement migré.

### **Coloration**

- Démouler les gels.
- Mettre chaque gel dans une boîte plastique avec couvercle.
- Ajouter 50 à 100 ml de Bleu de Coomassie.
- Agiter pendant 1 heure – vitesse 40.
- Jeter le Bleu de Coomassie.

### **Décoloration**

#### Au décolorant rapide

- Ajouter 50 à 100ml de décolorant rapide.
- Agiter 1 heure – vitesse 40.
- Jeter le décolorant.

#### Au décolorant de conservation

- Ajouter 50 à 100ml de décolorant de conservation.
- Agiter 1 nuit – vitesse 40.
- Jeter le décolorant.

### **Réhydratation**

- Ajouter 50 à 100ml H<sub>2</sub>O distillée pour réhydrater le gel (qui s'est rétracté car les décolorants contiennent de l'alcool).

ELECTROPHORESE-REACTIFS

Tris-HCl 1,5M pH=8,8

Pour 100ml de solution :

- Tris Amino *18,171g*
- H<sub>2</sub>O distillée *70ml*
- Mesurer le pH
- Ajuster à pH 8,8 avec HCl 35%
- Ajouter H<sub>2</sub>O distillé qsp 100ml

SDS 10%

- SDS *10g*
- H<sub>2</sub>O distillée qsp *100ml*

Tampon d'échantillon

- Tris- HCl 0,5M pH 6,8 *1,25ml*
- β Mercapto Ethanol *0,45ml*
- SDS 10% *2ml*
- Bleu de Bromophénol 0,05% *(p/v)*
- Glycérol *1ml*
- EDTA *3,72mg*
- H<sub>2</sub>O distillée qsp *5ml*

Décolorant rapide

- Ethanol *400ml*
- H<sub>2</sub>O distillée *500ml*
- CH<sub>3</sub>COOH *100ml*

Tris-HCl 1,5M pH=8,8

Pour 100ml de solution :

- Tris Amino *6,057g*
- H<sub>2</sub>O distillée *70ml*
- Mesurer le pH
- Ajuster à pH 8,8 avec HCl 35%
- Ajouter H<sub>2</sub>O distillé qsp 100ml

Persulfate d'ammonium 10%

- Persulfate d'ammonium *500mg*
- H<sub>2</sub>O distillée *5ml*

Tampon de migration 10X

- Tris *30g*
- SDS *10g*
- Glycine *144g*
- H<sub>2</sub>O distillée qsp *1000ml*

Décolorant de conservation

- Ethanol *100ml*
- H<sub>2</sub>O distillée *825ml*
- CH<sub>3</sub>COOH *75ml*

Acrylamide Bis 30%, Temed et Bleu de Coomassie

- Solution prête à l'emploi

## Références du Chapitre 2

- 1 Glover, K. J. *et al.* Structural Evaluation of Phospholipid Bicelles for Solution-State Studies of Membrane-Associated Biomolecules. *Biophys J.* **81**, 2163-2171 (2001).
- 2 Fischer, H., Polikarpov, I. & Craievich, A. F. Average protein density is a molecular-weight-dependent function. *Protein science : a publication of the Protein Society* **13**, 2825-2828, doi:10.1110/ps.04688204 (2004).
- 3 Chou, J. J., Baber, J. L. & Bax, A. Characterization of phospholipid mixedmicelles by translational diffusion. *J Biomol NMR* **29**, 299-308 (2004).
- 4 Wang, L., Guo, Y., Li, P. & Song, Y. Anion-specific effects on the assembly of collagen layers mediated by magnesium ion on mica surface. *The journal of physical chemistry. B* **118**, 511-518, doi:10.1021/jp405035x (2014).
- 5 Zanaboni, G., Rossi, A., Tina Onana, M. T. & Tenni, R. Stability and networks of hydrogen bonds of the collagen triple helical structure: influence of pH and chaotropic nature of three anions. *Matrix Biol* **19**, 511-520 (2000).
- 6 Kar, K. *et al.* Self-association of collagen triple helix peptides into higher order structures. *The Journal of biological chemistry* **281**, 33283-33290, doi:10.1074/jbc.M605747200 (2006).
- 7 Schmitt, F. O., Hall, C. E. & Jakus, M. A. Electron microscope investigations of the structure of collagen. *J Cell Comp Physiol*, 11-33 (1942).
- 8 Morimoto, K. *et al.* Characterization of type I collagen fibril formation using thioflavin T fluorescent dye. *Journal of biochemistry* **145**, 677-684, doi:10.1093/jb/mvp025 (2009).
- 9 Kadler, K. E., Holmes, D. F., Trotter, J. A. & Chapman, J. A. Collagen fibril formation. *Biochemistry* **316**, 1-11 (1996).
- 10 Ruben, G. C. Vertical Pt-C replication for TEM a revolution in imaging non-periodic macromolecules. *Micron* **29**, 359-396 (1998).
- 11 Kerfah, R., Plevin, M. J., Sounier, R., Gans, P. & Boisbouvier, J. Methyl-specific isotopic labeling: a molecular tool box for solution NMR studies of large proteins. *Curr Opin Struct Biol* **32**, 113-122, doi:10.1016/j.sbi.2015.03.009 (2015).
- 12 Meyer, B. & Peters, T. NMR Spectroscopy Techniques for Screening and Identifying Ligand Binding to Protein Receptors. *Angewandte Chemie* **42**, 864-890 (2003).
- 13 Angulo, J. & Nieto, P. M. STD-NMR: application to transient interactions between biomolecules-a quantitative approach. *Eur Biophys J* **40**, 1357-1369, doi:10.1007/s00249-011-0749-5 (2011).
- 14 Fischer, G. Peptidyl-Prolyl &/trans Isomerases and Their Effectors. *Angewandte Chemie* **33**, 1415-1436 (1994).
- 15 Kotch, F. W. & Raines, R. T. Self-assembly of synthetic collagen triple helices. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 3028-3033, doi:10.1073/pnas.0508783103 (2006).
- 16 Freund, I. & Deutsch, M. Second-harmonic microscopy of biological tissue. *Optics Letters* **11**, 94-96 (1986).
- 17 Fine, S. & Hansen, W. P. Optical second harmonic generation in biological systems. *Applied Optics* **10**, 2350-2353 (1971).
- 18 Strupler, M. *et al.* Second harmonic imaging and scoring of collagen in fibrotic tissues. *Optics Express* **15**, 4052-4063, doi:DOI: 10.1364/OE.15.004054. (2007).



- 19 Kottmann, R. M. *et al.* Second harmonic generation microscopy reveals altered collagen microstructure in usual interstitial pneumonia versus healthy lung. *Respiratory Research* **16**, doi:10.1186/s12931-015-0220-8 (2015).
- 20 Tiaho, F., Recher, G. & Rouede, D. Estimation of helical angles of myosin and collagen by second harmonic generation imaging microscopy. *Optics Express* **15**, 12286-12295, doi:DOI: 10.1364/OE.15.012286. (2007).
- 21 Andrew, E. R., Bradbury, A. & Eades, R. G. Nuclear Magnetic Resonance Spectra from a Crystal rotated at High Speed. *Nature* **182**, 1659-1659 (1958).
- 22 De Sa Peixoto, P., Laurent, G., Azais, T. & Mosser, G. Solid-state NMR study reveals collagen I structural modifications of amino acid side chains upon fibrillogenesis. *The Journal of biological chemistry* **288**, 7528-7535, doi:10.1074/jbc.M112.390146 (2013).
- 23 Wu, S.-T. Birefringence dispersions of liquid crystals. *Physical Review A* **33**, 1270-1274, doi:10.1103/PhysRevA.33.1270 (1986).
- 24 Jackson, D. S. & Fessler, J. H. Isolation and Properties of a Collagen Soluble in Salt Solution at Neutral pH. *Nature* **176**, 69-70 (1955).
- 25 Chandrakasan, G., Torchia, D. A. & Piez, K. A. Preparation of Intact Monomeric Collagen from Rat Tail Tendon and Skin and the Structure of the Non-helical Ends in Solutions". *J Bio Chem* **251**, 6062-6067 (1976).
- 26 Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).
- 27 Giraud-Guille, M. M. Liquid crystalline phases of sonicated type I collagen. *Biol Cell* **67**, 97-101 (1989).
- 28 Giraud-Guille, M. M. Liquid crystallinity in condensed type I collagen solutions. A clue to the packing of collagen in extracellular matrices. *Journal of molecular biology* **224**, 861-873 (1992).
- 29 Besseau, L. & Giraud-Guille, M. M. Stabilization of Fluid Cholesteric Phases of Collagen to Ordered Gelated Matrices. *Journal of molecular biology* **251**, 197-202 (1995).
- 30 Mosser, G., Anglo, A., Helary, C., Bouligand, Y. & Giraud-Guille, M. M. Dense tissue-like collagen matrices formed in cell-free conditions. *Matrix Biol* **25**, 3-13, doi:10.1016/j.matbio.2005.09.002 (2006).
- 31 Sousa Neto, V. O. *et al.* Study of the electrical conductivity and piezoelectricity in iron doped collagen film. *Journal of Materials Science* **4**, 43-51 (2002).
- 32 Desimone, M. F. *et al.* Silica-collagen bionanocomposites as three-dimensional scaffolds for fibroblast immobilization. *Acta Biomater* **6**, 3998-4004, doi:10.1016/j.actbio.2010.05.014 (2010).
- 33 Nassif, N. *et al.* Self-Assembled Collagen–Apatite Matrix with Bone-like Hierarchy. *Chemistry of Materials* **22**, 3307-3309, doi:10.1021/cm903594n (2010).
- 34 Gao, J., Dennis, J. E., Solchaga, L. A., Goldberg, V. M. & Caplan, A. I. Repair of osteochondral defect with tissue-engineered two-phase composite material of injectable calcium phosphate and hyaluronan sponge. *Tissue Engineering* **8**, 827-837 (2002).
- 35 Liu, X., Yue, Z., Higgins, M. J. & Wallace, G. G. Conducting polymers with immobilised fibrillar collagen for enhanced neural interfacing. *Biomaterials* **32**, 7309-7317, doi:10.1016/j.biomaterials.2011.06.047 (2011).
- 36 Prager, M. R. *et al.* Collagen- versus gelatine-coated Dacron versus stretch PTFE bifurcation grafts for aortoiliac occlusive disease: Long-term results of a prospective, randomized multicenter trial. *Surgery* **134**, 80-85 (2003).

- 37 Gasiowski, J. Z. & Collier, J. H. Directed intermixing in multicomponent self-assembling biomaterials. *Biomacromolecules* **12**, 3549-3558, doi:10.1021/bm200763y (2011).
- 38 Wang, A. Y., Mo, X., Chen, C. S. & Yu, S. M. Facile Modification of Collagen Directed by Collagen Mimetic Peptides. *Journal of the American Chemical Society* **127**, 4130-4131 (2005).

### ANNEXE 3 : Estimation du nombre de peptide par nano-objet

Les études de DLS révèlent la présence de deux populations :  $R_h = 2.5 \text{ nm}$  et  $R_h = 700 - 1200 \text{ nm}$

La densité du peptide  $\sim 1.5 \text{ g/cm}^3$  et la densité des assemblages :  $\sim 1.41 \text{ g/cm}^3$  (d'après la publication de *Craievich et al.*<sup>2)</sup>)

Le volume moléculaire (POG)<sub>7</sub> (M=1888g/mol)

$$V_{\text{mole}} = 1888/1.5 = 1258 \text{ cm}^3/\text{mol}$$

$$V_{(\text{POG})_7} = 1258 \cdot 10^{-6} \text{ m}^3 / 6.022 \cdot 10^{23} = 2.09 \cdot 10^{-27} \text{ m}^3 = 2.09 \text{ nm}^3$$

$$\rightarrow R_h \text{ théorique : } V = 4/3 \pi R^3$$

$$R_v = 0,8 \text{ nm}$$

$$\rightarrow R_h \text{ expérimental : } R_{h(\text{POG})_7} = 2,5 \text{ \AA}$$

Les données expérimentales sont en accord avec les calculs théoriques. Nous pouvons alors estimer un nombre moyen de peptide contenu dans un nano-objet:

$$R_h = 700 - 1200 \text{ nm pour les nano-objets}$$

$$\text{Soit un rapport de volume : } V / V_{(\text{POG})_7} = R^3 / R_{(\text{POG})_7}^3$$

$$R_h = 700 \rightarrow V / V_{(\text{POG})_7} = (700/2.5)^3 = 20 \cdot 10^6$$

$$R_h = 1200 \rightarrow V / V_{(\text{POG})_7} = (1200/2.5)^3 = 110 \cdot 10^6$$

Nous estimons un assemblage moyen de  $\sim 10^8$  peptides par nano-objet.

ANNEXE 4 : Exemple du code Mathematica utilisé pour ajuster les courbes expérimentales des 3 états (monomériques, triples hélices et nano-objets) aux quatre températures.

## Système d'équations différentielles pour $k_3$ & $k_4 = 0$

# DSolve résoud symboliquement le système d'équations différentielles

```
solution = DSolve[
  {M'[t] == -(k1 + k6) M[t] + k2 T[t] + k5 A[t],
   T'[t] == k1 M[t] - (k2 + k3) T[t] + k4 A[t],
   A'[t] == k6 M[t] + k3 T[t] - (k4 + k5) A[t],
   M[0] == 1,
   T[0] == 0,
   A[0] == 0},
  {M, T, A}, t
] /. {k4 -> 0, k3 -> 0};
```

# Les solutions apparaissent dans un vecteur dans l'ordre suivant : 1 pour les agrégats, 2 pour le monomère, 3 pour la triple hélice

### Fitting of the kinetics data

# Lecture des données expérimentales dans la matrice dat dans l'ordre monomère (ou PPII), triple hélice (TH), agrégats (AGG); Dimensions affiche la dimensions de la matrice dat

```
dat = Table[Join[PPII[[1]] /. {t_, y_} -> {1, t, y}, TH[[1]] /. {t_, y_} -> {2, t, y},
  AGG[[1]] /. {t_, y_} -> {3, t, y}], {1, 1, 4}];
Dimensions[dat]
{4, 1536, 3}
```

# Définition des solutions

```
solAGG[k1_, k2_, k5_, k6_, t_] := solution[[1]][[1]][[2]][[2]];
solPPII[k1_, k2_, k5_, k6_, t_] := solution[[1]][[2]][[2]][[2]];
solTH[k1_, k2_, k5_, k6_, t_] := solution[[1]][[3]][[2]][[2]];
```

```
# Définition du modèle de fit
fitmodel[set_, k1_, k2_, k5_, k6_, t_] := Which[set == 1, Evaluate@solPPII[k1, k2, k5, k6, t], set == 2, Evaluate@solTH[k1, k2, k5, k6, t],
set == 3, Evaluate@solAGG[k1, k2, k5, k6, t]]

# Le fit avec le protocole NonlinearModelFit; Timing donne la durée du calcul
Timing[
fit = Table[NonlinearModelFit[dat[[1]], {fitmodel[set, k1, k2, k5, k6, t]}, {k1 > 0, k2 > 0, k5 > 0, k6 > 0}], {k1, k2, k5, k6}, {set, t},
Method -> {NMinimize, Method -> "NelderMead"}, MaxIterations -> 50], {1, 1, 4}]
fitparams = Table[fit[[1]]["BestFitParameters"], {1, 1, 4}]

{1172.25, {FittedModel[ Which[set == 1,  $\frac{\llcorner 1 \gg}{\llcorner 1 \gg}$ ,  $\llcorner 2 \gg$ ,  $\llcorner 1 \gg$ ,  $-\frac{2(-2.0.780112^2 \cdot 0.377472 \times 0.238835 + \llcorner 56 \gg)}{((\llcorner 1 \gg)^2 - 4(\llcorner 1 \gg) + \llcorner 19 \gg \sqrt{\llcorner 1 \gg} + \llcorner 1 \gg + \llcorner 19 \gg \sqrt{\llcorner 1 \gg} + 0.238835 \sqrt{\llcorner 19 \gg^2 + \llcorner 11 \gg + \llcorner 20 \gg^2}) \llcorner 1 \gg}$  ]},
FittedModel[ Which[set == 1,  $\frac{\llcorner 1 \gg}{\llcorner 1 \gg}$ ,  $\llcorner 3 \gg$ ,  $-\frac{2(-2.0.311307^2 \cdot 0.310809 \times 0.0625201 + \llcorner 56 \gg)}{((\llcorner 1 \gg)^2 - 4(\llcorner 1 \gg) + \llcorner 19 \gg \sqrt{\llcorner 1 \gg} + \llcorner 1 \gg + \llcorner 20 \gg \sqrt{\llcorner 1 \gg} + 0.0625201 \sqrt{\llcorner 19 \gg^2 + \llcorner 11 \gg + \llcorner 20 \gg^2}) \llcorner 1 \gg}$  ]},
FittedModel[ Which[set == 1,  $\llcorner 4 \gg$ ,  $-\frac{2(-2.0.254037^2 \cdot 0.471008 \times 0.0688487 + \llcorner 56 \gg)}{((\llcorner 1 \gg)^2 - 4(\llcorner 1 \gg) + \llcorner 19 \gg \sqrt{\llcorner 1 \gg} + \llcorner 1 \gg + \llcorner 19 \gg \sqrt{\llcorner 1 \gg} + 0.0688487 \sqrt{\llcorner 19 \gg^2 + \llcorner 11 \gg + \llcorner 19 \gg^2}) \llcorner 1 \gg}$  ]},
FittedModel[ Which[set == 1,  $\llcorner 1 \gg$ , set == 2,  $-\llcorner 1 \gg$ , set == 3,  $-\frac{2(-2.0.128087^2 \cdot 1.31244 \times 0.0727894 + \llcorner 56 \gg)}{(\llcorner 6 \gg + 0.0727894 \sqrt{\llcorner 20 \gg^2 + \llcorner 11 \gg + \llcorner 18 \gg^2}) (\llcorner 19 \gg + 0.0727894 \sqrt{\llcorner 1 \gg})}$  ]} ]}],
{ {k1 -> 0.780112, k2 -> 0.377472, k5 -> 0.570893, k6 -> 0.238835}, {k1 -> 0.311307, k2 -> 0.310809, k5 -> 0.276045, k6 -> 0.0625201},
{k1 -> 0.254037, k2 -> 0.471008, k5 -> 0.556371, k6 -> 0.0688487}, {k1 -> 0.128087, k2 -> 1.31244, k5 -> 1.47293, k6 -> 0.0727894}}
```

# Relecture de données expérimentales dans la matrice alldata avec le même ordre PPII, TH, AGG

```
alldata = Table[{PPII[[1]], TH[[1]], AGG[[1]]}, {1, 1, 4}];
Dimensions[alldata]
{4, 3, 512, 2}
```

# Définition du modèle de modélisation incluant les solutions du système d'équations différentielles; les paramètres obtenus par le fit (fitparams) seront ensuite utilisés pour

```
mm = {solPPII[k1, k2, k5, k6, t], solTH[k1, k2, k5, k6, t], solAGG[k1, k2, k5, k6, t]};
```

# Affichage des données expérimentales et des résultats du fit

```
Table[Show[ListPlot[alldata[[1]], PlotRange -> All, Frame -> {True},
FrameTicks -> {{Automatic, None}, {Automatic, None}}, Joined -> False,
FrameLabel -> {"Temps (h)", "%"}, LabelStyle -> Directive[Black(*, Bold*), 25],
AxesOrigin -> {0, 1}, PlotLegends -> Placed[{"PPII", "TH", "AGG"}, {0.4, 1}],
PlotMarkers -> Automatic, ImageSize -> {800, 600}],
Plot[Evaluate[mm /. fitparams[[1]]], {t, 0.002, 13}], {1, 1, 4}]
```





# Chapitre 3

## Vers de nouveaux peptidomimétiques du collagène

La *peptidomimétique* est l'art de concevoir des variations structurales sur un peptide biologiquement actif afin d'en améliorer ses propriétés pharmacologiques.

Dans ce chapitre, nous expliquerons pourquoi nous avons choisi de synthétiser des peptides incorporant des groupements fluorés pour mimer le collagène. Nous introduirons alors les premières études réalisées sur des analogues de prolines fluorées. Nous décrirons la synthèse de quatre peptides de type  $\text{CH}_3\text{CO}-(\text{Xaa}-\text{Yaa}-\text{Gly})-\text{NH}_2$  incorporant ces analogues de prolines fluorées ainsi que leur caractérisation par RMN. L'étude structurale comparative de ces quatre courts peptides, nous fournira des éléments rationnels pour prédire leur intérêt pour la synthèse de CMP (POG)<sub>7</sub>. Nous étudierons leur influence sur les premiers peptides de ce type que nous avons synthétisés.



### 3.1. Contexte

#### 3.1.1 La proline et ses analogues

*Généralités* - La proline est le seul acide aminé naturel possédant une structure cyclique, le cycle pyrrolidine. Ce cycle à 5 chaînons peut prendre, principalement, deux conformations selon la position du carbone  $\gamma$ , on parle alors de plissement de cycle. On le définit par la valeur de l'angle  $\chi_1$ , celui-ci est dit  $C^\gamma$ -endo (*down puckering*) quand  $\chi_1 \approx 25^\circ$  et  $C^\gamma$ -exo (*up puckering*) quand  $\chi_1 \approx -25^\circ$  (Figure 1). Dans un contexte peptidique, et en absence de structure secondaire privilégiée, le résidu proline adopte préférentiellement (~60%) le plissement de cycle  $C^\gamma$ -endo.<sup>1</sup>

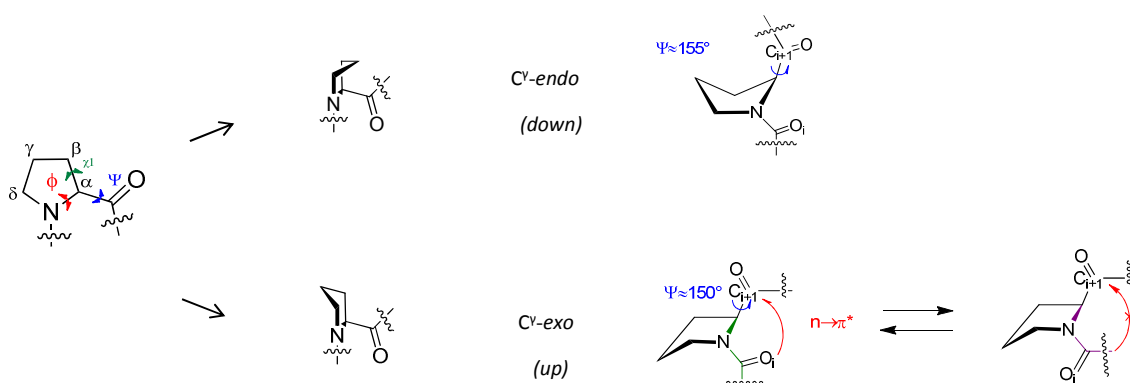


Figure 1: Illustration du plissement de cycle de la proline

Le cycle pyrrolidine induit par ailleurs une rigidité au squelette peptidique qui se traduit par une valeur de l'angle dièdre  $\Phi$  fixée entre  $-50^\circ$  et  $-90^\circ$ . L'angle  $\psi$ , quant à lui, peut prendre trois valeurs distinctes qui correspondent aux trois rotamères stables autour de la liaison  $C^\alpha$ - $C'$ . La première, comprise entre  $120^\circ$  et  $180^\circ$ , permet au squelette peptidique d'adopter une conformation étendue (PPII pour les angles  $\Phi$  contraints autour de  $-60^\circ$ ). La seconde (entre  $-50^\circ$  et  $0^\circ$ ) correspond à la structuration en hélice  $\alpha$ . Enfin, la troisième région (de  $40^\circ$  à  $90^\circ$ ), plus rarement observée pour les résidus prolines, définit la conformation en coude  $\gamma$  inverse, stabilisée par une liaison hydrogène entre le carbonyle  $CO_{(i)}$  du résidu précédent et le proton amide du résidu suivant  $NH_{(i+2)}$ .<sup>2</sup>

La liaison peptidique peut être de conformation *cis* ou *trans*. La présence de l'amide tertiaire du cycle pyrrolidine va influencer la conformation de la liaison peptidique qui le précède. Ainsi, dans les protéines, 5,8% des liaisons peptidiques Xaa-Yaa sont *cis* lorsque Yaa=proline, contre 0,06% pour les autres acides aminés.<sup>3</sup> De manière plus précise, il a été

trouvé des différences significatives dans le ratio *cis:trans* selon la nature de l'acide aminé Xaa.<sup>4,5</sup> De nombreuses études ont par ailleurs montré la corrélation entre la conformation de la liaison peptidique et le plissement de cycle pyrrolidine, C<sup>γ</sup>-*endo* favorise le conformère *cis* alors que le C<sup>γ</sup>-*exo* stabilise le conformère *trans*.<sup>5-7</sup> Cette observation est due à une forte interaction non-covalente *n*→*π*\* entre l'oxygène de CO<sub>i</sub> et le carbone de CO<sub>i+1</sub> (Figure 1).<sup>8,9</sup> Maccalum *et al.* ont prouvé que 89% des *cis*-proline adoptaient un plissement de cycle *endo* alors que les *trans*-prolines n'en favorisaient aucun. Par ailleurs, de nombreux analogues de proline ont été synthétisés en modifiant le groupement en C<sup>γ</sup> en vue d'influencer le plissement de cycle.<sup>10</sup>

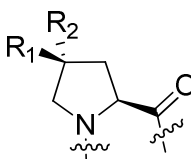
*Analogues de la proline* – Raines *et al.* ont démontré que des effets stéréoelectroniques peuvent être créés par substitution sur le cycle pyrrolidine lui-même, permettant ainsi de contrôler le plissement de cycle.<sup>11,12</sup> Des analogues de proline ont été synthétisés avec en position <sup>4</sup>C<sup>γ</sup>, un groupement –OH,<sup>12</sup> –Me,<sup>13</sup> –SH<sup>14</sup> ou encore un atome de –F,<sup>15</sup> –Cl.<sup>6</sup> Il est alors trouvé que le plissement de cycle C<sup>γ</sup>-*exo* est favorisé par effet stéréoelectronique lorsque le <sup>4</sup>C<sup>γ</sup> est substitué par un (4*R*)-OH, (4*R*)-F ou (4*R*)-Cl et est stabilisé par effet stérique quand il est substitué par (4*R*)-Me ou (4*R*)-SH. La présence de groupement favorisant le plissement de cycle C<sup>γ</sup>-*exo* va de ce fait favoriser l'interaction de Bürgi-Dunitz et donc stabiliser le conformère *trans*.<sup>16</sup> Lorsque tous les groupements ajoutés sont de stéréochimie 4*S*, le plissement de cycle C<sup>γ</sup>-*endo* est alors stabilisé par effet gauche. Wennemers *et al.* ont par la suite apporté des précisions en montrant que la présence de groupements donneurs de liaisons hydrogènes en position <sup>γ</sup> sur un cycle C<sup>γ</sup>-*endo* favorisait le conformère *trans*.<sup>5,17</sup> Les substitutions en positions <sup>3</sup>C<sup>β</sup> et <sup>5</sup>C<sup>δ</sup> ont pour leur part une influence directe sur l'isomérisation *cis-trans* de la liaison Xaa-Pro. Celle-ci est très marquée pour des alkylations en <sup>5</sup>C<sup>δ</sup> car l'encombrement des substituants déstabilise fortement la liaison peptidique *trans* en raison de gêne stérique avec le résidu précédent.<sup>9,18-21</sup>

*Inducteur de structure secondaire particulière* – Le résidu proline est rarement rencontré dans les structures de type hélice  $\alpha$  et feuillet  $\beta$  en raison de la présence d'une amine secondaire sur le cycle pyrrolidine et donc de l'impossibilité de former des liaisons hydrogènes stabilisatrices de ces structures.<sup>22</sup> Les prolines sont en revanche souvent retrouvées en amorce des hélices  $\alpha$  car elles possèdent l'angle  $\phi$  requis de  $-60^\circ$ , maintenu par le cycle pyrrolidine. Les peptides riches en proline adoptent généralement une structure dite « polyproline II ». Celle-ci consiste en une hélice gauche ayant des valeurs d'angles dièdres  $\Phi$  et  $\Psi$  égaux à  $-75^\circ$  et  $150^\circ$ , respectivement, pour tous les résidus du squelette peptidique, ainsi que l'ensemble des

liaisons peptidiques en *trans*.<sup>23</sup> Contrairement à l'hélice  $\alpha$ , l'hélice PPII ne possède aucune liaison hydrogène et est plus étendue : 3.1Å par résidu contre 1.5Å pour l'hélice  $\alpha$ . On rencontre également en solvant organique, des motifs « polyproline I » qui sont des hélices droites définies par des angles dièdres  $\Phi$  et  $\Psi$  de  $-75^\circ$  et  $-160^\circ$ , respectivement, ainsi que l'ensemble des liaisons peptidiques en *cis*.

*Mimes de triple hélice de collagène* – A partir de ces observations, de nombreux CMP ont été synthétisés. Le Tableau 1, issu d'une revue de *Raines et al.*, recense des CMP substitués en position  $^4C\gamma$ -Pro ainsi que la température de fusion  $T_m$  mesurée pour les triples hélices formées par des peptides (Xaa-Yaa-Gly)<sub>7</sub> ou (Xaa-Yaa-Gly)<sub>10</sub>. Ces valeurs de  $T_m$  permettent une comparaison directe de leur stabilité.<sup>10</sup>

(Xaa-Yaa-Gly) <sub>10</sub>	T <sub>m</sub> (°C)
(Pro-Pro-Gly) <sub>10</sub>	31-41
(Pro-Hyp-Gly) <sub>10</sub>	61-69
(Hyp-Pro-Gly) <sub>10</sub>	Pas d'hélice
(Hyp-Hyp-Gly) <sub>10</sub>	65
(Pro-hyp-Gly) <sub>10</sub>	Pas d'hélice
(hyp-Pro-Gly) <sub>10</sub>	Pas d'hélice
(Pro-Flp-Gly) <sub>10</sub>	91
(Flp-Pro-Gly) <sub>10</sub>	Pas d'hélice
(flp-Pro-Gly) <sub>10</sub>	58
(flp-Flp-Gly) <sub>10</sub>	30
(Pro-Clp-Gly) <sub>10</sub>	52
(Clp-Pro-Gly) <sub>10</sub>	Pas d'hélice
(clp-Pro-Gly) <sub>10</sub>	33
(clp-Clp-Gly) <sub>10</sub>	Pas d'hélice
(Pro-Mop-Gly) <sub>10</sub>	70



Flp : R<sub>1</sub>=H, R<sub>2</sub>=F  
flp : R<sub>1</sub>=F, R<sub>2</sub>=H

Hyp : R<sub>1</sub>=H, R<sub>2</sub>=OH  
hyp : R<sub>1</sub>=OH, R<sub>2</sub>=H

Mep : R<sub>1</sub>=H, R<sub>2</sub>=CH<sub>3</sub>  
mep : R<sub>1</sub>=CH<sub>3</sub>, R<sub>2</sub>=H

Clp : R<sub>1</sub>=H, R<sub>2</sub>=Cl  
clp : R<sub>1</sub>=Cl, R<sub>2</sub>=H

Mop : R<sub>1</sub>=H, R<sub>2</sub>=O-CH<sub>3</sub>

(Xaa-Yaa-Gly) <sub>7</sub>	T <sub>m</sub> (°C)
(Pro-Pro-Gly) <sub>7</sub>	Pas d'hélice
(Pro-Hyp-Gly) <sub>7</sub>	36
(Pro-Flp-Gly) <sub>7</sub>	45
(Flp-Pro-Gly) <sub>7</sub>	Pas d'hélice
(Pro-flp-Gly) <sub>7</sub>	Pas d'hélice
(flp-Pro-Gly) <sub>7</sub>	33
(flp-Flp-Gly) <sub>7</sub>	Pas d'hélice
(Pro-Clp-Gly) <sub>7</sub>	23
(clp-Pro-Gly) <sub>7</sub>	Pas d'hélice
(Pro-Mep-Gly) <sub>7</sub>	29
(mep-Pro-Gly) <sub>7</sub>	13
(mep-Mep-Gly) <sub>7</sub>	36

Tableau 1: Valeurs des températures de fusion des triples hélices de CMP<sup>10</sup>

Les peptides (Pro-Hyp-Gly)<sub>n</sub> représentent les CMP témoins, étant constitués des triplets les plus représentés dans la séquence du collagène sauvage. Des  $T_m$  de 36°C et ~65°C sont trouvées pour (Pro-Hyp-Gly)<sub>7</sub> et (Pro-Hyp-Gly)<sub>10</sub>, respectivement. La comparaison avec les peptides (Pro-Pro-Gly)<sub>10</sub> montrent que l'hydroxylation du résidu proline en position Yaa a un impact décisif sur la stabilité de la triple hélice ce qui n'est pas le cas en position Xaa. Par

ailleurs, il est important de noter que la stéréochimie de la 4-hydroxyproline est importante, la (2*S*,4*R*)-4-hydroxyproline en position Yaa augmente considérablement la stabilité de la triple hélice, ce qui n'est pas le cas de son stéréoisomère. L'influence du groupe hydroxyle sur la stabilité de la triple hélice a été l'objet de nombreux débats. Une hypothèse a mis en avant leur rôle prépondérant dans l'établissement d'un réseau de liaisons hydrogène les reliant aux carbonyles du squelette.<sup>24</sup> De plus, la première structure cristalline obtenue sur un peptide mime de collagène<sup>25</sup> a montré la présence d'une sphère d'hydratation autour de la triple hélice maintenue par des liaisons H entre les groupements -OH et l'eau. Une autre hypothèse explique l'importance des groupes OH sur le carbone <sup>4</sup>C<sub>γ</sub> en raison de la forte électronégativité de l'oxygène. Des effets stéréoélectroniques permettraient de pré-organiser la chaîne peptidique de manière propice à la formation de triple hélice. Afin de tester cette hypothèse, l'hydroxyproline a été remplacée par une fluoroproline (Flp ou flp), les atomes de fluor n'étant pas de bon accepteur de liaisons H.<sup>26</sup> Il s'est avéré que la triple hélice était plus stable en présence de l'atome de F qu'en présence du groupement -OH, la température de fusion de (Pro-Flp-Gly)<sub>10</sub> étant supérieure de 25°C à celle de (Pro-Hyp-Gly)<sub>10</sub> (cf. Tableau 1). La triple hélice semble ainsi être stabilisée par des effets stéréoélectroniques, même en absence de liaisons H. On note à nouveau que la stéréochimie de la substitution est cruciale pour obtenir cette stabilisation, le peptide (Pro-flp-Gly)<sub>7</sub> ne formant pas d'hélice. Pour conforter cette hypothèse, la proline a également été substituée par une chloroproline (Clp ou clp). La T<sub>m</sub> du peptide (Pro-Clp-Gly)<sub>10</sub>, bien qu'inférieure à la T<sub>m</sub> de (Pro-Hyp-Gly), est tout de même supérieure à celle de (Pro-Pro-Gly)<sub>10</sub>. Ces observations sont vraies pour les peptides comportant 10 répétitions mais également pour n=7, ce qui conforte l'idée que la substitution en position 4 de la proline influence la stabilité de la triple hélice par des effets stéréoélectroniques et non par la possibilité de faire des liaisons H avec l'eau. Enfin, pour mettre en évidence l'effet uniquement lié à l'encombrement stérique, la proline a été substituée par une méthylproline (Mep ou mep). Le méthyle, qui n'a pas de contribution électronique, a un effet significatif sur la stabilité des triples hélices. Cet effet apparaît stéréospécifique, puisque (Pro-Mep-Gly)<sub>7</sub> présente une T<sub>m</sub> de 29°C alors que ni (Pro-Pro-Gly)<sub>7</sub> ni (Pro-mep-Gly)<sub>7</sub> ne forment de triple hélice. L'étude des préférences conformationnelles de ces cycles pyrrolidines substitués fournit un élément important dans la compréhension de la stabilité des triples hélices. Il a en effet été trouvé que, dans l'analogie Mep de la proline, le méthyle stabilisait la conformation *up* du cycle pyrrolidine tandis que le mep stabilisait un plissement *down*.<sup>9</sup> Ainsi, l'alternance de conformation aux positions successives Xaa-Yaa semble être cruciale pour obtenir une triple hélice stable. L'acide aminé

en Xaa doit adopter un plissement *down* et celui en Yaa un plissement *up*. Les stabilités obtenues avec les dérivés halogénés obéissent aux mêmes règles, les atomes électronégatifs (Cl, O, F) jouant le rôle de stabilisateurs de conformation, par effet gauche, en fonction de la configuration du  $^4C\gamma$ . Ainsi, la conformation de la (2*S*,4*R*)-4-hydroxyproline (Hyp) est très fortement stabilisée en plissement *up* et celle de la (2*S*,4*S*)-4-hydroxyproline (hyp) en plissement *down*. L'échelle de la stabilité de ces plissements, et par extension, de la triple hélice est directement liée à l'électronégativité des éléments. On note toutefois que le peptide (flp-Flp-Gly)<sub>10</sub> possède une T<sub>m</sub> relativement basse (30°C) par rapport aux critères évoqués et à celle obtenues pour (flp-Pro-Gly)<sub>10</sub> (58°C) et (Pro-Flp-Gly)<sub>10</sub> (91°C). Cela pourrait être dû à une gêne stérique entre les deux cycles pyrrolidine successifs substitués.<sup>9</sup>

La connaissance de ces préférences permet de prévoir et de concevoir des peptidomimétiques de la triple hélice de collagène. A la lumière des données publiées, le fluor apparaît comme un atome de choix pour augmenter sa stabilité.

### 3.1.2 Pourquoi le Fluor ?

*Un peu d'histoire* – Le fluor a été isolé en 1886 par Henri Moissan, ce qui lui valut le Prix Nobel en 1906,<sup>\*</sup> ce fut une étape importante dans l'histoire de la chimie car cette expérience intervint après de nombreux essais infructueux d'électrolyse de l'acide fluorhydrique. Il existe très peu de produits naturels fluorés, ce qui explique pourquoi il était peu connu.<sup>27</sup> Au début du XX<sup>e</sup> siècle, le fluor intervient surtout en chimie inorganique et c'est seulement à partir des années 1930 que sont découverts les chlorofluorocarbures (CFC) utilisés comme fluides réfrigérants (aujourd'hui bannis à cause de leur contribution probable au trou de la couche d'ozone) et le Téflon. C'est en 1941 que le fluor connaît un essor important avec le projet Manhattan pour le lancement de la bombe nucléaire, il est alors nécessaire de produire du HF et du F<sub>2</sub> pour faire du UF<sub>6</sub> (utilisé pour séparer les différents isotopes de l'uranium par diffusion gazeuse). Et ce n'est qu'au début des années 50 que la chimie organique du fluor est née. Très vite, des découvertes décisives ont permis aux composés fluorés de devenir des médicaments majeurs : les anesthésiques généraux fluorés, les propriétés antitumorales du fluorouracile, et surtout les extraordinaires effets biologiques résultant de l'introduction du fluor dans les corticostéroïdes, découverts par J. Fried. En 2000, 18% du marché pharmaceutique sont des composés fluorés.<sup>28</sup> Il est également très présent dans les domaines

<sup>\*</sup> [http://www.nobelprize.org/nobel\\_prizes/chemistry/laureates/1906/](http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1906/)

des matériaux biocompatibles, phytosanitaires notamment en agrochimie (50% des parts du marché).

*Propriétés structurales* – L'atome de fluor est très souvent comparé à l'atome d'hydrogène du fait de son faible rayon de Van der Waals (1.35Å).<sup>29</sup> Sa petite taille, son importante électronégativité, un excellent recouvrement de ses orbitales 2s et 2p avec les orbitales correspondantes du carbone, la présence de trois paires non liantes, font qu'un atome de fluor porté par un carbone est toujours, sur le plan inductif, un substituant électroattracteur, et sur le plan mésomère, un substituant électrodonneur. Les liaisons sont fortement polarisées du carbone sp<sup>3</sup> ( $\delta^+$ ) vers le fluor ( $\delta^-$ ), ce qui confère à la liaison C-F une énergie plus importante qu'une liaison entre un carbone et un autre halogène.

A de rares exceptions près, les composés fluorés ont un point d'ébullition plus bas que leurs analogues non fluorés. Ils sont également accepteur de liaisons hydrogènes. La présence d'un fluor va augmenter l'acidité des groupements vicinaux et diminuer la basicité des amines vicinales, ce qui va permettre d'améliorer la biodisponibilité de certains composés. Dans le cadre des peptides, il a un effet sur la stabilité métabolique notamment vis-à-vis de l'hydrolyse.

C'est pour ces différentes propriétés que le fluor est largement employé en chimie médicinale. Il permet à la fois d'améliorer le profil métabolique des composés bioactifs et de créer des interactions particulières avec leur(s) cible(s) endogène(s). Dans le cadre de la conception de nouveaux mimes de collagène, nous avons choisi d'incorporer des groupes trifluorométhylés CF<sub>3</sub> qui induisent également des propriétés particulières que nous allons développer dans le paragraphe suivant.

### 3.1.3 Les pseudo-prolines trifluorométhylés

Les pseudoprolines ( $\Psi$ -Pro, Figure 2) ont été très utilisées en synthèse peptidique, notamment pour la synthèse de peptide hydrophobes car elles permettent d'inhiber les phénomènes d'auto-association/d'agrégation.<sup>30</sup> Suite à ces observations, des études par RMN ont permis de mettre en évidence que la présence de pseudoprolines abaissait de 1-2 kcal/mol la barrière d'énergie de l'isomérisation *cis-trans* comparé à la proline.<sup>31,32</sup> Dans le cadre de la conception de peptides CMP, le remplacement des résidus proline et/ou hydroxyproline par des analogues pseudoprolines apparaît très intéressant car le passage de l'isomère *cis* à *trans* est une étape limitante pour la formation de la triple hélice de collagène. Il a été suggéré que

la baisse de la barrière d'énergie serait due à la diminution de la densité d'électrons de la liaison peptidique (effet inductif de l'oxygène) et aux différents plissements de cycle imposés par la substitution d'hétéroatomes. *Kang et al.* ont alors réalisé des études comparatives entre des dipeptides Ac-Pro-NHMe, Ac-Hyp-NHMe et Ac- $\Psi$ Pro-NHMe par DFT (Density Functional Theory), les populations pour chaque état correspondant à un minimum local sont présentées dans le Tableau 2.<sup>33,34</sup> On observe alors que, d'après les calculs théoriques, l'insertion d'un hétéroatome en position 4 va peupler la conformation PPII (dénommée F) principalement en raison des contributions importantes des populations *trans-ppII-down* (*tFd*) et *trans-ppII-up* (*tFu*). En effet, 51% des Ac-Pro-NHMe et 41% des Ac-Hyp-NHMe sont en conformation PPII contre 69% des Ac- $\Psi$ Pro-NHMe. *Mutter et al.* ont montré que la nature des groupements ainsi que la stéréochimie du centre  $^5C_\delta$  avaient une grande influence sur l'isomérisation *cis-trans*<sup>31</sup> en ajoutant un groupement *p*-méthoxyphényle ((*S*)  $\rightarrow$  37% en *cis*, (*R*)  $\rightarrow$  5% en *cis*). L'enjeu pour nous a été de contrôler la population *cis-trans* ainsi que les plissements de cycle pyrrolidine.

C'est ainsi que notre laboratoire a débuté une collaboration avec l'équipe du Professeur Thierry Brigaud qui a développé une méthode de synthèse de 1,3-oxazolidines trifluorométhylées en position  $^5C_\delta$ . Celles-ci peuvent subir une N-acylation et des réactions d'amidification, et ceci sans modification de la stéréochimie, ce qui leur permet d'être résistante à l'hydrolyse<sup>35</sup>. Par conséquent, elles peuvent être considérées comme des analogues de proline (Figure 2).

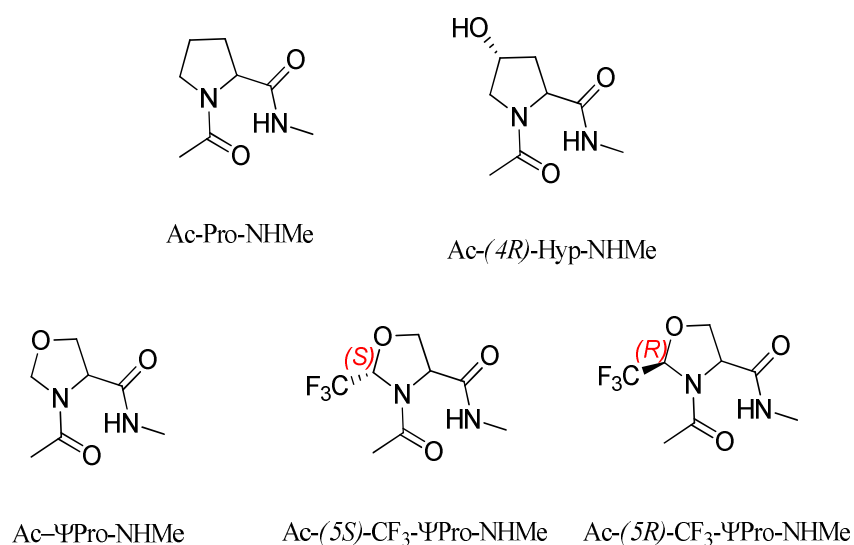


Figure 2: Pseudoprolines analogues de prolines et hydroxyprline

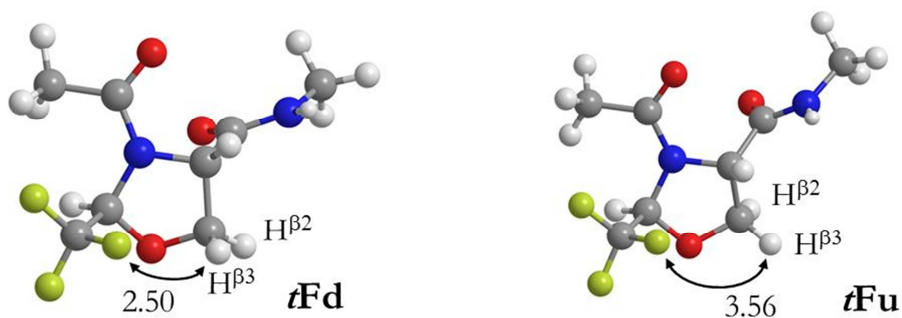
Des études comparatives par calculs DFT et par RMN ont alors été menées afin d'étudier l'isomérisation *cis-trans*, la présence de liaisons hydrogènes ainsi que le plissement du cycle  $\Psi$ Proline<sup>36</sup> et sont résumés dans le Tableau 2.<sup>37</sup>

Peptide	DFT										RMN
	<i>cAd</i>	<i>cAu</i>	<i>cFd</i>	<i>cFu</i>	<i>tAd</i>	<i>tAu</i>	<i>tCd</i>	<i>tCu</i>	<i>tFd</i>	<i>tFu</i>	<i>c</i>
<b>Pro</b>	19.2	1.3	6.6	5.7	23.9	3.1	0.8	0.0	17.6	21.6	
<b>(4<i>R</i>)-Hyp</b>	15.6	3.0	6.1	4.2	15.6	13.5	0.4	0.0	9.0	32.0	
<b><math>\Psi</math>Pro</b>	9.4	1.9	9.8	11.5	16.3	2.7	0.5	n.c.	27.1	20.7	34
<b>(5<i>S</i>)-CF<sub>3</sub>-<math>\Psi</math>Pro</b>	0.7	0.0	36.8	0.1	1.3	n.c.	n.c.	n.c.	<b>60.3</b>	0.7	43
<b>(5<i>R</i>)-CF<sub>3</sub>-<math>\Psi</math>Pro</b>	0.3	1.5	0.5	38.3	0.7	4.3	n.c.	n.c.	1.1	<b>53.4</b>	45

Tableau 2: Populations (%) de chaque conformère dans l'eau pour les différents peptides obtenues par calculs DFT ou d'après des mesures RMN (*c*=*cis*, *t*=*trans*, C=coude  $\gamma$ , A=hélices  $\alpha$ , et F=PPII, d=*down*, u=*up*, n.c.= calculs non convergent)

Ces résultats nous montrent que les estimations de populations *cis:trans* réalisées par RMN et DFT sont du même ordre, les calculs théoriques prédisant légèrement moins de conformations *cis*. D'après ces mesures, l'ajout d'un groupement CF<sub>3</sub> favorise largement une population particulière qui dépend de la stéréochimie du centre <sup>5</sup>C<sub>δ</sub>. La conformation *tFd* est adoptée à 60% par le motif (5*S*)-CF<sub>3</sub>- $\Psi$ Pro (Figure 3) alors que la (5*R*)-CF<sub>3</sub>- $\Psi$ Pro favorise la *tFu* (53%). La stéréochimie du carbone <sup>5</sup>C<sub>δ</sub> affecte uniquement le plissement du cycle à 5 chaînons, les conformations *trans* ( $\omega = 180^\circ$ ) et F ( $\phi = -70^\circ$ ,  $\psi = 150^\circ$ ) étant largement stabilisées dans les deux cas, ce qui est souhaitable dans le cadre de la conception de CMP. L'étude des effets électroniques a montré que les plissements de cycle sont stabilisés par des interactions électrostatiques entre les atomes chargés négativement F (-0.35 e) du groupe CF<sub>3</sub> et un des protons H<sub>β</sub> chargés positivement (+0,21 e). Ces interactions, relativement faibles, agissent en synergie avec deux effets d'hyperconjugaison impliquant chacune l'orbitale antiliante  $\sigma^*$ (C<sup>ε</sup>-F). En effet, la distance F---H<sub>β</sub> pour (5*S*)-CF<sub>3</sub>- $\Psi$ Pro est plus faible dans la conformation *tFd* (d=2.50Å) que *tFu* (d=3.56 Å). Le plissement de cycle *down* est donc stabilisé (Figure 3). Pour les mêmes raisons mais de manière symétrique, le cycle (5*R*)-CF<sub>3</sub>- $\Psi$ Pro est très fortement stabilisé avec un plissement de cycle *up*. Ces résultats indiquent que l'enchaînement (5*S*)-CF<sub>3</sub>- $\Psi$ Pro-(5*R*)-CF<sub>3</sub>- $\Psi$ Pro pourrait avantageusement remplacer la diade Pro-Hyp dans les triplets de collagène en renforçant l'alternance de plissement de cycle « *down-up* » requise pour la stabilité de la triple hélice.





**Figure 3:** Conformations *tFd* et *tFu* adoptées par  $(5S)$ - $CF_3$ - $\Psi$ Pro. La conformation est stabilisée par une interaction favorable entre un fluor et  $H^{\beta 3}$  uniquement pour un plissement de cycle *down*. En conséquence, *tFd* représente 60,3% des conformations adoptées par  $(5S)$ - $CF_3$ - $\Psi$ Pro et *tFu* n'est adopté qu'à 0,7%. Le plissement de cycle *up* est en revanche le plus stable pour l'analogue  $(5R)$ - $CF_3$ - $\Psi$ Pro en raison d'une proximité entre le fluor et  $H^{\beta 2}$

Les études de RMN menées sur  $Ac$ - $(5R)$ - $CF_3$ - $\Psi$ Pro-NHMe ont par ailleurs montrées que la barrière d'énergie de l'isomérisation *cis-trans* était particulièrement basse pour ce mime de tripeptide, au-delà de l'abaissement de 1-2 kcal/mol observé pour l'ensemble des  $\Psi$ Pro.<sup>29</sup> Une mesure de 15,5 kcal/mol a été trouvée pour cette barrière, alors qu'elle est de 19 kcal/mol pour le résidu proline. Cet effet, spécifique de  $(5R)$ - $CF_3$ - $\Psi$ Pro peut à nouveau s'expliquer par une contribution stéréoélectronique des atomes de fluor. Dans le chapitre précédent, nous avons évoqué la présence de nanoobjets solubles, pouvant provenir de l'assemblage désordonné de triples hélices mal repliées. Une raison possible expliquant la présence de tels édifices en solution provient de la liaison Gly-Pro, à 20% en conformation *cis*. Cette géométrie de la liaison peptidique, incompatible avec la formation de triple hélice, est relativement stable puisque la barrière d'énergie de l'isomérisation *cis-trans* est élevée pour la Proline. Il est attendu que le remplacement de la proline par un motif pseudoproline trifluorométhylé favorise l'isomérisation et ainsi (i) accélère la formation des triples hélices et (ii) déstabilise la formation des agrégats solubles.

Un modèle de modélisation moléculaire a été élaboré à l'aide du logiciel Discovery Studio (Biovia) en se basant sur la structure de la triple hélice [(Pro-Pro-Gly)<sub>10</sub>]<sub>3</sub> (PDB : 1k6f). L'incorporation de quatre groupes  $CF_3$  par chaîne a été faite pour obtenir les séquences suivantes au sein de la triple hélice [(Pro-Hyp-Gly)<sub>2</sub>-(( $5R$ )- $CF_3$ - $\Psi$ Pro-( $5S$ )- $CF_3$ - $\Psi$ Pro-Gly)<sub>2</sub>-(Pro-Hyp-Gly)<sub>6</sub>]. La minimisation a été effectuée dans le vide et a permis d'observer une structure de triple hélice stable, où les groupes  $CF_3$  portés par les  $^5C_8$  pointent vers l'extérieur sans causer de gêne stérique. Ceux-ci pourraient alors favoriser des interactions entre les triples hélices *via* des interactions fluor-fluor et former des édifices supramoléculaires plus stables (Figure 4).

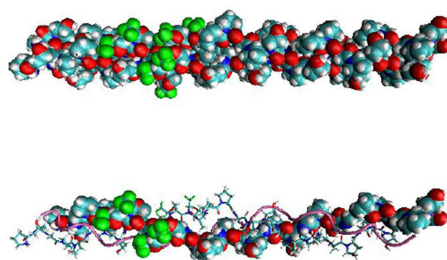


Figure 4: Illustration d'un modèle de triple hélice incorporant des pseudoprolines trifluorométhylées.

## 3.2. Synthèse et caractérisation des triplets

Pour mettre en évidence ces particularités structurales, nous avons synthétisé une famille de quatre triplets Ac-Xaa-Yaa-Gly-NHMe. Parmi eux, deux tripeptides correspondent à des références, utiles pour évaluer les effets des groupes CF<sub>3</sub>, les deux autres, incorporant des pseudoprolines fluorés. Des études RMN complètes ont ensuite été menées et seront présentées en partie 3.2.2

### 3.2.1 Synthèse

Nous avons synthétisé, dans un premier temps, deux pseudotripeptides de référence avec un groupement acétyle en *N*-terminal et amide en *C*-terminal afin de s'affranchir des phénomènes liés aux interactions électrostatiques (Figure 5). En effet, il a été montré que des interactions électrostatiques pouvaient contraindre un plissement de cycle et avaient une influence sur l'isomérisation *cis-trans*.<sup>36</sup>



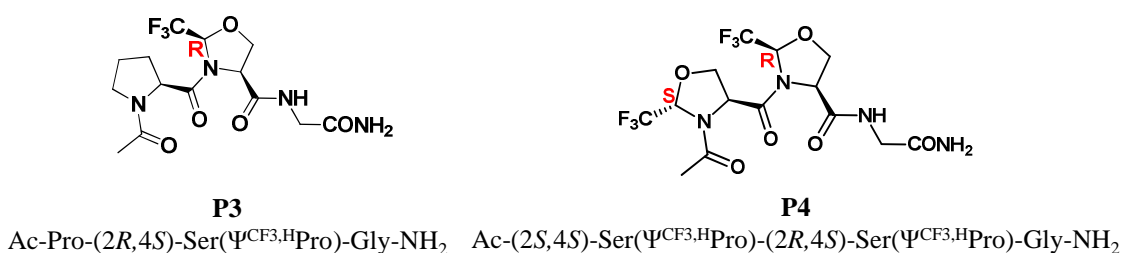
Figure 5: Triplets de référence

Nous avons donc choisi de travailler en stratégie Fmoc car le seul acide aminé qui nécessite la protection de sa chaîne latérale est l'hydroxyproline (hydroxyle en position 4). Cette fonction

est protégée par une fonction tertibutyle qui n'est pas sensible en milieu basique et qui sera déprotégée lors du clivage final en milieu acide.

Les peptides P1 et P2 (Figure 5) ont été préparés par couplage avec l'acide aminé approprié sur une résine amide MBHA (loading: 0.52mmol/g). Cette résine permet d'obtenir la terminaison amide (CO-NH<sub>2</sub>) en *N*-terminal. Chaque acide aminé a été préalablement activé par traitement avec une solution de HBTU ou HATU, DIEA dans la *N*-méthyl-2-pyrrolidone (NMP). La déprotection du groupement Fmoc a été faite par traitement avec une solution à 20% (v/v) de pipéridine dans la NMP. L'acétylation a été réalisée par traitement avec une solution à 20% (v/v) d'anhydride acétique dans la NMP. Les peptides ont été clivés de la résine avec une solution de TFA: triisopropylsilane: H<sub>2</sub>O (95:2.5:2.5). Après purification, nous avons obtenu 14 mg (44%) de **P1** et 29 mg (89%) de **P2**. Ces deux peptides nous ont alors servi de référence pour des études comparatives par RMN.

Dans un second temps, nous avons synthétisé deux autres tripeptides incorporant des groupements trifluorométhylés tel que la (2*R*,4*S*)-Ser( $\Psi^{CF_3,H}$ Pro) remplaçant l'hydroxyproline et (2*S*,4*S*)-Ser( $\Psi^{CF_3,H}$ Pro) remplaçant la proline (Figure 6).



**Figure 6: Triplets fluorés**

La synthèse de ces deux triplets a été réalisée en phase liquide du fait de l'incorporation d'acides aminés non naturels qui nécessite des conditions particulières de couplages. La présence de groupements électroattracteurs CF<sub>3</sub> induit une faible nucléophilie de la fonction amine ainsi qu'un encombrement stérique supérieur ce qui nécessite une activation par un chlorure d'acide. De plus, il est important de noter que la synthèse se fait de l'extrémité *N*-terminale vers *C*-terminale car, une fois acylées, les pseudoprolines trifluorométhylées présentent une grande stabilité.<sup>38</sup>

*Synthèse des pseudoprolines  $\Psi^{CF_3,H}$ Pro* - La première étape est de synthétiser les pseudoprolines, analogues de proline développés par Mutter *et al.*. La synthèse des pseudoprolines trifluorométhylées Ser( $\Psi^{CF_3,H}$ Pro) débute par la condensation d'une sérine

estérifiée avec l'hydrate du fluoral.<sup>39</sup> Deux synthèses ont été réalisées du fait des faibles rendements obtenus lors de la première synthèse en série méthylique (ANNEXE 5). Une seconde stratégie a été de travailler en série benzylique. Le groupement protecteur Boc est utilisé pour améliorer la solubilité de la sérine dans le toluène afin de faciliter l'étape de condensation. Il est alors clivé *in situ* au cours de la réaction. La condensation de la Boc-sérine ester benzylique sur l'hydrate de trifluoroacétaldéhyde en présence de *para*-toluène sulfonate de pyridinium (PPTS) permet d'obtenir la pseudoproline **1** sous forme d'un mélange diastéréoisomérique 81:19 (Schéma 1).

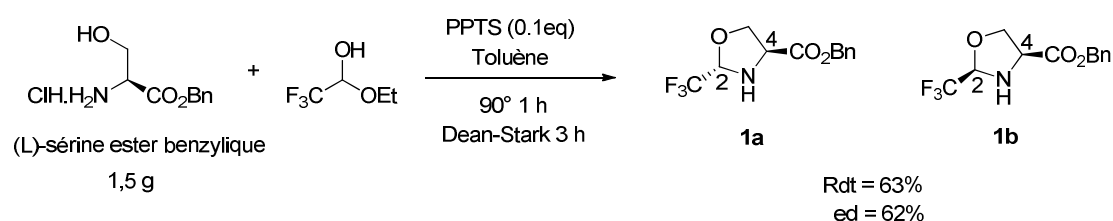


Schéma 1: Synthèse de la Ser( $\Psi^{\text{CF}_3,\text{H}}\text{Pro}$ )-OBn **1**

### Synthèse du peptide P3- Couplage Pro- $\Psi^{\text{CF}_3,\text{H}}\text{Pro}$

Nous avons montré au Laboratoire qu'après un couplage peptidique à partir du mélange diastéréoisomérique de pseudoprolines-1, il était possible d'obtenir un seul dipeptide incorporant une Ser( $\Psi^{\text{CF}_3,\text{H}}\text{Pro}$ ) de configuration (2*R*,4*S*) (Schéma 2). En effet, bien que la pseudoproline-1a soit majoritaire, lors de la réaction d'acylation, le couplage est rendu difficile par la présence des deux groupements ester benzylique et CF<sub>3</sub> en *trans* (encombrement de chacune des 2 faces). En revanche, la pseudoproline-1b, elle, n'a qu'une seule face d'encombrée, ce qui facilite l'approche du chlorure d'acyle et qui favorise la réaction d'acylation. Par ailleurs, la réaction de couplage se déroulant en absence de base, il y a équilibration entre les deux pseudoprolines via un mécanisme d'ouverture-fermeture catalysé en milieu acide (libération d'HCl lors de l'acylation).

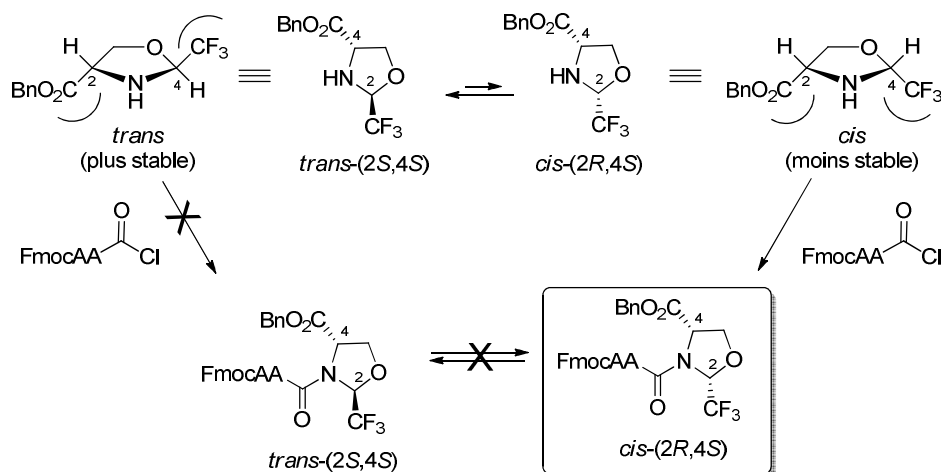


Schéma 2: Couplage peptidique à partir d'un mélange de pseudoproline 1 conduisant à un seul dipeptide

La synthèse du dipéptide Fmoc-Pro-(2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OBn **2** suit alors le schéma suivant (Schéma 3). La Fmoc-L-Pro-Cl est préparée selon la méthode décrite par Babu, par sonication en présence d'un large excès de chlorure de thionyle dans le DCM pendant 45 minutes puis évaporation.<sup>40</sup>

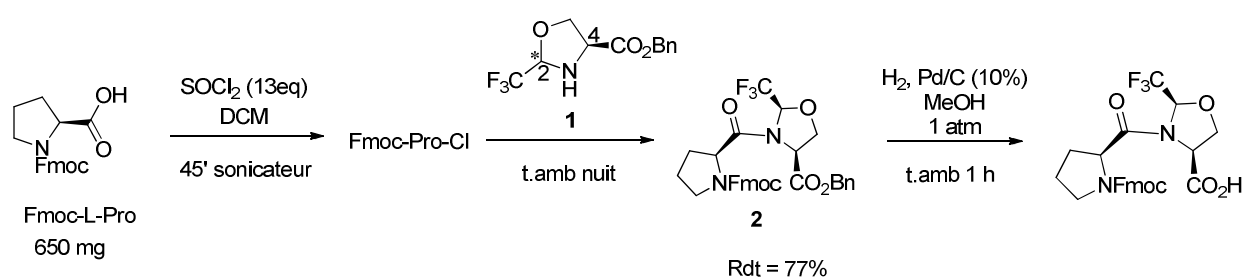


Schéma 3: Synthèse du dipeptide Fmoc-Pro-(2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)

La déprotection chimiosélective réalisée par hydrogénolyse en présence de palladium sur charbon à pression atmosphérique d'hydrogène est plus simple que celle utilisée en série méthylique et le rendement obtenu est plus élevé. De plus, le produit ne nécessite pas d'être purifié.

Synthèse du peptide P3- Couplage Pro-  $\Psi^{\text{CF}_3, \text{H}}$ Pro -Gly

L'acide est ensuite couplé avec la glycine amideafin d'obtenir le tripeptide Fmoc-Pro-(2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-Gly-NH<sub>2</sub>. Puis, afin d'obtenir le pseudotétrapeptide **P3**, le groupement amine est déprotégé et l'extrémité *N*-terminale est alors acétylée en présence d'anhydride acétique avec un rendement global de 16% sur six étapes (Schéma 4).

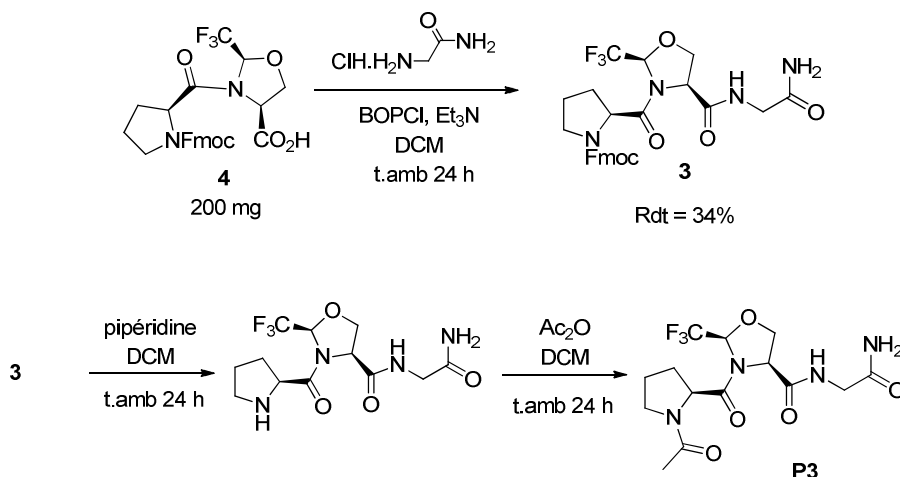


Schéma 4: Synthèse du pseudotétrapéptide Ac-Pro-(2R,4S)-Ser( $\Psi^{CF_3,H}Pro$ )-Gly-NH<sub>2</sub> P3

#### Synthèse du peptide P4- Acétylation $\Psi^{CF_3,H}Pro$

La première étape consiste à préparer Ac-(2S,4S)-Ser( $\Psi^{CF_3,H}Pro$ )-OBn afin de mimer le résidu proline. Contrairement aux réactions de couplage avec des acides aminés classiques, les  $CF_3$ - $\Psi$ Pro ne s'épimérisent pas en <sup>2</sup>C lors des réactions d'acétylation. Ceci peut s'expliquer par le faible encombrement stérique de l'anhydride acétique en comparaison avec un acide aminé *N*-protégé activé sous forme de chlorure d'acide.

Ainsi, le traitement de l'oxazolidinone (2S,4S)-1 par de l'anhydride acétique en présence d'une quantité catalytique de diiode conduit à l'oxazolidinone acétylée correspondante avec un rendement de 69% (Schéma 5).

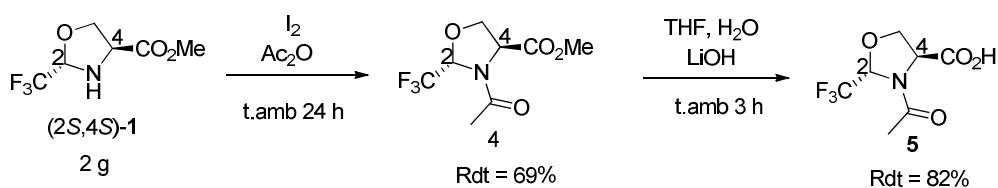


Schéma 5: Synthèse de Ac-(2S,4S)-Ser( $\Psi^{CF_3,H}Pro$ )-OH

#### Synthèse du peptide P4- Couplage de Ac-(2S,4S)-Ser( $\Psi^{CF_3,H}Pro$ )-(2R,4S)-Ser( $\Psi^{CF_3,H}Pro$ )-OBn

La réaction de couplage avec le second motif  $CF_3$ - $\Psi$ Pro, mime de l'hydroxyproline, nécessite une activation de l'oxazolidinone sous forme de chlorure d'acide. Celui-ci est préparé selon la méthode décrite précédemment et est immédiatement engagé dans la réaction de couplage

peptidique avec la (2*R*,4*S*)- Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro)-OBn. Le dipeptide **6** attendu est obtenu avec un rendement de 41% (Schéma 6).

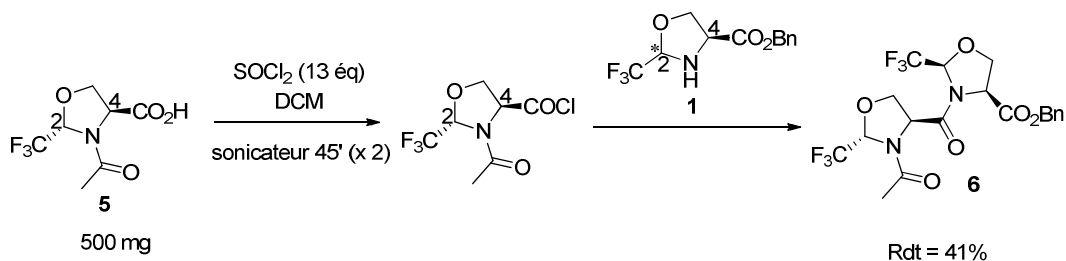


Schéma 6: Synthèse du dipeptide

Synthèse du peptide 4- Couplage de  $\Psi^{\text{CF}_3,\text{H}}$ Pro -  $\Psi^{\text{CF}_3,\text{H}}$ Pro -Gly

Le dipeptide **6** est ensuite déprotégé par hydrogénolyse en présence de palladium sur charbon à pression atmosphérique d' $\text{H}_2$ . L'acide correspondant est alors engagé directement dans la réaction de couplage peptidique avec la glycine amide en présence de BOP-Cl et de triéthylamine. Le tripeptide Ac-(2*S*,4*S*)-Ser ( $\Psi^{\text{CF}_3,\text{H}}$ Pro)-(2*R*,4*S*)-Ser ( $\Psi^{\text{CF}_3,\text{H}}$ Pro)-Gly-NH<sub>2</sub> **P4**, est ainsi obtenu avec un rendement de 17% (Schéma 7).

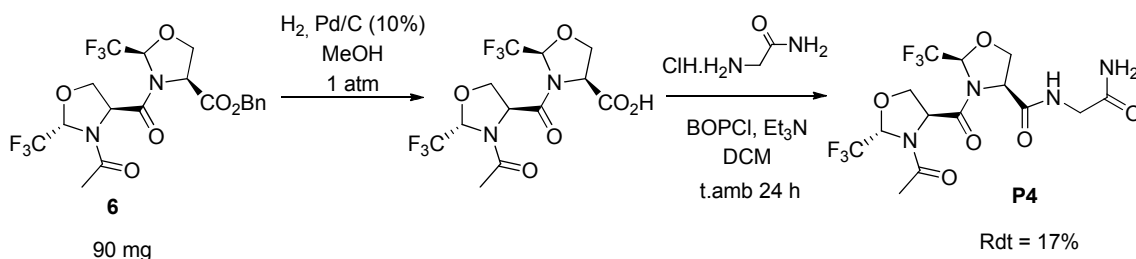


Schéma 7: Synthèse du pseudotétrapeptide Ac-(2*S*,4*S*)-Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro)-(2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro)-Gly-NH<sub>2</sub> **P4**

Nous disposons alors des quatre tripeptides nécessaires pour procéder à des études RMN comparatives.

### 3.2.2 Caractérisation par RMN

La triple hélice de collagène nécessite que toutes les liaisons peptidiques soient en *trans*. Elle est par ailleurs stabilisée par une alternance des plissements de cycle pyrrolidine (Xaa-*down*, Yaa-*up*), et la présence de liaisons hydrogène intramoléculeuses. La RMN permet d'appréhender ces caractéristiques structurales en solution. Nous avons donc enregistré différents types d'expériences (homo et hétéronucléaire) sur les quatre pseudotétrapeptides présentés dans la partie précédente et représenté sur la Figure 7.

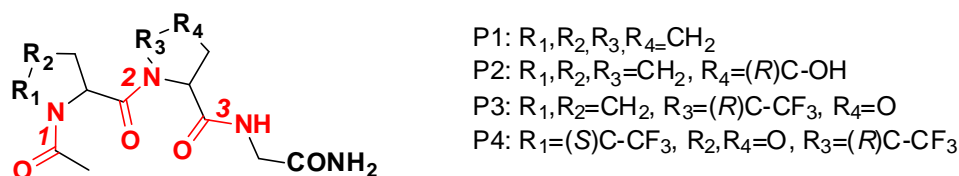


Figure 7: Schéma général des tripeptides

*Attribution* - Pour chaque peptide, différents conformères en échange ont été identifiés et quantifiés. Les attributions  $^1\text{H}$  et  $^{13}\text{C}$  ont été obtenues grâce à l'enregistrement des expériences TOCSY (TOtal Correlation SpectroscopY), NOESY (Nuclear Overhauser Effect SpectroscopY) ou ROESY (Rotating-frame nuclear Overhauser Effect SpectroscopY), et HSQC (Heteronuclear Single Quantum Coherence).

L'expérience TOCSY permet d'observer des corrélations entre l'ensemble des protons qui appartiennent au même système de spins, *via* des relais  $^3\text{J}$  ou  $^2\text{J}$ . Des protons séparés par plus de 5 liaisons peuvent être corrélés, tant que le système de spin n'est pas interrompu par un hétéroatome ou un carbone quaternaire. Cette expérience nous a permis d'identifier les différents systèmes de spins correspondant aux prolines et/ou à ses analogues (Figure 8).

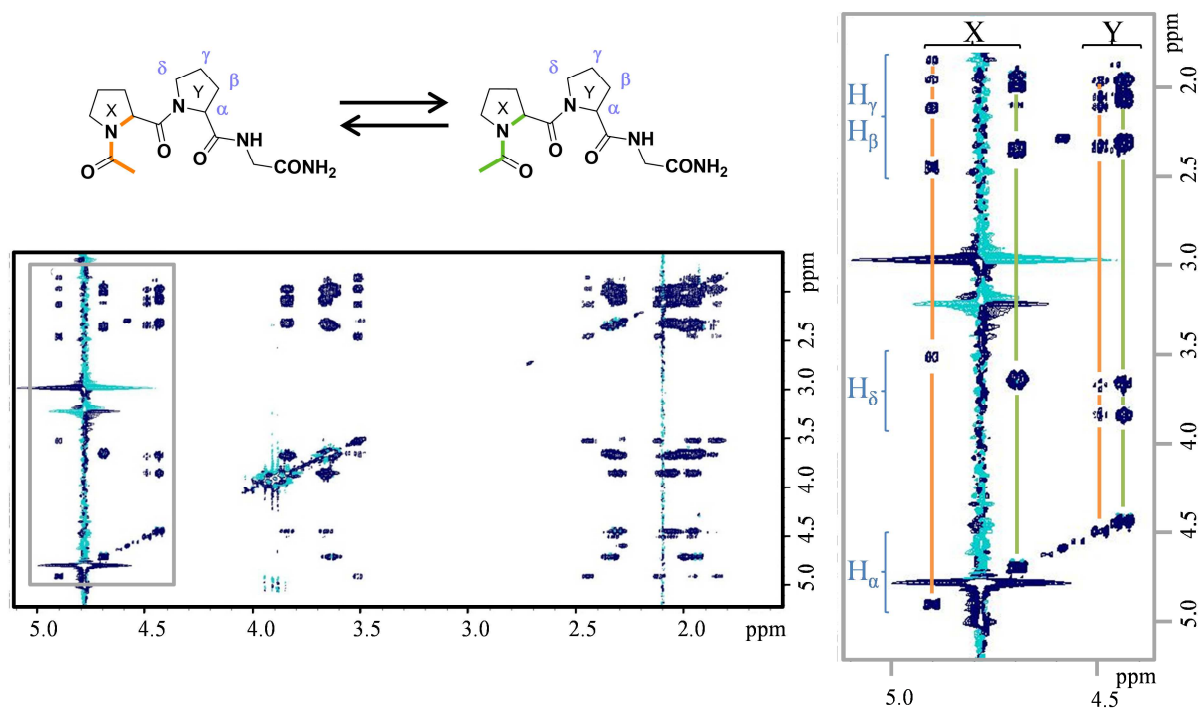


Figure 8: Isomérisie cis-trans du peptide P1 et spectre TOCSY enregistré à 500MHz, 10mM, 25°C,  $\text{D}_2\text{O}$



On observe 4 systèmes de spins différents correspondant à des résidus Proline. L'analyse du spectre NOESY<sup>33</sup> (Figure 9), qui reflète les proximités spatiales entre noyaux (couplage dipolaire), permet d'obtenir l'attribution séquentielle X-Y grâce aux corrélations  $H^{\alpha}_X-H^{\delta}_Y$ . Il montre également qu'une isomérisation a lieu au niveau de la première liaison peptidique Ac-Pro, et permet l'attribution des formes *cis* (minoritaire) et *trans* (majoritaire). Dans un régime d'échange lent, chaque proline est alors dédoublée, 2 correspondent à l'isomère *trans* et 2 pour l'isomère *cis*, (représentées en vert et en orange respectivement). On observe une plus grande différence de déplacement chimique pour les prolines X car elles sont les plus proches de l'isomérisation *cis-trans*. Une quantification a été faite par une analyse conjointe des spectres HSQC et des spectres NOESY. L'expérience HSQC est une expérience hétéronucléaire qui permet de relier le carbone à ses protons, elle permet de mettre en évidence les protons diastéréotopes, mais aussi, grâce à la grande dispersion des fréquences dans la dimension  $^{13}\text{C}$ , de différencier les différents protons superposés en 1D.

*Etude des populations cis-trans* – Nous avons enregistré des spectres NOESY de nos quatre tripeptides en solution dans le  $\text{D}_2\text{O}$ . En présence d'une liaison *cis*, nous avons pu observer un pic de corrélation entre le  $H^{\alpha}_i$  (ou  $\text{CH}_{3,\text{Ac}}$ )- $H^{\alpha}_{i+1}$  et en présence d'une liaison *trans*, nous avons pu observer une corrélation entre  $H^{\alpha}_i$  (ou  $\text{CH}_{3,\text{Ac}}$ )- $H^{\delta}_{i+1}$  (Figure 9).

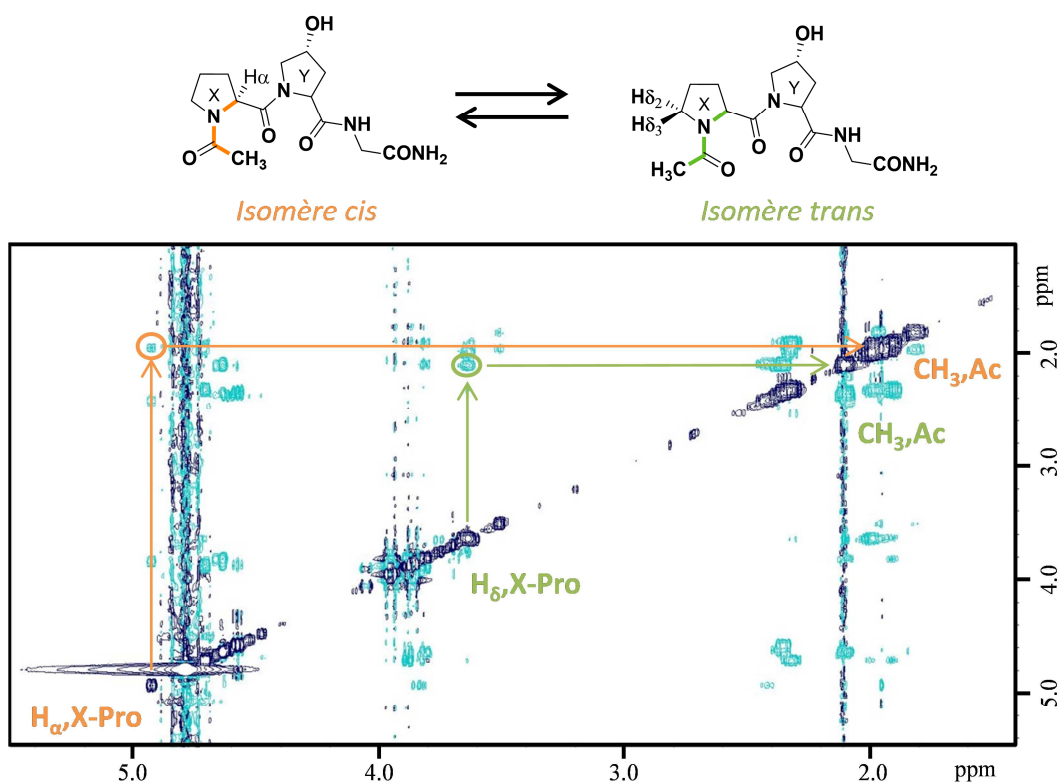


Figure 9: Spectre NOESY<sup>41</sup> enregistré sur le composé P2, 500MHz, 10mM, 25°C,  $\text{D}_2\text{O}$

Chaque tripeptide possède 3 liaisons peptidiques, il y aurait donc *a priori*  $3^2$  conformations possibles par peptides. Or, nous n'observons que deux conformères. En effet, la liaison peptidique qui porte le numéro 3 (Figure 7), est uniquement en *trans*, ce qui correspond aux observations faites pour la quasi-totalité des liaisons peptidiques Xaa-NonPro (donnant les % de *cis* de 0.06%).<sup>3</sup> En outre, la barrière d'énergie de l'isomérisation *cis-trans* pour un enchaînement Xaa-NonPro est de l'ordre de quelques kCal/mol, beaucoup plus basse que celle mesurée pour Xaa-Pro, ce qui conduit à l'impossibilité d'isoler la forme *cis* Xaa-Gly ultra-minoritaire (cas de l'échange rapide). Ainsi, nous n'avons que 4 conformations du squelette accessibles pour l'ensemble de nos peptides, définies par les isoméries des liaisons 1 et 2 (Figure 7). Celles-ci sont reportées dans le Tableau 3. La quantification des différentes populations a été réalisée par intégration des signaux spécifiques sur le spectre HSQC.

Peptide	Ac--- <sup>1</sup> ---Xaa--- <sup>2</sup> ---Yaa--- <sup>3</sup> ---Gly-NH <sub>2</sub>			
	<sup>(1)</sup> <i>cis</i> - <sup>(2)</sup> <i>trans</i>	<sup>(1)</sup> <i>trans</i> - <sup>(2)</sup> <i>trans</i>	<sup>(1)</sup> <i>trans</i> - <sup>(2)</sup> <i>cis</i>	<sup>(1)</sup> <i>cis</i> - <sup>(2)</sup> <i>cis</i>
<b>P1</b>	20	<b>80</b>	0	0
<b>P2</b>	20	<b>80</b>	0	0
<b>P3</b>	5	<b>50</b>	40	5
<b>P4</b>	10	<b>65</b>	20	5

Tableau 3: Populations *cis-trans* exprimées en %

Les peptides *P1* et *P2* possèdent uniquement une isomérisation sur la première liaison peptidique et adoptent très majoritairement la conformation *trans-trans*, requise pour la stabilisation de la triple hélice. La présence d'une population *cis-trans* est probablement surestimée car il ne s'agit pas d'une réelle liaison peptidique en première position mais d'un mime, le groupement acétyle étant moins encombré qu'un acide aminé.

Pour les peptides *P3* et *P4*, nous observons cette même isomérisation de la liaison 1 et dans les mêmes proportions, mais il y a également une isomérisation au niveau de la liaison peptidique 2. La substitution du carbone <sup>5</sup>C<sub>δ</sub> est en effet à l'origine de ce comportement car il déstabilise sensiblement la forme *trans* de la liaison Pro/Ψpro-ΨPro par encombrement stérique. Cependant, cette gêne reste limitée en comparaison avec le groupe isopropyle qui est considéré comme isostère du trifluorométhyle.<sup>9,18-21,32,42</sup> Nous avons pu étudier plus en détail la barrière d'énergie nécessaire pour passer d'une forme à une autre.

*Etude de la barrière d'énergie* – La barrière d'énergie reliant la forme *cis* à la forme *trans* est d'environ 20 kcal.mol<sup>-1</sup> dans le cas d'une liaison peptidique impliquant une proline. C'est une étape limitante lors de la formation de la triple hélice de collagène.<sup>1</sup> La mesure de cette barrière d'énergie peut être déterminée par RMN en réalisant une étude en température. En effet, lorsque l'on augmente la température l'interconversion entre les deux conformères est accélérée, ce qui se traduit par un changement du régime d'échange, de lent vers intermédiaire, puis de intermédiaire à rapide. Deux jeux de résonances distinctes sont observés pour un régime d'échange lent (le cas de nos peptides à température ambiante), les raies s'élargissent à mesure que la température est augmentée (pouvant même disparaître en raison d'une largeur de raie très importante : échange intermédiaire) et elles coalescent enfin à une température caractéristique marquant l'entrée dans un régime d'échange rapide (température de coalescence). Les résonances observées à cette température correspondent à la moyenne des déplacements chimiques des deux formes en échange. Il y a autant de températures de coalescence que de signaux en RMN <sup>1</sup>H. La température de coalescence est reliée à l'énergie d'activation grâce à l'équation d'Eyring:<sup>43</sup>

$$k_{av} = \frac{k_b T}{h} e^{-\frac{\Delta G^\ddagger}{RT}}$$

Où  $\Delta G^\ddagger$  correspond à l'énergie d'activation moyenne,  $h$  est la constante de Planck,  $R$  correspond à la constante universelle des gaz parfaits et  $k_{av}$ , la constante de vitesse moyenne.

En utilisant les deux relations :

$$k_{c-t} = (1 + \Delta p) \times k_{av} \text{ et } k_{t-c} = (1 - \Delta p) \times k_{av}$$

et 
$$k_{av} = \frac{\pi \Delta \nu}{\sqrt{2}}$$

où  $k_{c-t}$  correspond à la constante de vitesse du passage de la forme *cis* vers la forme *trans*,  $k_{t-c}$  correspond à la constante de vitesse du passage de la forme *trans* vers la forme *cis*,  $\Delta \nu$  correspond à la différence de déplacements chimiques en Hz et  $\Delta p$  correspond à la différence de population entre les formes *cis* et *trans*.

il est possible de calculer

$$\Delta G_{c-t}^\ddagger = RT \left[ 23.760 + \ln \left( \frac{T}{k_{c-t}} \right) \right]$$

et

$$\Delta G_{t-c}^\ddagger = RT \left[ 23.760 + \ln \left( \frac{T}{k_{t-c}} \right) \right]$$

Ainsi, nous pouvons calculer les énergies d'activation  $\Delta G_{c \rightarrow t}^\ddagger$  et  $\Delta G_{t \rightarrow c}^\ddagger$  en mesurant la température de coalescence des différents signaux et en calculant les ratio  $c/t$ . Pour les peptides *P1* et *P2*, la barrière d'énergie pour passer d'une forme à l'autre est supérieure à 75°C et il est alors impossible de déterminer la température de coalescence avec notre spectromètre RMN équipé d'une cryosonde. Pour les peptides *P3* et *P4*, les barrières d'énergies ont pu être calculées. Les valeurs trouvées sont répertoriées dans le Tableau 4.

Peptide	Groupements	Population	$\delta$ (ppm)	Tc (K)	$k_{c \rightarrow t}$ ou $k_{t \rightarrow c}$	$\Delta G$ (kcal/mol)	Moyenne (kcal/mol)
<b>P3</b>	<sup>(t)</sup> Acétyl	0,52	2,10	308	14,70	16,39	$\Delta G_{t \rightarrow c}^\ddagger = 16,19$ $\Delta G_{c \rightarrow t}^\ddagger = 16,14$
	<sup>(c)</sup> Acétyl	0,48	2,09		16,17	16,33	
	<sup>(t)</sup> H <sup>o</sup>	0,51	6,34	338	314,44	15,99	
	<sup>(c)</sup> H <sup>o</sup>	0,49	6,05		327,54	15,96	
	<sup>(t)</sup> NH	0,53	8,61	318	494,55	14,72	
	<sup>(c)</sup> NH	0,47	9,08		549,50	14,65	
<b>P4</b>	<sup>(t)</sup> Acétyl	0,79	2,24	318	18,82	16,78	$\Delta G_{t \rightarrow c}^\ddagger = 16,54$ $\Delta G_{c \rightarrow t}^\ddagger = 15,80$
	<sup>(c)</sup> Acétyl	0,21	2,28		70,04	15,95	
	<sup>(t)</sup> H <sup>o</sup>	0,74	6,13	318	40,92	16,29	
	<sup>(c)</sup> H <sup>o</sup>	0,26	6,20		114,58	15,64	

Tableau 4: Détermination des barrières d'énergie pour les peptides 3 et 4

Les constantes de vitesses calculées à partir des résonances amides sont éloignées de celles obtenues pour les autres types protons. Cela est directement lié à la dépendance en température des déplacements chimiques. Les valeurs  $\Delta\nu$  ont été ici déterminées à 298K, soit 20°C en dessous de leur température de coalescence et cet écart en fréquence pourrait être différent à 318K ce qui fausse cette estimation. Les dérives en température des autres protons sont en revanche négligées. Ces résultats nous montrent que la barrière d'énergie est réduite d'environ 4 kcal/mol par rapport aux valeurs correspondantes déterminées sur les résidus proline.<sup>35</sup> Les effets électroniques combinés des pseudoprolines trifluorométhylées et des motifs pseudoprolines expliquent cet abaissement de la barrière d'isomérisation. Ces résultats sont très intéressants pour la stabilisation de la triple hélice qui exige la présence de liaisons peptidiques uniquement en *trans*. En effet, malgré une proportion de liaisons peptidiques *cis* supérieure, les motifs  $\Psi$ pro- $\Psi$ Pro-Gly permettront une interconversion *c-t* beaucoup plus rapide que le triplet Pro-Hyp-Gly. Comme nous le verrons ci-après, la présence non négligeable à l'équilibre des conformères  $\Psi$ pro- $\Psi$ Pro *cis* peut être reliée à l'établissement

d'une liaison hydrogène au sein de ces courts tripeptides. Celle-ci implique le proton amide de la glycine, qui ne sera pas disponible dans l'agencement en triple hélice.

*Liaisons hydrogènes* - La présence de liaisons hydrogènes intramoléculaires et/ou intermoléculaires peut être mise en évidence par RMN grâce au calcul du coefficient de température. Il est défini par la pente de la courbe  $\delta_{NH} = f(T)$ , appelée coefficient de température  $\frac{\Delta\delta_{NH}}{\Delta T}$ . La présence de liaisons hydrogènes est caractérisée par  $\frac{\Delta\delta_{NH}}{\Delta T} > -6\text{ppb/K}$  (Figure 10).<sup>44</sup>

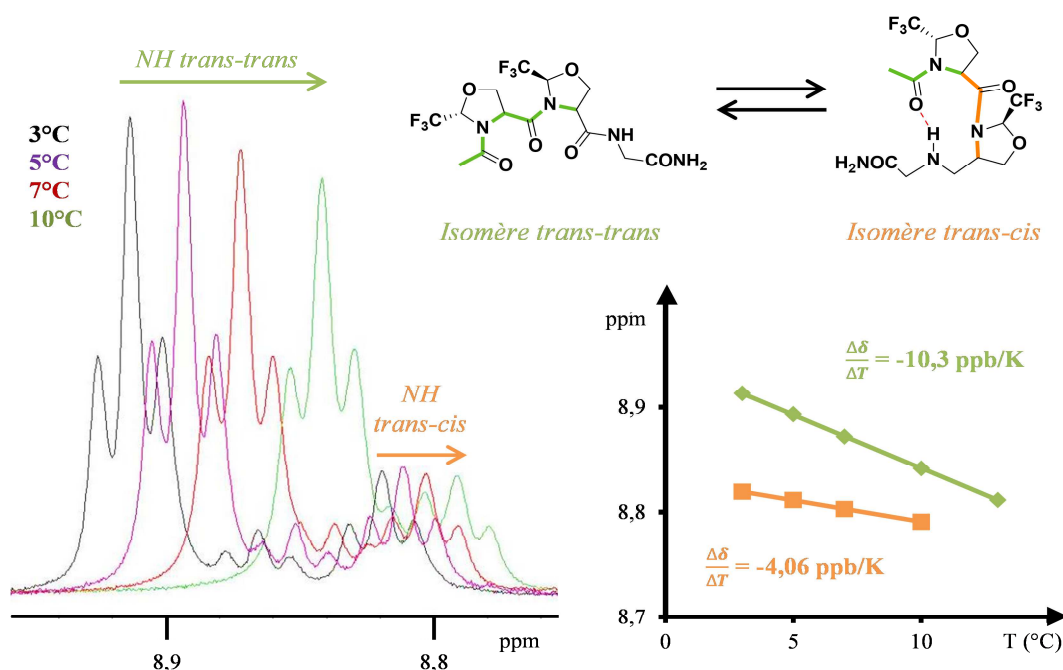


Figure 10: Mesure des coefficients de température pour les 2 conformères principaux de P4, 500MHz, 10mM, D<sub>2</sub>O

L'étude a été réalisée sur chaque tripeptide en enregistrant des expériences 1D-<sup>1</sup>H à différentes températures. Les valeurs ont été calculées pour le seul proton amide présent, positionné sur la glycine terminale, pour les conformères *trans-trans* et, le cas échéant, *trans-cis*. Les résultats sont présentés dans le Tableau 5.

	P1		P2		P3		P4	
Isomère	<i>trans-trans</i>	<i>trans-cis</i>	<i>trans-trans</i>	<i>trans-cis</i>	<i>trans-trans</i>	<i>trans-cis</i>	<i>trans-trans</i>	<i>trans-cis</i>
$\frac{\Delta\delta_{NH}}{\Delta T}$ (ppb/K)	-9,74	X	-10,35	X	-11,48	-4,69	-10,3	-4,06

Tableau 5: Calculs des coefficients de température sur les quatre tripeptides

On remarque la présence de liaisons hydrogène intramoléculaires entre le NH de la glycine et le carbonyle du groupement acétyle en position N-terminal pour les peptides P3 et P4. Elle est

présente seulement pour les conformères *trans-cis*, cette liaison hydrogène pourrait alors expliquer la stabilisation de cet isomère à hauteur de 40% et 20% pour les peptide *P3* et *P4*, respectivement. Pour conclure, la taille du groupement CF<sub>3</sub>-pseudoproline stabilise la liaison peptidique *cis* du fait de son encombrement stérique mais également par la promotion de l'établissement d'une liaison hydrogène. En effet, de nombreuses études ont montré que la conformation de la liaison peptidique était en lien étroit avec le plissement du cycle pyrrolidine.<sup>17,34</sup> La rigidification du cycle pyrrolidine apportée par le groupe trifluorométhyle dans les composés *3* et *4* permettrait à la liaison hydrogène de s'établir plus aisément que dans les composés *1* et *2*.

*Les plissements de cycle* – Les plissements de cycle ont pu être inspectés par RMN en mesurant les constantes de couplage  $^3J_{H\beta H\alpha}$ . En effet, lorsque les constantes de couplage  $^3J_{H\beta 2H\alpha} \approx ^3J_{H\beta 3H\alpha}$  nous sommes en présence d'un plissement de cycle *up* ( $\chi_1 \approx -25^\circ$ ) et dans le cas où  $^3J_{H\beta 2H\alpha} < ^3J_{H\beta 3H\alpha}$  nous sommes en présence d'un plissement de cycle *down* ( $\chi_1 \approx 25^\circ$ ). La représentation de Newman (Figure 11) permet de mettre en évidence ces différences de constantes de couplage. En effet, on peut voir que la constante de couplage  $^3J_{H\beta 2H\alpha}$  est faible dans le cas d'un plissement de cycle *down* car l'angle dièdre est d'environ  $90^\circ$  alors que dans le cas d'un plissement de cycle *up*,  $^3J_{H\beta 2H\alpha} \approx ^3J_{H\beta 3H\alpha}$ .

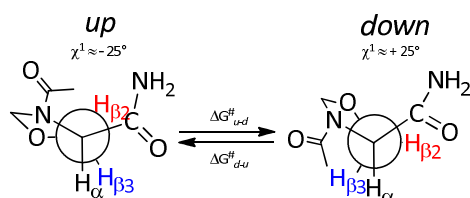


Figure 11: Représentation de Newman des plissements de cycle pyrrolidine

La mesure des constantes de couplage peut être effectuée sur le spectre  $^1H$  ou, si des recouvrements de résonances ne permettent pas cette analyse simple, sur le spectre CH<sub>2</sub>-TROSY. Sur le spectre 1D- $^1H$ , la multiplicité du H<sup>α</sup> est un doublet de doublet lorsque le plissement de cycle est *down* alors que le même H<sup>α</sup> apparaît sous forme de triplet lorsque que le plissement de cycle est *up* (Figure 12).

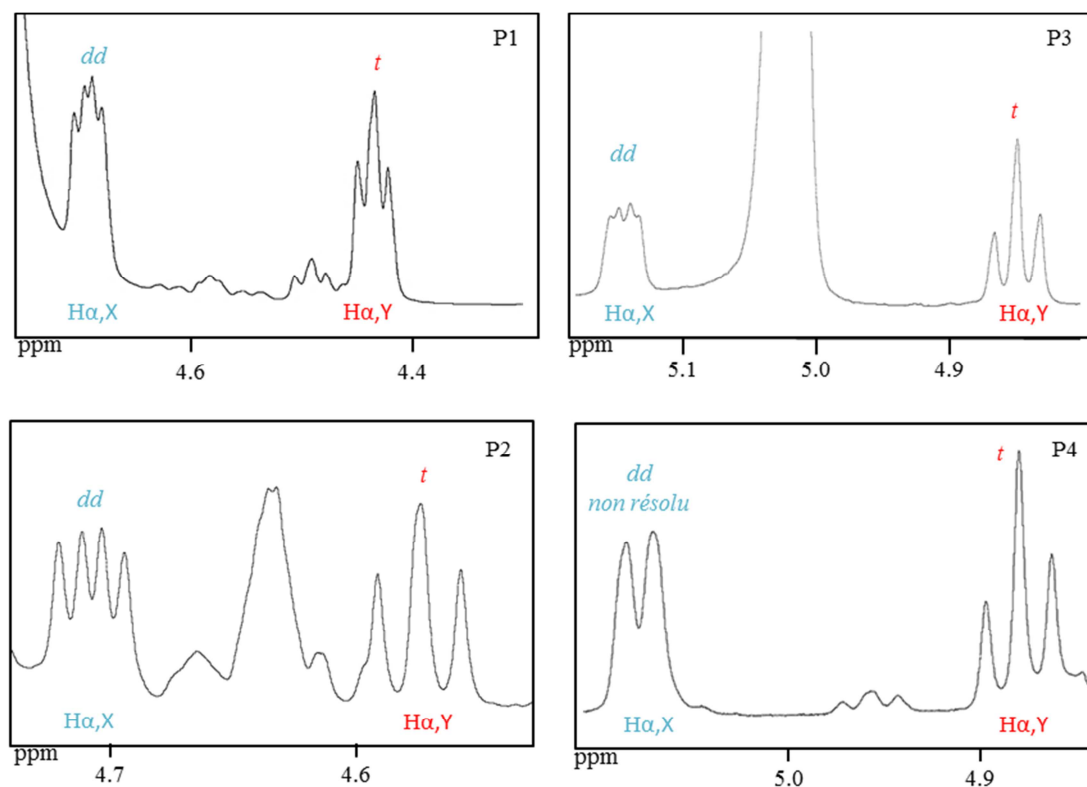


Figure 12: Spectre  $^1\text{H}$  des 4 peptides centré sur la zone des  $\text{H}\alpha$  des prolines, 500MHz, 10mM,  $\text{D}_2\text{O}$

Grace à ces expériences simples, nous pouvons déduire les plissements de cycle majoritaires. De manière plus fine, les mesures de couplage peuvent être effectuées avec une excellente précision grâce à l'utilisation de l'expérience  $\text{CH}_2\text{-TROSY}$ . On visualise bien pour les exemples présentés que la quantification exacte est difficile pour les peptides *P3* et *P4*. Le triplet observé pour *P2* fait également apparaître deux constantes de couplage voisines mais non identiques puisque la raie centrale du triplet apparaît élargie. Les conformères minoritaires sont également difficiles à analyser sur l'unique base des spectres 1D. L'expérience  $\text{CH}_2\text{-TROSY}^{45}$  permet de supprimer le couplage  $^2J_{\text{H}\beta_2\text{H}\beta_3}$  qui prédomine et masque les couplages  $^3J_{\text{H}\beta\text{H}\alpha}$  (cf. chapitre 1). La Figure 13 montre le gain de résolution entre un spectre hétéronucléaire  $^1\text{H}\text{-}^{13}\text{C}$  HSQC (en noir) et un spectre  $\text{CH}_2\text{-TROSY}$  (en bleu). On observe le décalage en fréquence du spectre  $\text{CH}_2\text{-TROSY}$  par rapport au spectre HSQC, dû à la sélection de la corrélation la plus fine. A la différence du spectre  $^1\text{H}$ , nous déduisons les constantes de couplage à partir des protons  $\text{H}\beta$ .

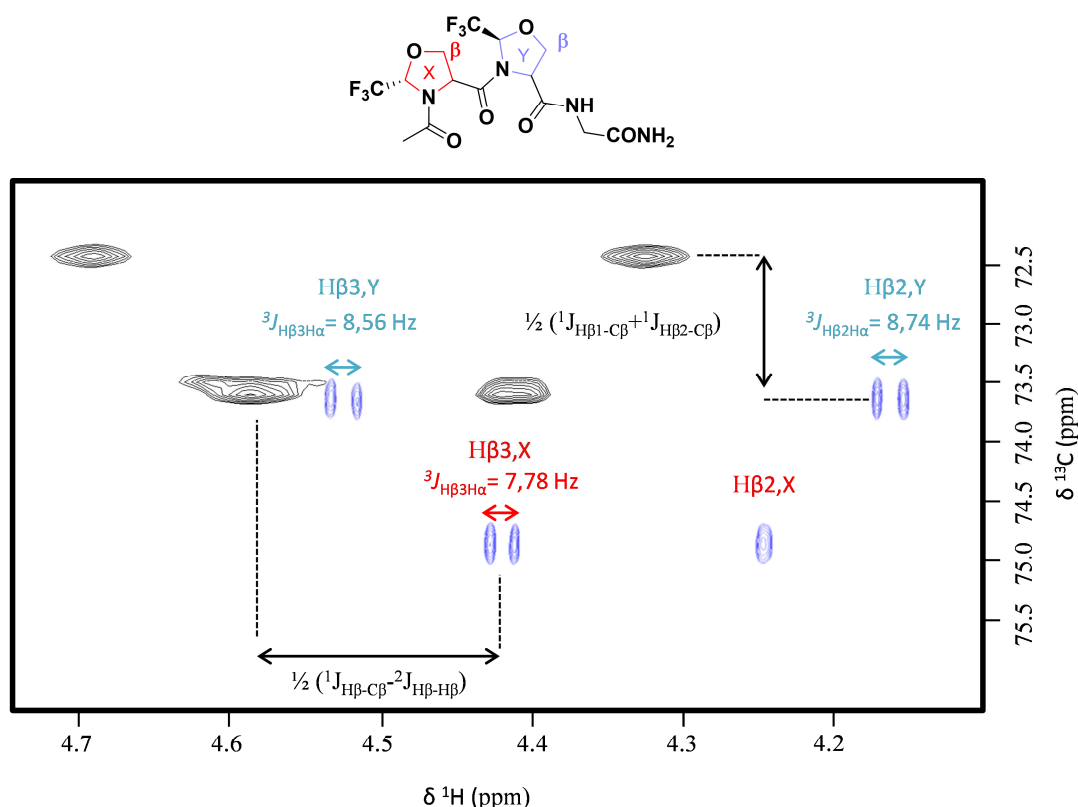


Figure 13: Spectres de corrélations  $^1\text{H}$ - $^{13}\text{C}$  centrés sur la glycine du conformère  $P4$  majoritaire. Spectre HSQC en noir, spectre  $\text{CH}_2$ -TROSY en bleu.  $P4$ , 10mM, 500MHz,  $\text{D}_2\text{O}$

A partir de ces expériences, nous avons pu mesurer les constantes de couplage  $^3J_{\text{H}^\beta\text{H}^\alpha}$  sur les  $\text{H}^\beta$  de chaque cycle pyrrolidine et quantifier les populations de *up* et *down* (Tableau 6). La quantification a été faite grâce aux constantes de couplage théoriques obtenues par des calculs quantiques sur des prolines et des pseudoprolines. Ces calculs différencient les constantes de couplages suivant le plissement de cycle mais également la conformation de la liaison peptidique (cf. Partie expérimentale §B p198).<sup>33</sup>

Peptide		Ac-Xaa-Yaa-Gly-NH <sub>2</sub>					
		<i>cis-trans</i>		<i>trans-trans</i>		<i>trans-cis</i>	<i>cis-cis</i>
<b>P1</b>	Puckering majoritaire	$^{(c)}d - ^{(t)}u/d$		$^{(t)}u/d - ^{(t)}u/d$		-	-
	% <i>down</i> (liaison 1   liaison 2)	88	37	57	45	-	-
<b>P2</b>	Puckering majoritaire	$^{(c)}d - ^{(t)}u$		$^{(t)}u/d - ^{(t)}u$		-	-
	% <i>down</i> (liaison 1   liaison 2)	88	ov	60	7	-	-
<b>P3</b>	Puckering majoritaire	-		$^{(t)}u - ^{(t)}u$		$^{(t)}u - ^{(c)}d$	
	% <i>down</i> (liaison 1   liaison 2)	-		28	5	15	70
<b>P4</b>	Puckering majoritaire	-		$^{(t)}d - ^{(t)}u$		$^{(t)}d - ^{(c)}u$	
	% <i>down</i> (liaison 1   liaison 2)	-		<b>85</b>	<b>10</b>	92	60

Tableau 6: Analyse des populations *up* et *down* pour chacun des peptides étudiés. *ov* (overlap) correspond à un chevauchement des signaux empêchant la détermination du puckering. *u/d* signifie qu'aucun plissement largement majoritaire ne ressort de l'analyse des couplages



On remarque que la présence d'une hydroxyproline (*P2*) à la place d'une proline (*P1*) stabilise le plissement de cycle *up*. Le peptide *P2* correspond au triplet le plus commun dans la triple hélice de collagène, on retrouve bien l'alternance de plissement de cycle *down-up*, bien que la proline ne soit que très légèrement favorisée en conformation *down*. Dans le cas de *P3*, l'hydroxyproline a été remplacée par une (2*R*,4*S*)-Ser( $\Psi_{\text{CF}_3, \text{H}}^{\text{Pro}}$ ), qui d'après les études précédentes semblait être un bon mime de celle-ci (cf. §3.2.2). Il est intéressant de noter que les plissements de cycle des motifs  $\Psi\text{Pro}$  sont inversés lorsque la liaison peptidique précédente change de géométrie. Entre le conformère *trans-trans* et le conformère *trans-cis*, on observe un changement de 95% à 30% de *up* pour le peptide *P3* et de 90% à 40% pour le peptide *P4*. *P4* possède une (2*R*,4*R*)-Ser( $\Psi^{\text{CF}_3, \text{H}}\text{Pro}$ ) à la place de la proline et une (2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}\text{Pro}$ ) à la place de l'hydroxyproline, ces mutations favorisent fortement les plissements de cycle requis pour la stabilisation de la triple hélice, *down* à 85% puis *up* à 90%, il apparaît alors comme le meilleur candidat mime de la triple hélice de collagène.

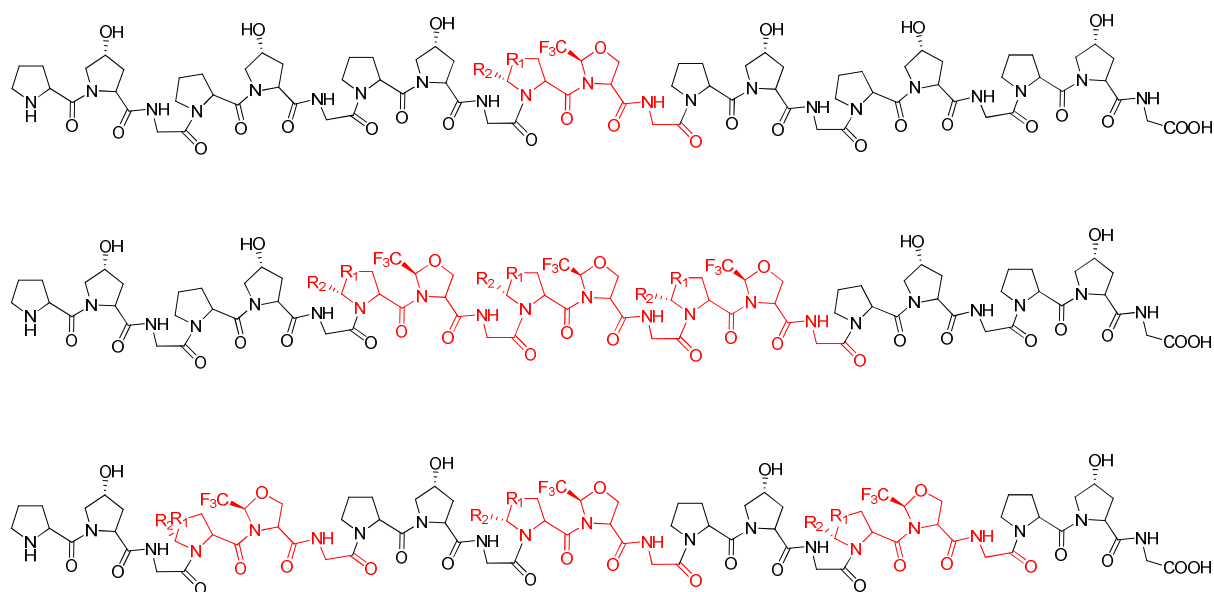
Ces études structurales ont permis de confirmer l'intérêt des pseudoprolines fluorées dans la conception de nouveaux CMP. La stabilisation de la conformation *up* en position Yaa a été démontrée pour les peptides *P3* et *P4*. L'alternance de plissement de cycle *down-up* est parfaitement respectée pour le peptide *P4* ce qui devrait stabiliser la structure en triple hélice de collagène. La synthèse de peptides plus longs incorporant ces triplets a donc été effectuée en vue d'étudier leur influence sur la stabilité de la triple hélice.

### 3.3. Etudes de peptides fluorés

La question posée ici est : le comportement de peptides contenant les motifs fluorés va-t-il être le reflet des études réalisées sur les pseudotérapeptides ?

#### 3.3.1 Synthèse

La stratégie adoptée pour répondre à cette question a consisté en la synthèse de trois peptides incorporant un motif Pro-(2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}\text{Pro}$ )-Gly (**PYPG**) ou (2*S*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}\text{Pro}$ )-(2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}\text{Pro}$ )-Gly (**YPYPG**) à des positions stratégiques (Figure 14).



R<sub>1</sub> = H (PΨPG) ou O (ΨΨΨPG)  
 R<sub>2</sub> = H (PΨPG) ou CF<sub>3</sub> (ΨΨΨPG)

Figure 14 : Peptides incorporant des motifs  $\Psi^{\text{CF}_3,\text{H}}\text{Pro}$

*Synthèse avec PΨPG* – Pour des raisons méthodologiques, nous avons commencé par la synthèse des peptides incorporant le triplet PΨPG. Ce triplet, protégé en N-terminal, qui n'est autre que le composé *P3*, a été synthétisé en grande quantité afin d'être incorporé en blocs lors de la synthèse. Le motif (2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3,\text{H}}\text{Pro}$ ) étant entouré de part et d'autre par la proline et la glycine, respectivement, évite l'épimérisation en <sup>2</sup>C de la CF<sub>3</sub>ΨPro. La synthèse a été réalisée à l'aide d'un synthétiseur automatique et est détaillée en partie expérimentale (cf. §C p.199). Chaque peptide a été obtenu avec des rendements d'environ 20%.

*Synthèse avec GΨΨΨP* – Cette synthèse s'est avérée plus compliquée car l'incorporation du triplet Fmoc-ΨΨΨPG présente deux difficultés : la nécessité d'utiliser un acide aminé activé sous forme de chlorure d'acide (incompatible avec la SPPS) pour palier la faible nucléophilie de la ΨP et l'épimérisation du centre <sup>2</sup>C (cf. §3.2.1). Afin de contourner ces difficultés, nous avons envisagé une stratégie consistant en la synthèse du bloc GΨΨΨP en partant du N-terminal vers le C-terminal. La principale difficulté est le couplage peptidique entre GΨP et ΨP, ce couplage nécessite l'activation du dipeptide sous forme de chlorure d'acide.

*Synthèse avec GΨΨΨP* - Acylation de la pseudoproline **1a**

Plusieurs conditions réactionnelles ont été testées (et sont détaillées en ANNEXE 6). L'addition de bromure de bromoacétyle sur la pseudoproline **1a** à 0°C sans solvant conduit au

produit d'acylation **7** attendu avec un rendement corrigé de 81% (Schéma 8). Le produit d'épimérisation est observé à l'état de trace et est séparé lors de la purification par chromatographie. L'oxazolidine **1a** de départ, qui n'a pas réagi, est protonnée, et précipite dans le milieu réactionnel. Celle-ci est alors récupérée par simple filtration puis régénérée par traitement basique avec une solution saturée en  $\text{NaHCO}_3$ .

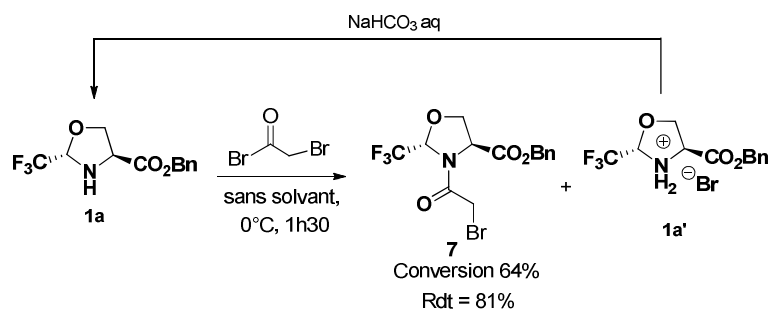


Schéma 8: Synthèse de l'oxazolidine bromoacétylée **7**

### Synthèse avec *GYPYP* - Synthèse du dipeptide **9**

La première étape est d'élaborer un résidu glycine en substituant le brome par une amine. Pour ce faire, une réaction de substitution de l'azoture de sodium sur le produit **7** est effectuée à température ambiante et l'azide **8** est obtenu avec un rendement de 83%. Un suivi de la réaction par RMN du  $^{19}\text{F}$  s'avère ici particulièrement utile, les composés **7** et **8** étant peu séparables par CCM. La fonction amine est ensuite générée de façon concomitante avec la déprotection de l'acide par hydrogénolyse de l'azide **8** en présence d'une quantité catalytique de palladium sur charbon (Schéma 9).

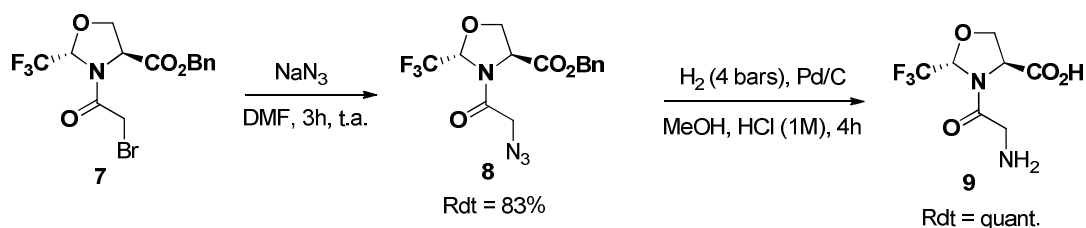


Schéma 9: Synthèse du dipeptide **9**

### Synthèse avec *GYPYP* - Synthèse du tripeptide **11**

Dans un premier temps, la fonction amine est protégée par un groupement Fmoc de façon classique par traitement du dipeptide **6** par du Fmoc-OSu en présence de carbonate de sodium dans un mélange dioxane/eau. Puis la pseudoproline **1b** est ajoutée au dipeptide **9**. L'étape la plus délicate est alors maîtrisée grâce à une double approche : activation du dipeptide à l'aide d'un chlorure d'acide et activation de l'amine secondaire par le chlorure de triméthylsilyle ( $\text{TMS-Cl}$ ) en présence de zinc (Schéma 10). L'amine silylée est alors engagée dans la réaction

de couplage peptidique en présence du dipeptide activé sous forme de chlorure d'acyle. Cela conduit au tripeptide protégé **11** attendu avec un rendement satisfaisant de 42%.

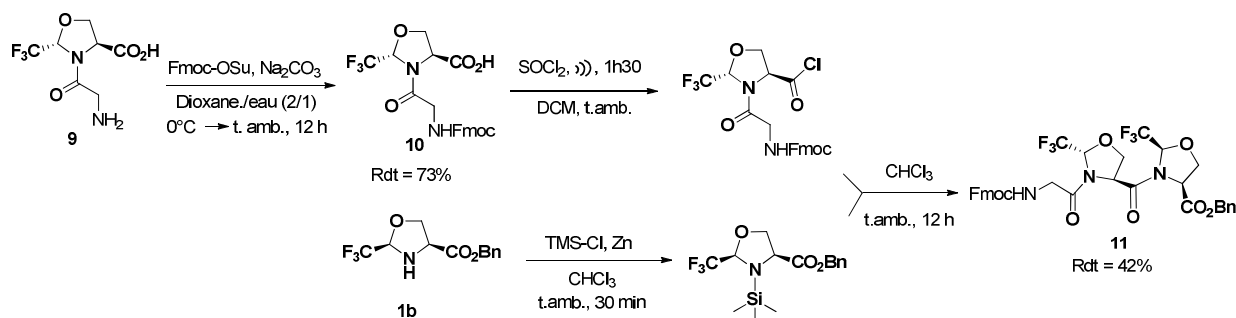


Schéma 10: Synthèse du tripeptide protégé **11** avec activation de l'amine

*Synthèse avec GYPYP* - Déprotection finale pour la préparation du tripeptide **9**

La déprotection de l'acide, par hydrogénolyse de l'ester benzylique sous atmosphère d'hydrogène et en présence d'une quantité catalytique de palladium sur charbon est réalisée. Cette étape nécessite un suivi de la réaction en temps réel par CCM en raison de la possibilité de déprotection de l'amine (Schéma 11).

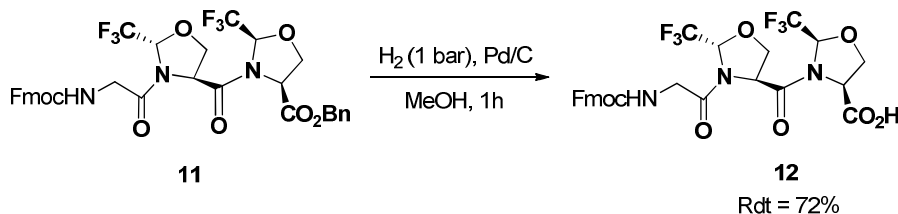


Schéma 11: Déprotection finale pour l'obtention du tripeptide **12**

*Synthèse avec GYPYP* – Le GYPYP étant prêt, il a ensuite été engagé dans la synthèse par blocs sur le synthétiseur automatique par activation aux micro-ondes (§C p.199). Le chromatogramme montre la présence de deux peptides de masse équivalente correspondant à la moitié de la masse attendue pour le peptide (POG)<sub>3</sub>-[(2*S*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)]-(2*R*,4*S*)-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-Gly]-(POG)<sub>3</sub>. Nous avons pensé dans un premier temps que l'utilisation des micro-ondes et du mélange DIC/Oxyma pur pouvait être responsable de cet échec. Nous avons relancé une synthèse sur phase solide automatisée mais en réalisant l'incorporation du tripeptide **12** par double couplage manuel en présence de HATU et de DIEA. Cet essai n'a pas non plus permis d'obtenir le CMP souhaité. L'hypothèse la plus probable serait que la liaison amide entre les deux pseudoprolines soit sensible aux conditions acides du clivage de la résine. Cela expliquerait la présence des deux fragments de masses proches sur le

chromatogramme. L'alternative maintenant est l'utilisation d'une résine de type 2-chlorotriyle qui permettrait un clivage dans des conditions plus douces.

Nous disposons alors de trois peptides incorporant le triplet PΨPG:

- CMPF01 = (POG)<sub>3</sub>- PΨPG-(POG)<sub>3</sub>,
- CMPF02 = (POG)<sub>2</sub>-(PΨPG)<sub>3</sub>-(POG)<sub>2</sub>
- CMPF03 = (POG)-(PΨPG)-(POG) -(PΨPG)- (POG)-(PΨPG)-(POG)

### 3.3.2 Analyse structurale

Deux types d'analyses ont été réalisés afin de comparer les peptides avec et sans groupements CF<sub>3</sub> : du dichroïsme circulaire afin de montrer si il y a formation ou non de la triple hélice, et le cas échéant, de déterminer sa stabilité, et des expériences de STD afin de mettre en évidence une éventuelle agrégation.

*Dichroïsme circulaire* – La structure secondaire polyproline II est caractérisée par un maximum à 225 nm par dichroïsme circulaire. Des études comparatives ont été menées sur les trois peptides contenant du fluor, ainsi que le témoin (POG)<sub>7</sub> (Figure 15).

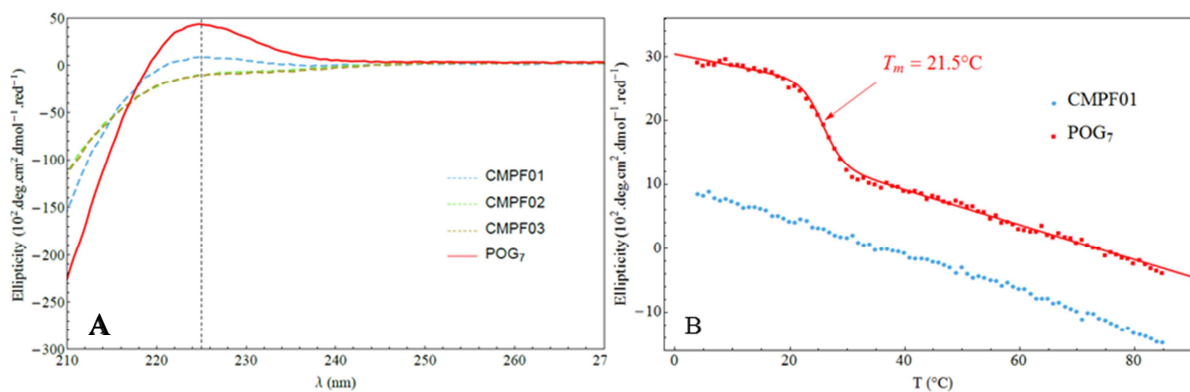


Figure 15: A. Spectre dichroïque des CMPFs et (POG)<sub>7</sub>, 100µM, 4°C, pH=7 (phosphate), B. Ellipticité à 225 nm en fonction de la température pour CMPF01 et (POG)<sub>7</sub>, 100µM, pH=7 (phosphate)

Le (POG)<sub>7</sub>, représenté en rouge sur la Figure 15A, possède un maximum à 225nm significatif de sa structuration en PPII. On observe une diminution de ce maximum caractéristique lorsque l'on introduit un groupement fluoré puis une absence de celui-ci après l'introduction de trois groupements fluorés. Le groupement (2*R*,4*S*)-Ser(Ψ<sup>CF<sub>3</sub>,H</sup>Pro) déstabilise donc la structure PPII et/ou la triple hélice. Le CMPF01 comportant une population PPII non nulle, une étude en température a été réalisée afin de mettre en évidence la présence d'une structuration en triple hélice (Figure 15B). Nous observons, en fonction de la température, une

chute décroissante du signal à 225 nm qui est caractéristique des monomères de PPII, sans observer d'effet de stabilisation par une triple hélice comme on peut le voir pour le (POG)<sub>7</sub>.

Le CMPF01 ne se structure donc pas en triple hélice dans ces conditions. La formation de la triple hélice étant dépendante de la concentration, nous avons voulu réaliser des études CD à une concentration plus élevée. Nous avons choisi de réaliser des échantillons à 10mM pour pouvoir comparer ces résultats aux analyses sur les peptides modèles présentées dans le chapitre 1.

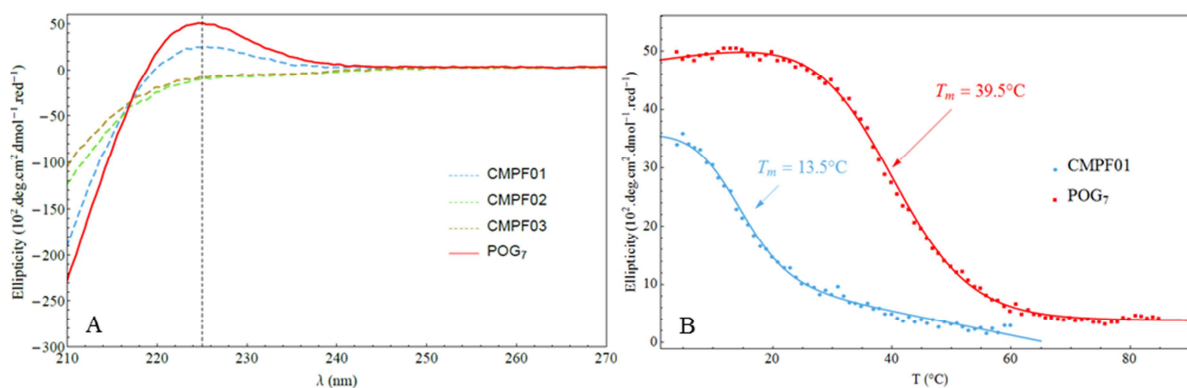


Figure 16: A. Spectre dichroïque des CMPFs et (POG)<sub>7</sub>, 10mM, 4°C, pH=7 (phosphate), B. Ellipticité à 225 nm en fonction de la température pour CMPF01 et (POG)<sub>7</sub>, 10mM, pH=7 (phosphate)

Le signal dichroïque est normalisé suivant la concentration et la longueur de la cuve. Nous observons une similitude entre spectres dichroïques à 4°C pour des échantillons à 100μM et 10mM, ce qui prouve que la concentration n'influe pas sur la structure secondaire des peptides CMPF02, CMPF03, et (POG)<sub>7</sub> (Figure 16A.). En revanche, la proportion d'hélice PPII paraît augmentée pour CMPF01. Pour évaluer si cette structuration était liée à une formation de triple hélice, une étude en température a été réalisée sur le CMPF01. A c=10mM, nous sommes en présence d'une triple hélice, beaucoup moins stable car sa température de fusion est de 13,5°C contre 39,5°C pour (POG)<sub>7</sub>.

Nous avons donc montré que l'incorporation de motif (2R,4S)-Ser(Ψ<sup>CF3,H</sup>Pro), qui ne permet pas de respecter pas la règle d'alternance des plissements de cycle introduite par Raines,<sup>10,46</sup> déstabilisait effectivement la triple hélice. Les analyses des peptides incorporant le motif (2R,4S)-Ser(Ψ<sup>CF3,H</sup>Pro)-(2R,4S)-Ser(Ψ<sup>CF3,H</sup>Pro) seront déterminantes pour montrer la nécessité de cette alternance de plissement de cycle dans la stabilisation de la triple hélice. Cependant, leur synthèse n'a pas pu être menée à terme à ce jour.

*RMN STD* – Nous avons voulu évaluer la capacité des peptides fluorés à s’autoassocier, ainsi que ce que nous avons observé pour les peptide (POG)<sub>n</sub>. Nous souhaitons renforcer les interactions supramoléculaires grâce à l’incorporation de groupes trifluorométhyles, exposés à l’extérieur de la triple hélice. Nous avons donc réalisé des expériences STD (cf. chapitre 1) sur le CMPF01. Dans un premier temps, à titre de comparaison nous avons travaillé avec un échantillon à 200μM.

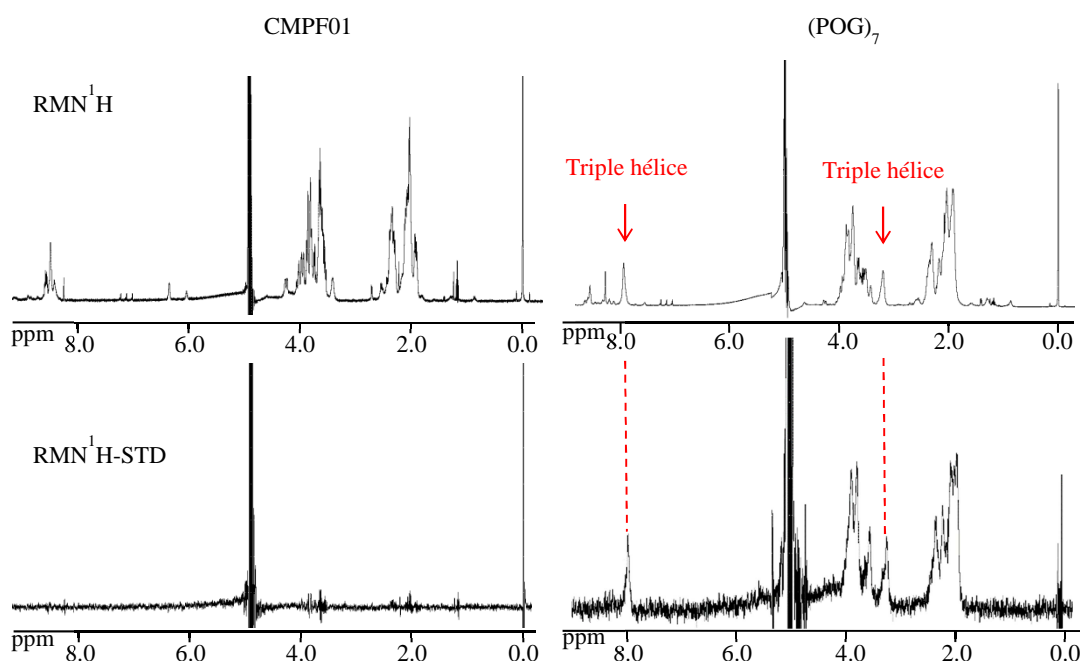


Figure 17: Spectre STD de CMPF01 et (POG)<sub>7</sub>, 200μM, HCl, 20°C

Les expériences enregistrées ne nous ont pas permis de mettre en évidence un signal STD sur le CMPF01 contrairement au (POG)<sub>7</sub> (Figure 17). Les études STD précédentes nous avaient montré que les nanoobjets pouvaient être identifiés par cette technique et étaient en interaction avec la triple hélice. Pour CMPF01, nous n’avons pas de triple hélice à cette concentration, nous ne pouvons donc pas conclure quant à la présence d’assemblage supramoléculaire dans ces conditions.

## Conclusions et perspectives

Nous avons vérifié l'influence de l'ajout d'acides aminés incorporant une ou deux prolines trifluorométhylés au sein d'un pseudotétrapeptide. Les préférences conformationnelles trouvées correspondent à celles qui avaient été établies pour les acides aminés non engagés dans un peptide. On observe néanmoins qu'ils influencent le plissement de cycle de l'acide aminé qui le précède. De plus, nous avons observé une augmentation de la population *cis* qui est en partie expliquée par la présence d'une liaison hydrogène entre le NH de la glycine et le CO du groupement précédent. Cependant, cette liaison hydrogène ne pourrait pas s'établir simultanément avec la liaison hydrogène  $\text{HN}_{\text{Gly}}-\text{CO}_{\text{Pro}}$  présente dans les triples hélices. L'analyse structurale de peptides composés de la répétition de 7 motifs triplets, incorporant le motif  $(2R,4R)\text{-Ser}(\Psi^{\text{CF}_3,\text{H}}\text{Pro})$  en lieu et place de l'hydroxyproline, nous a montré qu'il déstabilisait la triple hélice, ce qui était prévisible en raison du non-respect de l'alternance des plissements de cycle.<sup>10,46</sup>

Des optimisations de synthèse sont en cours pour achever la synthèse de CMP basés sur la répétition du motif  $\text{Gly-(2S,4R)-Ser}(\Psi^{\text{CF}_3,\text{H}}\text{Pro})\text{-(2R,4R)-Ser}(\Psi^{\text{CF}_3,\text{H}}\text{Pro})$ . A la lumière de nos études structurales, ce motif est le plus prometteur pour la synthèse de CMP. L'utilisation d'une résine permettant un clivage dans des conditions plus douces va être testée.



## Partie expérimentale

### *Solvants*

Tous les solvants ont été purifiés et séchés par des techniques classiques et distillés avant utilisation. Le dichlorométhane a été distillé sur hydrure de calcium sous atmosphère d'argon.

### *Chromatographie*

Les plaques CCM sont visualisées sous lumière 254 nm UV et révélées avec une solution d'acide phosphomolybdique à 10% dans de l'éthanol ou une solution de  $\text{KMnO}_4$ . Une purification par chromatographie sur gel de silice a été effectuée sur Macherey-Nagel 70-230 mesh ASTM silice.

### *Analyse - Synthèse*

Les spectres RMN  $^1\text{H}$ ,  $^{13}\text{C}$  et  $^{19}\text{F}$  ont été enregistrés sur un JEOL ECX-400 (400 MHz  $^1\text{H}$ , 100,5 MHz et 376,2 MHz  $^{13}\text{C}$   $^{19}\text{F}$ ). Les déplacements chimiques de  $^1\text{H}$ -RMN sont exprimés en parties par million par rapport au tétraméthylsilane ( $\delta = 0$ ) dans  $\text{CDCl}_3$ . Les déplacements chimiques de RMN  $^{13}\text{C}$  sont exprimés en parties par million en aval du  $\text{CDCl}_3$  comme standard interne ( $\delta = 77,0$ ). Les déplacements chimiques de RMN  $^{19}\text{F}$  sont exprimés en parties par million en aval à partir de  $\text{C}_6\text{F}_6$  comme étalon interne ( $\delta = -164,9$ ). Les données sont signalées comme suit : déplacement chimique ( $\delta$  ppm), la multiplicité (s = singulet, d = doublet, t = triplet, q = quartet, m = multiplet), constante de couplage (Hz). Les spectres infrarouge (FT-IR) ont été réalisés sur un Bruker Tensor. Les rotations spécifiques ont été mesurées sur un polarimètre JASCO P1010. Les spectres de masse à haute résolution ont été obtenus en utilisant une ionisation par électronébulisation en mode positif et un analyseur de masse quadripolaire.

### *Analyse par RMN- Analyse structurale*

Des spectres  $^1\text{H}$  NMR et  $^{13}\text{C}$  NMR ont été enregistrés sur un spectromètre Bruker Avance III 500MHz équipée d'une cryosonde. Les déplacements chimiques de  $^1\text{H}$ -RMN et  $^{13}\text{C}$  sont exprimés en parties par million par rapport au 4,4-diméthyl-4-silapentane-1-sulfate ( $\delta = 0$ ) dans  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (90 :10) ou  $\text{D}_2\text{O}$ . Les données sont signalées comme suit: déplacement chimique ( $\delta$  ppm), la multiplicité (s = singulet, d = doublet, t = triplet, q = quartet, m = multiplet), constante de couplage (Hz). Les échantillons (10mM) ont été préparés par dissolution du peptide dans 550 $\mu\text{l}$  de solvant (1% (v/v) DSS dans  $\text{D}_2\text{O}:\text{H}_2\text{O}$  (10:90)).

## A. Synthèse des peptides P1, P2, P3 et P4

### A1. Synthèse sur support solide

#### *Réactifs*

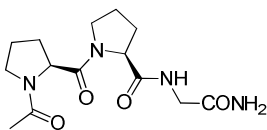
Les acides aminés utilisés lors des synthèses sont : Fmoc-Gly-OH, Fmoc-Pro-OH et Fmoc-Hyp(tBu)-OH.

#### *Procédure générale*

La résine utilisée est de type Fmoc Rink Amide MBHA (loading : 0.52mmol/g). Les synthèses sont réalisées sur 0.1mmol. La résine est mise à gonfler avec du CH<sub>2</sub>Cl<sub>2</sub> pendant 5 min puis dans la NMP pendant 5 min dans une seringue de 5 ml. Ensuite, un test Kaiser est réalisé<sup>47</sup> afin de vérifier que la résine était bien protégée. Ensuite, on déprotège la résine avec une solution de pipéridine/NMP (20 :80) pendant 3 min, puis 5 min et enfin 10 min. Puis la résine est lavée avec de la NMP. On vérifie que la résine est bien déprotégée avec le test de Kaiser. Puis chaque acide aminé (5éq), excepté l'hydroxyproline (2éq), est solubilisé dans 4,5éq de HBTU (ou 1.8éq HATU pour l'hydroxyproline) et 1.5 ml de NMP. Ensuite, 10éq de DIEA (4éq pour l'hydroxyproline) sont ajoutés à la solution d'acide aminé pour activer la fonction acide. Après 3 min, cette solution est ajoutée à la résine dans la seringue. Le couplage est réalisé en 30 min (60 min pour les couplages difficiles engageant l'hydroxyproline) et vérifié par le test Kaiser. Après couplage la solution est alors lavée avec de la NMP (2 fois minimum) puis on réalise la déprotection. Ce cycle est répété autant de fois qu'il y a d'acides aminés à coupler. Il est important de noter que le test Kaiser permet seulement de mettre en évidence les amines primaires libres, un autre test est alors réalisé pour détecter la présence d'amines secondaires : le test chloranil<sup>48</sup>. La dernière étape est l'acétylation, nous utilisons 1 à 2 ml d'une solution d'anhydride acétique dans la NMP (20 :80). La réaction est laissée pendant 30 min. Le test au chloranil est réalisé. La résine est lavée au CH<sub>2</sub>Cl<sub>2</sub> puis dans le MeOH et enfin séchée au dessiccateur. Enfin, le peptide est clivé de la résine avec 4 ml d'une solution de TFA:TIS:H<sub>2</sub>O (95 :2.5 :2.5) pendant 2h. Ensuite, le TFA est évaporé avec de l'air comprimé puis le peptide est précipité avec de l'éther froid (préalablement mis dans un bain de glace). Les peptides, étant petits, ne précipitent pas et sont donc dilués dans l'eau puis lyophilisés.

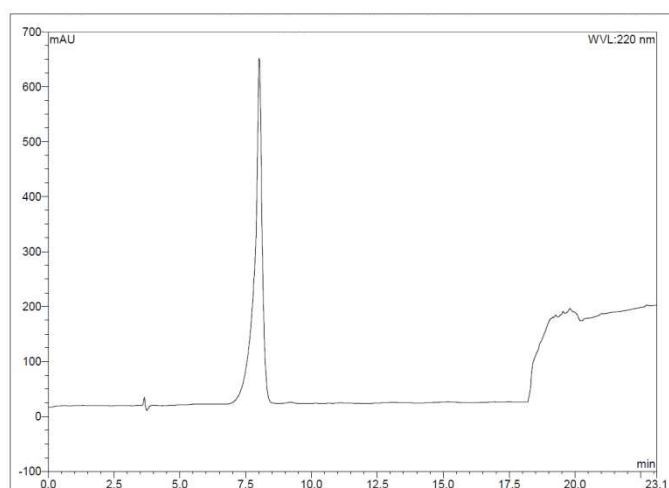
#### *Caractérisation des produits*

Les réactions ont été suivies par HPLC équipée d'une colonne analytique ACE 5 C18-300 (200\*4,6mm) avec un gradient de 95:5 à 80:20 A/B pour **1** et de 100:0 à 90:10 A/B pour **2** à un débit de 1ml/min en 30 min ; éluant A (H<sub>2</sub>O/TFA 0.1%) et éluant B (ACN/TFA 0.1%). La détection a été effectuée à 220nm. Les peptides ont été purifiés sur une colonne préparative Xbridge C18-5μM (19\*50mm). Des analyses de spectrométrie de masse ont été effectuées sur un instrument doté d'une source d'ionisation MALDI en mode négatif et d'un analyseur à temps de vol (TOF).

**Ac-Pro<sub>1</sub>-Pro<sub>2</sub>-Gly-CONH<sub>2</sub> (P1)**


	Résine	Couplage Glycine (30 min)			Couplage Proline 1 (30 min)			Couplage Proline 2 (30 min)			Acétylé
		Gly	DIEA	HBTU	Pro	DIEA	HBTU	Pro	DIEA	HBTU	
n (mmol)	0.1	0.5	1	0.45	0.5	1	0.45	0.5	1	0.45	
Equivalent	1	5	10	4.5	5	10	4.5	5	10	4.5	
m <sup>46</sup>	196	150	X	170	176	X	170	174	X	170	
V (μl)	X	X	173	X	X	173	X	X	173	X	2.10 <sup>3</sup>

13.76 mg de poudre blanche correspondant au peptide 1 ont été récupérés après purification (R=44%).



Conditions expérimentales :  
5 à 20 % en ACN en 30 min

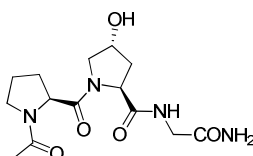
Conformère majoritaire (*trans-trans*)

<sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O, 298K): 1.93 (m, 1H, H<sub>β2</sub>-Pro<sub>1</sub>), 1.96 (m, 1H, H<sub>β2</sub>-Pro<sub>2</sub>), 1.99 (m, 2H, H<sub>γ</sub>-Pro<sub>1</sub>), 2.07 (m, 2H, H<sub>γ</sub>-Pro<sub>2</sub>), 2.31 (m, 1H, H<sub>β3</sub>-Pro<sub>2</sub>), 2.35 (m, 1H, H<sub>β3</sub>-Pro<sub>1</sub>), 3.65 (m, 2H, H<sub>δ</sub>-Pro<sub>1</sub>), 3.66 (m, 1H, H<sub>δ2</sub>-Pro<sub>2</sub>), 3.84 (m, 1H, H<sub>δ3</sub>-Pro<sub>2</sub>), 3.90 (m, 2H, H<sub>α</sub>-Gly), 4.44 (dd, J= 5.80, 8.40 Hz, 1H, H<sub>α</sub>-Pro<sub>2</sub>), 4.70 (dd, J= 4.95, 8.75 Hz, 1H, H<sub>α</sub>-Pro<sub>1</sub>); <sup>13</sup>C NMR (500 MHz, H<sub>2</sub>O:D<sub>2</sub>O (90:10), 298K): 7.11(s, 1H, H-CONH<sub>2</sub>), 7.43(s, 1H, H-CONH<sub>2</sub>), 8.55 (s, 1H, NH-Gly); <sup>13</sup>C NMR (125.8 MHz, D<sub>2</sub>O, 298K): 27.0 (CH<sub>2</sub>, C<sub>γ</sub>-Pro<sub>1</sub>), 27.5 (CH<sub>2</sub>, C<sub>γ</sub>-Pro<sub>2</sub>), 31.4 (CH<sub>2</sub>, C<sub>β</sub>-Pro<sub>1</sub>), 31.8 (CH<sub>2</sub>, C<sub>β</sub>-Pro<sub>2</sub>), 44.9 (CH, C<sub>α</sub>-Gly), 50.5 (CH<sub>2</sub>, C<sub>δ</sub>-Pro<sub>2</sub>), 51.4 (CH<sub>2</sub>, C<sub>δ</sub>-Pro<sub>1</sub>), 61.1 (CH, C<sub>α</sub>-Pro<sub>1</sub>), 63.4 (CH, C<sub>α</sub>-Pro<sub>2</sub>)

Conformère minoritaire (*trans-cis*)

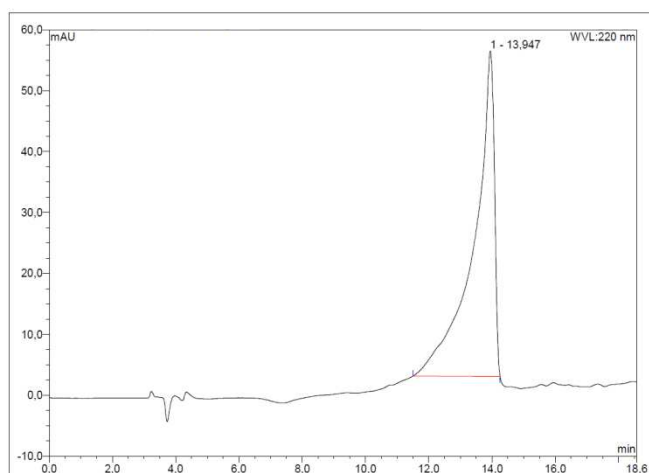
<sup>1</sup>H NMR (500 MHz, D<sub>2</sub>O, 298K): 1.85 (m, 1H, H<sub>γ2</sub>-Pro<sub>1</sub>), 1.96 (m, 1H, H<sub>γ3</sub>-Pro<sub>1</sub>), 1.96 (m, 1H, H<sub>β2</sub>-Pro<sub>2</sub>), 2.07 (m, 2H, H<sub>γ</sub>-Pro<sub>2</sub>), 2.12 (m, 1H, H<sub>β2</sub>-Pro<sub>1</sub>), 2.31 (m, 1H, H<sub>β3</sub>-Pro<sub>2</sub>), 2.44 (m, 1H, H<sub>β3</sub>-Pro<sub>1</sub>), 3.51 (m, 2H, H<sub>δ</sub>-Pro<sub>1</sub>), 3.66 (m, 1H, H<sub>δ2</sub>-Pro<sub>2</sub>), 3.84 (m, 1H, H<sub>δ3</sub>-Pro<sub>2</sub>), 3.90 (m, 2H, H<sub>α</sub>-Gly), 4.49 (dd, J= 6.08, 8.38 Hz, 1H, H<sub>α</sub>-Pro<sub>2</sub>), 4.90 (dd, J= 3.05, 8.95 Hz, 1H, H<sub>α</sub>-Pro<sub>1</sub>); <sup>1</sup>H NMR (500 MHz, H<sub>2</sub>O:D<sub>2</sub>O (90:10), 278K): 7.10(s, 1H, H-CONH<sub>2</sub>), 7.59(s, 1H, H-CONH<sub>2</sub>), 8.80 (s, 1H, NH-Gly); <sup>13</sup>C NMR (125.8MHz, D<sub>2</sub>O, 298K): 25.2 (CH<sub>2</sub>, C<sub>γ</sub>-Pro<sub>1</sub>), 27.5 (CH<sub>2</sub>, C<sub>γ</sub>-Pro<sub>2</sub>), 31.8 (CH<sub>2</sub>, C<sub>β</sub>-Pro<sub>2</sub>), 32.9 (CH<sub>2</sub>, C<sub>β</sub>-Pro<sub>1</sub>), 44.9 (CH, C<sub>α</sub>-Gly), 49.9 (CH<sub>2</sub>, C<sub>δ</sub>-Pro<sub>1</sub>), 50.5 (CH<sub>2</sub>, C<sub>δ</sub>-Pro<sub>2</sub>), 62.9 (CH, C<sub>α</sub>-Pro<sub>1</sub>), 63.6 (CH, C<sub>α</sub>-Pro<sub>2</sub>)

**Ac-Pro-Hyp-Gly-CONH<sub>2</sub> (P2)**



	Résine	Couplage Glycine (30 min)			Couplage Hydroxyproline (60min)			Couplage Proline (30 min)			Acétyl
		Gly	DIEA	HBTU	Hyp	DIEA	HATU	Pro	DIEA	HBTU	
n (mmol)	0.1	0.5	1	0.45	0.2	0.4	0.18	0.5	1	0.45	
Equivalent	1	5	10	4.5	2	4	1.8	5	10	4.5	
m <sup>46</sup>	194	146	X	171	82	X	68	159	X	170	
V (μl)	X	X	173	X	X	69	X	X	173	X	2.10 <sup>3</sup>

29.03 mg de poudre blanche ont été récupérées après purification (R=89%)



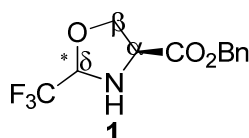
Conditions expérimentales :  
0 à 10 % en ACN en 30 min

Conformère majoritaire (*trans-trans*)

$^1\text{H}$  NMR (500 MHz,  $\text{D}_2\text{O}$ , 298K): 1.90 (m, 1H,  $\text{H}_{\beta 2}\text{-Pro}$ ), 1.99 (m, 1H,  $\text{H}_{\gamma}\text{-Pro}$ ), 2.10 (m, 1H,  $\text{H}_{\beta 2}\text{-Hyp}$ ), 2.31 (m, 1H,  $\text{H}_{\beta 3}\text{-Pro}$ ), 2.35 (m, 1H,  $\text{H}_{\beta 3}\text{-Hyp}$ ), 3.63 (m, 2H,  $\text{H}_{\delta}\text{-Pro}$ ), 3.82 (m, 1H,  $\text{H}_{\delta 2}\text{-Hyp}$ ), 3.86 (m, 1H,  $\text{H}_{\alpha}\text{-Gly}$ ), 3.88 (m, 1H,  $\text{H}_{\delta 3}\text{-Hyp}$ ), 3.95 (m, 1H,  $\text{H}_{\alpha}\text{-Gly}$ ), 4.57 (dd,  $J=7.76, 9.38$  Hz, 1H,  $\text{H}_{\alpha}\text{-Hyp}$ ), 4.63 (m, 1H,  $\text{H}_{\gamma}\text{-Hyp}$ ), 4.70 (dd,  $J=4.73, 8.78$  Hz, 1H,  $\text{H}_{\alpha}\text{-Pro}$ );  $^1\text{H}$  NMR (500 MHz,  $\text{H}_2\text{O}:\text{D}_2\text{O}$  (90:10), 278K): 7.13(s, 1H, H-CONH<sub>2</sub>), 7.43(s, 1H, H-CONH<sub>2</sub>), 8.73 (s, 1H, NH-Gly);  $^{13}\text{C}$  NMR (125.8MHz,  $\text{D}_2\text{O}$ , 298K): 27.0 ( $\text{CH}_2$ ,  $\text{C}_{\gamma}\text{-Pro}$ ), 31.2 ( $\text{CH}_2$ ,  $\text{C}_{\beta}\text{-Pro}$ ), 39.5 ( $\text{CH}_2$ ,  $\text{C}_{\beta}\text{-Hyp}$ ), 44.9 (CH,  $\text{C}_{\alpha}\text{-Gly}$ ), 51.4 ( $\text{CH}_2$ ,  $\text{C}_{\delta}\text{-Pro}$ ), 57.9 ( $\text{CH}_2$ ,  $\text{C}_{\delta}\text{-Hyp}$ ), 61.1 (CH,  $\text{C}_{\alpha}\text{-Pro}$ ), 62.1 (CH,  $\text{C}_{\alpha}\text{-Hyp}$ ), 72.8 (CH,  $\text{C}_{\gamma}\text{-Hyp}$ )

Conformère minoritaire (*trans-cis*)

$^1\text{H}$  NMR (500 MHz,  $\text{D}_2\text{O}$ , 298K): 1.81 (m, 1H,  $\text{H}_{\gamma 2}\text{-Pro}$ ), 1.96 (m, 1H,  $\text{H}_{\gamma 3}\text{-Pro}$ ), 2.08 (m, 1H,  $\text{H}_{\beta 2}\text{-Pro}$ ), 2.1 (m, 1H,  $\text{H}_{\beta 2}\text{-Hyp}$ ), 2.35 (m, 1H,  $\text{H}_{\beta 3}\text{-Hyp}$ ), 2.41 (m, 1H,  $\text{H}_{\beta 3}\text{-Pro}$ ), 3.50 (m, 2H,  $\text{H}_{\delta}\text{-Pro}$ ), 3.82 (m, 1H,  $\text{H}_{\delta 2}\text{-Hyp}$ ), 3.86 (m, 1H,  $\text{H}_{\alpha}\text{-Gly}$ ), 3.88 (m, 1H,  $\text{H}_{\delta 3}\text{-Hyp}$ ), 3.95 (m, 1H,  $\text{H}_{\alpha}\text{-Gly}$ ), 4.61 (dd,  $J= / /$  Hz, 1H,  $\text{H}_{\alpha}\text{-Hyp}$ ), 4.63 (m, 1H,  $\text{H}_{\gamma}\text{-Hyp}$ ), 4.92 (dd,  $J=3.05, 8.80$  Hz, 1H,  $\text{H}_{\alpha}\text{-Pro}$ );  $^1\text{H}$  NMR (500 MHz,  $\text{H}_2\text{O}:\text{D}_2\text{O}$  (90:10), 278K): 7.13(s, 1H, H-CONH<sub>2</sub>), 7.43(s, 1H, H-CONH<sub>2</sub>), 8.73 (s, 1H, NH-Gly);  $^{13}\text{C}$  NMR (125.8MHz,  $\text{D}_2\text{O}$ , 298K): 25.1 ( $\text{CH}_2$ ,  $\text{C}_{\gamma}\text{-Pro}$ ), 32.7 ( $\text{CH}_2$ ,  $\text{C}_{\beta}\text{-Pro}$ ), 39.5 ( $\text{CH}_2$ ,  $\text{C}_{\beta}\text{-Hyp}$ ), 44.9 (CH,  $\text{C}_{\alpha}\text{-Gly}$ ), 49.8 ( $\text{CH}_2$ ,  $\text{C}_{\delta}\text{-Pro}$ ), 57.9 ( $\text{CH}_2$ ,  $\text{C}_{\delta}\text{-Hyp}$ ), 62.3 (CH,  $\text{C}_{\alpha}\text{-Hyp}$ ), 62.9 (CH,  $\text{C}_{\alpha}\text{-Pro}$ ), 72.8 (CH,  $\text{C}_{\gamma}\text{-Hyp}$ ),

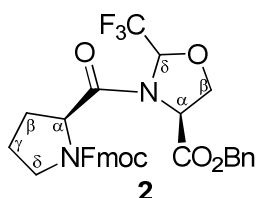
A2.Synthèse en phase liquide**Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OBn **1****

A une solution de chlorhydrate d'ester benzylique de (*S*)-sérine (1.4 g, 7.2 mmol, 1 equiv) et de toluène (24 mL) est ajoutée de l'hémiacétal éthylique de trifluoracétaldéhyde (1.08 g, 7.53 mmol, 1.05 equiv) et du PPTS (181 mg, 0.72 mmol, 0.1 equiv) à température ambiante. Le mélange résultant a été agité à 90°C pendant 1h, puis chauffé au reflux en utilisant un appareil de Dean-Stark pendant 3h à 140°C. Le mélange réactionnel a été filtré et évaporé sous pression réduite pour donner 1,25 g (63%) d'un mélange de **1a** et **1b**.

(*S,S*)-**1**:  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.39 (dd,  $J=8.2, 7.3$  Hz, 1 H, NH), 3.82 (dd,  $J=7.8, 6.4$  Hz, 1 H), 4.09 (ddd,  $J=7.8, 7.3, 6.4$  Hz, 1 H), 4.26 (t,  $J=7.8$  Hz, 1 H), 5.07 (dq,  $J=8.2, 5.5$  Hz, 1 H), 5.17 (d,  $J=11.9$  Hz, 1 H), 5.23 (d,  $J=11.9$  Hz, 1 H), 7.31-7.41 (m, 5H);  $^{13}\text{C}$  NMR (100.5 MHz,  $\text{CDCl}_3$ )  $\delta$  58.6, 67.8, 69.6, 87.7 (q,  $J=34.5$  Hz), 123.2 (q,  $J=283.7$  Hz), 128.4, 128.7, 128.8, 134.7, 171.2;  $^{19}\text{F}$  NMR (376.2 MHz,  $\text{CDCl}_3$ )  $\delta$  -85.0 (d,  $J=5.5$  Hz)

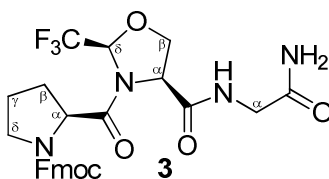
**(R,S)-1:**  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ )  $\delta$  3.15 (m, 1 H, NH), 4.10–4.20 (m, 3 H), 4.91 (dq,  $J = 7.8, 5.0$  Hz, 1 H), 5.18 (s, 2 H), 7.34–7.37 (m, 5H);  $^{13}\text{C}$  NMR (100.5 MHz,  $\text{CDCl}_3$ )  $\delta$  58.8, 67.5, 68.9, 87.8 (q,  $J = 34.5$  Hz), 122.8 (q,  $J = 283.7$  Hz), 128.4, 128.6, 128.7, 135.0, 170.4;  $^{19}\text{F}$  (376.2 MHz,  $\text{CDCl}_3$ )  $\delta$  - 84.1 (d,  $J = 5.2$  Hz).

### Fmoc-Pro-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OBn **2**



A une solution de Fmoc-L-proline (675 mg, 2 mmol, 1 équivalent) dans du dichlorométhane (11 ml), nous avons ajouté du chlorure de thionyle (2 ml, 26 mmol, 13 équiv). Le mélange résultant a été traité par ultrasons pendant 45 min. Le solvant a été évaporé et la Fmoc-Pro-Cl correspondante (1.1 équivalents) a été solubilisée dans du dichlorométhane (8 ml). Le mélange des oxazolidines **1** (500 mg, 1,82 mmol, 1 équivalent) a été ajouté via une canule et le mélange résultant a été agité à température ambiante pendant une nuit puis évaporé. Le produit brut a été purifié par flash chromatographie (cyclohexane / acétate d'éthyle 80:20) pour donner 660 mg (57%) de dipeptide pur Fmoc-Pro-Ser ( $\Psi^{\text{CF}_3, \text{H}}$ Pro) -OBn **2** caractérisé comme un isomère *trans* dans  $\text{CDCl}_3$  rotation à 300 K :  $R_f = 0.63$  (cyclohexane/ethyl acetate 60:40);  $[\alpha]_{\text{D}}^{23} -51.9$  ( $c$  1.35,  $\text{CHCl}_3$ ); IR (neat) 2954, 1741, 1686  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (500 MHz,  $\text{CDCl}_3$ )  $\delta$  1.90–2.06 (m, 1 H,  $\text{H}_\gamma$  Pro-Ha), 2.1–2.19 (m, 2 H,  $\text{H}_\beta$  Pro-Ha), 2.23–2.33 (m, 1 H,  $\text{H}_\gamma$  Pro-Hb), 3.51–3.57 (m, 1 H,  $\text{H}_\delta$  Pro-Ha), 3.67–3.73 (m, 1 H,  $\text{H}_\delta$  Pro-Hb), 4.23 (t,  $J = 7.33$  Hz, 1 H, Fmoc CH), 4.32 (dd,  $J = 8.5, 6.9$  Hz, 1 H,  $\text{H}_\beta$   $\Psi$ Pro-Ha), 4.34 (d,  $J = 7.1$  Hz, 2 H Fmoc  $\text{CH}_2$ ), 4.36–4.4 (m, 1 H,  $\text{H}_\alpha$  Pro-H), 4.50 (t,  $J = 8.5$  Hz, 1 H,  $\text{H}_\beta$   $\Psi$ Pro-Hb), 5.1 (t,  $J = 7.8$  Hz, 1 H,  $\text{H}_\alpha$   $\Psi$ Pro-H), 5.16 (s, 2 H, OBn), 6.21 (q,  $J = 4.8$  Hz, 1 H,  $\text{H}_\delta$   $\Psi$ Pro-H), 7.29 (s, 5 H, OBn), 7.32 (t,  $J = 7.3$  Hz, 2 H, Fmoc arom), 7.40 (t,  $J = 7.3$  Hz, 2 H, Fmoc arom), 7.58 (t,  $J = 9.1$  Hz, 2 H, Fmoc arom), 7.76 (d,  $J = 7.8$  Hz, 2 H, Fmoc arom);  $^{13}\text{C}$  (100.5 MHz,  $\text{CDCl}_3$ )  $\delta$  24.9 ( $\text{CH}_2$ ,  $\text{C}_\gamma$  Pro), 27.0, 30.3 ( $\text{CH}_2$ ,  $\text{C}_\beta$  Pro), 47.1 ( $\text{CH}_2$ ,  $\text{C}_\delta$  Pro), 47.2 (CH, Fmoc CH), 57 (CH,  $\text{C}_\alpha$  Pro), 58.5 (CH,  $\text{C}_\alpha$   $\Psi$ Pro), 67.7 ( $\text{CH}_2$ , Fmoc  $\text{CH}_2$ ), 67.8 ( $\text{CH}_2$ , OBn), 69 ( $\text{CH}_2$ ,  $\text{C}_\beta$   $\Psi$ Pro), 85.1 (q,  $J = 30.7$  Hz, CH  $\text{C}_\delta$   $\Psi$ Pro), 120.1 (2  $\times$  CH, Fmoc arom), 125.2 (2  $\times$  CH, Fmoc arom), 127.1 (2  $\times$  CH, Fmoc arom), 127.8 (2  $\times$  CH, Fmoc arom), 141.4 (2  $\times$  C, Fmoc arom), 143.9 (2  $\times$  C, Fmoc arom), 155.4 (C, C=O), 168.2 (C, C=O), 172.9 (C, C=O);  $^{19}\text{F}$  NMR (376.2 MHz,  $\text{CDCl}_3$ )  $\delta$  - 81.8 (d,  $J = 4.3$  Hz), - 81.5 (d,  $J = 4.3$  Hz), - 85 (d,  $J = 5.2$  Hz).

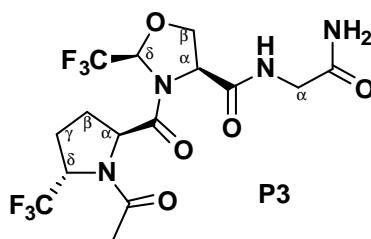
### Fmoc-Pro-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-Gly-NH<sub>2</sub> **3**



*Préparation du Fmoc-Pro-Ser( $\Psi$ Pro<sup>CF<sub>3</sub>,H</sup>)-OH* - A une solution de dipeptide **2** (680 mg, 1,17 mmol) dans du méthanol (7 ml) a été ajouté du palladium sur charbon 10% (10% p / p, 68 mg) sous une atmosphère de H<sub>2</sub> à température ambiante pendant 1h. Le mélange réactionnel a été ensuite filtré sur Büchner, lavé avec du méthanol et évaporé sous pression réduite pour donner l'acide correspondant. Le solide a été repris avec du pentane et de l'éther diéthylique a été ajouté jusqu'à ce que l'acide précipite. Le solide est ensuite séché pour donner 630 mg (quantitatif) de l'acide correspondant utilisé directement sans autre purification.

*Réaction de couplage* - A une solution de Fmoc-Pro-Ser ( $\Psi$ CF<sub>3</sub>, HPro)-OH (185 mg, 0,367 mmol, 1 équivalent) dans du dichlorométhane (33 ml), nous avons ajouté du chlorhydrate d'amide de glycine (45 mg, 0,40 mmol, 1,1 équiv), Et<sub>3</sub>N (0,16 ml, 1,14 mmol, 3,1 équiv). Après agitation pendant 10 min à température ambiante, BOP-Cl (117 mg, 0,46 mmol, 1,25 équivalent) a été ajoutée. Le mélange réactionnel a été agité pendant une nuit à température ambiante, puis désactivé avec HCl 1 M. Les phases ont été séparées et la phase aqueuse a été extraite avec du dichlorométhane (3 x 30 ml). Les phases organiques combinées ont été lavées avec une solution saturée de NaHCO<sub>3</sub> aqueux (2 x 10 ml), NaCl (10 ml) puis séchée sur MgSO<sub>4</sub>, filtrée et concentrée sous pression réduite. Une purification par chromatographie flash (95:5 de dichlorométhane/méthanol) a donné 70 mg (34%) du tripeptide pur **3** sous forme de mélange inséparable d'isomères 69/31 dans CDCl<sub>3</sub> à 298 K : *R<sub>f</sub>* = 0.3 (dichloromethane/methanol 95:5); IR (neat) 3325, 1675 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>)  $\delta$  1.85–2.35 (m, 4 H, 2  $\times$  H $\gamma$ , 2  $\times$  H $\beta$  Pro-H), 3.5–3.77 (m, 2 H, H $\delta$  Pro-H), 3.87–3.95 (m, 2 H, H $\alpha$  Gly-H), 4.19–4.24 (m, 1 H, Fmoc CH), 4.28–4.60 (m, 4 H, Fmoc CH<sub>2</sub>, H $\alpha$  Pro-H, H $\beta$   $\Psi$ Pro-Ha), 4.65–4.75 (m, 1 H, H $\beta$   $\Psi$ Pro-Hb), 4.85–4.95 (m, 1 H, H $\alpha$   $\Psi$ Pro-H), 5.5–6.4 (m, 3 H, H $\delta$   $\Psi$ Pro-H, NH<sub>2</sub> Gly-H), 7.1–7.25 (m, 1 H, NH Gly-H *mino*), 8.6–8.8 (m, 1 H, NH Gly-H *majo*); <sup>19</sup>F NMR (376.2 MHz, CDCl<sub>3</sub>)  $\delta$  – 81.7 and – 81.5.

### Ac-Pro-Ser( $\Psi$ <sup>CF<sub>3</sub>,H</sup>Pro)-Gly-NH<sub>2</sub> **P3**



*Déprotection du groupement Fmoc du tripeptide 3* - A une solution de tripeptide **3** (70 mg, 0,125 mmol, 1 équivalent) dans du dichlorométhane (3 ml), a été ajouté de la pipéridine (106 mg, 1,25 mmol, 10 équiv) à température ambiante. Le mélange résultant a été agité à température ambiante pendant 2h et a été évaporé sous pression réduite. Une purification par chromatographie sur gel de silice (95:5 de dichlorométhane/méthanol) a donné 40 mg du tripeptide correspondant directement utilisé dans l'étape suivante.

### *Synthèse du tripeptide Ac-Pro-Ser( $\Psi$ <sup>CF<sub>3</sub>,H</sup>Pro)-Gly-NH<sub>2</sub> **P3***

A une solution de tripeptide Pro-Ser( $\Psi$ CF<sub>3</sub>,HPro)-Gly-NH<sub>2</sub> (42 mg, 0,125 mmol, 1 équivalent) dans du dichlorométhane (4,8 ml) a été ajouté de l'anhydride acétique (13  $\mu$ L, 0,138 mmol, 1,1 équiv). Le mélange réactionnel a été agité à température ambiante pendant 1h, puis évaporé sous pression réduite. Le produit brut a été solubilisé dans de l'éther

diéthylique, filtré et évaporé sous pression réduite pour donner 48 mg (100%) de Ac-Pro-Ser ( $\Psi$ CF<sub>3</sub>, HPro)-Gly-NH<sub>2</sub> **P3** comme un mélange 1:1 *trans-trans* et *trans-cis* dans CDCl<sub>3</sub> rotation à 298 K. La présence de conformères minoritaires *cis-trans* et *cis-cis* a également été observée:  $R_f = 0.7$  (dichlorométhane/méthanol 90:10); IR (neat) 3300, 2950, 1735, 1678, 1627 cm<sup>-1</sup>; HRMS (ESI-TOF) calcd for C<sub>14</sub>H<sub>19</sub>F<sub>3</sub>N<sub>4</sub>O<sub>5</sub>Na 403.1205, found: 403.12036.

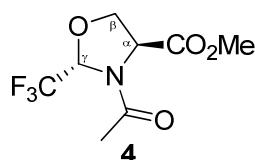
Conformère majoritaire (*trans-trans*)

<sup>1</sup>H NMR (500 MHz, H<sub>2</sub>O/D<sub>2</sub>O 9:1, 298 K)  $\delta$  2.01 (H $\beta$  Pro-Ha), 2.014 (H $\gamma$  Pro-Ha), 2.144 (H $\gamma$  Pro-Hb), 2.308 (H $\beta$  Pro-Hb), 3.662 (2 H, H $\delta$  Pro-H), 3.938 (d, 2 H, H $\alpha$  Gly-H), 4.307 (t, J = 8.9 Hz, 1 H, H $\beta$   $\Psi$ Pro-Ha), 4.591 (m, 1 H, H $\alpha$  Pro-H), 4.641 (m, 3 H, H $\beta$   $\Psi$ Pro-Hb), 4.869 (t, J = 8.5 Hz, 1 H, H $\alpha$   $\Psi$ Pro-H), 6.341 (dd, J = 5 Hz, 1 H, H $\delta$   $\Psi$ Pro-H), 7.228 (d, 1 H, NH<sub>2</sub>-Ha), 7.52 (d, 1H, NH<sub>2</sub>-Hb), 8.722 (s, 1 H, NH); <sup>13</sup>C NMR(500 MHz, H<sub>2</sub>O/D<sub>2</sub>O 9:1, 298 K)  $\delta$  27.639 (C $\gamma$  Pro), 32.527 (C $\beta$  Pro), 44.95 (C $\alpha$  Gly), 51.396 (C $\delta$  Pro), 61.325 (C $\alpha$  Pro) 61.843 (C $\alpha$   $\Psi$ Pro), 72.195 (C $\beta$   $\Psi$ Pro), 87.504 (C $\delta$   $\Psi$ Pro);

Conformère minoritaire (*trans-cis*)

<sup>1</sup>H NMR (500 MHz, H<sub>2</sub>O/D<sub>2</sub>O 9:1, 298 K)  $\delta$  1.965 (m, 1 H, H $\beta$  Pro-Ha), 2.011 (H $\gamma$  Pro-Ha), 2.143 (H  $\gamma$  Pro-Hb), 2.38 (dd, J = 6.5, 1 H, H $\beta$  Pro-Hb), 3.686 (s, 1 H, H $\delta$  Pro-Ha), 3.686 (s, 1 H, H $\delta$  Pro-Hb), 3.914 (d, 1 H, H $\alpha$  Gly-H), 4.107 (d, J = 14.2 Hz, 1 H, H $\alpha$  Gly-H), 4.713 (t, J = 7.4 Hz, 1 H, H $\alpha$  Pro-H)  $\delta$  4.52 (t, J = 8.2 Hz, 1 H, H $\beta$   $\Psi$ Pro-Ha), 4.639 (m, 1 H, H $\beta$   $\Psi$ Pro-Hb), 5.115 (1 H, H $\alpha$   $\Psi$ Pro-H), 6.046 (dd, J = 2.9 Hz, 1 H, H $\delta$   $\Psi$ Pro-H), 7.127 (d, 1 H, NH<sub>2</sub> Gly-Ha), 7.633 (d, 1 H, NH<sub>2</sub> Gly-Hb), 9.08 (s, 1 H, NH Gly-H). <sup>13</sup>C NMR(500 MHz, H<sub>2</sub>O/D<sub>2</sub>O 9:1, 298 K)  $\delta$  44.924 (C $\alpha$  Gly), 27.639 (C $\gamma$  Pro), 31.331 (C $\beta$  Pro), 51.396 (C $\delta$  Pro), 61.955 (C $\alpha$  Pro) 61.678 (C $\alpha$   $\Psi$ Pro), 73.269 (C $\beta$   $\Psi$ Pro), 87.345 (C $\delta$   $\Psi$ Pro).

**Ac-Ser( $\Psi$ <sup>CF<sub>3</sub>,H</sup>Pro)-OMe **4****

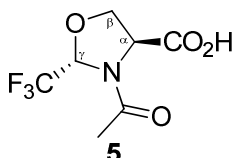


Une solution de (2S, 4S)-**1a** (2 g, 10 mmol, 1 équivalent) et d'iode (255 mg, 1 mmol, 0,1 équiv) dans de l'anhydride d'acide (9,4 ml, 100 mmol, 10 équivalents) a été agitée à température ambiante pendant 24h, suivie par analyse RMN <sup>19</sup>F. Par la suite, le dichlorométhane (40 ml) a été ajoutée et la phase organique a été lavée avec une solution aqueuse 1 M de NaHCO<sub>3</sub>. La phase aqueuse a été extraite avec du dichlorométhane (3 x 40 ml) et les extraits organiques combinés ont été séchés sur du MgSO<sub>4</sub>, filtrés et concentrés sous pression réduite. La purification du produit brut par chromatographie flash (cyclohexane/acétate d'éthyle 70:30) a donné 1,68 g (69%) d'un mélange inséparable d'isomères *cis/trans* oxazolidine acétylé (2S, 4S)-**4** dans CDCl<sub>3</sub> à 298 K: huile incolore;  $R_f = 0.22$  (70:30 cyclohexane/ethyl acetate);  $[\alpha]_D^{23} -53.8$  (c 1.7, CHCl<sub>3</sub>); IR (neat) 2960, 1749, 1678, 1384, 1342, 1281, 1206, 1175, 1147, 1118, 943, 844 cm<sup>-1</sup>; <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>) (*trans*)  $\delta$  2.21 (s, 3 H), 3.78 (s, 3 H), 4.24 (d, J = 9.6 Hz, 1 H), 4.50 (dd, J = 8.24, 9.6 Hz, 1 H), 4.55–4.63 (m, 1 H), 5.60 (q, J = 4.6 Hz, 1 H); (*cis*)  $\delta$  2.06 (s, 3 H), 3.85 (s, 3 H), 4.38 (d, J = 7.8 Hz, 1 H), 4.55–4.65 (m, 2 H), 5.87 (q, J = 4.6 Hz, 1 H); <sup>13</sup>C NMR (100.5 MHz, CDCl<sub>3</sub>)



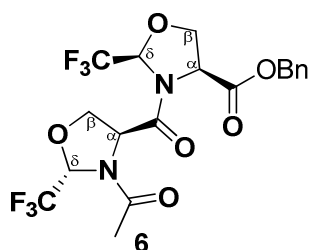
(*trans*)  $\delta$  22.3, 52.7, 58.1, 70.5, 84.8 (q,  $J = 35.4$  Hz), 122.9 (q,  $J = 287.5$  Hz), 168.9, 169.9; (*cis*)  $\delta$  22.7, 53.3, 59.0, 72.1, 84.6 (q,  $J = 34.5$  Hz), 122.9 (q,  $J = 287.5$  Hz), 170.0, 170.2;  $^{19}\text{F}$  NMR (376.2 MHz,  $\text{CDCl}_3$ ) (*trans*)  $\delta$  -81.7 (d,  $J = 4.3$  Hz) (*cis*)  $\delta$  -81.4 (d,  $J = 4.3$  Hz); MS (EI)  $m/z = 241$  [ $\text{M}^+$ ], 182 [ $\text{M}^+ - \text{CO}_2\text{Me}$ ] (50), 172, 140 (100), 130, 112; HRMS (EI) calcd for  $\text{C}_8\text{H}_{10}\text{F}_3\text{NO}_4$  241.0562, found 241.0561.

### Ac-Ser ( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OH **5**



A une solution de (2S, 4S)-**4** (1,61 g, 6,68 mmol, 1 équivalent) dans du THF (38 ml) à 0 ° C, a été ajoutée une solution aqueuse de LiOH 2M (3,8 ml, 7,34 mmol, 1,1 équiv). Le mélange réactionnel a été agité à température ambiante pendant 3 h puis évaporé sous pression réduite. De l'éther diéthylique a été ajouté et le mélange réactionnel a été extrait avec de l'eau. Les phases aqueuses ont été combinées, traitées avec une solution de HCl 1M et le produit a été extrait avec 3 x 40 ml d'acétate d'éthyle, séché sur du  $\text{MgSO}_4$ , filtré et évaporé sous pression réduite pour donner 1,316 g (86%) (2S, 4S)-**5** d'un mélange inséparable d'isomères *cis/trans*.  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ) (*majo*)  $\delta$  1.88 (s, 3 H), 4.15 (dd,  $J = 1.37, 7.8$  Hz, 1 H), 4.34–4.44 (m, 2 H), 5.72 (q,  $J = 5.0$  Hz, 1 H) (*mino*)  $\delta$  2.03 (s, 3 H), 4.22 (dd,  $J = 8.2, 2.3$  Hz, 1 H), 4.34–4.44 (m, 2 H), 5.82 (q,  $J = 5.04$  Hz, 1 H);  $^{19}\text{F}$  NMR (376.2 MHz,  $\text{CDCl}_3$ ) (*majo*)  $\delta$  -81.2 (d,  $J = 5.2$  Hz), -81.4 (d,  $J = 5.2$  Hz) (*mino*)  $\delta$  -81.6 (d,  $J = 4.3$  Hz), -81.4 (d,  $J = 4.3$  Hz).

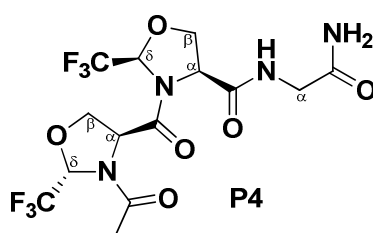
### Ac-Ser ( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-Ser ( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OBn **6**



A une solution de (2S, 4S)-**5** (450 mg, 1,98 mmol, 1 équivalent) dans du dichlorométhane (12 ml) a été ajouté du chlorure de thionyle (1,9 ml, 25,8 mmol, 13 équiv). Le mélange résultant a été traité par ultrasons pendant 1h30. Le solvant a été évaporé et le chlorure d'acyle correspondant (1,1 équiv) a été solubilisé dans du dichlorométhane (15 ml). L'oxazolidine **1b** (495 mg, 1,8 mmol, 1 équivalent) a été ajouté via une canule et le mélange résultant a été agité à température ambiante pendant une nuit puis évaporé. Le produit brut a été purifié par chromatographie flash (cyclohexane/acétate d'éthyle 60:40) pour donner 395 mg (41%) de dipeptide pur **6**, isomère *trans*, dans  $\text{CDCl}_3$  à 300 K :  $R_f = 0.4$  (cyclohexane:ethyl acetate 60/40)  $[\alpha]_{\text{D}}^{23} -46.36$  ( $c$  1.25,  $\text{CHCl}_3$ ); IR (neat) 2924, 2854, 1736, 1698, 1670  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR

(400 MHz, CDCl<sub>3</sub>, 298 K)  $\delta$  2.15 (s, 3 H, Ac), 4.08–4.11 (m, 2 H, H $\alpha$   $\Psi$ Pro1-Ha and H $\alpha$   $\Psi$ Pro2-Ha), 4.33 (d, 2 H,  $J$  = 7.79 Hz, H $\beta$   $\Psi$ Pro1-Ha and H $\beta$   $\Psi$ Pro2-Ha), 4.51 (d, 2 H,  $J$  = 7.79 Hz, H $\beta$   $\Psi$ Pro1-Hb and H $\beta$   $\Psi$ Pro2-Hb), 5.05–5.19 (m, 2 H, H $\alpha$   $\Psi$ Pro1-Hb and H $\alpha$   $\Psi$ Pro2-Hb), 5.16 (s, 2 H, OBn), 5.62 (q, 1 H,  $J$  = 5.04 Hz, H $\delta$   $\Psi$ Pro2-H), 5.96 (q, 1 H,  $J$  = 4.58 Hz, H $\delta$   $\Psi$ Pro1-H), 7.23–7.38 (m, 5 H, OBn); <sup>13</sup>C NMR (100.5 MHz, CDCl<sub>3</sub>, 298 K)  $\delta$  57.1 (CH, C $\alpha$   $\Psi$ Pro1), 57.5 (CH, C $\alpha$   $\Psi$ Pro2), 67.8 (CH<sub>2</sub>, C $\beta$   $\Psi$ Pro2), 69.1 (CH<sub>2</sub>, C $\beta$   $\Psi$ Pro1), 70.9 (CH<sub>2</sub>, C $\beta$   $\Psi$ Pro2), 85.2 (q, CH,  $J$  = 18.2 Hz, C $\delta$   $\Psi$ Pro1), 85.3 (q, CH,  $J$  = 19.1 Hz, C $\delta$   $\Psi$ Pro2), 121.4 (q, C,  $J$  = 36.2 Hz, CF<sub>3</sub>), 124.5 (q, C,  $J$  = 40.5 Hz, CF<sub>3</sub>), 128.5 (C, OBn), 134.8 (C, OBn), 167.8 (C, C=O), 169.5 (C, C=O), 169.7 (C, C=O); <sup>19</sup>F NMR (376.5 MHz, CDCl<sub>3</sub>, 298 K)  $\delta$  – 81.9 (d,  $J$  = 4.33 Hz, CF<sub>3</sub>), – 82.1 (d,  $J$  = 4.33 Hz, CF<sub>3</sub>).

### Ac-Ser( $\Psi$ <sup>CF<sub>3</sub>,H</sup>Pro)-Ser( $\Psi$ <sup>CF<sub>3</sub>,H</sup>Pro)-Gly-NH<sub>2</sub> P4



*Déprotection du groupement Bn du tripeptide 6* - A une solution de dipeptide **6** (355 mg, 0,7 mmol) dans du méthanol (6 ml), a été ajouté du palladium sur charbon 10% (10% p/p, 36 mg) sous une atmosphère de H<sub>2</sub> à température ambiante pendant 1h. Le mélange réactionnel a ensuite été filtré sur Büchner, lavé avec du méthanol et évaporé sous pression réduite pour donner l'acide correspondant. Le solide a été repris avec du pentane. De l'éther diéthylique a été ajouté jusqu'à ce que l'acide précipite. Le solide est ensuite séché pour donner 281 mg (quantitatif) de l'acide correspondant et a été utilisé sans autre purification.

*Couplage peptidique* - A une solution de dipeptide Ac-Ser ( $\Psi$ CF<sub>3</sub>, HPro) -Ser ( $\Psi$ CF<sub>3</sub>, HPro) (280 mg, 0,71 mmol, 1 équivalent) dans du dichlorométhane (66 ml), a été ajouté du chlorhydrate de l'amide de glycine (86 mg, 0,78 mmol, 1,1 equiv), Et<sub>3</sub>N (340  $\mu$ L, 2,4 mmol, 3,1 equiv). Après agitation pendant 10 min à température ambiante, BOP-Cl (226 mg, 0,89 mmol, 1,25 équivalent) a été ajouté. Le mélange réactionnel a été agité pendant une nuit à température ambiante, puis désactivé avec 1M de HCl. Les phases ont été séparées et la phase aqueuse a été extraite avec du dichlorométhane (3 x 30 ml). Les phases organiques combinées ont été lavées avec une solution saturée de NaHCO<sub>3</sub> aqueux (2 x 10 ml), NaCl (10 ml), puis séchées sur du MgSO<sub>4</sub>, filtrées et concentrées sous pression réduite. La purification par chromatographie flash (95:5 de dichlorométhane/méthanol) a donné 54 mg (34%) du tripeptide pur **P4** sous forme de deux isomères *trans-trans* et *cis-trans* (65:20) CDCl<sub>3</sub> à 298 K. La présence d'isomères minoritaires *cis-trans* et *cis-cis* a également été observée dans CDCl<sub>3</sub> à 298 K;  $R_f$  = 0.4 (dichlorométhane/méthanol 90:10); IR (neat) 3320, 2010, 1653 cm<sup>-1</sup>; HRMS (ESI-TOF) calcd for C<sub>14</sub>H<sub>16</sub>F<sub>6</sub>N<sub>4</sub>O<sub>6</sub>Na 473.0872, found: 473.08615.

#### Conformère majoritaire (*trans-trans*)

<sup>1</sup>H NMR (500 MHz, H<sub>2</sub>O/D<sub>2</sub>O 9:1, 298 K)  $\delta$  3.937 (d, 2 H, H $\alpha$  Gly-H), 4.419 (H $\beta$   $\Psi$ Pro(2*S*,4*S*)-Ha), 4.326 (H $\beta$   $\Psi$ Pro(2*R*,4*S*)-Ha), 4.42 (H $\beta$   $\Psi$ Pro(2*R*,4*S*)-Hb), 4.589 (H $\beta$   $\Psi$ Pro(2*S*,4*S*)-Hb), 4.865 (H $\alpha$   $\Psi$ Pro(2*R*,4*S*)-H), 5.062 (m, 2H, H $\alpha$   $\Psi$ Pro(2*S*,4*S*)-H), 6.137 (dd,

$J = 4.85$  Hz, 1 H,  $H\delta$   $\Psi$ Pro(2*S*,4*S*)-H), 6.349 (dd,  $J = 4.8$  Hz, 1H,  $H\delta$   $\Psi$ Pro(2*R*,4*S*)-H), 8.8933 (t,  $J = 5.9$  Hz, 1 H, NH Gly-H);  $^{13}\text{C}$  NMR(500 MHz,  $\text{H}_2\text{O}/\text{D}_2\text{O}$  9:1, 298 K)  $\delta$  59.419 (C $\alpha$   $\Psi$ Pro(2*S*,4*S*)), 61.785 (C $\alpha$   $\Psi$ Pro(2*R*,4*S*)), 72.424 (C $\beta$   $\Psi$ Pro(2*R*,4*S*)), 73.605 (C $\beta$   $\Psi$ Pro(2*S*,4*S*)), 87.456 (C $\delta$   $\Psi$ Pro(2*S*,4*S*)), 87.456 (C $\delta$   $\Psi$ Pro(2*R*,4*S*))

### Conformère minoritaire (*trans-cis*)

$^1\text{H}$  NMR (500 MHz,  $\text{H}_2\text{O}/\text{D}_2\text{O}$  9:1, 298 K)  $\delta$  4.478 ( $H\beta$   $\Psi$ Pro(2*R*,4*S*)-Ha), 4.5 (m, 1 H,  $H\beta$   $\Psi$ Pro(2*S*,4*S*)-Ha), 4.56 (1H,  $H\beta$   $\Psi$ Pro(2*S*,4*S*)-Hb), 4.636 ( m, 1 H,  $H\beta$   $\Psi$ Pro(2*R*,4*S*)-Hb), 5.065 (H $\alpha$   $\Psi$ Pro(2*S*,4*S*)-H), 5.148 (H $\alpha$   $\Psi$ Pro(2*R*,4*S*)-H), 8.8116 (t,  $J = 6$  Hz, 1 H, NH Gly-H);  $^{13}\text{C}$  NMR(500 MHz,  $\text{H}_2\text{O}/\text{D}_2\text{O}$  9:1, 298 K) (*trans-cis*)  $\delta$  59.853 (C $\alpha$   $\Psi$ Pro(2*S*,4*S*)), 61.52 (C $\alpha$   $\Psi$ Pro(2*R*,4*S*)), 73.586 (C $\beta$   $\Psi$ Pro(2*S*,4*S*)), 73.03 (C $\beta$   $\Psi$ Pro(2*R*,4*S*)).

## B. Calculs des populations *up/down*

Les calculs des constantes de couplage par des méthodes quantiques sont résumés dans les tableaux suivants. Les constantes de couplage  $JH^\nu H^\delta$  permettent également d'avoir accès au plissement de cycle.

<b>Proline</b>			<i>cis</i>	<i>trans</i>
<i>up</i>	$H^\alpha$	$H^\alpha-H^{\beta 2} (\chi 1-120)$	8,19	8,28
		$H^\alpha-H^{\beta 3} (-\chi 1)$	7,22	7,16
	$H^{\delta 2}$	$H^{\gamma 2}-H^{\delta 2} (-\chi 3)$	7,40	7,47
		$H^{\gamma 3}-H^{\delta 2} (120-\chi 3)$	9,65	9,57
	$H^{\delta 3}$	$H^{\gamma 2}-H^{\delta 3} (120+\chi 3)$	0,98	0,98
		$H^{\gamma 3}-H^{\delta 3} (-\chi 3)$	7,40	7,47
<i>down</i>	$H^\alpha$	$H^\alpha-H^{\beta 2} (\chi 1-120)$	2,38	2,43
		$H^\alpha-H^{\beta 3} (\chi 1)$	6,53	6,75
	$H^{\delta 2}$	$H^{\gamma 2}-H^{\delta 2} (\chi 3)$	8,53	8,28
		$H^{\gamma 3}-H^{\delta 2} (120-\chi 3)$	1,12	1,12
	$H^{\delta 3}$	$H^{\gamma 2}-H^{\delta 3} (120+\chi 3)$	9,63	9,94
		$H^{\gamma 3}-H^{\delta 3} (\chi 3)$	8,53	8,28

<b><math>\Psi</math>-Proline</b>			<i>cis</i>	<i>trans</i>
<i>up</i>	$H^\alpha$	$H^\alpha-H^{\beta 2} (\chi 1-120)$	9,3	9,27
		$H^\alpha-H^{\beta 3} (-\chi 1)$	7,17	7,19
<i>down</i>	$H^\alpha$	$H^\alpha-H^{\beta 2} (\chi 1-120)$	1,37	1,47
		$H^\alpha-H^{\beta 3} (\chi 1)$	6,03	6,37

Dans le tableau suivant sont reportées les valeurs des constantes de couplages  $3JH^{\alpha}-H^{\beta}$  pour chacun des peptides.

Peptides		P1		P2		P3		P4	
Pop.	Cycle	$JH^{\alpha}-H^{\beta 2}$	$JH^{\alpha}-H^{\beta 3}$	$JH^{\alpha}-H^{\beta 2}$	$JH^{\alpha}-H^{\beta 3}$	$JH^{\alpha}-H^{\beta 2}$	$JH^{\alpha}-H^{\beta 3}$	$JH^{\alpha}-H^{\beta 2}$	$JH^{\alpha}-H^{\beta 3}$
Maj.	X	4,95	8,75	4,73	8,78	6,65	6,65	2,45	7,8
	Y	5,80	8,40	7,76	9,38	8,87	9,09	8,56	8,74
Min.	X	3,05	8,95	3,05	8,8	7,43	7,43	2,1	7,61
	Y	6,075	8,38	?	?	3,16	9,03	4,52	7,6

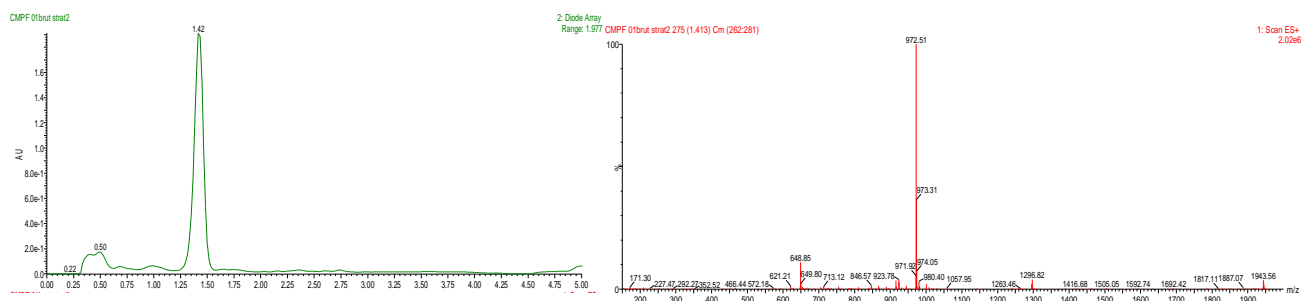
## C. Synthèse sur support solide des CMPF

### C1. Synthèse avec le triplet PYPG

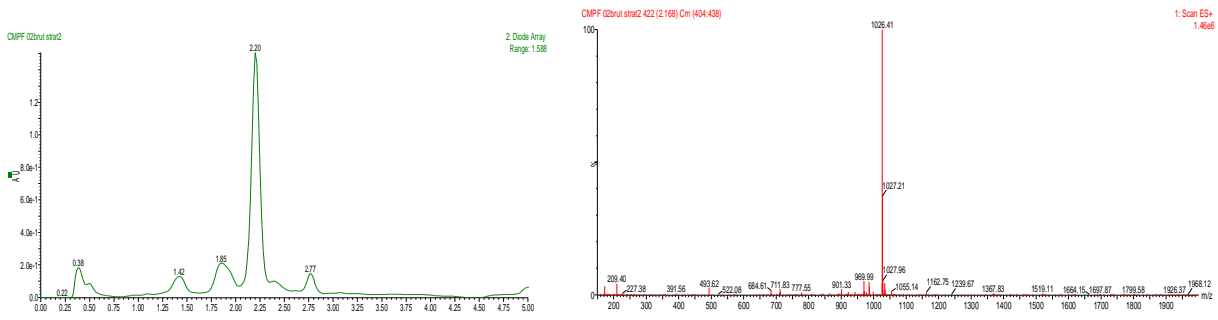
Les synthèses ont été réalisées en parallèle sur un synthétiseur automatique SyroWave sans micro-onde. La résine initialement utilisée était une Fmoc-Gly-CT avec un loading 0.51mmol/g mais le rendement de synthèse étant faible, nous avons choisi de travailler avec une résine Fmoc-Gly-Wang avec un loading 0.36 mmol/g. Nous avons travaillé avec 200mg de résine (0.072mmol, 1éq). Les acides aminés et le bloc PYPG sont solubilisés dans le DMF et ajoutés un à un. Chaque acide aminé (4éq) est couplé pendant 40min excepté les prolines, sur qui, il est effectué un double couplage (2\*4éq, 2\*40min). Nous avons remarqué que le couplage Pro-Hyp était difficile. Le bloc PYPG (2éq) est également couplé pendant 40min. Une déprotection en deux temps est réalisée, d'abord avec une solution 40% pipéridine/DMF pendant 3min suivi d'une déprotection avec une solution de 20% pipéridine/DMF pendant 15 min. Le peptide est ensuite clivé de la résine dans les conditions classiques TFA/TIS/H<sub>2</sub>O (95/2.5/2.5) (10mL).

Chaque peptide a été analysé par UPLC/MS sur Acquity UPLC BEH C18 (1.7micron, 2.1\*50mm) avec un gradient de 90:10 à 50:50 A/B en 5min; Débit 0.6ml/min ; éluant A (H<sub>2</sub>O/TFA 0.1%) et éluant B (ACN/TFA 0.1%).

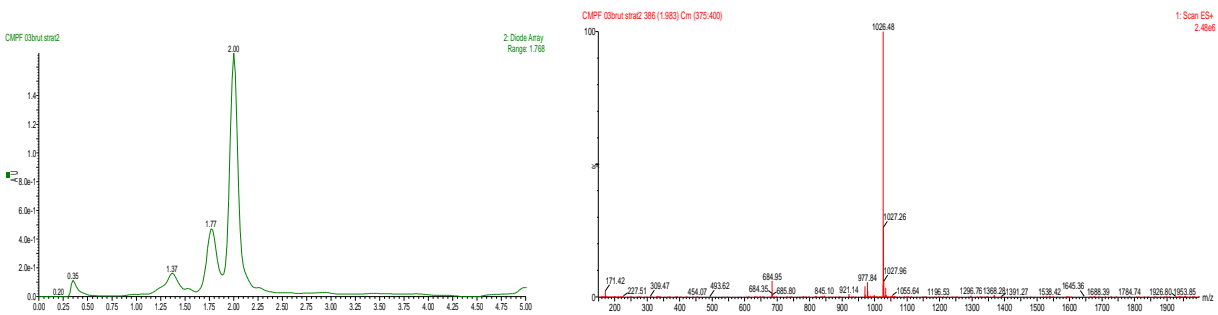
### CMPF01



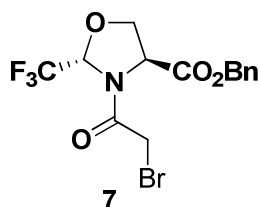
CMPF02



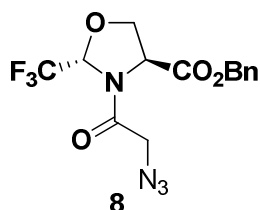
CMPF03



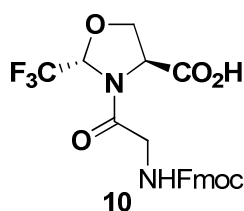
Les peptides ont ensuite été purifiés sur colonne préparative X-Bridge C18 (5µM, 19\*50mm) avec un gradient de 90:10 à 75:25 A/B en 10 min pour CMPF01 et un gradient de 90:10 à 50:50 A/B en 10 min pour CMPF02 et CMPF03; débit 10ml/min. Le peptide CMPF01 a été obtenu avec un rendement de 18% (25mg) et le CMPF02 a été obtenu avec un rendement de 33% (48mg) et le CMPF03 n'a pas été purifié entièrement.

C2. Synthèse avec le triplet  $\Psi\Psi\Psi$ Synthèse of the N-Bromoacetyl-(2S)-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OBn **7**

Le bromure de bromo-acétyle (2,08 ml, 24 mmol, 1 équiv.) a été ajouté directement sur l'oxazolidine **1a** (6,545 g, 1 équ., 23,8 mmol) et agité pendant 90 min sous atmosphère d'argon. Le mélange a ensuite été dilué dans du DCM (100 ml) puis filtré, afin de récupérer 2,312 g de la matière de départ protonée sous forme d'un solide blanc brut. La phase organique a ensuite été lavée avec de la saumure (2 x 200 ml), séchée avec du  $\text{MgSO}_4$  et filtrée. Après évaporation sous pression réduite, l'huile jaune brute a été purifiée par chromatographie flash sur gel de silice (95:5 cyclohexane /acétate d'éthyle) pour donner 4,714 g (rendement de 81%, avec un taux de conversion de 64%) de N-bromoacétyl oxazolidine **7** sous forme d'un mélange d'isomères inséparables à 50:50 dans  $\text{CDCl}_3$ ; huile incolore; IR : 2926, 1744, 1686, 1377, 1282, 1179, 1149  $\text{cm}^{-1}$ ;  $^1\text{H}$  NMR (400 MHz,  $\text{CDCl}_3$ ): (50:50 mixture of rotational isomers)  $\delta$  (ppm) 3.59 (d,  $J = 12.0$  Hz, 1 H,  $\text{CH}_2\text{-Br}$ , Ha isomer 1), 3.69 (d,  $J = 12.0$  Hz, 1 H,  $\text{CH}_2\text{-Br}$ , Hb isomer 1), 3.89 (d,  $J = 10.7$  Hz, 1 H,  $\text{CH}_2\text{-Br}$ , Ha isomer 2), 3.96 (d,  $J = 10.7$  Hz, 1 H,  $\text{CH}_2\text{-Br}$ , Hb isomer 2), 4.19 (m, 1 H,  $\text{H}_\beta$   $\Psi$ Pro-Ha isomer 1), 4.20-4.52 (m, 2 H,  $\text{H}_\beta$   $\Psi$ Pro-Hb isomer 1 and  $\text{H}_\beta$   $\Psi$ Pro-Ha isomer 2), 4.58 (m, 1 H,  $\text{H}_\beta$   $\Psi$ Pro-Hb isomer 2), 5.12-5.17 (m, 2 H,  $\text{CH}_2$  Bn isomer 1), 5.20-5.24 (m, 2 H,  $\text{CH}_2$  Bn isomer 2), 5.83 (m, 1 H,  $\text{H}_\delta$   $\Psi$ Pro-H isomer 1), 5.88 (m, 1 H,  $\text{H}_\delta$   $\Psi$ Pro-H isomer 2), 7.32-7.42 (m, 10 H, Bn H arom isomers 1 and 2);  $^{13}\text{C}$  NMR (100.5 MHz,  $\text{CDCl}_3$ ): (50:50 mixture of rotational isomers)  $\delta$  (ppm) 25.9 ( $\text{CH}_2$ ,  $\text{CH}_2\text{-Br}$  isomer 1), 26.2 ( $\text{CH}_2$ ,  $\text{CH}_2\text{-Br}$  isomer 2), 58.4 (CH,  $\text{C}_\alpha$   $\Psi$ Pro), 67.6 ( $\text{CH}_2$ , Bn  $\text{CH}_2$  isomer 1), 68.3 ( $\text{CH}_2$ , Bn  $\text{CH}_2$  isomer 1), 70.0 ( $\text{CH}_2$ ,  $\text{C}_\beta$   $\Psi$ Pro isomer 1), 72.1 ( $\text{CH}_2$ ,  $\text{C}_\beta$   $\Psi$ Pro isomer 2), 83.6-85.6 (2 CH,  $\text{C}_\delta$   $\Psi$ Pro), 122.6 (q,  $J = 287.5$  Hz, 2  $\text{CF}_3$ ), 128.2-129.2 (10 CH, Bn arom isomers 1 and 2), 134.1 (C, Bn arom isomer 1), 134.7 (C, Bn arom isomer 2) 164.4 (C,  $\text{C}=\text{O}$  isomer 1), 166.7 (C,  $\text{C}=\text{O}$  isomer 2), 168.4 (C,  $\text{C}=\text{O}$  isomer 1), 169.3 (C,  $\text{C}=\text{O}$  isomer 2); RMN  $^{19}\text{F}$  (376.2 MHz,  $\text{CDCl}_3$ ): (50:50 mixture of rotational isomers)  $\delta$  (ppm) -81.2 (s,  $\text{CF}_3$ , isomer 1), -81.4 (s,  $\text{CF}_3$ , isomer 2).; HRMS (ESI-TOF): Calculated for  $\text{C}_{14}\text{H}_{13}\text{BrF}_3\text{NO}_4\text{Na}$  417.9878, found 417.9870.

Synthèse de N-azidoacetyl-(2S)-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OBn **8**

L'azoture de sodium (0,780 g, 1,8 équiv. 12 mmol) a été ajouté avec précaution à une solution du bromo-4 oxazolidine (2,65 g, 1 équiv. 6,7 mmol) dans du DMF (15 ml) sous atmosphère d'argon. La solution a été agitée pendant 3h à température ambiante, avant d'être lavée avec de l'eau (2 x 125 ml), puis extrait avec du DCM (2 x 125 ml). La phase organique est séchée avec du  $\text{MgSO}_4$  puis filtrée. Après évaporation sous pression réduite, l'huile rouge brute résultante a été purifiée par chromatographie flash sur gel de silice (95:5 cyclohexane/acétate d'éthyle) pour donner 1,93 g (rendement 83%) du produit **8** sous forme d'un mélange d'isomères inséparables 57:43 dans le  $\text{CDCl}_3$ ; huile incolore; IR : 2110, 1747, 1689, 1401, 1186, 1153, 905, 725  $\text{cm}^{-1}$ ; RMN  $^1\text{H}$  (400 MHz,  $\text{CDCl}_3$ ) : (57:43 mixture of rotational isomers)  $\delta$  (ppm) 3.60-4.70 (m, 8 H,  $\text{H}_\beta$   $\Psi\text{Pro-H}$  and  $\text{CH}_2\text{-N}_3$  *majo and mino*), 4.70-4.74 (m, 2 H,  $\text{H}_\alpha$   $\Psi\text{Pro-H}$  *majo and mino*), 5.20-5.28 (m, 4 H,  $\text{CH}_2$  Bn *majo and mino*), 5.68 (m, 1 H,  $\text{H}_\delta$   $\Psi\text{Pro-H}$  *mino*), 5.93 (m, 1 H,  $\text{H}_\delta$   $\Psi\text{Pro-H}$  *majo*), 7.28-7.54 (m, 10 H, Bn H arom *majo and mino*); RMN  $^{13}\text{C}$  (100.5 MHz,  $\text{CDCl}_3$ ) : (50:50 mixture of rotational isomers)  $\delta$  (ppm) 51.2 ( $\text{CH}_2$ ,  $\text{CH}_2\text{-N}_3$  *mino*), 51.8 ( $\text{CH}_2$ ,  $\text{CH}_2\text{-N}_3$  *majo*), 58.1 (CH,  $\text{C}_\alpha$   $\Psi\text{Pro}$  *majo*), 58.5 (CH,  $\text{C}_\alpha$   $\Psi\text{Pro}$  *mino*), 67.9 ( $\text{CH}_2$ , Bn  $\text{CH}_2$  *mino*) 68.6 ( $\text{CH}_2$ , Bn  $\text{CH}_2$  *majo*), 69.9 ( $\text{CH}_2$ ,  $\text{C}_\beta$   $\Psi\text{Pro}$  *mino*), 72.3 ( $\text{CH}_2$ ,  $\text{C}_\beta$   $\Psi\text{Pro}$  *majo*), 84.0 (CH,  $\text{C}_\delta$   $\Psi\text{Pro}$  *mino*), 85.2 (CH,  $\text{C}_\delta$   $\Psi\text{Pro}$  *majo*), 122.7 (q,  $J = 287.5$  Hz, 2  $\text{CF}_3$ ), 128.2-129.2 (10 CH, Bn arom *majo and mino*), 134.1 (C, Bn arom *majo*), 134.8 (C, Bn arom *mino*), 166.3 (C,  $\text{C}=\text{O}$  *mino*), 167.7 (C,  $\text{C}=\text{O}$  *majo*), 168.6 (C,  $\text{C}=\text{O}$  *mino*), 169.2 (C,  $\text{C}=\text{O}$  *majo*); RMN  $^{19}\text{F}$  (376.2 MHz,  $\text{CDCl}_3$ ) : (57:43 mixture of rotational isomers)  $\delta$  (ppm) -81.3 (s,  $\text{CF}_3$  *majo*), -81.7 (s,  $\text{CF}_3$  *mino*); HRMS (ESI-TOF) : Calculated for  $\text{C}_{14}\text{H}_{13}\text{F}_3\text{N}_4\text{O}_4\text{Na}$  381.0781, found 381.0781.

Synthèse du Fmoc-Gly-(2S)-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OH **10**

Le palladium sur charbon (10% p/p, 76 mg, 0,06 mmol, 0,1 équivalent) a été ajouté à une solution de **8** (0.460g, 1 équiv. 0,64 mmol) et de HCl 3M (5 ml) dans du méthanol (10 ml), sous atmosphère de  $\text{H}_2$  (4 bars) à température ambiante pendant 4 heures. Le mélange est ensuite filtré, rincé avec du méthanol puis évaporé sous pression réduite pour donner le produit **9** sous forme de cristaux blancs avec un rendement quantitatif. Le produit brut a été

utilisé directement sans purification supplémentaire sous forme de mélange inséparable d'isomères 71:29 dans D<sub>2</sub>O.

Le dipeptide **9** (2,954 g, 10,6 mmol, 1 équivalent) a été dilué dans de l'eau (50 ml) et traité avec du Na<sub>2</sub>CO<sub>3</sub> (2,810 g, 26,5 mmol, 2,5 équiv.). Après dissolution complète des solides, la solution a été refroidie à 0 ° C et un mélange de Fmoc-OSu (4,290 g, 12,72 mmol, 1,2 équiv.) dans du dioxane (85 ml) a été ajoutée goutte à goutte pendant 1 h. On a laissé le mélange réactionnel se réchauffer lentement jusqu'à température ambiante et l'agitation est poursuivie pendant 15 h. La solution a été ensuite évaporée sous pression réduite afin de se débarrasser du dioxane. La solution aqueuse résultante a été lavée avec EtOAc (2 x 200 ml), acidifié avec de l'HCl 3M jusqu'à atteindre pH ~ 2, et ensuite extraite avec du EtOAc (2 x 200 ml). La phase organique a ensuite été séchée avec du MgSO<sub>4</sub> et évaporée sous pression réduite pour donner le dipeptide brut **10** sous forme de solide blanc brut. Après trituration dans le pentane pendant 1 h, le solide blanc résultant a été purifié par chromatographie flash sur gel de silice (99:1 dichlorométhane/méthanol) pour donner 3,790 g (73%) de produit pur **10** ; RMN <sup>1</sup>H (400 MHz, MeOD) : δ (ppm) 3.98-4.59 (m, 8H), 5.80 (d, 1H, H<sub>δ</sub> ΨPro-H), 7.25-7.78 (m, 8H, Fmoc-H-Arom) ; RMN <sup>13</sup>C (100.5 MHz, MeOD) : δ (ppm) 53.2 (CH<sub>2</sub>, C<sub>γ</sub>-Gly), 56.3 (CH, Fmoc), 69.5 (CH, C<sub>α</sub> ΨPro), 75.5 (CH<sub>2</sub>, Fmoc), 83.4 (CH<sub>2</sub>, C<sub>β</sub> ΨPro), 94.0 (CH, C<sub>δ</sub> ΨPro), 133.1 (C, J = 283.7 Hz, CF<sub>3</sub>), 129.9 - 137.4 (CH, Fmoc Arom), 150.9 (C<sub>q</sub>, Fmoc arom), 153.5 (C<sub>q</sub>, Fmoc arom), 166.2 (C<sub>q</sub>, (C=O)Fmoc), 180.2 (C<sub>q</sub>, (C=O)NR<sub>2</sub>), 182.5 (C<sub>q</sub>, COOH) ; RMN <sup>19</sup>F (376.2 MHz, MeOD) : δ (ppm) -81.2 (s, CF<sub>3</sub> *mino*), -81.3 (s, CF<sub>3</sub> *majo*)

### Synthesis of the Fmoc-Gly-(2S)-Ser(Ψ<sup>CF<sub>3</sub>,H</sup>Pro)-(2R)-Ser(Ψ<sup>CF<sub>3</sub>,H</sup>Pro)-OBn **11**



**11**

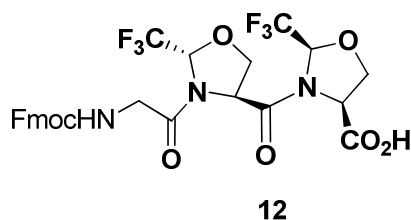
*Préparation du chlorure d'acyle* – Le chlorure de thionyle (0,61 ml, 8,4 mmol, 13 équiv.) a été ajouté avec précaution à un mélange de produit **10** (0,3 g, 0,65 mmol, 1 équ.) dans du DCM (10 ml) sous atmosphère d'argon. Le mélange réactionnel est soumis à ultrasons pendant 90 min (37 kHz, 60%, T <30 ° C) avant d'être évaporé sous pression réduite. Puis il est lavé avec du pentane et séché sous vide poussé. Le chlorure d'acyle résultant a été utilisé sous forme de poudre blanche dans la réaction de couplage sans autre purification.

*Activation de l'amine* – Du zinc (42 mg, 0,65 mmol, 1 équ.) a été ajouté à l'oxazolidine **1b** (180 mg, 0,65 mmol, 1 équiv.) diluée dans du CHCl<sub>3</sub> (4 ml) et le mélange a été agité pendant 10 min avant d'ajouter TMS-Cl (0,084 ml, 0,65 mmol, 1 équ.). Après 15 minutes, le mélange a été filtré puis ajouté directement sur la poudre blanche obtenue précédemment sous atmosphère d'argon. Après 15 h d'agitation, le mélange a été évaporé sous pression réduite avant purification par chromatographie flash sur gel de silice (95:5 cyclohexane/acétate d'éthyle) pour donner 350 mg (42%) du produit **11** correspondant à l'isomère *trans* pur dans CDCl<sub>3</sub> ; solide blanc ; RMN <sup>1</sup>H (400 MHz, CDCl<sub>3</sub>) : δ (ppm) 4.05 (dd, J = 16.8, 4.4 Hz, 1 H, H<sub>α</sub> Gly-H Ha), 4.10 (dd, J = 16.8, 5.2 Hz, 1 H, H<sub>β2</sub> ΨPro-Ha), 4.18 (dd, J = 13.2, 5.6 Hz, 1 H,



$H_{\alpha}$  Gly-H Hb), 4.22 (t,  $J = 5.2$  Hz, 1 H, Fmoc CH), 4.35 (m, 3 H, Fmoc  $CH_2$ -Ha,  $H_{\beta 1}$   $\Psi$ Pro-Ha and  $H_{\beta 2}$   $\Psi$ Pro-Hb), 4.40 (m, 1 H, Fmoc  $CH_2$ -Hb), 4.51 (t,  $J = 7.2$  Hz, 1 H,  $H_{\beta 1}$   $\Psi$ Pro-Hb), 5.08 (t,  $J = 8.0$  Hz, 1 H,  $H_{\alpha 1}$   $\Psi$ Pro-H), 5.12 (d,  $J = 12.0$  Hz, 1H, Bn  $CH_2$ -Hb), 5.17 (d,  $J = 12.0$  Hz, 1H, Bn  $CH_2$ -Hb), 5.43 (t,  $J = 5.0$  Hz, 1 H, NH), 5.71 (q,  $J = 5.0$  Hz, 1 H,  $H_{\delta 1}$   $\Psi$ Pro-H), 5.92 (q,  $J = 5.04$  Hz, 1 H,  $H_{\delta 2}$   $\Psi$ Pro-H), 7.42-7.27 (m, 8H, CH, Bn arom and Fmoc arom), 7.58 (d,  $J = 7.3$  Hz, 2 H, Fmoc arom), 7.76 (d,  $J = 7.33$  Hz, 2 H, Fmoc arom) ; RMN  $^{13}C$  (100.5 MHz,  $CDCl_3$ ) :  $\delta$  (ppm) 43.3 ( $CH_2$ ,  $C_{\alpha}$  Gly), 47.0 (CH, Fmoc CH), 56.9 (CH,  $C_{\alpha 1}$   $\Psi$ Pro), 57.8 (CH,  $C_{\alpha 2}$   $\Psi$ Pro), 67.3 ( $CH_2$ , Fmoc  $CH_2$ ), 84.8 (2 C,  $C_{\delta 1}$   $\Psi$ Pro and  $C_{\delta 2}$   $\Psi$ Pro), 120.1 (2 CH, Fmoc arom), 122.5 (q,  $J = 285.5$  Hz, C,  $CF_3$ ), 122.7 (q,  $J = 287.4$  Hz, C,  $CF_3$ ), 125.2 (2 CH, Fmoc arom), 127.2 (2 CH, Fmoc arom), 127.8 (2 CH, Fmoc arom), 128.4 (2 CH, Bn arom), 128.8 (3 CH, Bn arom), 134.78 (C, Bn arom), 141.39 (2 C, Fmoc arom), 143.8 (2 C, Fmoc), 156.2 (C, C=O), 167.6 (C, C=O), 168.4 (C, C=O), 169.1 (C, C=O) ; RMN  $^{19}F$  (376.2 MHz,  $CDCl_3$ ) :  $\delta$  (ppm) -81.5 (s,  $CF_3$  *mino*), -81.7 (s,  $CF_3$  *majo*)

### Synthesis of the Fmoc-Gly-(2*S*)-Ser( $\Psi^{CF_3,H}$ Pro)-(2*R*)-Ser( $\Psi^{CF_3,H}$ Pro)-OH **12**



Le palladium sur charbon (10% p/p, 50 mg, 0,1 équivalent) a été ajouté à une solution de **11** (0.320g, 0,44 mmol, 1 éq.) dans du méthanol (6 ml). L'ensemble est placé sous agitation et atmosphère de  $H_2$  (1 bar) à la température ambiante pendant 1 heure. Le mélange est ensuite filtré, rincé avec du méthanol puis évaporé sous pression réduite. Le solide brut résultant a été purifié par chromatographie éclair sur gel de silice blanche (100:0 dichlorométhane/méthanol) pour donner 200 mg (77%) du produit **12**, sous forme de mélange de deux diastéréoisomères, caractérisé par spectroscopie de masse ; RMN  $^1H$  (400 MHz,  $CDCl_3$ ) :  $\delta$  (ppm) 4 - 4.88 (m, 11H), 5.57 (m, 1H, NH), 5.82 (m, 1H,  $H_{\delta 1}$   $\Psi$ Pro-H), 5.90 (m, 1H,  $H_{\delta 2}$   $\Psi$ Pro-H), 7.76 (d,  $J = 7.79$  Hz, 2H, Fmoc arom), 7.55 (d,  $J = 7.79$  Hz, 2H, Fmoc arom), 7.42-7.27 (m, 4H, Fmoc arom) ; RMN  $^{19}F$  (376.2 MHz,  $CDCl_3$ ) :  $\delta$  (ppm) 81.67 - 81.92 (s,  $CF_3$  *majo*), 81.31 - 81.77 (s,  $CF_3$  *mino*)

### Références du chapitre 3

- 1 Kang, Y. K. & Choi, H. Y. Cis-trans isomerization and puckering of proline residue. *Biophysical chemistry* **111**, 135-142, doi:10.1016/j.bpc.2004.05.006 (2004).
- 2 Lovell, S. C., Davis, I. W., Bryan Arendall III, W., Bakker, P., Word, J. M., Prisant, M. G., Richardson, J. S., Richardson, D. C. Structure validation by Ca geometry : phi, psi and Cb deviation. *Proteins: Structure, Functions, and Genetics* **50**, 437-450 (2003).
- 3 Morris, A.-L., MacArthur, M. W., Huntchinson, E. G. & Thornton, J. M. Stereochemical quality of protein structure coordinates. *Proteins: Structure, Functions, and Genetics* (1992).
- 4 MacArthur, M. W. & Thornton, J. M. Influence of proline residues in protein conformation. *Journal of molecular biology* (1991).
- 5 Kuemin, M. *et al.* Tuning the cis/trans conformer ratio of Xaa-Pro amide bonds by intramolecular hydrogen bonds: the effect on PPII helix stability. *Angewandte Chemie* **49**, 6324-6327, doi:10.1002/anie.201001851 (2010).
- 6 Shoulders, M. D., Guzei, I. A. & Raines, R. T. 4-chloroproline: synthesis, conformational analysis, and effect on the collagen triple helix. *Biopolymers* **89**, 443-454, doi:10.1002/bip.20864 (2008).
- 7 Milner-White, E. J., Bell, L. H. & Maccallum, P. H. Pyrrolidine ring puckering in cis and trans-proline residues in proteins and polypeptides Different puckers are favoured in certain situation. *Journal of molecular biology* (1992).
- 8 Bartlett, G. J., Choudhary, A., Raines, R. T. & Woolfson, D. N. n-->pi\* interactions in proteins. *Nature chemical biology* **6**, 615-620, doi:10.1038/nchembio.406 (2010).
- 9 Hinderaker, M. P. & Raines, R. T. An electronic effect on protein structure. *Protein science : a publication of the Protein Society* **12**, 1188-1194, doi:10.1110/ps.0241903 (2003).
- 10 Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annual review of biochemistry* **78**, 929-958, doi:10.1146/annurev.biochem.77.032207.120833 (2009).
- 11 Shoulders, M. D., Satyshur, K. A., Forest, K. T. & Raines, R. T. Stereoelectronic and steric effects in side chains preorganize a protein main chain. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 559-564, doi:10.1073/pnas.0909592107 (2010).
- 12 Bretscher, L. E., Jenkins, C. L., Taylor, K. M., DeRider, M. L. & Raines, R. T. Conformational stability of collagen relies on a stereoelectronic effect. *Journal of the American Chemical Society* (2001).
- 13 Shoulders, M. D., Hodges, J. A. & Raines, R. T. Reciprocity of steric and stereoelectronic effects in the collagen triple helix. *Journal of the American Chemical Society* (2006).
- 14 Cadamuro, S. A. *et al.* Conformational properties of 4-mercaptoproline and related derivatives. *Angewandte Chemie* **47**, 2143-2146, doi:10.1002/anie.200704310 (2008).
- 15 Lin, Y. J. & Horng, J. C. Impacts of terminal (4R)-fluoroproline and (4S)-fluoroproline residues on polyproline conformation. *Amino acids* **46**, 2317-2324, doi:10.1007/s00726-014-1783-2 (2014).
- 16 Choudhary, A., Gandla, D., Krow, G. R., Raines, R. T. Nature of Amide Carbonyl-Carbonyl Interactions in Proteins. *Journal of the American Chemical Society* **131**, 7244-7246 (2009).
- 17 Erdmann, R. S. & Wennemers, H. Importance of ring puckering versus interstrand hydrogen bonds for the conformational stability of collagen. *Angewandte Chemie* **50**, 6835-6838, doi:10.1002/anie.201008118 (2011).

- 18 Delaney, N. G. & Madison, V. Novel conformational distributions of methylproline peptides. *Journal of the American Chemical Society* **104**, 6635-6641 (1982).
- 19 Swarbrick, M. E., Gosselin, F. & Lubell, W. D. Alkyl Substituent Effects on Pipecolyl Amide Isomer Equilibrium: Efficient Methodology for Synthesizing Enantiopure 6-Alkylpipercolic Acids and Conformational Analysis of Their N-Acetyl N'-Methylamides. *The Journal of organic chemistry* **64**, 1993-2002 (1999).
- 20 Beausoleil, E. & Lubell, W. D. Steric Effects on the Amide Isomer Equilibrium of Prolyl Peptides. Synthesis and Conformational Analysis of N-Acetyl-5-tert-butylproline N'-Methylamides. *Journal of the American Chemical Society* **118**, 12902-12908 (1996).
- 21 Beausoleil, E. & Lubell, W. D. An Examination of the Steric Effects of 5-tert-Butylproline on the Conformation of Polyproline and the Cooperative Nature of Type II to Type I Helical Interconversion. *Biopolymers* **53**, 249-256 (2000).
- 22 Saraswathi, S., Fernandez-Martinez, J. L., Kolinski, A., Jernigan, R. L. & Kloczkowski, A. Distributions of amino acids suggest that certain residue types more effectively determine protein secondary structure. *J Mol Model* **19**, 4337-4348, doi:10.1007/s00894-013-1911-z (2013).
- 23 Adzhubei, A. A., Sternberg, M. J. & Makarov, A. A. Polyproline-II helix in proteins: structure and function. *Journal of molecular biology* **425**, 2100-2132, doi:10.1016/j.jmb.2013.03.018 (2013).
- 24 Suzuki, E., Fraser, R. D. B. & MacRae, T. P. Role of hydroxyproline in the stabilization of the collagen molecule via water molecules. *Int. J. Biol. Macromol.* **2**, 54-56 (1980).
- 25 Bella, J., Eaton, M., Brodsky, B., Berman, H. M. Crystal and molecular structure of a collagen-like peptide at 1.9 Å resolution. *Science (New York, N.Y.)* **266**, 75-81 (1994).
- 26 Dunitz, J. D. & Taylor, R. Organic Fluorine Hardly Ever Accepts Hydrogen Bonds. *Chem Eur J* **3**, 89-98 (1997).
- 27 Chan, K. K. J. & O'Hagan, D. Chapter Eleven: The Rare Fluorinated Natural Products and Biotechnological Prospects for Fluorine Enzymology. *Methods in Enzymology* **516**, 219-235, doi:10.1016/B978-0-12-394291-3.00003-4 (2012).
- 28 Bégué, J.-P. & Bonnet-Delpon, D. Chimie bioorganique et médicinale du fluor. (2012).
- 29 Arnaud, P., Jamart, B., Bodiguel, J. & Brosse, N. Chimie organique: Cours et applications. *Dunod Ed.* (2004).
- 30 Wöhr, T. *et al.* Pseudo-Prolines as a Solubilizing, Structure-Disrupting Protection Technique in Peptide Synthesis. *Journal of the American Chemical Society* **118**, 9218-9227 (1996).
- 31 Keller, M. *et al.* Enhancing the Proline Effect: Pseudo-Prolines for Tailoring Cis/Trans Isomerization. *Journal of the American Chemical Society* **120**, 2714-2720 (1998).
- 32 Mutter, M., Wöhr, T., Gioria, S., Keller, M. Pseudo-Prolines: Induction of Cis/Trans-Conformational Interconversion by Decreased Transition State Barriers. *Biopolymers : Peptide Science* **51**, 121-128 (1999).
- 33 Kang, Y. K. & Park, H. S. Conformational preferences of pseudoproline residues. *The journal of physical chemistry. B* **111**, 12551-12562 (2007).
- 34 Kang, Y. K. Cis-trans isomerization and puckering of pseudoproline dipeptides. *The journal of physical chemistry. B* **106**, 2074-2082 (2002).
- 35 Chaume, G., Barbeau, O., Lesot, P. & Brigaud, T. Synthesis of 2-trifluoromethyl-1,3-oxazolidines as hydrolytically stable pseudoprolines. *The Journal of organic chemistry* **75**, 4135-4145, doi:10.1021/jo100518t (2010).

- 36 Feytens, D. *et al.* Local control of the cis-trans isomerization and backbone dihedral angles in peptides using trifluoromethylated pseudoprolines. *The journal of physical chemistry. B* **116**, 4069-4079, doi:10.1021/jp300284u (2012).
- 37 Kang, Y. K., Jhon, J. S. & Park, H. S. Conformational preferences of proline oligopeptides. *The journal of physical chemistry. B* **110**, 17645-17655 (2006).
- 38 Chaume, G., Lensen, N., Caupene, C. & Brigaud, T. Convenient Synthesis of N-Terminal Tfm-Dipeptides from Unprotected Enantiopure  $\alpha$ -Tfm-Proline and  $\alpha$ -Tfm-Alanine (English). *European journal of organic chemistry (Print)*, 5717-5724 (2009).
- 39 Mutter, M., Nazi, A., Sato, T., Sun, X., Wahl, F., Wöhr, T. Pseudo-prolines (psi Pro) for accessing "inaccessible" peptides. *Pept. Res.* **8**, 145-153 (1995).
- 40 Kantharaju, Patil, B. & Babu, V. S. Synthesis of Fmoc-amino acid chlorides assisted by ultrasonication, a rapid approach. *Lett Pept Sci* **9**, 227-229, doi:10.1007/BF02538388 (2002).
- 41 Kumar, A. Two-dimensional nuclear Overhauser effect experiment in a protein: the first NOESY (1979–80)†. *Magnetic Resonance in Chemistry* (2003).
- 42 Leroux, F. Atropisomerism, biphenyls, and fluorine: a comparison of rotational barriers and twist angles. *Chembiochem* **5**, 644-649, doi:10.1002/cbic.200300906 (2004).
- 43 Zimmer, K. D., Shoemaker, R. & Ruminski, R. R. Synthesis and characterization of a fluxional Re(I) carbonyl complex fac-[Re(CO)3(dpop')Cl] with the nominally tridentate ligand dipyrido(2,3-a:3',2'-j)phenazine (dpop'). *Inorganica Chimica Acta* **359**, 1478-1484, doi:10.1016/j.ica.2005.11.042 (2006).
- 44 Halab, L., Lubell, W.D. Use of Steric Interactions To Control Peptide Turn Geometry. Synthesis of Type VI  $\gamma$ -Turn Mimics with 5-tert-Butylproline. *The Journal of organic chemistry* **64**, 3312-3321 (1999).
- 45 Miclet, E. *et al.* Relaxation-Optimized NMR Spectroscopy of Methylene Groups in Proteins and Nucleic Acids. (2006).
- 46 Holmgren, S. K., Taylor, K. M., Bretscher, L. E. & Raines, R. T. Code for collagen's stability deciphered. *Nature* **392** (1998).
- 47 Wellings, D. A. Standard Fmoc Protocols. *Methods in enzymology* **289**, 44-67 (1997).
- 48 Boas, U. & Mirsharghi, S. Color test for selective detection of secondary amines on resin and in solution. *Organic letters* **16**, 5918-5921, doi:10.1021/ol502936d (2014).
- 49 Bochicchio, B., Laurita, A., Heinz, A., Schmelzer, C. E. & Pepe, A. Investigating the role of (2S,4R)-4-hydroxyproline in elastin model peptides. *Biomacromolecules* **14**, 4278-4288, doi:10.1021/bm4011529 (2013).

## ANNEXE 5 : Synthèse du peptide modèle *P3* Ac-Pro-(2*R*,4*S*)-Ser( $\Psi^{CF_3,H}$ Pro)-Gly-NH<sub>2</sub> en stratégie ester méthylique

### 1. Synthèse

*Synthèse de Ser( $\Psi^{CF_3,H}$ Pro)-OMe 2* - Dans un premier temps, la sérine est protégée sous forme d'ester méthylique par traitement d'une solution de chlorure d'acétyle dans du méthanol. La fonction amine est ensuite protégée par un groupement *tert*-butoxycarbonyl<sup>49</sup> par ajout d'anhydride de di-*tert*-butyldicarbonate (Boc<sub>2</sub>O) en présence de triéthylamine dans le THF pour conduire à la Boc-sérine ester méthylique **1** avec un rendement de 52% sur deux étapes (Schéma 1). Le groupement protecteur Boc est utilisé pour améliorer la solubilité de la sérine dans le toluène afin de faciliter l'étape de condensation. Il est alors clivé *in situ* au cours de la réaction.

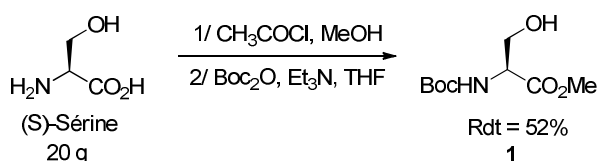


Schéma 1: Préparation de la Boc-sérine ester méthylique **1**

La condensation de la Boc-sérine ester méthylique **1** sur le trifluoroacétaldéhyde hydrate en présence de *para*-toluène sulfonate de pyridinium (PPTS) permet d'obtenir la pseudoproline **2** sous forme d'un mélange diastéréoisomérique 84:16 (Schéma 2). Après purification par chromatographie, les oxazolidines majoritaires (2*S*,4*S*)-**2** et minoritaires (2*R*,4*S*)-**2** sont obtenues avec des rendements respectifs de 54% et de 10%.

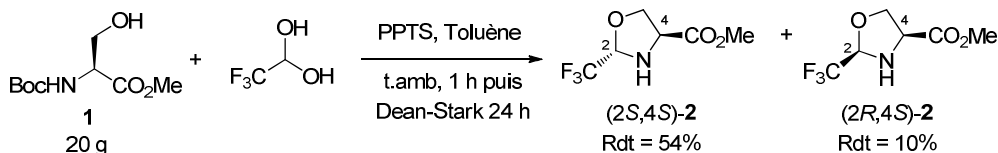
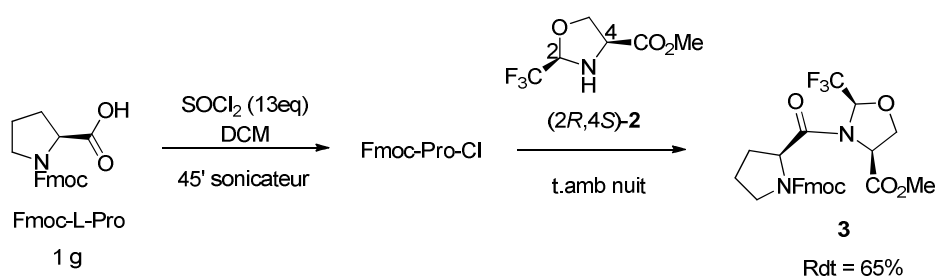


Schéma 2: Synthèse des pseudoprolines Ser( $\Psi^{CF_3,H}$ Pro)-OMe **2**

*Synthèse de Fmoc-Pro-(2*R*,4*S*)Ser( $\Psi^{CF_3,H}$ Pro)-OH 4* - Le couplage est réalisé avec la pseudoproline (2*R*,4*S*)-**2**. Cependant, il a été montré que le couplage peptidique à partir d'un mélange d'oxazolidines (2*R*,4*S*)-**2** et (2*S*,4*S*)-**2** conduisait à l'obtention du seul dipeptide incorporant une Ser( $\Psi^{CF_3,H}$ Pro) de configuration (2*R*,4*S*), une épimérisation du centre en <sup>2</sup>C ayant lieu avant la réaction d'amidification.

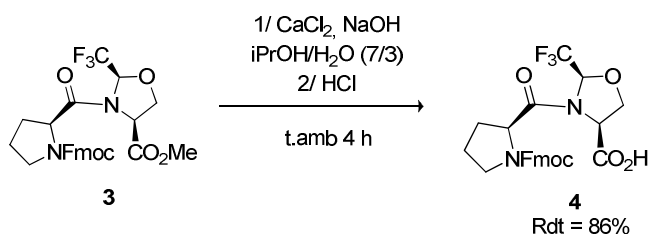
Le couplage de la pseudoproline (2*R*,4*S*)-**2** avec la Fmoc-L-Pro est effectué sans base après activation de la proline sous forme de chlorure d'acide. La Fmoc-L-Pro-Cl est préparée selon la méthode décrite par Babu, par « sonication » en présence d'un large excès de chlorure de thionyle dans le DCM pendant 45 minutes puis évaporation. Le dipeptide Fmoc-Pro-(2*R*,4*S*)-Ser( $\Psi^{CF_3,H}$ Pro)-OMe **3** est ainsi obtenu avec un rendement de 65% (Schéma 3). Lors de la purification, il est difficile de séparer parfaitement le dipeptide attendu du Fmoc-Pro-Cl qui n'a pas réagi. En effet, les deux composés ont des rapports frontaux quasi identiques.



**Schéma 3: Synthèse du dipeptide Fmoc-Pro-(2R,4S)-Ser( $\Psi^{CF_3,H}$ Pro)-OMe 3**

Afin de coupler avec la glycine, le dipeptide **3** est saponifié dans des conditions douces afin de ne pas cliver le groupement Fmoc. L'utilisation de soude dans un mélange isopropanol/eau en présence de chlorure de calcium conduit à l'acide **4** (Schéma 4).

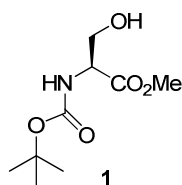
A l'issue de la purification, l'acide **4** attendu est obtenu avec un rendement de 86%, en mélange avec des traces de Fmoc-Pro-OH, vraisemblablement issu de l'hydrolyse du Fmoc-Pro-Cl. C'est pour éviter ces difficultés que la synthèse en série benzylique a été réalisée.



**Schéma 4: Synthèse de l'acide 4**

## 2. Partie expérimentale

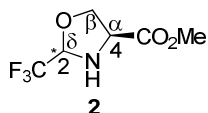
### Boc-L-Ser-OMe 1



A une solution de méthanol (280 ml) à 0°C, nous avons ajouté du chlorure d'acétyl (27,2 ml, 0,38 mol, 2 équivalents). Le mélange est agité à température ambiante pendant 5 min. Ensuite, la L-sérine (20,2 g, 0,19 mol, 1 éq.) est ajoutée. Le mélange résultant est agité, chauffé à reflux pendant 2 h, refroidi à température ambiante, puis évaporé sous pression réduite. L'ester méthylique de sérine brut a été dilué dans du THF (412 ml). De la triéthylamine (54 ml, 0,38 mol, 2 éq.) et du Boc<sub>2</sub>O (46 g, 0,21 mol, 1,1 éq.) ont été ajoutés successivement à 0°C. Le mélange résultant a été chauffé au reflux à 50°C pendant 20 h puis refroidi à 0°C et désactivé avec 200 ml d'eau. Le mélange résultant a été agité pendant 30 min. Les phases ont été séparées et la phase organique a été extraite avec de l'acétate d'éthyle (3 x 290 ml), lavé avec une solution de HCl 0,1 M (2 x 75 ml) et de la saumure (2 x 75 ml). Les phases organiques combinées ont été séchées sur MgSO<sub>4</sub>, filtrées et évaporées sous pression réduite. Le résidu

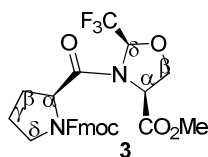
brut a été purifié par chromatographie sur gel de silice (70/30 cyclohexane/acétate d'éthyle) pour donner 21,6 g (52%) de (*S*)-Boc-serine ester de méthyle **1**;  $R_f = 0.16$  (60:40 cyclohexane/acétate d'éthyl);  $^1\text{H NMR}$  (400 MHz,  $\text{CDCl}_3$ )  $\delta$  1.45 (s, 9 H, *t*Bu), 2.36 (t, 1 H,  $J = 5.0$  Hz, OH), 3.78 (s, 3 H, OMe), 3.91–3.97 (m, 2 H,  $\text{H}_\beta$  Ser-H), 4.39 (m, 1 H,  $\text{H}_\alpha$  Ser-H), 5.47 (d, 1 H,  $J = 5.5$  Hz, NH).

### Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OMe **2**



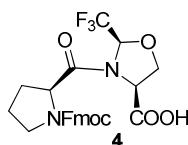
A une solution de (*S*)-Boc-sérine ester de méthyle **1** (21 g, 98 mmol, 1 équivalent) dans du toluène (50 ml) à température ambiante, nous avons ajouté de l'hydrate de trifluoroacétaldéhyde (9,8 ml, 118,5 mmol, 1,2 éq.) et du pyridinium-*p*-toluènesulfonate (PPTS) (5 g, 19,8 mmol, 0,2 éq.). Le mélange résultant a été agité à 90°C pendant 1h puis 800ml de toluène ont été ajoutés. Le mélange réactionnel a été chauffé au reflux en utilisant un appareil de Dean-Stark à 140°C pendant 24 h. Le mélange réactionnel a ensuite été refroidi à 0°C, filtré et le toluène a été évaporé sous pression réduite. Le résidu brut a été purifié par chromatographie sur gel de silice (90/10 éther de pétrole/acétate d'éthyle) pour donner 10,6 g de (*2S,4S*)-**2** (54%) sous forme d'une huile incolore et 1,9 g de (*2R,4S*)-**2** (10%) sous forme d'un solide blanc. (*2S,4S*)-**2**: huile incolore;  $R_f = 0.31$  (80:20 cyclohexane/acétate d'éthyle);  $[\alpha]_D^{23} -50.4$  (c 4.95,  $\text{CHCl}_3$ ); IR (neat) 3338, 2962, 1740, 1439, 1287, 1223, 1158, 1131, 665  $\text{cm}^{-1}$ ;  $^1\text{H NMR}$  (600 MHz,  $\text{CDCl}_3$ )  $\delta$  3.28 (m, 1 H, NH), 3.79 (s, 3 H, OMe), 3.83 (ddd,  $J = 7.8, 6.2, 1.0$  Hz, 1 H), 4.02 (dd,  $J = 7.8, 6.2$  Hz, 1 H), 4.23 (t,  $J = 7.8$  Hz, 1 H), 5.04 (q,  $J = 5.7$  Hz, 1 H);  $^{13}\text{C NMR}$  (100.5 MHz,  $\text{CDCl}_3$ )  $\delta$  52.9, 58.5, 69.7, 87.8 (q,  $J = 34.5$  Hz), 123.2 (q,  $J = 283.6$  Hz), 171.9;  $^{19}\text{F NMR}$  (376.2 MHz,  $\text{CDCl}_3$ )  $\delta$  -85.0 (dd,  $J = 5.7, 1.0$  Hz); MS(EI)  $m/z = 199$  [ $\text{M}^+$ ], 140 [ $\text{M}^+ - \text{CO}_2\text{Me}$ ] (100), 130, 112, 92, 70; HRMS (EI) calcd for  $\text{C}_6\text{H}_8\text{F}_3\text{NO}_3$  199.0456, found 199.0457. (*2R,4S*)-**2**: solide blanc; mp 65 °C;  $R_f = 0.10$  (80:20 cyclohexane/acétate d'éthyle);  $[\alpha]_D^{23} -17.61$  (c 4.9,  $\text{CHCl}_3$ ); IR (neat) 3309, 1737, 1462, 1285, 1231, 1156, 1127, 669  $\text{cm}^{-1}$ ;  $^1\text{H NMR}$  (600 MHz,  $\text{CDCl}_3$ )  $\delta$  3.18 (m, 1 H, NH), 3.71 (s, 3 H, OMe), 4.08 (ddq,  $J = 6.5, 6.0, 0.7$  Hz, 1 H), 4.11 (t,  $J = 6.0$  Hz, 1 H), 4.15 (ddq,  $J = 6.5, 6.0, 1.1$  Hz, 1 H), 4.89 (q,  $J = 5.2$  Hz, 1 H);  $^{13}\text{C NMR}$  (100.5 MHz,  $\text{CDCl}_3$ )  $\delta$  52.6, 58.6, 68.8, 87.6 (q,  $J = 34.5$  Hz), 122.9 (q,  $J = 283.6$  Hz), 171.1;  $^{19}\text{F NMR}$  (376.2 MHz,  $\text{CDCl}_3$ )  $\delta$  -84.1 (ddd,  $J = 5.2, 1.1, 0.7$  Hz); MS (EI)  $m/z = 199$  [ $\text{M}^+$ ], 140 [ $\text{M}^+ - \text{CO}_2\text{Me}$ ] (100), 130, 112, 92, 70; HRMS (EI) calcd for  $\text{C}_6\text{H}_8\text{F}_3\text{NO}_3$  199.0456; found 199.0458.

### Fmoc-Pro-Ser( $\Psi^{\text{CF}_3, \text{H}}$ Pro)-OMe **3**



A une solution de Fmoc-L-proline (1,16 g, 3,45 mmol, 1 équivalent) dans du dichlorométhane (18 ml) a été ajouté du chlorure de thionyle (5,3 g, 45 mmol, 13 éq.). Le mélange résultant a été traité par ultrasons pendant 45min. Le solvant a été évaporé et le Fmoc-Pro-Cl correspondant (1.1 éq.) a été solubilisé dans du dichlorométhane (10 ml). L'oxazolidine (2*R*,4*S*)-**2** (625 mg, 3,14 mmol, 1 éq.) a été ajouté via une canule et le mélange résultant a été agité à température ambiante pendant une nuit, puis évaporé. Le produit brut a été purifié par chromatographie flash (80:20 cyclohexane/acétate d'éthyle) pour donner 1,14 g (65%) du dipeptide pur Fmoc-Pro-Ser( $\Psi$ ProCF<sub>3</sub>, H)-OMe **3** comme un isomère *trans* dans CDCl<sub>3</sub> à 300 K: solide blanc; mp 115 °C; *R*<sub>f</sub> = 0.3 (70:30 cyclohexane/acétate d'éthyle); [ $\alpha$ ]<sub>D</sub><sup>23</sup> -56.1 (*c* 1.0, CHCl<sub>3</sub>); IR (neat) 2956, 1763, 1689 cm<sup>-1</sup>; <sup>1</sup>H NMR (500 MHz, CDCl<sub>3</sub>)  $\delta$  1.94–2.05 (m, 1 H, H $\gamma$  Pro-Ha), 2.10–2.26 (m, 2H, H $\beta$  Pro-H), 2.27–2.39 (m, 1 H, H  $\gamma$  Pro-Hb), 3.52–3.60 (m, 1 H, H $\delta$  Pro-Ha), 3.67–3.74 (m, 1 H, H $\delta$  Pro-Hb), 3.77 (s, 3 H, OMe), 4.23 (t, *J* = 7.1 Hz, 1 H, Fmoc CH), 4.32 (dd, *J* = 8.5, 6.9 Hz, 1 H, H $\beta$   $\Psi$ Pro-Ha), 4.34 (d, *J* = 7.1 Hz, 2 H, Fmoc CH<sub>2</sub>-H), 4.36–4.40 (m, 1 H, H $\alpha$  Pro-H), 4.50 (t, *J* = 8.7 Hz, 1 H, H $\beta$   $\Psi$ Pro-Hb), 5.06 (t, *J* = 7.8 Hz, 1 H, H $\alpha$   $\Psi$ Pro-H), 6.23 (q, *J* = 4.8 Hz, 1 H, H $\delta$   $\Psi$ Pro-H), 7.31 (t, *J* = 7.3 Hz, 2 H, Fmoc arom), 7.40 (t, *J* = 7.3 Hz, 2 H, Fmoc arom), 7.56 (t, *J* = 8.5 Hz, 2 H, Fmoc arom), 7.77 (d, *J* = 7.3 Hz, 2 H, Fmoc arom); <sup>13</sup>C NMR (125 MHz, CDCl<sub>3</sub>)  $\delta$  24.8 (CH<sub>2</sub>, C $\gamma$  Pro), 30.9 (CH<sub>2</sub>, C $\beta$  Pro), 47.1 (CH<sub>2</sub>, C $\delta$  Pro), 47.2 (CH, Fmoc CH), 52.8 (CH<sub>3</sub>, OMe), 56.9 (CH, C $\alpha$  Pro), 58.3 (CH, C $\alpha$   $\Psi$ Pro), 67.7 (CH<sub>2</sub>, Fmoc CH<sub>2</sub>), 68.9 (CH<sub>2</sub>, C $\beta$   $\Psi$ Pro), 85.1 (q, *J* = 30.7 Hz, CH, C $\delta$   $\Psi$ Pro), 120.1 (2  $\times$  CH, Fmoc arom), 122.9 (q, *J* = 286.6 Hz, CF<sub>3</sub>), 125.2 (2  $\times$  CH, Fmoc arom), 127.1 (2  $\times$  CH, Fmoc arom), 127.8 (2  $\times$  CH, Fmoc arom), 141.4 (2  $\times$  CH, Fmoc arom), 155.4 (C, C=O), 168.8 (C, C=O), 173.0 (C, C=O); <sup>19</sup>F NMR (376.2 MHz, CDCl<sub>3</sub>)  $\delta$  -81.9 (d, *J* = 5.2 Hz); ESIMS *m/z*: 519.28 [M + H]<sup>+</sup>; Anal. Calcd for C<sub>26</sub>H<sub>25</sub>F<sub>3</sub>N<sub>2</sub>O<sub>6</sub> (518.17): C, 60.23; H, 4.86; N, 5.40, found: C, 60.22; H, 4.84; N, 5.51.

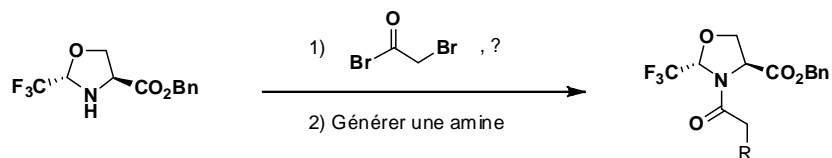
#### Fmoc-Pro-Ser( $\Psi$ Pro<sup>CF<sub>3</sub>,H</sup>)-OH **4**



La saponification du dipeptide **3** (1 g, 1,93 mmol, 1 équivalent) a été réalisée par addition de NaOH (108 mg, 2,70 mmol, 1,4 éq.) à une solution 0,8 M CaCl<sub>2</sub> dans iPrOH-H<sub>2</sub>O 7:3 (50ml). Le mélange réactionnel a été agité pendant 4 h à température ambiante, désactivé avec HCl 1M, puis concentré sous pression réduite et dilué avec H<sub>2</sub>O. La solution aqueuse a été extraite avec de l'acétate d'éthyle (3  $\times$  50 ml). Les phases organiques combinées ont été lavées avec H<sub>2</sub>O (20 ml), séchées sur du MgSO<sub>4</sub>, filtrées et concentrées sous pression réduite. Une chromatographie éclair (95:5 de dichlorométhane/méthanol) a donné 963 mg (86%) de l'acide **4** correspondant.



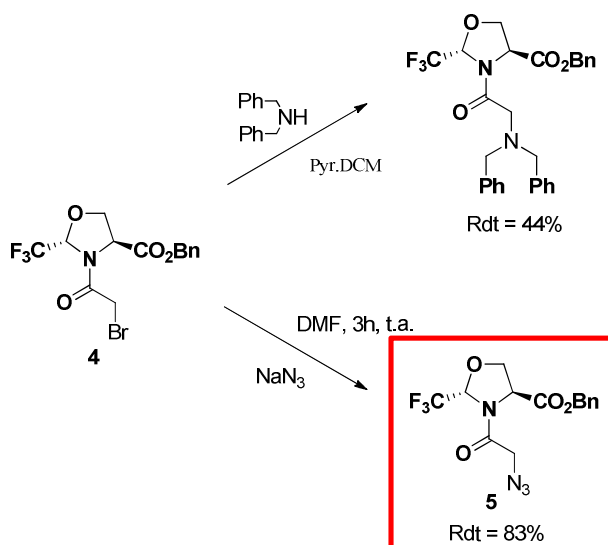
## ANNEXE 6: Optimisation du couplage G-ΨPro

3. Optimisation de l'étape 1

Pour piéger l'équivalent d'acide chlorhydrique généré lors de l'acylation, la réaction a été réalisée en présence de base. Quelle que soit la base utilisée (pyridine, bicarbonate de sodium), ou le nombre d'équivalents, la réaction procède avec épimérisation. Nous avons également testé différents solvants (THF, DCM) en jouant notamment sur la polarité. Là encore, les résultats ont été peu concluants. Enfin, la température joue quant à elle un rôle très important. En effet, il a été démontré qu'une légère hausse de température favorisait fortement l'épimérisation, et donc la formation du produit acylé comportant la pseudoproline de stéréochimie *syn*.

4. Optimisation de l'étape 2

L'élaboration du résidu glycine a tout d'abord été envisagée en réalisant la substitution du brome par la dibenzylamine. Cette stratégie ne s'est pas avérée satisfaisante, le dipeptide attendu étant obtenu avec un faible rendement de 44%. La réaction de substitution par l'azoture de sodium a ensuite été testée. Un premier essai à 80°C dans le DMF a conduit à des produits de dégradation. Puis la réaction a été réalisée à température ambiante, le produit a été obtenu avec un rendement satisfaisant de 83%.







# Conclusion et perspectives

Le développement de CMP a débuté dans les années 1960 en vue de comprendre la stabilité de la triple hélice de collagène.<sup>1</sup> A travers cet ouvrage, nous avons donc contribué à ces travaux à plus d'un titre. Nous avons montré que de simples spectres en RMN du proton permettaient d'étudier les échanges monomère-triple hélice, que ce soit en condition d'équilibre thermodynamique ou hors équilibre, pour caractériser les cinétiques d'assemblage. Ce travail a nécessité le développement de méthode de synthèse nous permettant d'avoir des quantités suffisantes de peptides pour pouvoir ensuite les engager dans des études biophysiques. En utilisant un marquage spécifique au  $^{13}\text{C}$  et  $^{15}\text{N}$  des glycines, nous avons pu mettre en évidence par RMN hétéronucléaire la présence des deux populations majoritaires, monomère et triple hélice, mais également de nombreux conformères minoritaires. Cette approche nous a permis d'inspecter pour chacun d'eux les paramètres pouvant affecter leur stabilité (concentration, température) alors que le dichroïsme circulaire ne détecte que la population PPII. A notre connaissance, personne n'avait encore utilisé cette méthode de marquage spécifique pour caractériser cet équilibre jusqu'à la publication de *Hartgerink et al.*<sup>2</sup> parue au début de cette année. Nos études sont cependant complémentaires. Ils ont pu montrer des différences structurales aux extrémités des triples hélices ainsi que la nécessité d'avoir au moins deux triplets à chaque extrémité pour stabiliser la triple hélice. Nous avons également fait ces observations et nous avons par ailleurs obtenu des informations structurales locales par le développement de l'expérience  $\text{CH}_2\text{-TROSY}$ . Grâce à cette expérience, nous avons pu déterminer les constantes de couplages  $^3J_{\text{HN-H}\alpha}$  sur toute la longueur du peptide, celles-ci étant directement liées à l'angle  $\phi$  par la relation de Karplus. A ce jour, les valeurs obtenues ne sont pas en bon accord avec les études de dynamique moléculaire ce qui reste préoccupant. L'établissement de liaisons hydrogène, l'échange des protons amides, ou la faible fraction d'amides deutérés pourront être pris en compte pour rapprocher au mieux expérience et théorie. Le développement de deux expériences à 3 dimensions a été réalisé : s3-HACAN et s3-HACACO. Celles-ci ont été appliquées aux peptides  $^4(\text{POG})_7$  et  $^3(\text{POG})_7$ , elles permettent d'obtenir pour chaque glycine et chaque conformère de nombreuses constantes de couplages, jusqu'au 12 pour l'expérience s3-HACAN. L'enregistrement de ces

expériences 3D sur les peptides marqués aux extrémités sera réalisé dans un avenir très proche afin de confronter les valeurs de couplages mesurées aux analyses par dynamique moléculaire. Nous attendons ces derniers résultats pour publier ces développements méthodologiques en RMN.

Par ailleurs, nous avons mis en évidence la présence de deux types d'assemblages supramoléculaires qui résultent, selon nous, soit d'un assemblage imparfait des monomères, reliés localement par des amorces de triples hélices, soit d'un assemblage plus organisé issu des interactions latérales et axiales (tête-queue) de triples hélices. Ce modèle est fondé sur nos études combinées de DLS, dichroïsme circulaire, SHG et RMN liquide/solide. La littérature rend compte de la présence de ces différents types d'assemblages par la formation de gels<sup>3-6</sup> à forte concentration ou de fibres<sup>7</sup> mais ceux-ci sont très peu étudiés d'un point de vue moléculaire. Le modèle que nous avons établi nous a permis d'obtenir des paramètres thermodynamiques pour les assemblages de (POG)<sub>7</sub> solubles : constantes de vitesse, variations d'enthalpie et d'entropie des assemblages en triple hélice et des assemblages désordonnés de haut poids moléculaire. La modélisation des cinétiques met en évidence une absence d'échange conformationnel direct entre les espèces triples hélices et les espèces oligomériques solubles. Pour finaliser cette partie du travail, nous prévoyons de réaliser des études complémentaires à différentes concentrations et pH. Il serait par ailleurs très intéressant de réaliser des études similaires sur le peptide (POG)<sub>10</sub>, qui à forte concentration, forme des assemblages insolubles fortement organisés. Nous souhaiterions également caractériser à l'échelle moléculaire ces assemblages. Les études de SHG, microscopie électronique (ombrage) RMN solide indiquent de manière convergente que ces espèces insolubles ont un degré d'ordre élevé mais nous ne disposons pas pour le moment d'information à l'échelle moléculaire. Des essais supplémentaires seront effectués pour obtenir des quantités de fibres marqués <sup>13</sup>C, <sup>15</sup>N qui soient compatibles avec une étude par RMN du solide. Nous poursuivons par ailleurs nos efforts pour trouver des conditions expérimentales permettant d'observer et de caractériser ces assemblages par des techniques de microscopie à très haute résolution. Il serait intéressant de voir comment se comporte le peptide (POG)<sub>7</sub> à plus forte concentration, celui-ci forme-t-il des assemblages supramoléculaires insolubles tout comme le (POG)<sub>10</sub> ?

Les résultats que nous avons obtenus nous ont permis de définir des paramètres critiques pour la conception de molécules CMP. Celles-ci doivent pouvoir former rapidement des triples hélices stables, qui au-delà, doivent s'autoassocier et former des réseaux organisés.

Sous sa forme monomérique, le peptide CMP de référence (POG)<sub>n</sub> explore plusieurs conformations en raison de l'équilibre *cis-trans* des liaisons Gly-Pro. Les peptides contenant une ou plusieurs liaisons *cis* ne sont pas compatibles avec la formation d'une triple hélice mais peuvent cependant s'autoassocier pour former des oligomères de haut poids moléculaire. Un premier élément que nous avons envisagé dans la conception de nouveau CMP a été d'abaisser la barrière d'énergie *cis-trans* grâce à l'emploi de motifs oxazolidine et la conjugaison de groupements trifluorométhylés fortement électroattracteurs. L'effet de ces pseudoprolines fluorés devrait (i) accélérer la formation des triples hélices, (ii) inhiber la formation d'agrégats. Par ailleurs, nous avons montré que les motifs (2*S*,4*R*)-Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro) et (2*R*,4*R*)-Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro) avaient la capacité de contraindre les plissements de cycle des motifs oxazolidines. Cette propriété est cruciale pour la conception de CMP car il a été montré que l'alternance *down-up* était déterminante pour stabiliser la triple hélice de collagène. Le motif (2*S*,4*R*)-Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro) est un très bon mime de l'hydroxyproline puisqu'il stabilise fortement le plissement *up* mais au sein du tripeptide Pro- $\Psi$ Pro-Gly (P3), il influence également le plissement de cycle de la proline précédente en *up*. Ceci peut s'expliquer par l'encombrement du groupe CF<sub>3</sub> ainsi que par la faible contrainte de plissement du cycle proline. L'enchaînement (2*S*,4*R*)-Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro)-(2*R*,4*R*)-Ser( $\Psi^{\text{CF}_3,\text{H}}$ Pro) quant à lui, possède bien l'alternance *down-up* requise pour former la triple hélice de collagène. L'étude RMN approfondie des propriétés structurales par RMN de différents tripeptides incorporant ces motifs fait l'objet d'une publication qui sera soumise prochainement à Organic & Biomolecular Chemistry. La synthèse et l'étude structurale de plus longs peptides fluorés ont été entreprises. Nous ne sommes parvenus pour le moment qu'à insérer le triplet Pro- $\Psi$ Pro-Gly dans ces nouveaux CMP. Nous avons pu voir que la triple hélice était alors fortement déstabilisée, voire complètement inexistante, ce qui était prédictible du fait du non-respect de l'alternance des plissements de cycle.<sup>8,9</sup> La synthèse de peptides incorporant le motif  $\Psi$ Pro- $\Psi$ Pro-Gly est en cours de développement. Les avancées méthodologiques effectuées pour les études des peptides (POG)<sub>n</sub> (n=7 ou 10) seront mises à profit pour étudier ces nouveaux CMP. Nous espérons de ces peptides qu'ils augmentent la stabilité de la triple hélice mais également qu'ils promeuvent des assemblages supramoléculaires grâce à des interactions latérales fortes médiées par les atomes de fluor. En effet, nous avons pu visualiser par modélisation moléculaire que les atomes de fluor pointent vers l'extérieur de la triple hélice. Ce positionnement serait donc propice à promouvoir des interactions inter-triples hélices et stabiliser des assemblages supramoléculaires.

Il est encore difficile d'envisager, à ce jour, des biomatériaux ne contenant que des peptides synthétiques. C'est pourquoi nous avons également testé la capacité des  $(\text{POG})_n$  à moduler les phases denses de collagène naturel ce qui pourrait permettre la conception de nouveaux biomatériaux aux propriétés originales. En effet, *Wang et al.* ont montrés qu'il était possible de modifier les phases de collagène par dopage avec des peptidomimétiques.<sup>10</sup> Nous avons engagé des études de concentration, à différents ratio peptide:collagène et observé par microscopie ou RMN les interactions éventuelles. Jusqu'à présent, nous n'avons obtenu aucune évidence expérimentale nous permettant d'affirmer que  $(\text{POG})_{10}$  ou  $(\text{POG})_7$  s'intercalent dans les phases de collagène. L'utilisation de CMP fluorés pour doper des phases de collagène représente un autre champ d'investigation possible, les atomes de fluor étant bien connus pour augmenter les affinités des petites molécules pour les protéines.

## Références de la conclusion

- 1 Fields, G. B. & Prockop, D. J. Perspectives on the synthesis and application of triple-helical, collagen-model peptides. *Biopolymers : Peptide Science* **50**, 345-357 (1996).
- 2 Acevedo-Jake, A. M., Jalan, A. A. & Hartgerink, J. D. Comparative NMR analysis of collagen triple helix organization from N- to C-termini. *Biomacromolecules* **16**, 145-155, doi:10.1021/bm501281a (2015).
- 3 Yamazaki, C. M., Asada, S., Kitagawa, K. & Koide, T. Artificial collagen gels via self-assembly of de novo designed peptides. *Biopolymers* **90**, 816-823, doi:10.1002/bip.21100 (2008).
- 4 O'Leary, L. E., Fallas, J. A., Bakota, E. L., Kang, M. K. & Hartgerink, J. D. Multi-hierarchical self-assembly of a collagen mimetic peptide from triple helix to nanofibre and hydrogel. *Nature chemistry* **3**, 821-828, doi:10.1038/nchem.1123 (2011).
- 5 Suehiro, T. *et al.* Temperature-dependent higher order structures of the (Pro-Pro-Gly)(1)(0)-modified dendrimer. *Biopolymers* **95**, 270-277, doi:10.1002/bip.21576 (2011).
- 6 He, L. & Theato, P. Collagen and collagen mimetic peptide conjugates in polymer science. *European Polymer Journal* **49**, 2986-2997, doi:10.1016/j.eurpolymj.2013.05.033 (2013).
- 7 Kar, K. *et al.* Self-association of collagen triple helix peptides into higher order structures. *The Journal of biological chemistry* **281**, 33283-33290, doi:10.1074/jbc.M605747200 (2006).
- 8 Shoulders, M. D. & Raines, R. T. Collagen structure and stability. *Annual review of biochemistry* **78**, 929-958, doi:10.1146/annurev.biochem.77.032207.120833 (2009).
- 9 Holmgren, S. K., Taylor, K. M., Bretscher, L. E. & Raines, R. T. Code for collagen's stability deciphered. *Nature* **392** (1998).
- 10 Wang, A. Y., Mo, X., Chen, C. S. & Yu, S. M. Facile Modification of Collagen Directed by Collagen Mimetic Peptides. *Journal of the American Chemical Society* **127**, 4130-4131 (2005).