



HAL
open science

Mimes peptidiques de glycosaminoglycanes : synthèses, structures et activités biologiques

Rodrigue Marquant

► **To cite this version:**

Rodrigue Marquant. Mimes peptidiques de glycosaminoglycanes : synthèses, structures et activités biologiques. Chimie organique. Sorbonne Université, 2021. Français. NNT : 2021SORUS305 . tel-03905719

HAL Id: tel-03905719

<https://theses.hal.science/tel-03905719v1>

Submitted on 19 Dec 2022

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

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



Sorbonne Université

Ecole Doctorale de Chimie Moléculaire Paris Centre

Laboratoire des BioMolécules – UMR7203

Mimes peptidiques de glycosaminoglycanes : synthèses, structures et activités biologiques

Par Rodrigue Marquant

Thèse de doctorat de Chimie

Dirigée par Dr Jean-Maurice Mallet et Pr Solange Lavielle

Présentée et soutenue publiquement le 17 Décembre 2021

Devant un jury composé de :

Dr Gilles GUICHARD	Rapporteur
Dr Carine VAN HEIJENOORT	Rapporteuse
Pr Florence PILARD	Examinatrice
Dr Fabienne BURLINA	Examinatrice
Pr Alain PROCHIANTZ	Examineur
Dr Jean-Maurice MALLET	Directeur de thèse
Dr Solange LAVIELLE	Invitée

*A mon épouse et meilleure amie, Camille,
qui supporte mes bouderies avec amour...
et un soupçon de fatalisme ^^*

Remerciements

Je tiens tout d'abord à remercier très chaleureusement le Docteur Carine Van Heijenoort, Directrice de Recherche à l'Institut de Chimie des Substances Naturelles de l'Université Paris-Saclay et le Docteur Gilles Guichard, Directeur de Recherche à l'Institut de Chimie et Biologie des Membranes et des Nano-objets de Bordeaux, d'avoir accepté d'être les rapporteurs de ce travail. Je remercie également le Professeur Florence Pilard du Laboratoire de Glycochimie, des Antimicrobiens et des Agroressources de l'Université de Picardie Jules Verne d'Amiens, le Docteur Fabienne Burlina, Directrice de Recherche au Laboratoire des BioMolécules (LBM) de Sorbonne Université à Paris et le Professeur Alain Prochiantz du Centre Interdisciplinaire de Recherche en Biologie du Collège de France à Paris, de me faire l'honneur de participer à ce jury.

Avant de remercier comme il se doit toutes les personnes qui m'ont accompagné, de près ou de loin, dans cette aventure, je souhaiterais faire un bond dans le temps, d'une petite vingtaine d'année en arrière, au moment où le jeune homme que j'étais alors a dû faire un choix : continuer ses études vers la recherche académique et un DEA (on dira aujourd'hui un Master recherche) ou opter pour un DESS (Master professionnel). J'ai choisi le DESS et la possibilité d'entrer plus rapidement dans la vie active, et, si je pouvais effectivement remonter le temps, je ne pense pas que je changerais quoi que ce soit. Ce choix a en partie défini la personne que je suis, me confrontant, certes, à un certain nombre de difficultés mais, surtout, me faisant rencontrer des personnes formidables.

La première de ces personnes que je souhaite remercier est le Dr Serge Bouaziz, de l'Université de Paris. Il a été le premier à me faire confiance, malgré mon manque d'expérience et alors même que j'avais fait une croix sur la possibilité de refaire un jour de la Science. Je peux dire à présent que c'est grâce à lui si, par la suite, tout s'est bien enchaîné pour moi. Travailler avec Serge m'a permis de faire une autre rencontre déterminante dans ma vie professionnelle aussi bien que personnelle, en la personne du Pr Solange Lavielle. A l'époque Directrice du laboratoire qui allait devenir MON labo, Solange n'a cessé de m'accompagner tout au long de ma vie et elle continue encore aujourd'hui à le faire. Si j'ai eu la chance de pouvoir réaliser cette thèse c'est en grande partie parce qu'elle a su me conseiller, me guider pour que je laisse davantage entrer la recherche dans mon travail. Si je n'ai pas abandonné

face aux difficultés rencontrées pour réaliser un travail de thèse tout en assurant les tâches quotidiennes inhérentes à mon métier d'ingénieur, c'est une fois de plus grâce à son soutien. Je sais la chance que j'ai d'avoir rencontré une personne, une amie, telle que Solange et je suis fier de dire que je suis le dernier thésard qu'elle aura encadré.

Je remercie très chaleureusement le Dr Jean-Maurice Mallet de m'avoir proposé le projet de recherche qui est devenu mon sujet de thèse. Merci également d'avoir accepté d'être mon directeur de thèse et en cela de m'avoir permis de développer ma curiosité scientifique et mon envie de voir plus loin, d'avoir su me laisser la liberté nécessaire à mon épanouissement tout en m'apportant les données nécessaires à mon apprentissage. Je suis toujours en admiration devant la multitude d'idées qui germent de l'esprit de Jean-Maurice, ce n'est pas toujours très simple de le suivre mais comme c'est stimulant !

Si cette thèse a pu voir le jour au sein du LBM c'est également grâce au Dr Sandrine Sagan puis au Pr Olivier Lequin, qui, en tant que Directrice/teur de l'unité, ont amicalement accepté que soit consacrée une partie de mon temps de travail à ce projet. De même, je n'aurais pas pu travailler sur mon sujet de thèse si mes collègues, Roba Moumné et Cillian Byrne, n'avaient pas accepté de prendre en charge une partie de mon travail de gestion des laboratoires. Merci de tout cœur.

Deux autres personnes ont largement contribué à me faciliter la vie, et même à l'embellir, durant ces quatre années de thèse. Il s'agit de Agathe Brudo et Taleen Peker, mes deux thésardes préférées. Merci pour votre investissement sans faille pour le labo et surtout merci pour votre gentillesse, vos petites attentions, vos sourires, votre bonne humeur, vos confidences, votre soutien... Je n'aurais pas pu rêver meilleures partenaires de thèse que vous deux pour m'accompagner.

Côté sciences (même si le côté humain n'est jamais très loin), je tiens à remercier les nombreuses personnes qui se sont investies dans mon travail et ont contribué à enrichir cette thèse. Je pense à Christophe Piesse et Tahar Bouceba, de l'Institut de Biologie Paris-Seine (IBPS) de Sorbonne Université, à qui je dois, respectivement, la synthèse de plusieurs de mes peptides et les expériences réalisées sur le Biacore. Merci à Gilles Clodic, ingénieur de la plateforme de spectrométrie de masse de l'IBPS, pour son travail sur la caractérisation des peptides par MS et à Gérard Bolbach pour m'avoir expliqué les subtilités des analyses des composés sulfatés

par MS. Merci à Nicolas Pietrancosta pour les expériences de modélisation, à Françoise Illien pour m'avoir, entre autres, apporté son aide pour les mesures de fluorescence et à Thierry Drujon et Claire Lacombe pour le dichroïsme circulaire. Je remercie également Delphine Ravault, qui a eu la lourde tâche de me familiariser avec l'expression des protéines (et surtout la préparation des gels d'électrophorèse...), et Ludovic Carlier, dont la patience n'a d'égale que la gentillesse et qui a su me réconcilier, au moins en partie, avec la RMN.

Un grand merci à nos collaborateurs du Collège de France : Pr Alain Prochiantz, Dr Ariel Di Nardo, Rachel Gibel-Russo et particulièrement Damien Testa dont une partie du travail de thèse, réalisée avec les sulfopeptides, me permet de présenter de beaux résultats biologiques.

Grâce au travail de Thomas Palacin, post-doctorant au laboratoire, une nouvelle génération de mimes de GAGs va faire son apparition, alors merci à Thomas d'avoir pris la relève !

Un laboratoire ne fonctionnant pas sans gestion administrative, je tiens également à remercier Eliane Moulinié et Karine Gherdi, non seulement pour leur travail et leur efficacité, mais aussi pour leur sympathie et les discussions que nous avons eu (même s'il est vrai que l'une des deux est peut-être un peu plus bavarde que l'autre...).

J'ai une pensée toute particulière pour Gérard Chassaing, avec qui j'ai toujours plaisir à discuter même si nous avons à présent moins l'occasion de le faire, pour Laurent Cattiaux, mon collègue doctorant et ami de longue date puisque nos vies ne cessent de se croiser, et pour Sophie Cribier, qui sait toujours comment me faire plaisir avec ses tablettes de chocolat au lait.

Enfin, parce qu'avoir un mari et un père doctorant ce n'est pas drôle tous les jours, je remercie pour leur patience ma femme et mes enfants et je leur envoie tout mon amour.

Abréviations

1, 2 ou 3D	1, 2 ou 3 dimensions
A	Absorbance
AA	Acide aminé
Ac	Acétyl
AcOH	Acide acétique
ADN	Acide désoxyribonucléotidique
Ahx	Acide amino-6-hexanoïque
Ala, A	Alanine
APS	Persulfate d'ammonium
Arg, R	Arginine
Arg ₉	Nonaarginine (RRRRRRRRR)
ARN	Acide ribonucléotidique
Asn, N	Asparagine
Asp, D	Acide Aspartique
B	Bateau
BBL	Building blocks library
BNP	Peptide cérébral natriurétique
BPTI	Inhibiteur de la trypsine pancréatique bovine
Boc	<i>t</i> -Butyloxycarbonyl
BSA	Albumine de sérum bovin
Bzl	Benzylque (de type)
c	Concentration
C	Chaise
C'	Acide Cystéique
Cbz, Z	Carbobenzoxy
ChABC	Chondroïtinase ABC
CI ₅₀	Concentration inhibitrice médiane
CIRB	Centre Interdisciplinaire de Recherche en Biologie
COMU	Hexafluorophosphate de (1-Cyano-2-ethoxy-2-oxoethylidenaminoxy)diméthylamino-morpholino-carbenium
CP	Période critique
CPC	Cation-Polaire-Cation
CPP	Peptide pénétrant
CS	Sulfate de chondroïtine
CSD	Déplacement chimique secondaire
CSP	Perturbation de déplacement chimique
CSPG	Protéoglycanes de sulfate de chondroïtines
CSP _{max}	CSP maximale obtenue pour 100 % de protéine liée
C-ter	C-terminal
Cys, C	Cystéine
Da	Dalton
DB	Dot-Blot
δ_b / δ_f	Déplacements chimiques sous la forme liée / libre
$\delta^1\text{H}_N / \delta^{15}\text{N}_H$	Déplacements chimiques des noyaux $^1\text{H} / ^{15}\text{N}$
$\Delta\delta^1\text{H}_N$	Variation du déplacement chimique du proton amide

$\Delta\delta^{15}\text{N}_\text{H}$	Variation du déplacement chimique de l'azote
DC	Dichroïsme circulaire
DCC	<i>N,N'</i> -dicyclohexylcarbodiimide
DCM	Dichlorométhane
Dha	Déhydroalanine
DIC	<i>N,N'</i> -diisopropylcarbodiimide
DIEA	<i>N,N'</i> -diisopropyléthylamine
DMF	<i>N,N'</i> -diméthylformamide
DMSO	Diméthylsulfoxyde
DO	Densité optique
DS	Dermatane sulfate
DSS	Acide 4,4-diméthyl-4-silapentane-1-sulfonique
DTT	Dithiothréitol
E	Enveloppe
ECM	Matrice extracellulaire
EDC	(1-éthyl-3-(3-diméthylaminopropyl)carbodiimide
EDT	Ethanedithiol
EDTA	Éthylènediaminetétraacétique
ELISA	Dosage immuno-enzymatique sur support solide
EMSA	Electrophoretic mobility shift assay ou retard sur gel
En2	Engrailed 2
ϵ_th	Coefficient d'extinction molaire théorique
EtOH	Ethanol
FG	Filtration sur gel
FGF	Facteur de croissance des fibroblastes
FITC	Isothiocyanate de fluorescéine
Fluo	Fluorescéine
FPLC	Chromatographie liquide de protéine rapide
Fmoc	9-fluorénylméthoxy-carbonyl
Fuc	L-Fucose
GAG	Glycosaminoglycane
Gal	D-Galactose
GalNAc	<i>N</i> -Acétyl-D-galactosamine
gg	Gauche-gauche
Glc	D-Glucose
GlcA	Acide D-Glucuronique
GlcNAc	<i>N</i> -Acétyl-D-glucosamine
Gln, Q	Glutamine
Glu, E	Acide Glutamique
Gly, G	Glycine
GP	Groupe protecteur
GPI	Glycosylphosphatidylinositol
gt	Gauche-trans
H	Demi-chaise
HA	Acide hyaluronique
HABA	Acide 4'-hydroxyazobenzène-2-carboxylique
HATU	Hexafluorophosphate de (diméthylamino)- <i>N,N</i> -diméthyl(1H-

	[1,2,3]triazolo[4,5-b]pyridin-3-yloxy)méthaniminium
HBTU	Hexafluorophosphate de 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tétraméthyluronium
HCCA	α -cyano-4-hydroxycinnamique
HexaCSE	Analogue hexasaccharidique de CS-E
His, H	Histidine
HOBt	Hydroxybenzotriazole
Hp	Héparine
HPLC	Chromatographie liquide haute performance
HRP	Peroxydase de Raifort
HS	Sulfate d'héparane
HSPG	Protéoglycanes de sulfate d'héparane
HSQC	Corrélation quantique unique hétéronucléaire
IBPS	Institut de Biologie Paris-Seine
Ile, I	Isoleucine
IPTG	Isopropyl- β -D-1-thiogalactopyranoside
IsoAsp	Acide β -aspartique
ITC	Calorimétrie de titrage isotherme
K _d	Constante de dissociation à l'équilibre
K _{off}	Constante cinétique de dissociation
K _{on}	Constante cinétique d'association
KS	Kératane sulfate
l	Longueur du trajet parcouru par la lumière dans la solution
LB	Lysogeny broth ou bouillon lysogène
LBM	Laboratoire des BioMolécules
Leu, L	Leucine
[L]t	Concentration totale du ligand
Lys, K	Lysine
MALDI-TOF	Désorption et ionisation laser assistées par une matrice, couplées à un détecteur de temps de vol
Man	D-Mannose
MAP	Peptide antigenic multiple
MBHA	4-méthylbenzhydrylamine
MeOH	Méthanol
MERS-CoV	Coronavirus du syndrome respiratoire du Moyen-Orient
Met, M	Méthionine
Mmt	Méthoxytrityl
MS	Spectrométrie de masse
MW	Masse moléculaire
m/z	Rapport masse sur charge
NCL	Ligation chimique native
nd	Non déterminé
NeuAc	Acide <i>N</i> -acétylneuraminique
NHS	<i>N</i> -hydroxysuccinimide
NLS-HD	Homéodomaine étendu en <i>N</i> -terminal
NMP	<i>N</i> -Méthyl-2-pyrrolidone
<i>N</i> -ter	<i>N</i> -terminal
OBzl	4-benzyl ester

OtBu	Tert-butyl ester
Otx2	Orthodenticle homeobox 2
Oxyma Pure	Cyanohydroxyiminoacétate d'éthyle
PA	Peptide amphiphile
PBS	Tampon Phosphate Salin
PDB	Protein data bank
PF	Polarisation de fluorescence
PG	Protéoglycane
pGlu	Acide pyroglutamique
pl	Point isoélectrique
pMeOBzl	4-methoxybenzyl
PMSF	Fluorure de phénylméthylsulfonyle
PolyGlu	n-Butyl-poly(L-glutamic acid) ₂₀ sodium salt
Phe, F	Phénylalanine
pmg	Pyrèneméthylguanidinium
PNN	Réseaux périneuronaux
ppm	Parties par million
Pro, P	Proline
PSS-Na	Poly(styrène sulfonate) de sodium
[P]t	Concentration totale de protéine
PV	Parvalbumine
PVDF	Poly(fluorure de vinylidène)
RMN	Résonance Magnétique Nucléaire
rpm	Rotation par minute
R _T	Temps de rétention
RU	Unité de résonance
RX	Radiocristallographie
S	Chaise croisée
SA	Streptavidine
SARS-CoV-2	Coronavirus du syndrome respiratoire aigu sévère
SC-peptide	Scrambled peptide
SDS	Dodécylsulfate de sodium
SDS-PAGE	Électrophorèse sur gel de polyacrylamide en présence de dodécylsulfate de sodium
Sema-3A	Sémaphorine-3A
Ser, S	Sérine
Shh	Sonic Hedgehog
SNC	Système nerveux central
SPPS	Synthèse peptidique en phase solide
SPR	Résonance plasmonique de surface
T	Torsion
tBu	<i>tertio</i> -butylique (de type)
TEMED	<i>N,N,N',N'</i> -tétraméthyl-1-,2-diaminométhane
TFA	Acide trifluoroacétique
TFE	Trifluoroéthanol
tg	Trans-gauche
Thr, T	Thréonine

TIS	Triisopropylsilane
Trp, W	Tryptophane
Trt	Trityl
Tyr, Y	Tyrosine
UV	Ultraviolet
Val, V	Valine
WB	Western Blot
WFA	Agglutinine de Wisteria Floribunda
Xyl	D-Xylose

Sommaire

Sommaire	1
CHAPITRE I : Introduction.....	5
A. Deux familles de molécules du vivant.....	7
1. Les protéines et peptides	7
a) Structure et nomenclature des acides aminés	7
b) Biosynthèse de peptides et protéines	9
c) Structure des peptides et protéines	10
d) Synthèse chimique des peptides	13
e) Résumé de manière schématique de quelques faits que nous estimons majeurs dans l’histoire de la chimie des peptides/protéines.	14
2. Les glucides et les glycosaminoglycanes	15
a) Structure et nomenclature des monosaccharides.....	15
b) Polysaccharides et glycosaminoglycanes.....	18
c) Structure des osides et des GAGs	22
d) Synthèse chimique des oligosaccharides et des GAGs	27
e) Quelques faits majeurs dans l’Histoire de la chimie des glycanes	34
B. Les « GAG-Binding Proteins » ou Protéines de liaison aux Glycosaminoglycanes	34
1. Les GAG-Binding proteins et la matrice extracellulaire.....	35
2. Analyse des interactions entre GAGs et GAG-binding proteins	41
3. Les interactions entre les protéines à homéodomaine et les chondroïtines sulfates dans le système nerveux central.....	42
a) Organisation du réseau CSPG dans le SNC	42
b) Plasticité et période critique.....	44
c) Protéines à Homéodomaine et Otx2	46
C. Objectif de ce travail	49

CHAPITRE II : SYNTHÈSE DES MIMES DE GAG..... 55

A. Introduction : des sucres aux peptides.....56

B. Principe de la synthèse peptidique sur support solide61

C. Stratégie de préparation des mimes peptidiques de CS-E66

1. Séquence modèle pour les essais préliminaires, un dimère de térapeptide.67

2. Essais préliminaires, recherche d'un protocole.71

3. Les premiers mimes peptidiques synthétisés.74

 a) Protocole de synthèse des premiers mimes..... 75

 b) Oxydation des premiers mimes. 77

 c) Etude des conditions de conservation/stockage des premiers mimes. 79

4. Protocole de synthèse de pentamères et hexamères de mimes peptidiques.81

5. Partie expérimentale.....84

CHAPITRE III : ÉTUDES BIOCHIMIQUES ET BIOLOGIQUES DES MIMES

PEPTIDIQUES DE GAGS..... 89

A. Introduction90

B. Les expériences92

C. Les résultats92

1. Expérience de dot blots.....92

2. Expérience de pull-down.....93

3. Expérience d'infusion de sulfopeptides *in vivo*.95

D. Conclusion générale de ce chapitre98

CHAPITRE IV : ETUDES DES INTERACTIONS *IN VITRO* PAR RMN ET SPR DE TYPE

BIACORE.	117
A. Introduction	118
B. Etude des interactions par RMN	119
1. Surexpression et purification des protéines pour l'étude structurale	119
a) Surexpression et purification du domaine NLS-HD de la protéine Otx2 humaine.....	119
b) Surexpression et purification du domaine NLS-HD de la protéine Engrailed 2 (En2) de poulet	126
2. RMN et études structurales	127
a) La Résonance Magnétique Nucléaire, les séquences et stratégies utilisées.....	128
b) Résultats des études structurales par RMN	133
3. Discussion.....	155
4. Conclusion - Perspectives.....	160
C. Etude des interactions par Biacore	161
1. Synthèse de nouveaux analogues	161
2. Le Biacore, principe et résultats.....	164
a) Principe de la technologie Biacore	164
b) Résultats des expériences réalisées sur Biacore	167
3. Discussion.....	174
D. Conclusion	175
E. Partie expérimentale	176
1. Production de protéines recombinantes et RMN	176
a) Production des protéines recombinantes	176
b) Etudes RMN	181
2. Synthèse de nouveaux analogues et Biacore	185
a) Synthèse de Ac-Ahx-(EC'PC') ₆ -NH ₂ et Ac-Ahx-(EC'LC') ₆ -NH ₂	185
b) Biacore	186

CHAPITRE V : TRAVAUX ET RESULTATS PRELIMINAIRES	189
A. Introduction	190
B. Travaux en cours et résultats préliminaires	190
1. Synthèse d'un analogue linéaire avec agrafe.....	190
2. Synthèse de polymères non linéaires.....	193
a) Polymères de type dendrimère	193
b) Polymères de type branchés.....	197
3. Exploration de nouveaux protocoles de mesures d'affinité.....	198
a) Fluorescence du Tryptophane	199
b) Polarisation de fluorescence.....	199
C. Partie expérimentale	207
1. Synthèse des analogues linéaires avec agrafe	207
2. Synthèse des polymères de type dendrimère	207
3. Purifications par HPLC préparative	208
4. Caractéristiques HPLC et spectrométrie de masse.....	208
5. Fluorescence du Tryptophane.....	209
6. Polarisation de fluorescence	209
 CHAPITRE VI : CONCLUSION GENERALE ET PERSPECTIVES	 211
 Bibliographie	 214

CHAPITRE I : Introduction

L'intitulé de ce mémoire : « ***Mimes peptidiques de glycosaminoglycanes (GAGs) : synthèses, structures et activités biologiques*** » impose avant tout une présentation des deux partenaires, tout particulièrement leurs nomenclatures et caractéristiques : les peptides (et par extension les protéines), constitués d'acides aminés, et les glycosaminoglycanes, constitués de diholosides (ou disaccharides). Les acides aminés d'une part et les oses d'autre part représentent les briques, unités de base, de deux grandes familles de molécules du Vivant : les peptides/protéines et les glucides. Ceci constituera la première partie de ce chapitre introductif. Les deux autres grandes familles sont les oligonucléotides et -sides (ADN et ARN) et les lipides.

Dans une deuxième partie, je présenterai les « GAG-binding proteins », les protéines de liaison aux GAGs, et leurs interactions avec les GAGs, en me focalisant plus particulièrement sur les chondroïtines sulfates. Je présenterai l'organisation de la matrice extracellulaire dans le système nerveux central ainsi que les protéines à homéodomaine et le rôle de l'une d'entre elles dans la vision, la protéine codée par le gène orthodenticle homéobox 2, Otx2.

Puis, pour conclure, je présenterai l'objectif de mon travail de recherches.

A. Deux familles de molécules du vivant

1. Les protéines et peptides

a) Structure et nomenclature des acides aminés

Un acide aminé est une molécule possédant une fonction amine (-NH₂), une fonction acide carboxylique (-COOH) et un squelette carboné, tous les trois portés par le même carbone, dit C_α, **Figure 1**.

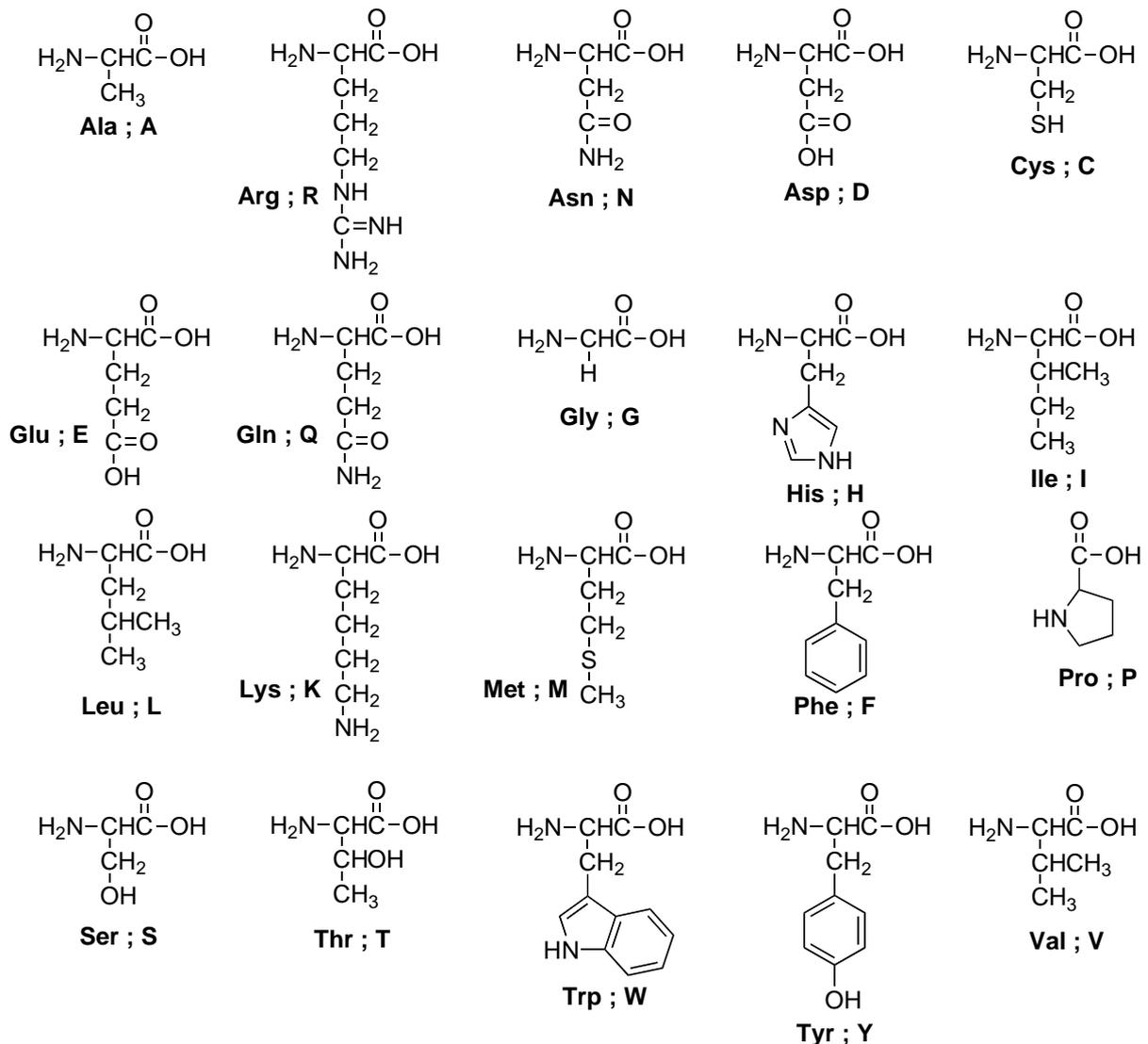


Figure 1. Structures et formules des vingt acides aminés protéinogéniques.

Même si de très nombreux acides aminés (AAs), autour de 400-500, ont été isolés dans différentes espèces, chez l'Homme ont été identifiés 20 acides aminés dits protéinogéniques car incorporés (un codon = un acide aminé) dans la séquence d'une protéine lors de la

traduction de l'ARN messager par les ribosomes, **Figure 2**. À ces 20 AA il convient d'ajouter la Sélénocystéine et la Pyrrolysine (codées indirectement *via* des codons stop).

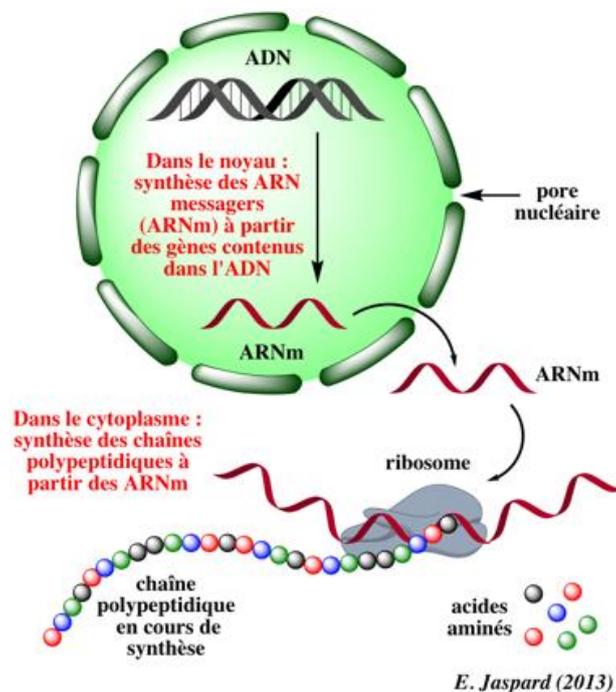


Figure 2. Schéma des étapes de la synthèse des protéines chez les Eucaryotes[1].

Il existe deux codes de nomenclature des acides aminés : un code à trois lettres utilisé pour écrire la formule d'un peptide et un code à une lettre quand cette séquence devient plus longue (protéine). Huit des vingt aminés présents chez les mammifères sont dits essentiels, car l'organisme ne sait pas les fabriquer, ils doivent être apportés par l'alimentation, ce sont : l'Histidine (His, H), l'Isoleucine (Ile, I), la Leucine (Leu, L), la Lysine (Lys, K), la Méthionine (Met, M), la Phénylalanine (Phe, F), la Thréonine (Thr, T), le Tryptophane (Trp, W) et la Valine (Val, V). Les autres acides aminés, dits non essentiels, peuvent être apportés par l'alimentation et/ou fabriqués par l'organisme par une suite de réactions enzymatiques, ce sont : l'Arginine (Arg, R), l'Alanine (Ala, A), l'Asparagine (Asn, N), l'acide Aspartique (Asp, D), l'acide Glutamique (Glu, E), la Cystéine (Cys, C), la Glutamine (Gln, Q), la Glycine (Gly, G), la Proline (Pro, P), la Sérine (Ser, S) et la Tyrosine (Tyr, Y).

Tous ces acides aminés sont de configuration L (dans la nomenclature/projection de Fischer. Ils sont tous (S), sauf la Cystéine et la Sélénocystéine qui sont (R) et la Glycine qui est achirale, **Figure 3**. L'Isoleucine et la Thréonine possèdent toutes les deux un deuxième carbone

asymétrique sur leur carbone C β (ou C2) de leur chaîne latérale, ces acides aminés naturels sont respectivement (2*S*, 3*S*) et (2*S*, 3*R*).

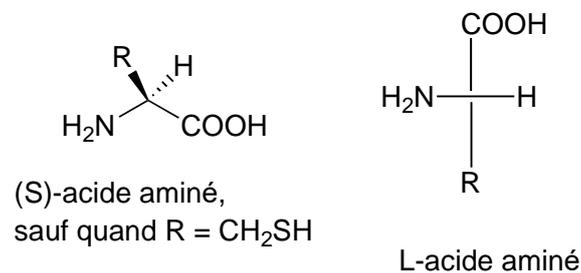


Figure 3. Structure et configuration des acides aminés protéinogéniques.

b) Biosynthèse de peptides et protéines

Chez l'Homme, les peptides biologiquement actifs, comme les protéines, sont biosynthétisés dans la cellule sous forme de précurseurs protéiques par une synthèse dite ribosomale (ADN \rightarrow ARNm \rightarrow protéines), **Figure 2**. Les peptides et protéines (au-delà de 50 acides aminés on parlera de protéines) sont obtenus par condensation d'acides aminés *via* une liaison amide, la biosynthèse se faisant du résidu *N*-terminal (*N*-ter, numéroté 1) au résidu *C*-terminal (*C*-ter), **Figure 4**.

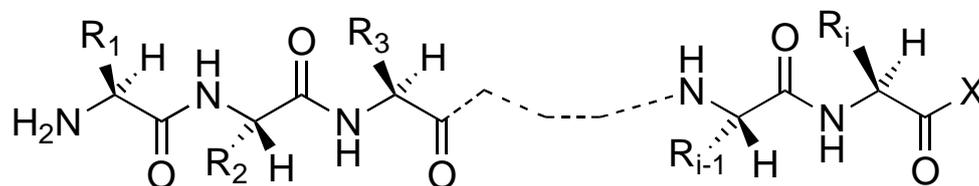


Figure 4. Chaîne polypeptidique.

Le précurseur protéique est ensuite coupé par des enzymes dans le cas des peptides. Le *C*-terminal du peptide final peut être sous forme acide carboxylique ($-\text{COO}^-$) ou, dans bon nombre d'hormones peptidiques, sous forme amide (CONH_2) alors que les protéines ont toujours leur *C*-terminal sous forme acide carboxylique. Les peptides sont alors soit libérés de la cellule pour agir sur d'autres cellules voisines, soit transportés, adsorbés en général sur des protéines dites de transport, vers leur site d'action. Les peptides jouent un rôle fondamental dans la signalisation et le métabolisme humain et, de fait, présentent un énorme potentiel thérapeutique, mais cet aspect ne sera pas développé dans ce manuscrit.

c) Structure des peptides et protéines

Les peptides et protéines adoptent, dans la très grande majorité des cas, des structures tridimensionnelles spécifiques (dites aussi structures tertiaires) qui sont définies par leurs séquences en acides aminés (structure primaire)[2], [3], à l'exception des protéines intrinsèquement désordonnées qui nécessitent un ou des partenaires en interaction pour adopter une structure tridimensionnelle définie[4]. Les structures tertiaires spécifiques, stabilisées par des liaisons de faible énergie (van der Waals, ionique, liaison hydrogène, dipôle, hydrophobe, π/π , π/cation), sont définies par l'agencement des acides aminés du peptide/de la protéine en éléments de structures secondaires, dont les plus courants sont les brins β , les hélices[5] et les coudes, **Figure 5**.

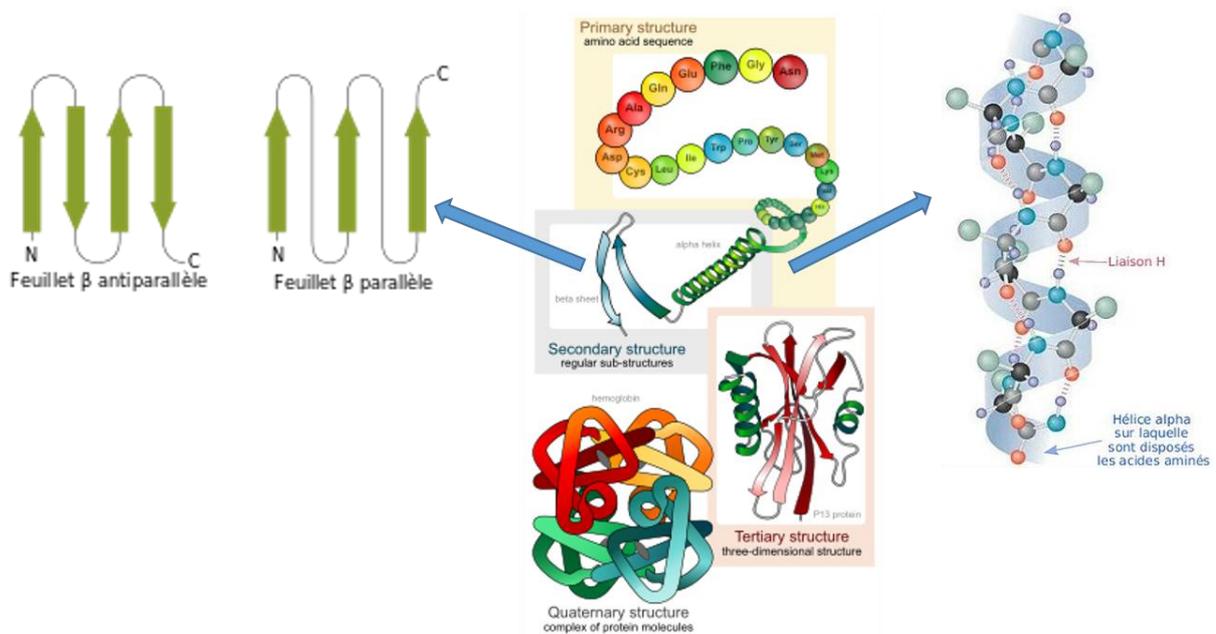


Figure 5. Eléments de structures secondaires identifiés dans les structures de peptides et protéines et structures tertiaires et quaternaires de protéines[6]–[8].

Dans une protéine, les régions structurées en brins β sont stabilisées par des liaisons hydrogènes entre les C=O et les NH du squelette peptidique de deux brins β distants. Ils forment alors un feuillet β . Il existe deux types d'agencement de brins β : antiparallèle et parallèle. Les brins β sont dits parallèles quand les numéros des acides aminés augmentent dans les deux brins. Les valeurs d'angles dièdres ϕ , angle autour de la liaison -NH-C $_{\alpha}$, et ψ , angle autour de la liaison C $_{\alpha}$ -C=O, sont alors respectivement de -119° et $+113^{\circ}$ en moyenne,

comme représentés dans un diagramme de Ramachandran[9], **Figure 6**. Les brins β sont dits antiparallèles quand les numéros des acides aminés des deux brins ont des sens de numérotations opposés. Les angles ϕ et ψ ont alors des valeurs proches de -139° , 135° .

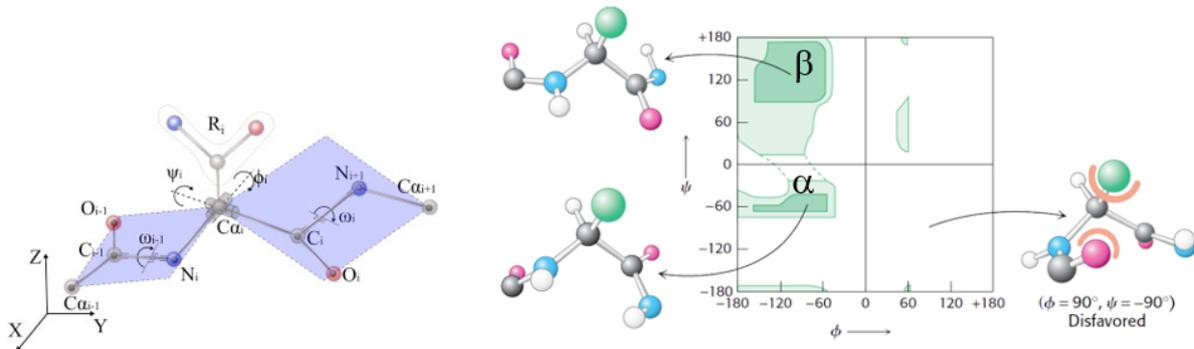


Figure 6. Diagramme de Ramachandran[9]–[11].

Il existe essentiellement deux types d'hélices, les hélices α (ou 3_6) et les hélices 3_{10} . Le pas de l'hélice α implique 3,6 acides aminés consécutifs par tour d'hélice avec une liaison H entre le C=O de l'acide aminé i et le NH de l'acide aminé $i+4$. Le pas de l'hélice 3_{10} implique 3 acides aminés consécutifs par tour d'hélice avec une liaison H entre le C=O de l'acide aminé i et le NH de l'acide aminé $i+3$, formant ainsi un pseudocycle à 10 atomes. Les valeurs d'angles dièdres ϕ et ψ sont autour de respectivement -57° , -47° pour l'hélice α et de -49° , -26° pour l'hélice 3_{10} , comme représentés dans un diagramme de Ramachandran, **Figure 6**.

Le troisième élément de structure secondaire correspond aux coudes, essentiellement β et plus rarement γ . Dans un coude β seuls quatre acides aminés contigus sont impliqués, cette structure est stabilisée par une liaison hydrogène entre le C=O de l'acide aminé i et le NH de l'acide aminé $i+3$. Dans le coude γ trois acides aminés contigus sont impliqués, cette structure est aussi stabilisée par une liaison hydrogène, entre le C=O de l'acide aminé i et le NH de l'acide aminé $i+2$. La définition du coude β a évolué en raison de la multitude de structures de protéines disponibles dans la PDB (Protein Data Bank) pour devenir : quatre acides aminés contigus (AA_i , AA_{i+1} , AA_{i+2} , AA_{i+3}) dont la distance entre le $C_{\alpha i}$ et le $C_{\alpha i+3}$ est inférieure à 0,7 nm, **Figure 7**, et avec des valeurs d'angles ϕ , ψ définies pour les acides aminés $i+1$ et $i+2$.

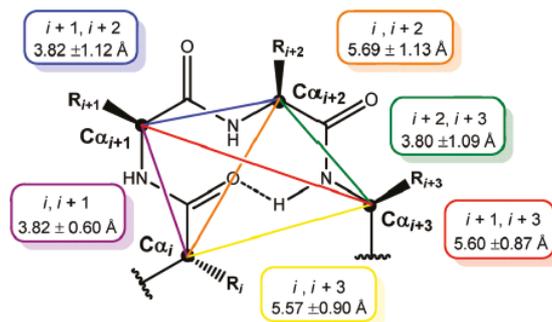


Figure 7. Moyenne des distances entre les différents C_{α} dans un coude β (moyenne à partir de l'étude de 10 245 coudes β dans la PDB)[12].

Suivant les valeurs d'angles ϕ , ψ adoptées, ces coudes β sont classés en différents types. Ces 30 dernières années on pouvait trouver des tableaux incluant jusqu'à 12 sous-types de coudes β , comme, par exemple, le **Tableau 1** extrait d'une revue récente où l'on peut découvrir une tentative de normalisation[13].

Coude β	Fréquence (%)
I	38,21
II	11,81
VIII	9,84
I'	4,10
II'	2,51
VI _b	0,88
VI _{a1}	0,73
VI _{a2}	0,20
IV ^{ori}	31,72

Tableau 1. Fréquence des coudes β classiques. Les types VI_{a1}, VI_{a2} et VI_b ont une Proline avec une liaison amide *cis* en position $i+2$. Le type IV noté IV^{ori} a été introduit pour le distinguer des autres[13].

Si la structure tertiaire d'une protéine correspond à l'agencement dans l'espace des structures secondaires que nous venons de décrire, la structure quaternaire correspond à l'assemblage entre elles de plusieurs sous-unités protéiques (présentant chacune une structure tertiaire), **Figure 5**. Dans tous les cas, ce sont des liaisons de faible énergie, comme listées ci-dessus, qui sont mises en jeu pour stabiliser ces édifices.

On peut maintenant analyser dans les détails les structures tridimensionnelles (3D) de nombreuses protéines cristallisées (PDB), malheureusement les descriptions de peptides

cristallisés sont très rares dans la littérature, à l'exception de peptides incorporant des acides aminés contraints synthétiques. Cet état de fait témoigne de la flexibilité des peptides et donc de l'équilibre de conformères présents en solution pour une séquence peptidique. Nous reviendrons sur ce point lors de l'étude des interactions mimes/homéoprotéines. Les quelques rares cas publiés montrent qui plus est des structures 3D différentes en fonction du solvant et des conditions de cristallisation[14]–[17].

d) Synthèse chimique des peptides

Les premiers balbutiements de synthèse peptidique datent du début du XX^{ème} siècle avec les travaux de T. Curtius, E. Fischer et E. Fourneau (1901). Les premières synthèses totales en 1954 de peptides à activité biologique ont été celles de deux hormones nonapeptidiques cycliques isolées de l'hypophyse : l'ocytocine et son homologue la vasopressine, deux peptides amidés en C-terminal (Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly-CONH₂ et Cys-Tyr-Phe-Gln-Asn-Cys-Pro-Arg-Gly-CONH₂)[18], [19]. La synthèse d'un autre « peptide/protéine » au potentiel thérapeutique considérable, l'insuline, 51 acides aminés, isolée dans les années 1920 par F. Banting, C. Best, J. Collip et J. Macleod, est à mettre au crédit de plusieurs équipes chinoises de Shanghai et Pékin. Elle a été décrite dans une publication de 1966 dans Scientia Sinica[20]. Cependant, il ne s'agit que d'un exploit historique (voire politique) avec 200 étapes et un rendement TRES faible.

Le principe de la synthèse chimique d'un peptide peut être résumé en une stratégie itérative de « protection-activation-couplage-déprotection ». Depuis les années cinquante, l'évolution des jeux de groupes protecteurs et la conception de nouveaux agents d'activation ont permis de très grandes avancées, mais sans nul doute l'étape décisive dans le domaine la synthèse chimique d'un peptide fut celle de l'introduction de la chimie sur support solide proposée par Bruce Merrifield au milieu des années soixante. Ces stratégies de synthèse seront décrites dans le détail dans le chapitre II-B « Principe de la synthèse peptidique sur support solide ».

En guise de conclusion de cette partie introductive sur les acides aminés et peptides, nous dirons que de nos jours l'exploit de la synthèse d'un nonapeptide qui a conduit au prix Nobel de V. du Vignaud au milieu du XX^{ème} siècle apparaît dérisoire puisque cinquante ans plus tard des peptides (protéines) de plus de 100 acides aminés ont été obtenus soit par une

synthèse chimique dite linéaire soit par ligation chimique de blocs judicieusement choisis. Mais dès maintenant nous précisons que certains peptides demeurent retors, pour des raisons différentes. Des mises au point spécifiques souvent longues voire laborieuses sont nécessaires. L'insuline thérapeutique est produite depuis la fin des années 1970 par génie génétique et aucune synthèse chimique n'est à ce jour compétitive, même si de nombreuses publications sur le sujet sont apparues[21].

e) Résumé de manière schématique de quelques faits que nous estimons majeurs dans l'histoire de la chimie des peptides/protéines.

1901 : première description et synthèse d'une liaison peptidique dans le dipeptide glycyl-glycine par Emil Fischer, considéré comme le fondateur de la synthèse peptidique[22].

1932 : Max Bergmann et Leonidas Zervas proposent le groupe carbobenzoxy (Cbz, Z) comme le premier groupe de protection d'amine en synthèse organique appliqué à la synthèse peptidique[23].

1954 : Synthèse d'une hormone polypeptidique, l'ocytocine, par Vincent du Vigneaud[24]. Vincent du Vigneaud reçoit le prix Nobel de chimie en 1955.

1963-1967 : Mise au point de stratégies de synthèse peptidique en phase solide (SPPS), par Bruce Merrifield[25]–[28]. Bruce Merrifield reçoit le prix Nobel de chimie en 1984[29].

1994 : Développement de la ligation chimiosélective, en particulier la ligation chimique native (NCL) par Stephen Kent et collaborateurs[30], [31].

1998 : Développement de stratégies semi-synthétiques, qui ont repoussé la limite de la taille des protéines susceptibles d'être synthétisées à quelques centaines d'acides aminés[32].

2000 – 2002 : Réactions bioorthogonales de modification de peptides et protéines, incluant des exemples de réactions de Staudinger ou de cycloaddition[33]–[35]. Barry Sharpless reçoit le prix Nobel de chimie en 2001[36].

2. Les glucides et les glycosaminoglycanes

a) Structure et nomenclature des monosaccharides

Remarque liminaire : Dans cette famille de molécules du Vivant il existe déjà plusieurs noms pour nommer l'unité de base. On anticipe que les descriptions de l'unité de base et de son (ses) agencement(s) en polysaccharides seront plus compliquées que pour les acides aminés et peptides ! Nous tenterons de faire succinct et clair.

Les oses ou monosaccharides, autrefois également appelés carbohydrates ou hydrates de carbone, ont pour formule générale $C_x(H_2O)_n$, n étant un nombre entier de 3 à 9. Ils possèdent une fonction aldéhyde (aldose) ou cétone (cétose). Ces sucres de base sont ensuite modifiés par oxydation (fonction acide), échange O->N pour donner des sucres aminés, ou désoxygénés pour donner des positions deoxy, conduisant à une grande quantité de motifs possibles (en particulier chez les bactéries). Chez les mammifères neuf monosaccharides sont majoritairement présents : le D-Glucose (Glc), la N-Acétyl-D-glucosamine (GlcNAc), le D-Galactose (Gal), la N-Acétyl-D-galactosamine (GalNAc), le D-Mannose (Man), le D-Xylose (Xyl), l'acide D-Glucuronique (GlcA), le L-Fucose (Fuc) et l'acide N-acétylneuraminique (NeuAc), représentés **Figure 8**.

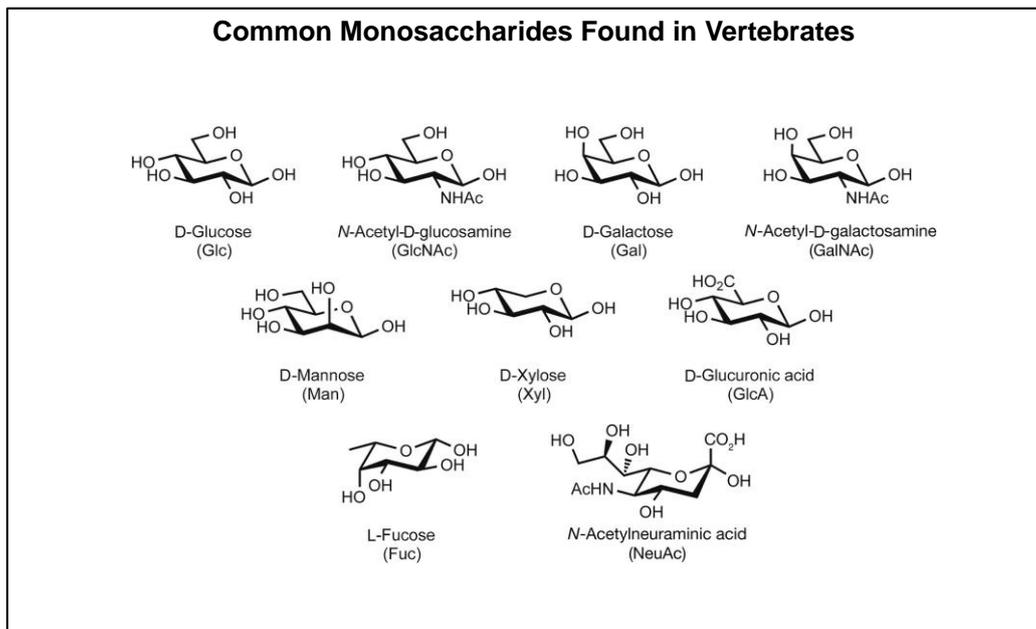


Figure 8. Monosaccharides présents chez les mammifères[37].

Un code couleur pour « représenter/écrire » ces oses a été proposé plus tard pour simplifier l'écriture de sucres complexes, **Figure 9**.

Symbolic Representations of Common Monosaccharides and Linkages	
● Galactose (Gal)	★ Xylose (Xyl)
■ N-Acetylgalactosamine (GalNAc)	◆ N-Acetylneuraminic acid (Neu5Ac)
▽ Galactosamine (GalN)	◆ N-Glycolylneuraminic acid (Neu5Gc)
● Glucose (Glc)	◆ 2-Keto-3-deoxynononic acid (Kdn)
■ N-Acetylglucosamine (GlcNAc)	▲ Fucose (Fuc)
▽ Glucosamine (GlcN)	◆ Glucuronic acid (GlcA)
● Mannose (Man)	◆ Iduronic acid (IdoA)
■ N-Acetylmannosamine (ManNAc)	◆ Galacturonic acid (GalA)
▽ Mannosamine (ManN)	◆ Mannuronic acid (ManA)
	?
	?

Figure 9. Code couleur pour représenter les monosaccharides[38].

On doit à Emil Fischer à la fin du XIX^{ème} siècle un travail remarquable qui a conduit à une première normalisation de la nomenclature des monosaccharides et de leurs stéréochimies, ces molécules possédant de nombreux centres stéréogènes. Donnons quelques clefs d'interprétation de ce travail avec comme exemple le D-Glucose (**Figure 10**). Le D-Glucose est l'isomère naturel du glucose, c'est un hexose, C_n avec n = 6, possédant une fonction aldéhyde, donc un aldose. En représentation de Fischer dans la forme D, le groupe alcool (-OH) du D-Glucose porté par le carbone n-1 (ici C5) est à droite. Le D-Glucose en solution s'équilibre entre deux formes cycliques représentées, dans la **Figure 10**, soit en représentation de Haworth soit en représentation en perspective. La fonction alcool en C5 réagit avec la fonction aldéhyde pour former un dérivé du pyrane (cycle à 5 carbones et 1 oxygène) dit glucopyranose. L'attaque pouvant se produire sur l'une ou l'autre des faces de la double liaison aldéhydrique, on trouvera en solution deux espèces qui ne diffèrent que par leur configuration en position anomérique (C1). L'anomère est dit « α » (on parlera alors de α-D-Glucose) si la fonction hydroxyle en C1 se trouve au-dessous du plan (représentation de Haworth) ou du plan moyen (représentation en perspective), ou « β » si elle se trouve au-dessus de ce plan.

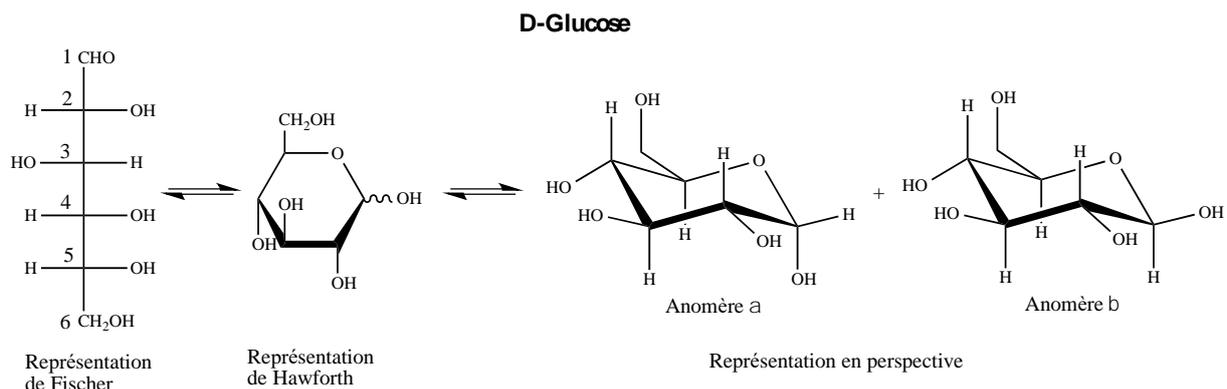


Figure 10. Les différentes représentations d'un monosaccharide, le D-Glucose, un hexose (6 carbones et une fonction aldéhyde). Représentation de Fischer[39], représentation de Haworth[40] et représentation en perspective avec les anomères α et β.

Les glucopyranoses peuvent exister sous plusieurs formes même s'ils adoptent des conformations privilégiées, **Figure 11**.

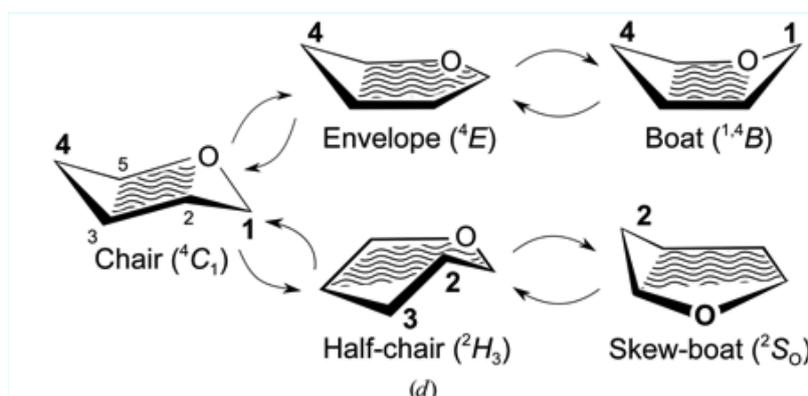


Figure 11. Interconversions conformationnelles d'un pyranose. Nomenclature IUPAC : les conformations sont identifiées par une lettre majuscule, chaise (C), enveloppe (E), bateau (B), chaise croisée (S), demi-chaise (H) et torsion (T), avec les atomes du côté supérieur ou inférieur du plan moyen respectivement indiqué en exposant et en indice[41].

La forme chaise ⁴C₁ (plus souvent appelée C1) est la conformation la plus stable du D-glucopyranose et la forme chaise ¹C₄ celle du L-glucopyranose, **Figure 12**. Les structures ⁴C₁ et ¹C₄ correspondent aux nomenclatures proposées par R.E. Reeves : l'*exposant* devant le C identifie le carbone situé au-dessus du plan moyen du cycle, et l'*indice* qui suit le C identifie le carbone situé au-dessous[42].

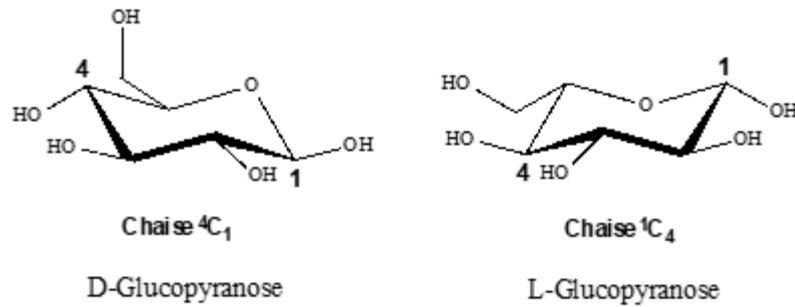


Figure 12. Nomenclatures pour les conformations privilégiées du D-et L-glucopyranose.

b) Polysaccharides et glycosaminoglycanes

Les glucides peuvent exister chez l'Homme sous forme de sucres simples (oses ou monosaccharides), importants entre autres comme source d'énergie pour l'organisme, et de polymères d'oses (osides), appelés oligosaccharides (conventionnellement pour un nombre de résidus inférieur à 11), polysaccharides ou glycanes, connectés entre eux par des liaisons glycosidiques pouvant être α ou β . Le terme glycane peut également être utilisé pour décrire la portion glucidique d'un glycoconjugué tel qu'une glycoprotéine (association covalente d'un glycane et d'une protéine) ou un glycolipide (association covalente d'un glycane et d'un lipide).

La nomenclature des polysaccharides est restée complexe, même si à plusieurs reprises des tentatives de simplification ont été avancées. R. W. Jeanloz, un des papes de la chimie des sucres et polysaccharides, a proposé dans les années 1950-1960 un début de normalisation. Pour le citer : « *The large number of recent publications dealing with amino sugar containing polysaccharides, free or linked to proteins, have employed a plethora of descriptive terms, such as, "mucopolysaccharides, mucopolyuronides, mucoproteins, mucoids, mucin, glycoproteins, glycopolypeptides, polysaccharides of connective tissue, polysaccharides of blood, etc."* They have brought to light the unfortunate state of confusion prevailing in the classification and nomenclature of these substances »[43]. C'est ainsi que, depuis cette mise au point, les mucopolysaccharides sont nommés glycosaminoglycanes (GAGs), même si les pathologies associées à des problèmes de métabolisme dans cette famille de molécules continuent d'être classées à ce jour sous le terme générique mucosaccharidoses, maladies provoquées par un déficit d'enzymes impliquées dans les étapes d'hydrolyse des glycosaminoglycanes !

Les GAGs sont des polysaccharides linéaires, en général polysulfatés et donc chargés négativement, qui ont des poids moléculaires variable de 1 à 8000 kDa. Chez l'Homme, les GAGs sont essentiellement localisés à la surface de la membrane cellulaire (glycocalyx) et dans la matrice extracellulaire (ECM) des tissus.

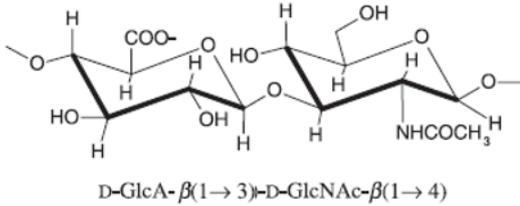
Il existe deux grandes classes de GAGs : les GAGs sulfatés ($-O-SO_3^-$) et les GAGs non sulfatés. Les GAGs sulfatés regroupent : i) les chondroïtines sulfates (CS), ii) les dermatanes sulfates (DS), iii) les kératanes sulfates (KS), iv) l'héparine (Hp) et le sulfate d'héparane (HS). L'acide hyaluronique (HA) est quant à lui un GAG non sulfaté.

Les chaînes de glycosaminoglycanes correspondent à des unités répétitives de disaccharides composées d'un acide uronique (acide D-glucuronique ou acide L-iduronique) ou d'un galactose et d'une hexosamine (D-galactosamine ou D-glucosamine). Les GAGs diffèrent donc selon le type d'unité hexosamine, hexose et acide hexuronique qu'ils contiennent, mais également par la configuration (α ou β) de la liaison glycosidique entre ces unités, **Tableau 2**[44].

Glycosaminoglycanes

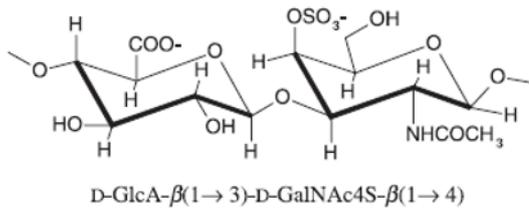
Caractéristiques

Acide hyaluronique



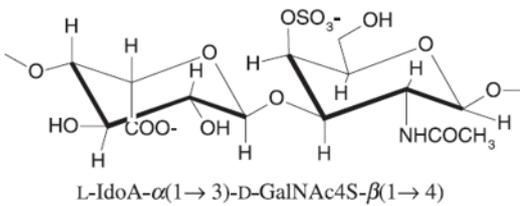
Non sulfaté - Poids moléculaire 4–8000 kDa
Attache non covalente aux protéines ECM
Existe dans les bactéries
Trouvé dans le liquide synovial, humeur vitreuse, ECM du tissu conjonctif lâche
Excellents lubrificateurs et amortisseurs

Chondroïtine sulfate (CS)



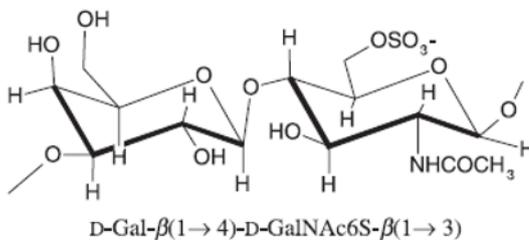
Poids moléculaire 5–50 kDa
GAG le plus abondant dans le corps
Trouvé dans cartilage, tendon, ligament, aorte
Se lie aux protéines (comme le collagène) pour former des agrégats de protéoglycane

Dermatane sulfate (DS)



Poids moléculaire 15–40 kDa
Trouvé dans la peau, les vaisseaux sanguins, les valves cardiaques

Keratan sulfates I and II (KS)



Poids moléculaire 4–19 kDa
Le plus hétérogène des GAGs
KS I se trouve dans la cornée
KS II se trouve dans le cartilage agrégé avec CS

Tableau 2. Unités disaccharidiques présentes dans divers glycosaminoglycanes[44].

Au sein d'une même classe de GAG il existe aussi une diversité, que nous allons détailler pour les chondroïtines sulfates (CS), puisque nous les étudierons plus particulièrement dans les chapitres suivants. Les chondroïtines sulfates ont été isolées à partir de cartilage par Fischer et Boedecker en 1861[45]. Ce n'est que 50 ans plus tard que Levene and Forge détermineront la structure de base des deux saccharides présents dans les CS (**Tableau 2**) : l'acide D-glucuronique (GlcA) et la N-acétylgalactosamine (GalNAc)[46]. Mais la sulfatation de ce disaccharide de base, l'une des principales modifications observées sur les

CS, va créer une diversité qui conduira à des sous-classes de CS, en fonction de la position de cette sulfatation : sur le carbone C4 du GalNAc pour CS-A, sur le C6 du GalNAc pour CS-C, sur le C2 du GlcA et le C6 du GalNAc pour CS-D et sur le C4 et le C6 du GalNAc pour CS-E, **Figure 13**[47].

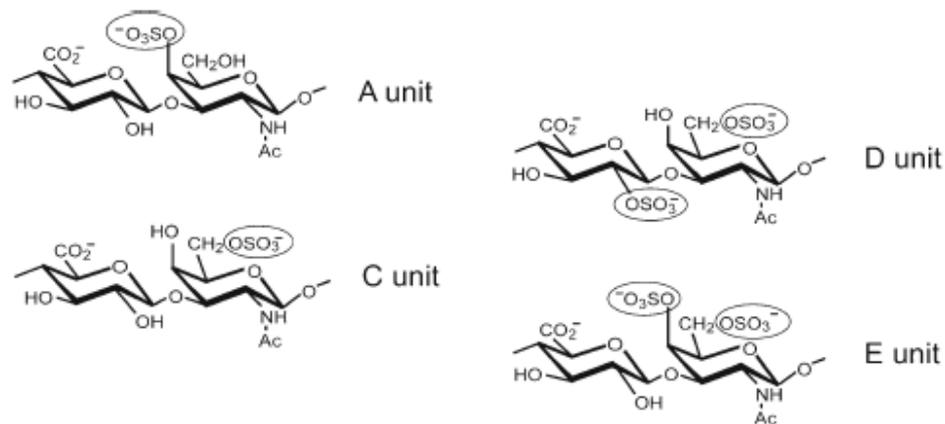


Figure 13. Structure des disaccharides des différents chondroïtines sulfate[47].

La biosynthèse des GAGs n'est pas un phénomène codé comme la biosynthèse des peptides et protéines. Chez l'Homme les CS sont produits par l'action combinée (concertée) de 22 enzymes (glycosyltransférases, épimérasés et sulfotransférases). Qui plus est, il a été montré que la composition en CS varie d'une espèce à l'autre, qu'elle est dépendante du type cellulaire et qu'elle varie aussi durant le développement. Dans le système nerveux central, par exemple, la composition en CS dans la matrice extracellulaire change : la CS-C est la plus abondante durant l'embryogenèse, alors que la CS-A est la plus abondante chez l'adulte. Cette variabilité de degré de sulfatation des CS, une réaction enzymatique réalisée dans l'appareil de Golgi, conduit à des interactions spécifiques avec pléthore de protéines, interactions difficiles à prédire. Ces modifications participent au mystère des GAGs, à leurs rôles biologiques et à ce que certains appellent le « code sucre »[48], [49].

La glycosylation de protéines (*O*-glycosylation et/ou *N*-glycosylation), par des glycanes sur des positions spécifiques (respectivement Ser/Thr et Asn) est une modification post-traductionnelle des protéines très fréquemment observée, on parle alors de glycoprotéine. Dans la majorité des cas, ces glycanes sont des oligosaccharides non sulfatés mais lorsqu'il s'agit de GAGs sulfatés (polysaccharides linéaires pouvant atteindre une centaine de résidus),

la glycoprotéine est alors appelée protéoglycane (PG). Les protéoglycane comprennent un ou plusieurs GAGs, plus ou moins sulfatés, composés de 25 à 100 unités disaccharidiques. La modification post-traductionnelle de protéines par les GAGs correspond généralement à une *O*-glycosylation de chaîne latérale de Sérine(s), dans certaines séquences Gly-Ser-Gly, *via* un lien impliquant un tétrasaccharide Xyl-Gal-Gal-GlcA, **Figure 14**.

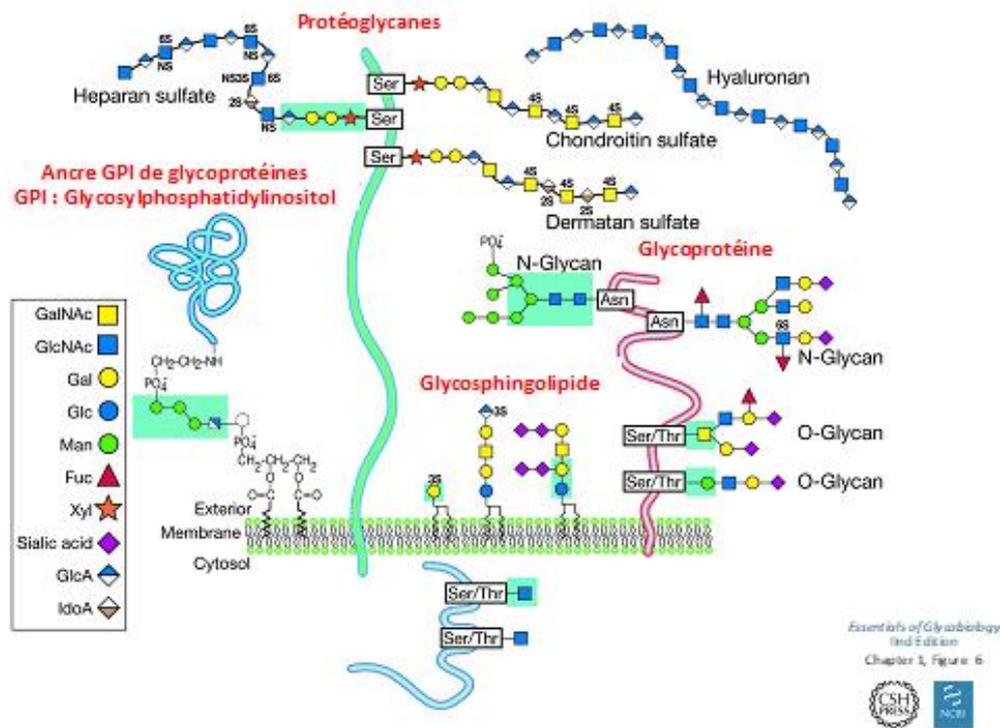


Figure 14. Glycanes, protéoglycane et glycoprotéines chez les mammifères[50].

Les glucides en général et les GAGs en particulier présentent donc une diversité exceptionnelle, inégalée dans les autres familles de biomolécules du Vivant, qui rend délicats leur isolement, caractérisation, synthèse chimique et utilisation.

c) Structure des osides et des GAGs

Les structures primaire, secondaire et tertiaire des osides (oligo et polysaccharides) correspondent aux différents niveaux de structuration, comme pour les protéines mais avec leurs spécificités. La structure I^{aire} correspond à l'enchaînement des résidus monosaccharidiques dans la chaîne, la structure II^{aire} définit l'organisation de ces structures I^{aire} et finalement la structure III^{aire} correspond à l'agencement des structures II^{aire} dans l'espace. Ces structures II^{aire} et III^{aire} sont stabilisées par des liaisons de faible énergie : liaisons

hydrogène, les interactions dipolaires et ioniques et la solvation de ces molécules polyhydroxylées[51].

Structure des osides.

Les osides, liés par une liaison *O*-glycosidique, sont des molécules souples et mobiles. Leur forme est décrite par les valeurs des angles dièdres qui correspondent aux rotations autour de la liaison glycosidique, les angles ϕ et ψ , par analogie avec les peptides et protéines. Toutes les valeurs (ϕ , ψ) ne sont pas explorées en raison de gêne stérique. La liaison glycosidique impliquant un alcool primaire en C6 présente un autre degré de liberté, l'angle dièdre ω , qui définit la libre rotation autour de la liaison O5-C5-C6-O6, **Figures 15 à 17**.

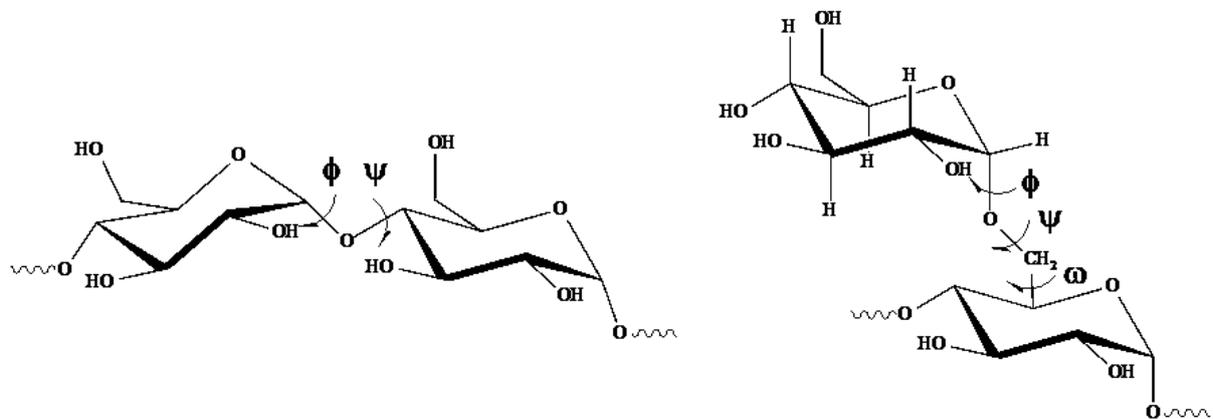
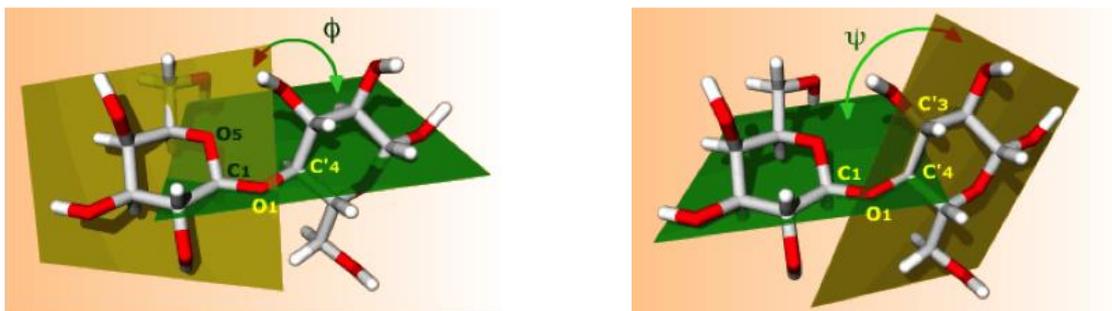


Figure 15. Angles de torsion de la liaison glycosidique (ϕ , ψ et ω).

Sur la **Figure 16** sont représentés respectivement les angles ϕ et ψ du D-Galactose et du D-Glucose dans une molécule de Lactose.



Exemple du D-Lactose

L'angle ϕ mesure la rotation du résidu de D-galactose. Il correspond à l'intersection du plan vert passant par O5, C1 et O1, et de l'autre plan brun contenant C1, O1 et C'4. L'angle ϕ mesure la rotation des deux plans selon l'axe C1-O1.

L'angle ψ mesure la rotation du résidu de D-glucose. Il correspond à l'intersection du plan vert passant par C1, O1 et C'4, et de l'autre plan brun contenant O1, C'4 et C'3. L'angle ψ mesure la rotation des deux plans selon l'axe O1-C'4.

Figure 16. Angles ϕ et ψ du D-Galactose et du D-Glucose dans une molécule de Lactose[52].

Les trois conformations les plus courantes autour de la liaison C5-C6 sont dites gg (gauche-gauche), gt (gauche-trans) et tg (trans-gauche), comme représentés **Figure 17**. Le groupe CH₂ est prochiral, les deux hydrogènes sont notés pro-R et pro-S. Les conformations les plus stables sont gt et gg.

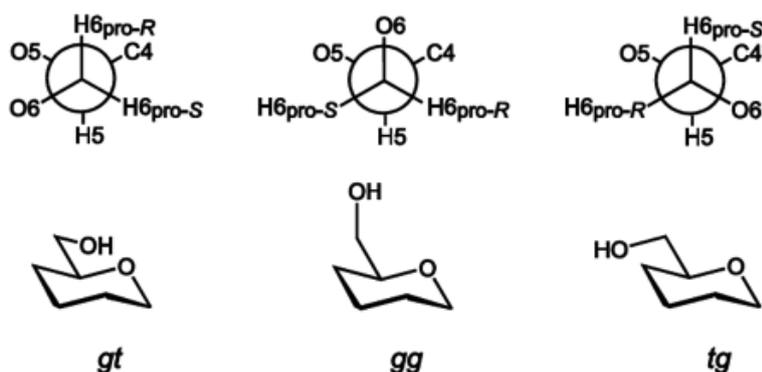


Figure 17. Les trois rotamères décalés décrivant l'angle de torsion ω [53].

Les structures III^{aire} des polysaccharides varient en fonction de leur composition en monosaccharides, de leur poids moléculaire, des configurations anomériques, de la position de liaison (1-4 ou 1-6 par exemple) et aussi de leur densité de charge (tout particulièrement en sulfates). Cette variabilité peut apparaître vertigineuse quand il s'agit d'envisager une ou

des structure(s) III^{aire} privilégiée(s). Néanmoins, il a été observé qu'une chaîne polysaccharidique peut adopter des structures III^{aire} hélicoïdales et des structures en forme de rubans, même si ces hélices, avec des pas allant de 7 et 47 Å, sont beaucoup plus diverses que celles identifiées dans les structures 3D de protéines.

La symétrie des hélices est décrite avec deux codes de nomenclature : i) U_v avec U correspondant au nombre d'unités répétées dans v tour d'hélice ou ii) (n,h) n étant le nombre d'unités répétées par tour d'hélice et h correspondant au pas de l'hélice. Dans la nomenclature (n,h) , pour indiquer la chiralité de l'hélice, on rajoute un signe négatif devant h si l'hélice est gauche[54]. Les structures en ruban correspondent à des structures relativement rigides, aplaties, avec un caractère amphiphile.

En solution, il a été observé que ces structures hélicoïdales ont une forte propension à former des structures hiérarchiques supérieures, avec des enroulements d'hélices qui peuvent ainsi conduire à la formation de gel. Sur la **Figure 18** sont représentées des hélices simple, double et triple adoptées par l'amylose, polymère de glucose lié par liaison α et du curdlane, polysaccharide formé de 400 à 500 résidus de D-glucose liés par des liaisons β [55]. La cavité centrale de ces hélices est hydrophobe, en raison de la présence des liaisons axiales C-H du sucre, alors que l'extérieur avec les groupes hydroxyles du sucre est très hydrophile.

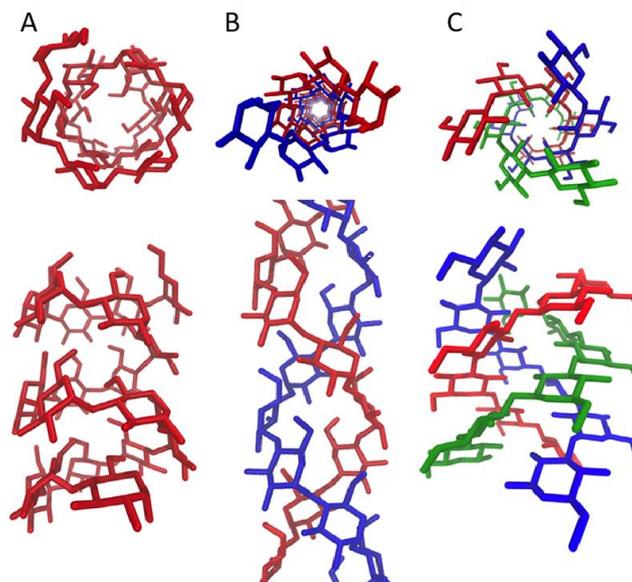


Figure 18. Vues de dessus et de côté d'hélices simple, double et triple de polysaccharides. A, hélice gauche 6_1 de V-amylose ; B, double hélice gauche 6_1 de A-amylose ; C, triple hélice droite 6_1 de curdlane[55].

Structure des glycosaminoglycanes.

La littérature relatant les structures d'héparine, d'héparane sulfate et d'analogues synthétiques de compositions bien définies est riche de données, ces molécules et analogues ayant un rôle incontournable dans la coagulation, **Figure 19**[56], mais pour les autres GAGs il y a moins de données.

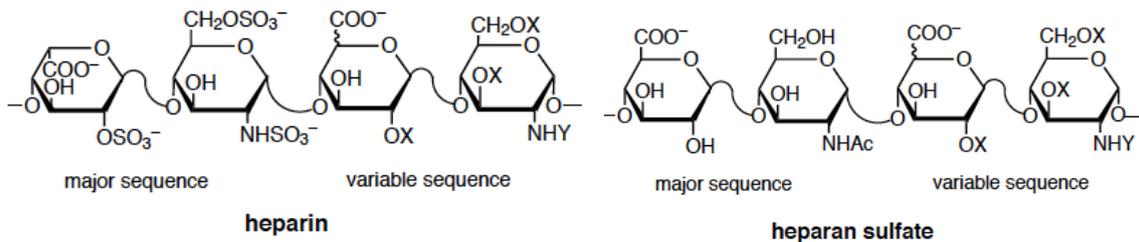


Figure 19. Disaccharides, unités de répétition, majeurs et mineurs dans la séquence de l'héparine et des héparane sulfates, avec X = H ou SO₃⁻, Y = Ac, SO₃⁻ ou H[56].

D'une façon générale, les GAGs conduisent à des structures hélicoïdales avec des motifs répétés de disaccharides de taille de 8 à 10 Å qui vont conduire à la présentation des groupements sulfates le long et autour de l'axe hélicoïdal, **Figures 20 et 21**.

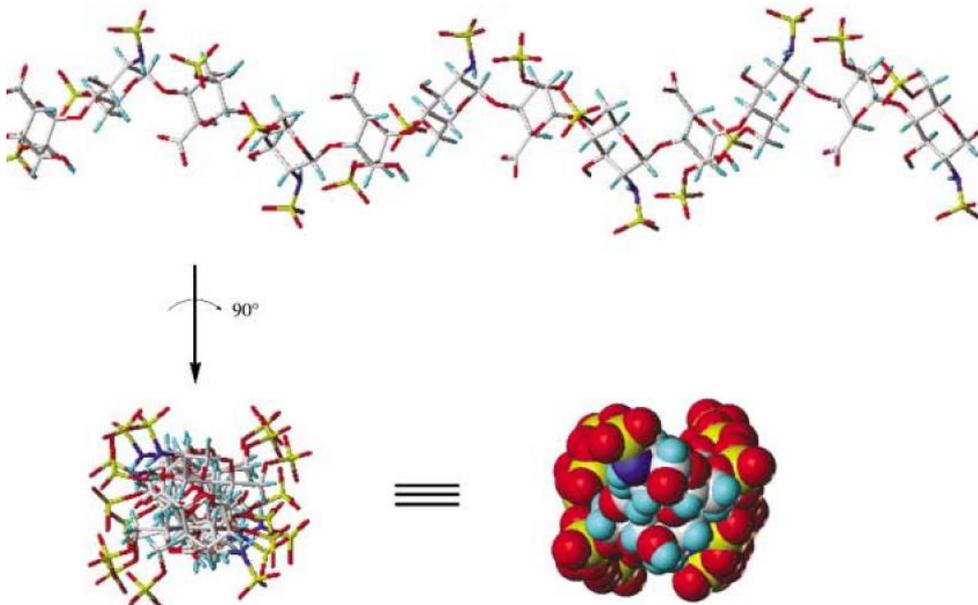


Figure 20. Structure hélicoïdale d'une séquence d'héparine dodécasaccharide. Les atomes de soufre sont indiqués en jaune, d'oxygène en rouge, d'azote en bleu et d'hydrogène en cyan[56].

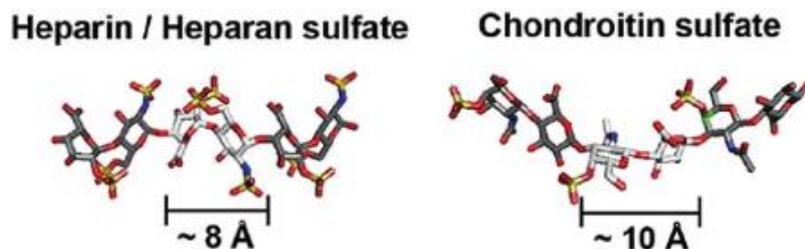


Figure 21. Architecture des deux types de GAGs naturels[57].

d) Synthèse chimique des oligosaccharides et des GAGs

Remarque liminaire : Je pense maîtriser les concepts de synthèse peptidique et leur mise en application. Au fil des années, j'ai synthétisé pour commencer des petites protéines, puis des peptides bien plus courts mais très hydrophobes, puis des peptides hydrophiles, voire trop hydrophiles ! Même si les synthèses de ces molécules du Vivant et de leurs analogues peuvent être réalisées en phase solide de manière automatique, certains peptides ont pu dévoiler « sur le tas » différents problèmes de synthèse ou de purification auxquels j'ai su faire face. Par contre, la chimie des sucres est bien au-delà de ma formation de chimiste/biochimiste et de mes activités de laboratoire. C'est un art en soi que je ne maîtrise absolument pas. Mon directeur de thèse est considéré comme un expert dans le domaine et mes camarades de paillasse m'ont fait part de leurs souffrances, et leur joie aussi, devant la synthèse d'oligosaccharides même limités en unités saccharidiques. C'est avec ignorance et donc ce sera avec modestie que je vais tenter de parler de la chimie des sucres et je me focaliserai sur la synthèse de chondroïtines sulfates qui sont au cœur de ce travail.

Pour atteindre la complexité des structures oligosaccharidiques, et, encore au-dessus, celle des glycosaminoglycanes, il a fallu développer au préalable certaines stratégies de base de chimie en solution. Cette chimie est une chimie, entre autres, de groupes protecteurs permanents vs. temporaires, de réactivité de sucres donneurs vs. sucres accepteurs. Les **Figures 22 à 24** résument les grandes questions et stratégies de synthèse chimique des réactions de glycosylation. La réaction de glycosylation met en jeu un sucre donneur, qui possède un bon groupe partant, et un sucre accepteur qui lui possède une fonction nucléophile, un alcool, Cf. **Figure 22** et pour un exemple avec la réaction de Koenigs et Knorr, **Figure 23**.

Glycosylation

Principe général

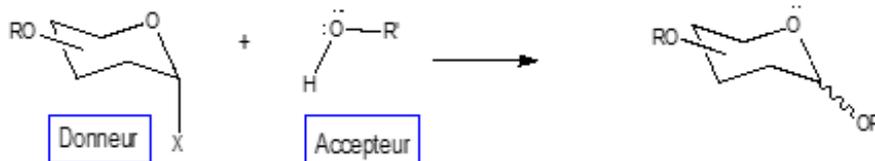


Figure 22. La réaction de glycosylation met en jeu un sucre donneur et un sucre accepteur.

Réaction de Koenigs-Knorr (un exemple de stéréosélectivité)

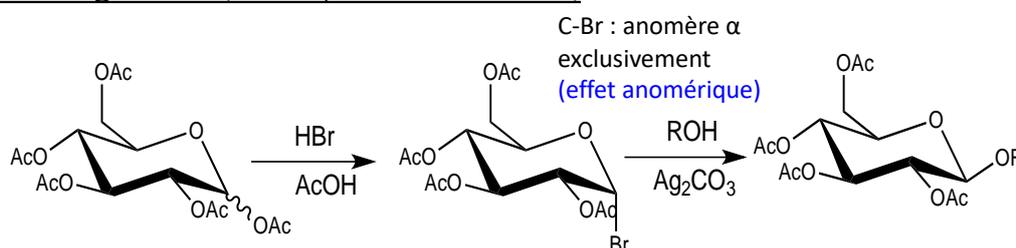
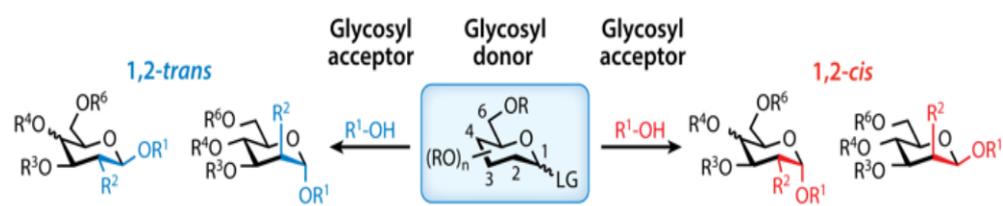


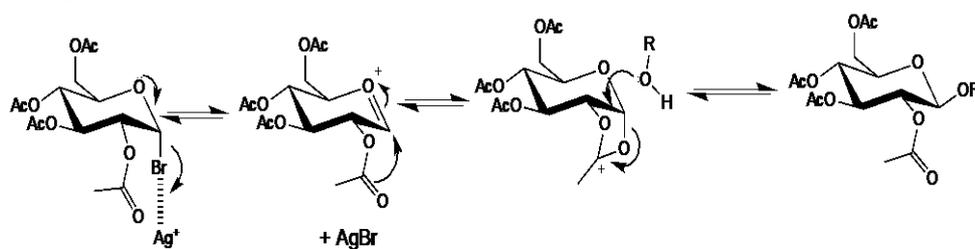
Figure 23. Réaction de glycosylation : principe général de la réaction de Koenigs et Knorr.

La stéréosélectivité de l'attaque entre les deux sucres donneurs et accepteur dépend fortement du solvant, de la température, du groupe partant et aussi de la présence ou non d'un groupe participant en C2, comme indiqué avec deux exemples sur la **Figure 24** : a) un groupe O-Acétyle en C2 et la réaction de Koenigs-Knorr (Br^- comme groupe partant et la présence de sel d'argent et b) un auxiliaire chiral en C2 qui possède un éthyle substitué par une fonction nucléophile, ici le (1S)-phényl-2-(phénylsulfanyl)éthyle, qui générera la formation d'un sulfonium cyclique. La forme la plus stable laisse la position axiale libre pour l'attaque de l'accepteur de glycosyle.

Stéréosélectivité



a) 1,2-trans Glycosylation



b) 1,2-cis Glycosylation

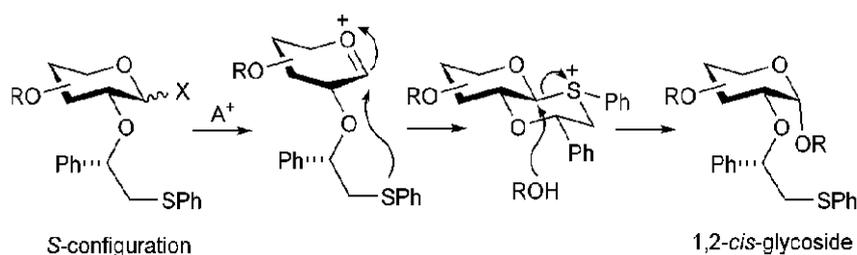


Figure 24. Stéréosélectivité de la réaction de glycosylation.

Depuis les travaux pionniers de l'équipe de P. H. Seeberger au début de ce siècle[58], la synthèse sur support solide et automatique peut être utilisée pour des oligosaccharides. La question centrale de cette démarche demeure la disponibilité des blocs protégés de départ, loin d'être tous commerciaux comme c'est le cas pour la majorité des dérivés d'acides aminés. Très souvent cette approche est limitée aux laboratoires spécialistes de ce domaine de chimie qui constituent eux-mêmes leurs « stocks » de monosaccharides protégés[59]–[61]. Dans la revue de 2015, il est écrit qu'avec les blocs représentés **Figure 25** environ 85 % des glycanes pourraient être préparés[62].

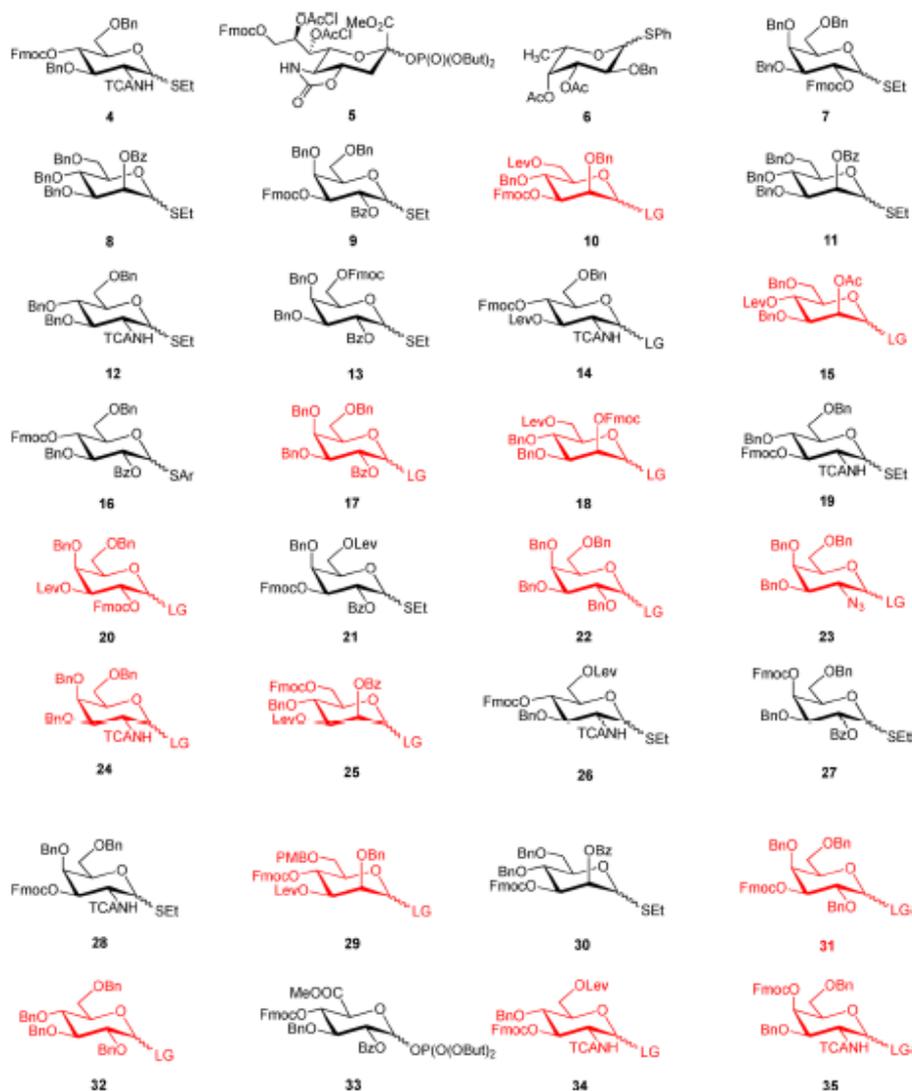


Figure 25. « Building blocks » pour l'assemblage de glycanes de mammifères. Les blocs en noir ont déjà été synthétisés et utilisés, tandis que la synthèse des blocs en rouge est toujours à réaliser[62].

De même, des stratégies de synthèse enzymatique ont été proposées pour la préparation d'oligosaccharides et de GAGs, sachant que le profil de sulfatation des GAGs ajoute un degré supplémentaire de difficultés à leur obtention sous forme pure/homogène[63], [64].

Au-delà de ces principes de base, j'ai lu, pour me « forger une idée », des articles et/ou revues récents que j'ai listés ci-dessous, par ordre chronologique.

Mende, M. et al. *Chemical Synthesis of Glycosaminoglycans*. *Chem. Rev.* 116, 8193–8255 (2016)[63]. Cette revue résume les progrès des synthèses chimiques et enzymatiques de GAGs de 2000 à 2015.

Kulkarni, S. S. et al. *“One-Pot” Protection, Glycosylation, and Protection–Glycosylation Strategies of Carbohydrates*. *Chem. Rev.* 118, 8025–8104 (2018)[60]. Cette revue résume les méthodologies de protection et de glycosylation en « un pot » pour préparer des glycanes complexes ainsi que l'automatisation de ces synthèses.

Panza, M., Pistorio, S. G., Stine, K. J. & Demchenko, A. V. *Automated Chemical Oligosaccharide Synthesis : Novel Approach to Traditional Challenges*. *Chem. Rev.* 118, 8105–8150 (2018)[61]. Cette revue décrit les avancées de l'automatisation de la synthèse des oligosaccharides décrite pour les oligosaccharides au début des années 2000 par le groupe de Seeberger[58].

Cheng, C.-W. et al. *Hierarchical and programmable one-pot synthesis of oligosaccharides*. *Nat. Commun.* 9, (2018)[65] et Zhang, Z. et al. *Programmable One-Pot Oligosaccharide Synthesis*. *J. Am. Chem. Soc.* 121, 734–753 (1999)[59]. Ces articles décrivent un programme « Optimizer » (1999) puis le logiciel software « Auto-CHO » (2018) pour déterminer le meilleur schéma de synthèse pour un oligosaccharide ciblé, à partir de librairies de blocs « disponibles » (BBL : building blocks library) et tenant compte des réactivités des intermédiaires pour le bon agencement des liaisons glycosidiques.

Li, J. et al. *Enzymatic Synthesis of Chondroitin Sulfate E to Attenuate Bacteria Lipopolysaccharide-Induced Organ Damage*. *ACS Cent. Sci.* 6, 1199–1207 (2020)[64] et Li, J., Su, G. & Liu, J. *Enzymatic Synthesis of Homogeneous Chondroitin Sulfate Oligosaccharides*. *Angew. Chem. Int. Ed.* 56, 11784–11787 (2017)[66]. Les auteurs décrivent leur stratégie enzymatique pour la synthèse de CS-A and CS-C. Puis, dans l'article de 2020, les auteurs détaillent la synthèse de CS-E homogènes de différentes tailles jusqu'au 19-mer, ce qui représente à ce jour le plus long CS synthétisé, et, qui plus est, plusieurs centaines de mg ont été ainsi obtenus, **Figure 26**.

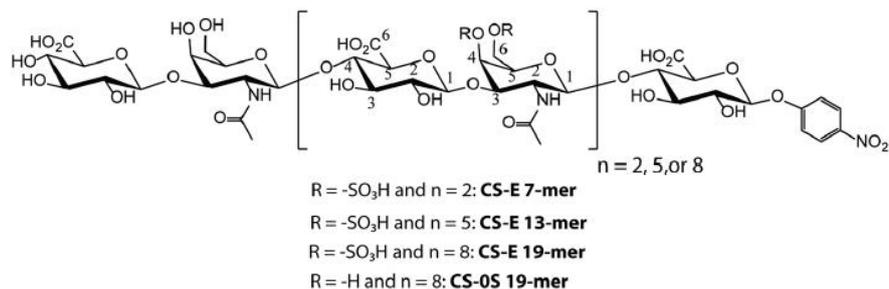


Figure 26. Formules des CS-E de différentes tailles du 7-mer au 19-mer[64].

Xiao, X. et al. *One-Pot Relay Glycosylation. J. Am. Chem. Soc.* 142, 5498–5503 (2020)[67]. Dans cette publication, les auteurs construisent plusieurs liaisons glycosidiques avec les bonnes stéréochimies et sélectivités en « un pot ».

La synthèse de GAGs réalisée par synthèse purement chimique est longue, laborieuse et coûteuse. J’ai pu me rendre compte de l’énergie nécessaire pour obtenir des analogues de chondroïtine sulfate ciblés pour une étude biologique programmée avec des collègues biologistes du Collège de France (équipe du Pr Alain Prochiantz). En effet, Guillaume Despras, doctorant dans le laboratoire sous la direction du Dr J.-M. Mallet, aura passé plus de deux ans de sa thèse à mettre au point une stratégie de synthèse pour obtenir les analogues décrits **Figure 27**, à l’échelle de quelques mg[68]. Ces analogues de chondroïtine sulfate, où un groupe neutre NHAc est remplacé par un OH, mais possédant le même profil de sulfatation, ont été préparés par « oligomérisation » d’un bloc disaccharide, le lactose, décrit **Figure 28**.

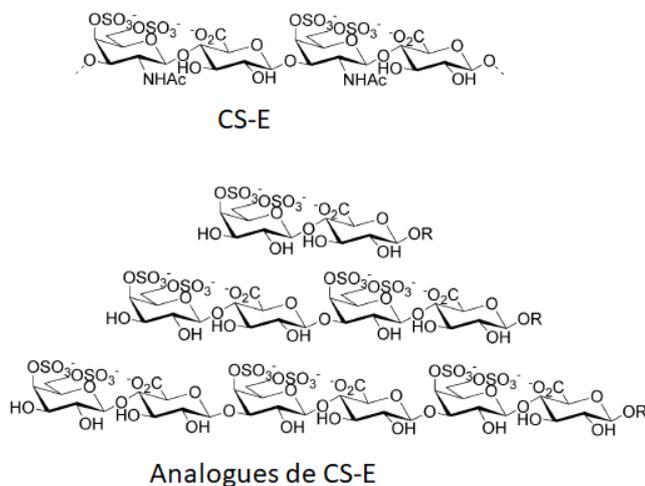


Figure 27. Formule des CS-E (dessus) et des analogues ciblés (dessous)[68].

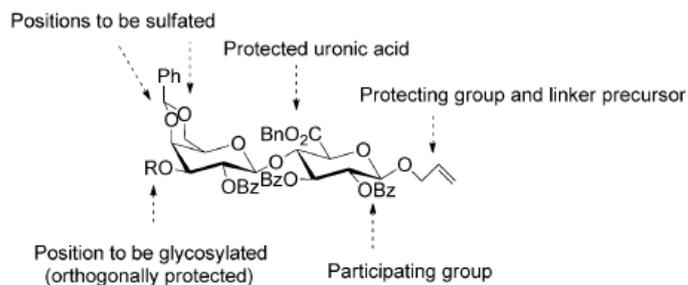


Figure 28. Bloc D-lactose de départ[68].

La stratégie de synthèse qui a été finalement adoptée est résumée **Figure 29**.

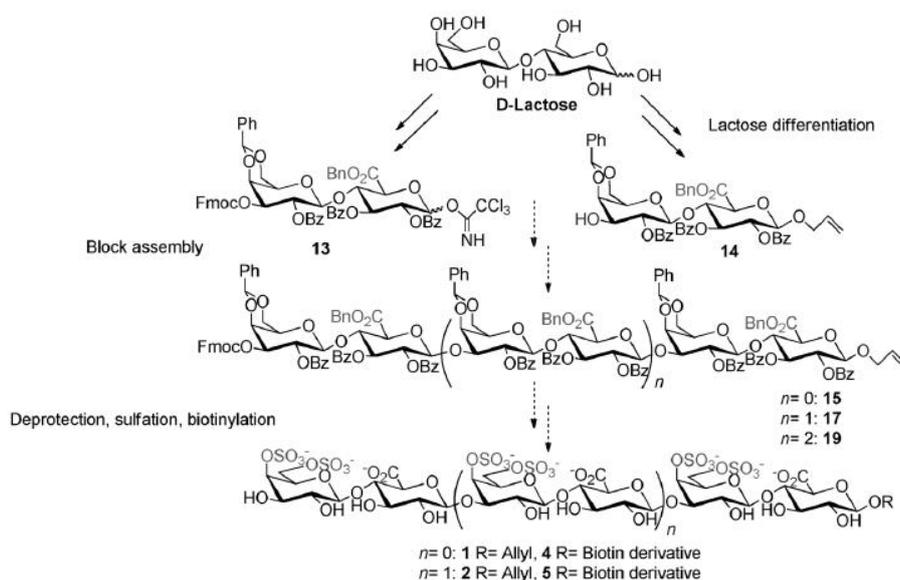
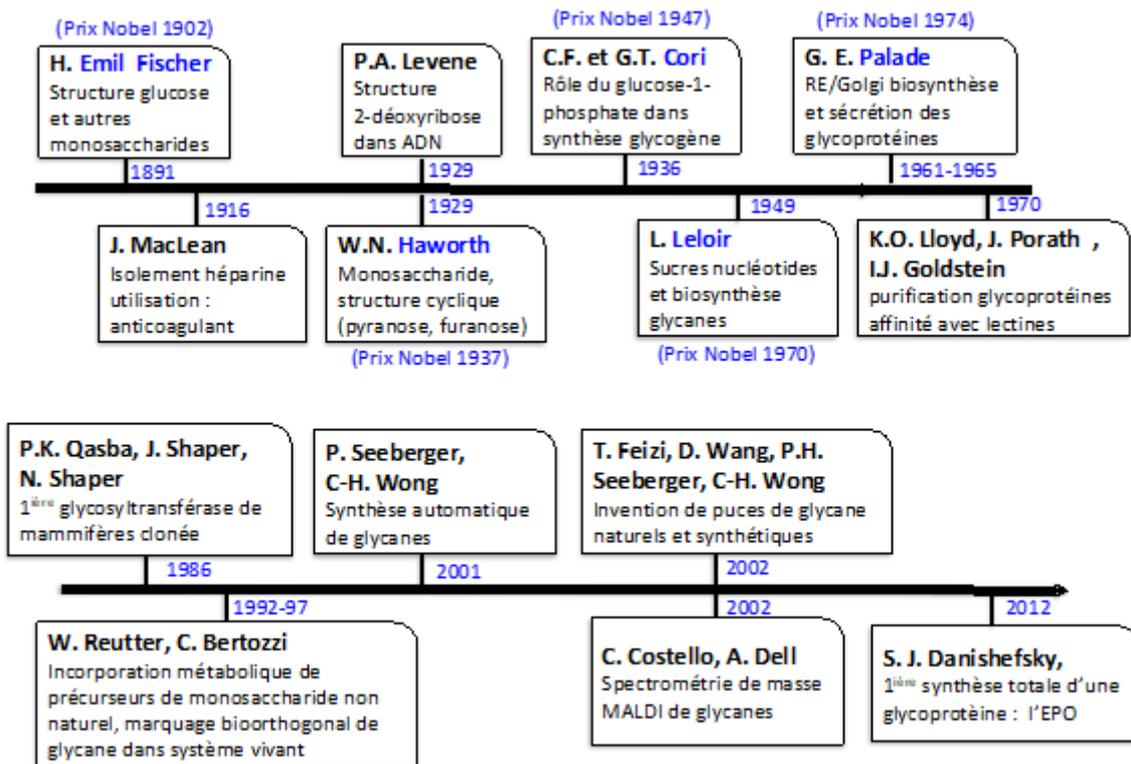


Figure 29. Stratégie de synthèse des CS-E ciblées à partir du bloc lactose[68].

Ces analogues laborieusement obtenus ayant présenté des activités biologiques intéressantes, il a été décidé d'engager un travail pour préparer des mimes peptidiques de ces oligosaccharides, *Cf.*, dans ce chapitre, « C. Objectif de ce travail ».

e) Quelques faits majeurs dans l'Histoire de la chimie des glycanes



Adapté de « Essentials of Glycobiology. 2nd edition[50].

B. Les « GAG-Binding Proteins » ou Protéines de liaison aux Glycosaminoglycanes

Pour rappel : i) Les GAGs sont des polymères acides non ramifiés, de masses moléculaires comprises entre 10^3 et 8.10^6 Da. Ils sont constitués de 20 à plus de 100 unités disaccharidiques, plus ou moins sulfatées (-O-SO₃⁻). Les blocs disaccharides sont composés d'un acide uronique ou d'un galactose lié à une molécule d'hexosamine, potentiellement sulfatés en différentes positions. Les GAGs sulfatés se retrouvent généralement liés à une protéine pour former des protéoglycanes (PG) et ces PG sont dits HSPG pour les protéoglycanes de sulfate d'héparane (pour Heparan Sulfate ProteoGlycan) et CSPG pour les protéoglycanes de sulfate de chondroïtines (Chondroitin Sulfate ProteoGlycan). Les GAGs sont présents à la surface de toutes cellules et dans la matrice extracellulaire. On les retrouve également dans le milieu intracellulaire.

Pour rappel : ii) Les différents groupes de GAGs, classés en fonction de leurs unités disaccharidiques de base, sont : héparine/héparane sulfate, chondroïtine sulfate (CS), dermatane sulfate (DS), kératane sulfate et acide hyaluronique.

Pour rappel : iii) Puisque la suite de ce travail sera focalisée sur les chondroïtine sulfates, les formules des différentes classes de CS sont représentées sur la figure ci-dessous, **Figure 30**.

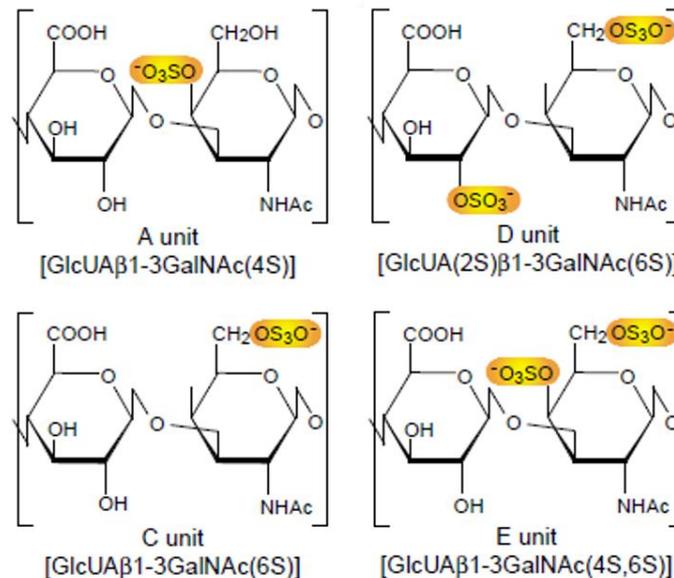


Figure 30. Les quatre unités disaccharidiques de base trouvées dans les chaînes de CS. Les unités disaccharidiques A, C, D et E sont caractéristiques respectivement de CS-A, CS-C, CS-D et CS-E [69].

1. Les GAG-Binding proteins et la matrice extracellulaire

Les interactions d'une cellule avec son environnement impliquent l'intervention de protéines localisées à sa surface, que ce soit pour interagir avec une autre cellule ou avec la matrice extracellulaire (ECM) qui l'entoure. L'ECM, qui maintient les cellules dans un espace défini, est constitué de protéines, telles que collagènes, fibronectines ou laminines et de polysaccharides liés (par exemple pour former les protéoglycanes, PG) ou libres (comme l'acide hyaluronique, HA), **Figures 31a et b**[70], [71]. Les proportions de ces composés varient selon le type de tissu[72].

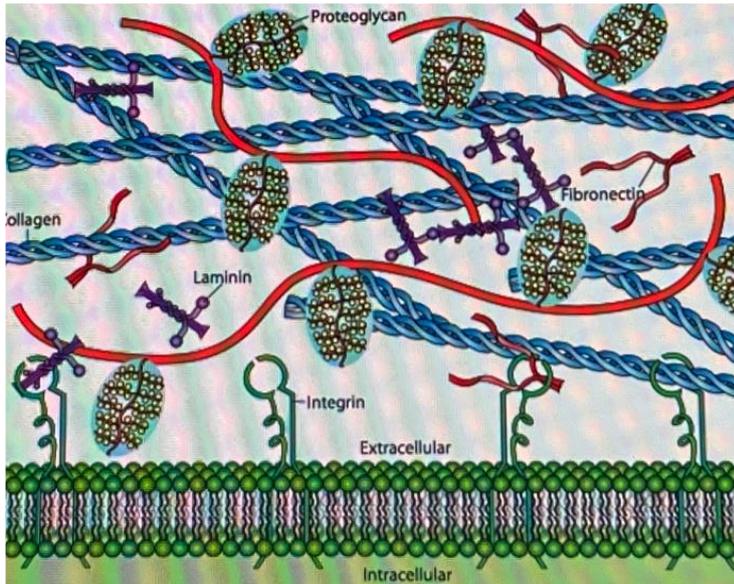


Figure 31a. La matrice extracellulaire. Représentation des composants majoritaires : le collagène, les protéoglycanes (avec leur sphère d'hydratation symbolisée), les fibronectine et laminine[70].

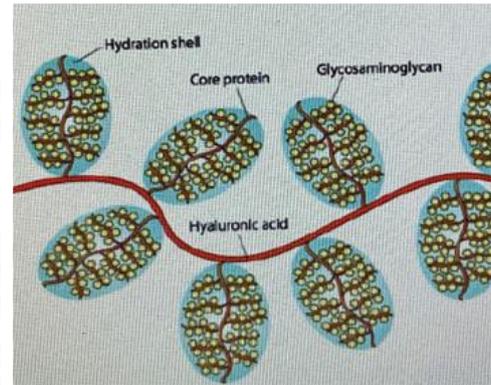


Figure 31b. Les protéoglycanes sont constitués de multiples chaînes de GAGs attachées à une protéine. Ces protéines de cœur sont parfois elles-mêmes liées à une molécule d'acide hyaluronique. Les sucres chargés négativement sont représentés en jaune. La sphère d'hydratation est en bleu[71].

L'assemblage des GAG pour former les PG est réalisé dans le Golgi, mais les mécanismes de contrôle de l'insertion et de la taille des unités disaccharidiques sont loin d'avoir été totalement élucidés. Les PG présentent une extraordinaire diversité de glycosylation, allant de l'ajout de seulement quelques sucres à plus d'une centaine. Les PG peuvent être sécrétés dans la matrice extracellulaire, comme pour les membranes basales (perlécan, agrine), ou liés à la membrane. Parmi les PG liés à la membrane, certains sont des protéines transmembranaires, les syndécans, d'autres sont liés à la surface cellulaire *via* une ancre glycosylphosphatidylinositol (GPI), les glypicanes.

Dans le système nerveux central (SNC), 10 à 20 % du volume du cerveau adulte est occupé par une combinaison de protéines et de polysaccharides qui constituent la matrice extracellulaire, dans laquelle les GAGs de type HS et CS sont les composants majoritaires. Les principaux CSPG sont les protéines de la famille lectican (neurocan, brévican, agrécan et versican), qui interagissent aussi avec l'acide hyaluronique (HA) et les lectines, **Figure 32**[73].

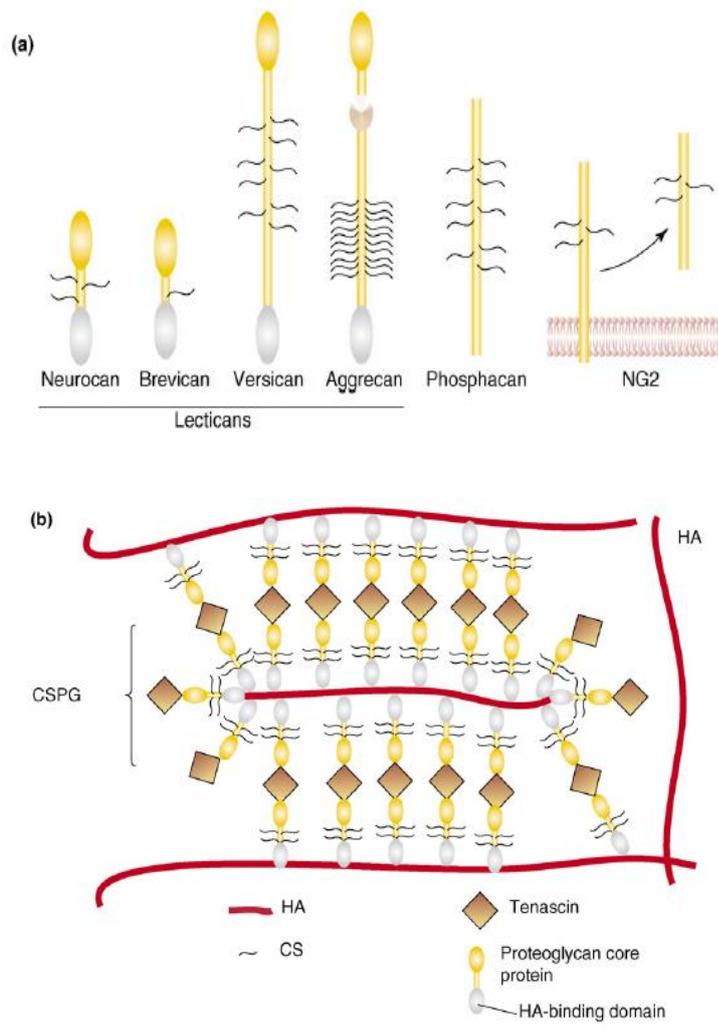


Figure 32. Les CSPGs et HA constituent une grande partie de l'ECM dans le SNC.
 a) Structures de quelques CSPG dans le système nerveux central (SNC). b) Exemple de structuration d'un CSPG avec un HA central[73].

L'ECM entourant les cellules a longtemps été considéré comme une sorte « de glu » pour les molécules d'eau et pour les cations, du fait de sa charge fortement négative. Il a été montré que l'ECM est en réalité un réseau structurel hautement dynamique qui subit en permanence un remodelage sous l'influence d'enzymes qui le dégradent dans des conditions normales et pathologiques. Les GAGs sont en fait les sites de fixation de nombreuses protéines, appelées "GAG-binding proteins", telles que des cytokines, des facteurs de croissance, des enzymes, etc... Les GAGs sont indubitablement impliqués dans divers processus physiologiques mais aussi pathologiques. La dérégulation de la composition et de la structure de l'ECM est associée au développement et à la progression de pathologies, telles

que les cancers, les infections bactériennes et virales, l'inflammation, la maladie d'Alzheimer etc... **Figure 33.**

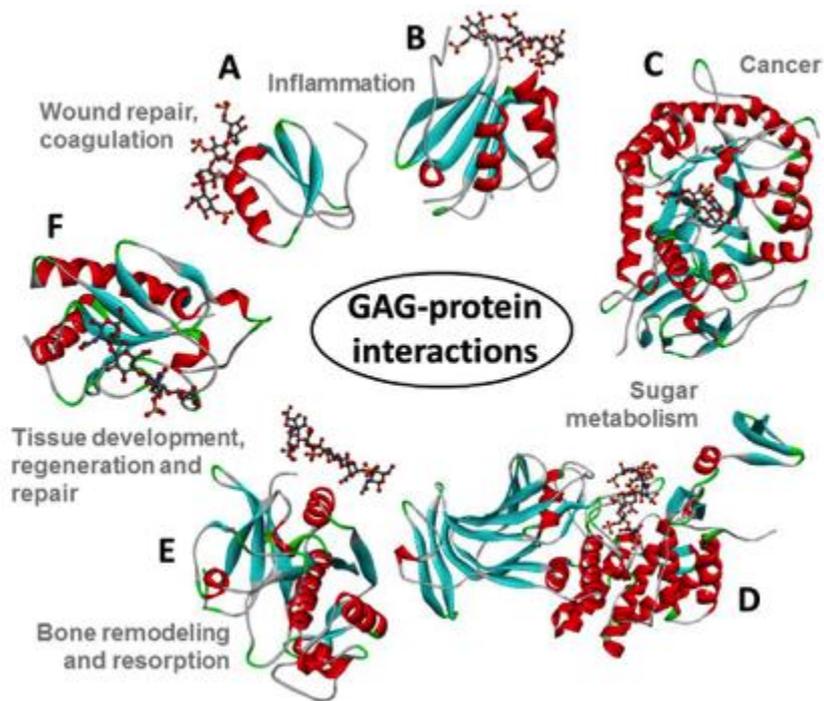


Figure 33. Représentations structurales de quelques complexes GAG-protéine représentatifs et des activités liées à ces complexes. Extraction de la « Protein Data Bank » (PDB). (A) monomère de protéine inflammatoire de macrophage 1-alpha (CCL3) + Héparine (Hp) 4-mer ; (B) dimère du facteur plaquettaire 4 (CXCL4) + fondaparinux (Hp 5-mer de haute affinité pour l'antithrombine) ; (C) complexe d'héparanase humaine + tétrasaccharide Hp ; (D) D-glucuronyl C5-épipimérase + Hp 6-mer ; (E) monomère de cathepsine K + DS 6-mer et (F) Monomère de Sonic Hedgehog (Shh) + CS-A 4-mer[74].

Le nombre de combinaisons de structures possibles pour les GAGs apparaît inconcevable puisque peuvent varier la longueur des chaînes de GAGs, les degrés de sulfatations, les épimérisations... Néanmoins, les interactions entre les GAGs et les GAG-binding proteins sont spécifiques, comme en témoignent des études structure-activité. L'activité biologique des GAGs est très souvent associée à la séquence des modifications par sulfatation et ce sont des séquences spécifiques hautement sulfatées qui sont à l'origine des interactions avec des facteurs de croissance ou avec d'autres protéines impliquées dans la communication cellulaire[75], [76]. Il a même récemment été publié que l'héparine se lie à la glycoprotéine spike du coronavirus du syndrome respiratoire du Moyen-Orient (MERS-CoV) et du coronavirus du syndrome respiratoire aigu sévère (SARS-CoV-2), responsable de la pandémie COVID-19[77].

Les méthodes pour identifier et caractériser les interactions des GAGs avec les différents partenaires ont récemment été inventoriées, **Tableau 3**. À cette liste on ajoutera les approches de glycoprotéomiques, telles que décrites par l'équipe de C. Bertozzi[78].

Méthode	Type	Débit	Principe
Chromatographie d'affinité	M	H/I	Ligands ou GAGs immobilisés sur colonne
Co-électrophorèse d'affinité	M/S	I	Retard sur gel au travers d'un gel imprégné de protéines
Ultracentrifugation analytique	S	L	Sédimentation à l'équilibre à différents rapports saccharides/protéines
Dichroïsme circulaire	S	I/L	Changement de rotation de la lumière polarisée dans le plan lors de la liaison
Compétition par ELISA (Dosage immuno-enzymatique sur support solide)	M	H	Les ligands en solution et immobilisés entrent en compétition pour la liaison
Computationale	S	L	Calcule la structure du complexe et l'énergie de liaison
Spectroscopie de fluorescence	S	H	Changement conformationnel, la liaison du ligand induit un changement de fluorescence
Spectrométrie de masse couplée à mobilité ionique	G	I/L	Mesure la forme du complexe et sa stoechiométrie
Calorimétrie par titrage isotherme	S	I	Mesure directement l'enthalpie de la liaison et les valeurs de K_d
Diffusion de la lumière laser	S	I	Intensités de diffusion intrinsèques du complexe GAG-protéine utilisées pour calculer la stoechiométrie
RMN (Résonance Magnétique Nucléaire)	S	L	Déplacement chimique, constante de couplage et effet nucléaire Overhauser pour déterminer les points de contact, les distances et la conformation
Résonance plasmonique de surface	M	H/I	Changement d'indice de réfraction induit par la masse en temps réel pour la mesure directe des constantes d'association et de dissociation
Rayons X	M	L	Structure cristalline à l'état solide

Tableau 3. Méthodes pour mesurer les interactions GAGs-protéines. M : phase mixte, S : solution, G : gazeuse ; H : élevé, I : intermédiaire, L : faible[37].

Certaines de ces stratégies/techniques permettent de déterminer une affinité, K_d , entre partenaires, d'autres permettent de visualiser les structures des complexes formés, comme indiqué dans cette référence[37], au chapitre 38 « *Proteins that bind sulfated glycosaminoglycans* ». Quoi qu'il en soit, il ne faut pas oublier que bien des stratégies/techniques ont été conduites *in vitro*, voire *in silico*, donc en dehors du contexte du mélange « infini » de structures de GAGs et des protéines, cations, molécules d'eau de solvataion... que l'on peut trouver dans la matrice extracellulaire.

Les approches computationnelles sont des outils importants pour analyser la spécificité des interactions des GAGs (CS, DS, HS/Hp, HA et KS) avec leurs partenaires. L'équipe de S. Ricard-Blum a récemment construit un réseau de ces interactions, après analyse de l'interactome GAGs/GAG-binding proteins d'une librairie de 832 biomolécules (827 protéines et 5 GAGs)[79]. Les résultats de cette analyse, présentés **Figure 34a, 34b**, montrent qu'il existe une spécificité de reconnaissance entre les GAGs et les différents partenaires protéiques (ou GAG-binding proteins). Les GAGs doivent adopter des structures privilégiées dans ces complexes avec des angles dièdres ϕ et ψ définis. L'organisation en domaines sulfatés et domaines *N*-acétylés contribue également à la spécificité des interactions avec les GAG-binding proteins.

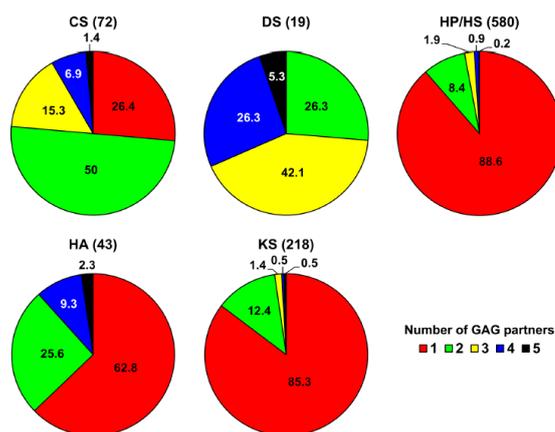


Figure 34a. Partenaires protéiques spécifiques et non-spécifiques des cinq GAGs. Les % de GAG-binding proteins qui se lient à un (rouge), deux (vert), trois (jaune), quatre (bleu) ou cinq (noir) GAGs sont indiqués pour chaque GAG. Le nombre total de partenaires de chaque GAG est indiqué entre parenthèses[79].

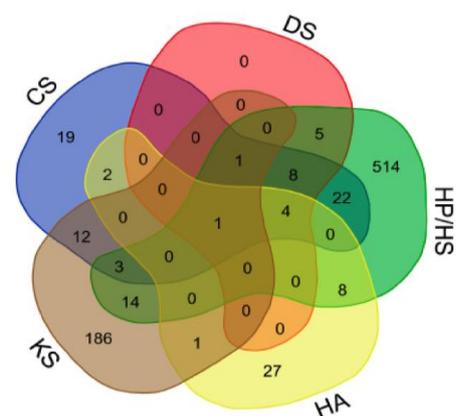


Figure 34b. Diagramme de Venn. Indique les nombres : i) de partenaires spécifiques de chaque GAG et ii) de GAG-binding proteins communes à deux, trois, quatre ou cinq GAGs[79].

2. Analyse des interactions entre GAGs et GAG-binding proteins

C'est encore Emil Fischer qui a émis dès 1894 l'hypothèse du modèle clef-serrure pour interpréter l'action spécifique des enzymes vis-à-vis de leurs substrats. Cette hypothèse est : *"Um ein Bild zu gebrauchen, will ich sagen, dass Enzym und Glucosid wie Schloss und Schlüssel zu einander passen müssen, um eine chemische Wirkung auf einander aus üben zukönnen"*. "Pour utiliser une image, je veux dire que l'enzyme et le glucoside doivent correspondre comme une serrure et une clé afin de pouvoir exercer un effet chimique l'un sur l'autre"[80].

Les enzymes ne reconnaîtraient pas des molécules dont la conformation serait différente, même si la différence est minime, une enzyme et son substrat seraient complémentaires comme une clé et une serrure. Il est cependant difficile d'envisager les GAGs dans un rôle « de clef » rigide considérant la flexibilité inhérente des chaînes oligosaccharidiques et les solvatations de ces molécules hydrophiles. Il est plus concevable qu'un mécanisme de type « induced-fit » (ajustement induit), avec une adaptation conformationnelle de la protéine qui se déforme pour « accepter » la conformation du substrat[81], voire un modèle plus dynamique, impliquant une adaptation mutuelle des deux partenaires[82], décrive mieux ces interactions.

Du fait de la diversité des structures des GAGs, il existe dans la littérature peu de structures de complexes élucidées et encore moins de données disponibles sur les cinétiques d'interactions entre GAGs et GAG-binding proteins. Quelques structures de GAG-binding proteins ont été déterminées, plus particulièrement des complexes impliquant des dérivés et/ou analogues de l'héparine et des mini-héparines.

Les interactions des GAGs avec les GAG-binding proteins impliquent des types d'interactions variés : des interactions ioniques mais aussi des forces de van der Waals, des liaisons hydrogène ou des interactions hydrophobes avec le squelette glucidique. Des études de ces complexes ont montré que les interactions ioniques, suspectées comme étant la base des interactions de ces molécules fortement chargées négativement, ne contribuent en fait qu'en partie à la liaison des deux entités. Par exemple, 30 % de l'énergie libre de liaison de l'héparine au β FGF (β -fibroblast growth factor) est à mettre au titre d'interactions ioniques tandis que les forces non ioniques, telles que des liaisons hydrogène et des interactions

hydrophobes, contribuent pour le reste à l'affinité entre l'héparine de bas poids moléculaire et le β FGF[44]. Dans le cas de l'interaction de l'héparine avec le peptide cérébral natriurétique (BNP), l'énergie libre de liaison provenant des interactions ioniques est presque négligeable (6 %), avec une contribution majeure provenant de liaisons hydrogène (94 %) entre les acides aminés polaires du BNP et l'héparine[44].

L'analyse de ces complexes a conduit ainsi à déterminer des séquences d'acides aminés consensus dans les protéines comme domaines d'ancrage pour la liaison de GAGs, avec bien évidemment l'implication d'acides aminés basiques. Cardin et Weintraub, en 1989, ont analysé 49 régions dans 21 protéines de liaison à l'héparine et ont proposé que les sites de liaison des protéines possèdent la séquence XBBXB ou XBBBXXB, où B est un résidu basique, une Lysine ou une Arginine, et X un résidu « hydrophobe », les plus courants étant Asn, Ser, Ala, Gly, Ile, Leu et Tyr[56], [83]. Ces acides aminés peuvent être localisés dans des régions en hélice α ou dans des régions structurées en brin β . En 2012, il a été proposé que le motif minimal de reconnaissance pouvait être décrit comme un agencement cation-polaires-cation (CPC), correspondant à deux acides aminés cationiques (Arg ou Lys) et un acide aminé polaire, le plus fréquemment Asn, Gln, Thr, Tyr ou Ser, et plus rarement Arg ou Lys. Ce motif de clip (CPC) fonctionnerait comme une agrafe permettant l'attachement de l'héparine, et d'autres GAGs sulfatés, aux protéines de liaison à l'héparine[84].

3. Les interactions entre les protéines à homéodomaine et les chondroïtines sulfates dans le système nerveux central

a) Organisation du réseau CSPG dans le SNC

La matrice extracellulaire (ECM) dans le système nerveux central (SNC) représente de 10 à 20 % de son volume total. Les réseaux périneuronaux (PNN) sont des assemblages d'ECM qui enveloppent préférentiellement les interneurons exprimant la parvalbumine (cellules PV), **Figure 35**.

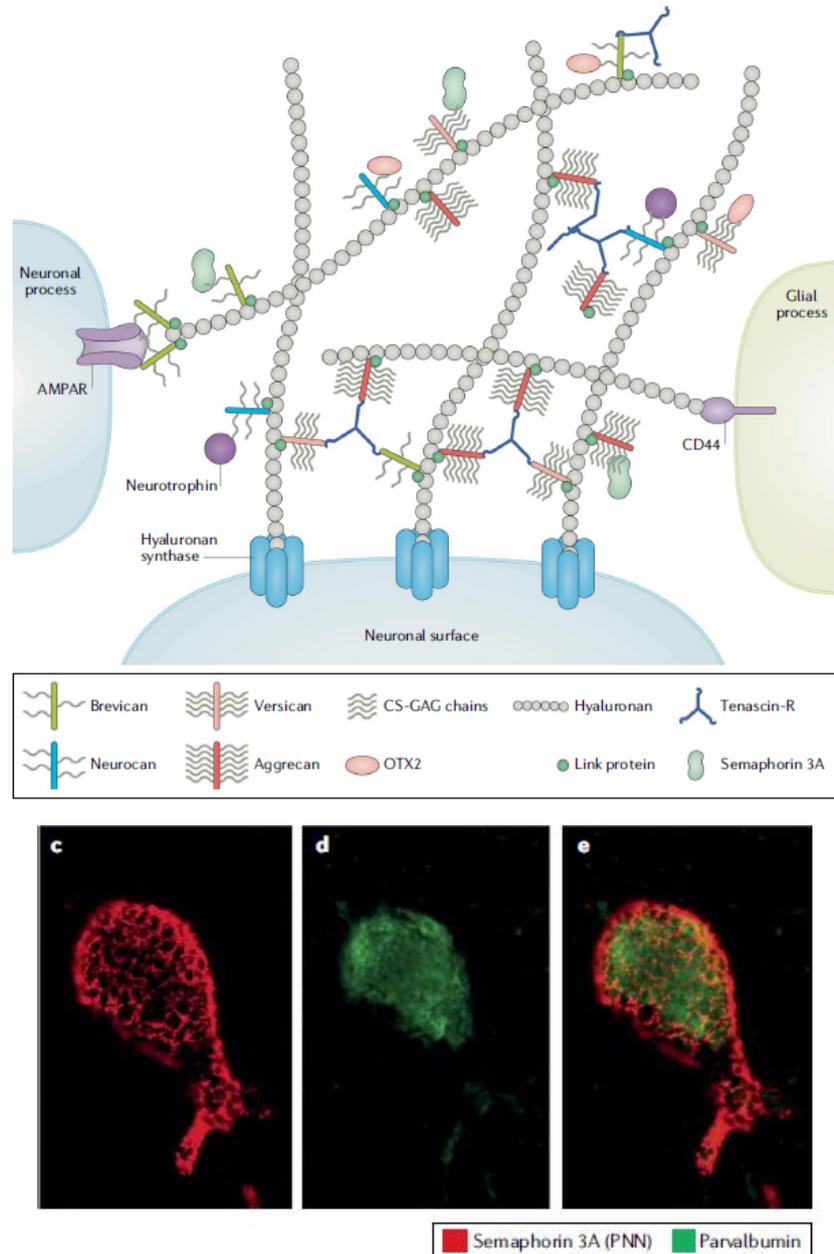


Figure 35. Composition, localisation et aspect des réseaux périneuronaux (PNN). c-e : Un interneurone exprimant la parvalbumine (en vert) provenant d'un cortex cérébral de rat, entouré d'un PNN révélé par un anticorps anti-sémaphorine 3A, montrant la structure en réseau du PNN[85].

Cette matrice extracellulaire est enrichie en glycosaminoglycans, chargés négativement, nécessaires aux rôles physiologiques des PNNs. Il a été établi que les cellules PV et le PNN sont altérés dans une variété de troubles neurologiques. Le développement et l'organisation finale du PNN sont nécessaires pour un certain nombre de processus au sein du SNC, notamment la régulation de la fonction cellulaire GABAergique, la protection des

neurones contre le stress oxydatif et la fermeture des fenêtres de plasticité de la période critique de développement. En effet il a été montré qu'un des mécanismes régulant la fermeture de cette période critique (courte période durant laquelle un circuit neuronal donné est hautement plastique) est la mise en place d'une matrice condensée de PNN autour des interneurons inhibiteurs exprimant la parvalbumine dans la couche IV du cortex cérébral[85].

Les CSPG sont la composante majeure en PG de l'ECM dans le SNC, le rapport chondroïtine sulfates (CS) total vs héparane sulfates (HS) dans le SNC est de 9:1 et il est de 7:3 dans la matrice du réseau périneuronal (PNN). Des changements drastiques s'opèrent dans les profils de sulfatation des chaînes CS pendant la formation des PNN. La sulfatation en position 6 est dominante dans le cerveau en formation, la 4-O-sulfatation devient majoritaire dans le cerveau adulte pour produire du C4S. Dans l'ensemble, une augmentation substantielle du rapport 4-sulfatation/6-sulfatation (C4S/C6S) est observée au cours du développement cérébral. Dans le développement postnatal du cerveau, ces profils différentiels de sulfatation spécifiques sont cruciaux pour la mise en place des périodes critiques.

b) Plasticité et période critique

Nous focaliserons cette présentation sur l'aspect que nous étudierons dans la suite de ce travail : Otx2, un facteur de transcription, et la vision binoculaire.

L'apprentissage de la vision binoculaire est un exemple de cette période critique. La plasticité de dominance oculaire chez la souris est déterminée dans le temps comme indiqué **Figure 36**. Chez la souris, après l'ouverture des yeux à 14 jours après la naissance (P14), l'ouverture de plasticité se réalise à P20 avec un pic de plasticité autour de P30 et se ferme à P40. La majorité des données impliquent les interneurons PV dans le contrôle de cette plasticité. Cette période correspond en effet à la maturation des neurones corticaux inhibiteurs GABAergiques à parvalbumine (cellules PV)[86].

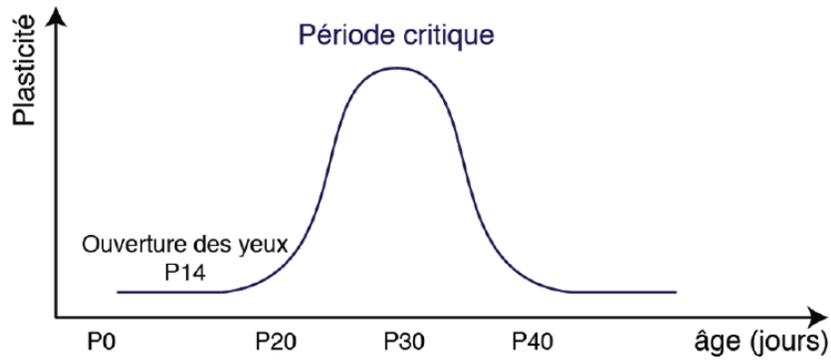


Figure 36 : Fenêtre temporelle de plasticité chez la souris.

Il a été montré que cette maturation, et, par elle, le contrôle de la plasticité, est induite par l'activité d'un facteur de transcription spécifique : l'homéoprotéine Otx2, **Figure 37**.

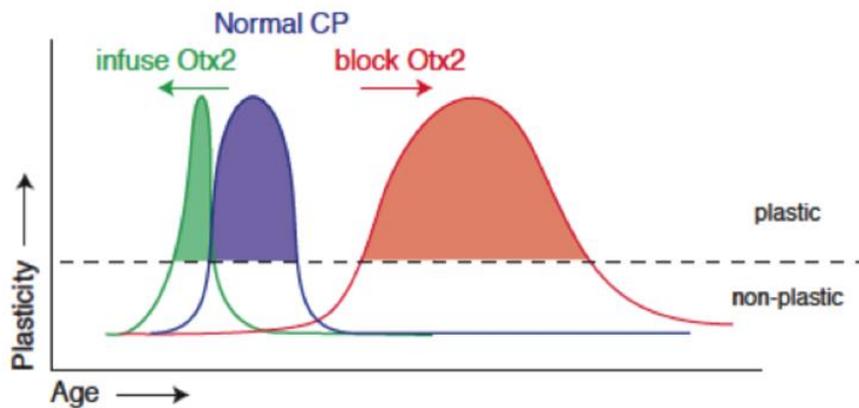


Figure 37 : Otx2 contrôle la période critique de dominance oculaire. La période critique (CP) de plasticité (en bleue) peut être avancée par infusion d'Otx2 (en vert) ou retardée en bloquant Otx2 (en rouge)[87].

Otx2 est synthétisée en dehors du cortex cérébral, elle est sécrétée, puis transportée dans le milieu extracellulaire et internalisée spécifiquement par les cellules à PV[88]–[90] puisque, dans les conditions physiologiques ou après son infusion, Otx2 est retrouvée à plus de 90 % dans les cellules GAD67 positives et majoritairement dans les interneurons PV [88]–[90].

c) Protéines à Homéodomaine et Otx2

Les protéines à homéodomaine

Les homéoprotéines sont des facteurs de transcription. En se fixant à des séquences spécifiques d'ADN, elles contrôlent la transcription d'un gène. Les protéines à homéodomaine, présentes exclusivement chez les Eucaryotes, adoptent une structure tridimensionnelle générique lors de cette fixation à l'ADN. Elles présentent en particulier une structure canonique, dite homéodomaine, hélice-tour-hélice d'environ 60 acides aminés, dans laquelle 3 hélices α sont reliées par des boucles courtes. L'hélice C-terminale, hélice 3, est la plus longue, elle est riche en Arg et Lys qui établissent des liaisons avec l'ADN, **Figure 38**.

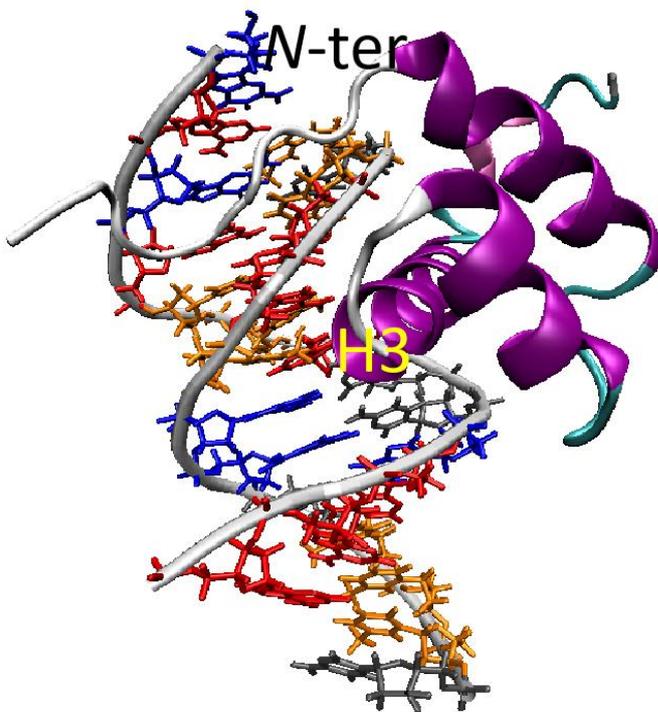


Figure 38. Homéodomaine de l'homéoprotéine Antennapedia de *Drosophila melanogaster* (en violet), liée à un fragment d'ADN (structure déterminée par RMN). L'hélice 3 (H3) et l'extrémité N-terminale (N-ter) non structurée sont liées respectivement au grand sillon et au petit sillon de l'ADN. PDB: 1AHD[91].

Au début des années 1990, le groupe du Pr Alain Prochiantz a découvert la capacité des homéoprotéines et des homéodomains à : i) être sécrétés d'un neurone et ii) à « traverser » les membranes biologiques. Cette hypothèse allait à l'encontre du dogme de « l'imperméabilité » de la bicouche lipidique vis à vis des molécules hautement hydrophiles comme les homéoprotéines. La première observation du rôle de ces protéines à homéodomaine dans la signalisation cellulaire a été faite avec l'homéodomaine d'*Antennapedia*. Cet homéodomaine pouvait pénétrer spontanément dans des neurones

vivants *in vitro*, sans altérer la membrane plasmique et, une fois internalisé, il conservait son affinité pour l'ADN, conduisant à une excroissance des neurites[92], [93]. Sur les 10 homéoprotéines testées, toutes ont montré cette capacité à passer d'une cellule à l'autre *in vitro*[94]. Un domaine minimal de 16 acides aminés de long, appelé pénétratine et issu de l'homéodomaine d'*Antennapedia*, a été démontré comme étant la séquence nécessaire et suffisante permettant la translocation, **Figure 39**. D'une façon générale cette séquence correspond à la troisième hélice de l'homéodomaine.

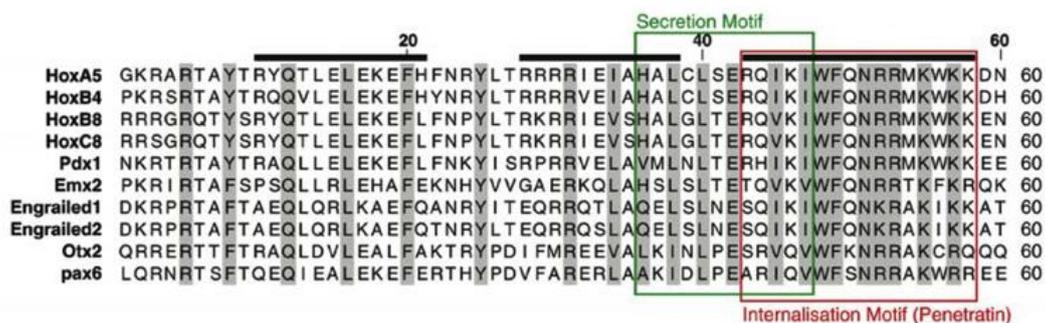


Figure 39. Séquences minimales permettant l'internalisation et la sécrétion des homéoprotéines, alignement de dix séquences. La séquence d'internalisation (en rouge) correspond à l'hélice 3 de l'homéodomaine et la séquence nécessaire pour la sécrétion (en vert) correspond à la fin de l'hélice 2, le début de l'hélice 3 et le coude entre les deux hélices[94].

L'internalisation de la pénétratine se produit à 4 °C ainsi qu'à 37 °C et n'est par conséquent pas dépendante d'une liaison à un récepteur, peu favorable à 4 °C[95]. Cette démonstration a ouvert la porte aux recherches sur les peptides pénétrants ou « cell penetrating peptides » (CPP) qui sont un des thèmes de recherches du Laboratoire des Biomolécules depuis de nombreuses années[95]. Nous ne développerons pas cet aspect, ce n'est pas l'objet de ce travail.

L'homéoprotéine Otx2

Pour rappel, il a été démontré que la protéine Otx2 pénétrait spécifiquement dans les cellules PV, induisant leur maturation et ainsi le contrôle de la plasticité. D'autres expériences ont montré que la concentration en Otx2 dans ces cellules était fonction de l'état du PNN et en particulier de la présence des CSPG. Cette concentration en protéine Otx2 pouvait être réduite après déstructuration de la matrice extracellulaire, soit par son hydrolyse avec la chondroïtinase ABC (ChABC), soit par la modification du statut rédox des cellules PV. La

fenêtre de plasticité est donc en partie contrôlée par l'interaction de Otx2 avec des GAGs, plus précisément avec des CS, présents dans le réseau périneuronnel enveloppant les cellules PV[86], [96].

Ces observations suggéraient l'existence, dans la séquence de la protéine Otx2, d'un motif de reconnaissance aux GAGs, de type XBBXB ou XBBBXXB, où B est un résidu basique, une Lysine ou une Arginine, et X un résidu « hydrophobe ». Un tel motif, de séquence RKQRRER, a bien été identifié, au niveau de la première hélice de l'homéodomaine, en position N-terminale d'Otx2, comme site de liaison entre l'homéoprotéine et les GAGs, **Figure 40**[86].

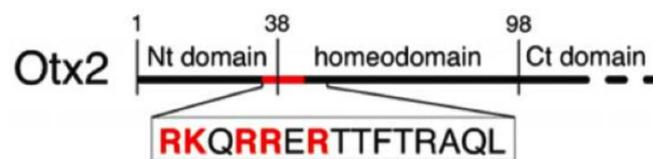


Figure 40. Domaine de liaison aux GAGs dans la séquence de Otx2[86].

Le profil de sulfatation des CS est important pour leur reconnaissance par le domaine de liaison aux GAGs d'Otx2. Des expériences de calorimétrie de titration isotherme, réalisées entre différents CS commerciaux et un peptide synthétique, le « **RK-peptide** » correspondant au motif de reconnaissance identifié (**Figure 40**), ont démontré la sélectivité de sa liaison avec des CS disulfatées CS-D et CS-E[86]. Chez des souris, la diminution *in vivo* du ratio C4S/C6S, a entraîné la réduction de l'accumulation d'Otx2 dans les cellules PV, provoquant un retard de la fin de la période critique[96].

Les équipes du Pr A. Prochiantz ont montré que l'infusion du **RK-peptide** dans le cortex visuel d'une souris adulte, entraînant ainsi une compétition avec la protéine Otx2 endogène pour l'interaction avec les CS du PNN, permettait d'induire une réouverture de la période de plasticité[90]. Cette même réouverture a été observée après blocage de la protéine Otx2 endogène par infusion d'un hexasaccharide synthétique analogue de CS-E, produit par l'équipe du Dr J.-M. Mallet[68].

Ces résultats indiquent qu'il est possible de jouer sur la plasticité du cortex cérébral adulte en modifiant la quantité d'Otx2 qui s'accumule dans les cellules PV. Perturber les

périodes critiques en modulant les interactions Otx2-GAG apparaît donc comme une approche innovante pour de possibles applications thérapeutiques.

C. Objectif de ce travail

Ci-dessous, je vais récapituler, au travers de deux figures, les informations connues concernant la période critique de plasticité, avec ses tenants et ses aboutissements, afin de présenter l'objectif de mon travail et résumer mes travaux de recherche.

Tout d'abord une figure plus « visuelle » de la période critique de plasticité, empruntée à la publication de Hou *et al.*, **Figure 41**[97].

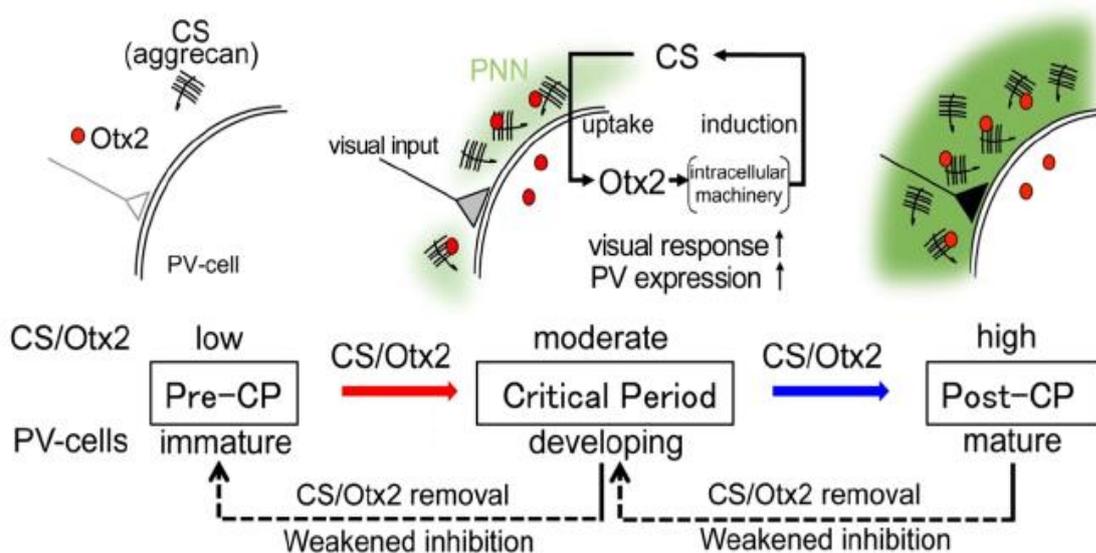


Figure 41. Modèle schématique : accumulation de CS-Otx2 et maturation des cellules PV. La présence initiale de CS et Otx2 (flèche rouge) favorise le début de la période critique de plasticité, une accumulation supplémentaire (flèche bleue) conduit à la fermeture de cette période critique (CP)[97].

La suite des événements qui conduisent à la vision binoculaire peut être décomposée en plusieurs étapes. La protéine, Otx2, **Figure 42**, sécrétée par les plexus choroïdes, est transférée pendant le développement postnatal vers le cortex visuel. Otx2 est captée préférentiellement dans les couches IV du cortex visuel par l'aggrécan substitué par des molécules de CS, préférentiellement de type CS sulfatés en C6, présents autour des interneurons GABAergiques à parvalbumine (PV)[86], [96], [97]. Otx2 est internalisée par ces neurones PV par un mécanisme n'impliquant pas de récepteur protéique[90]. Dans la couche

IV du cortex visuel une majorité des neurones contenant Otx2 sont des interneurons inhibiteurs et plus de 70 % d'entre eux sont PV-positifs[87]. Comme indiqué **Figure 41**, l'accumulation d'Otx2 dans les cellules PV se fait parallèlement à la maturation de ces cellules et à la production des réseaux périneuronaux (PNN) qui entourent ces neurones (en vert sur la **Figure 41**). L'accumulation de PV et d'agrégan va conduire à une rigidité de l'environnement des cellules PV au stade mature avec fermeture de la période critique (post-CP).

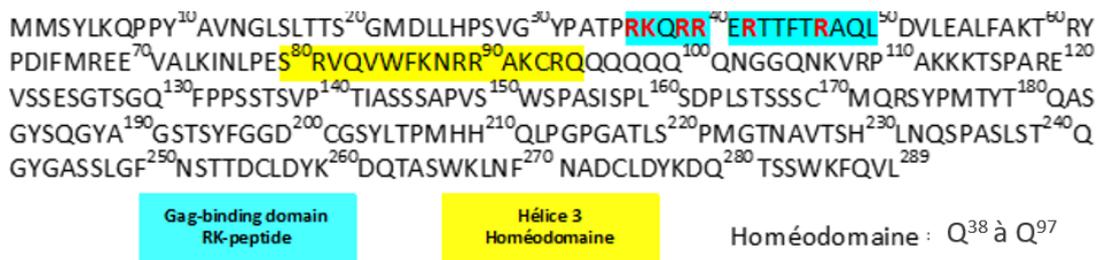


Figure 42. Séquence de Otx2 chez l'Homme[98].

Dans cette boucle de rétrocontrôle, l'activité visuelle induit donc la formation initiale des PNN au début de la période critique (CP) autour des cellules PV, permettant l'internalisation d'Otx2 dans ces cellules. Pendant cette CP, l'augmentation de la concentration en Otx2 dans les cellules PV active l'assemblage des PNN. À leur tour, les PNN assurent l'accumulation spécifique d'Otx2 dans les cellules PV. Chez l'adulte, le transfert constant d'Otx2 dans les cellules PV, en raison de cette boucle de rétroaction positive entre l'homéoprotéine et les PNN, maintient un état mature, non plastique.

Il a été montré que l'état plastique (au centre de la **Figure 41**) peut être activé ou désactivé, dans un sens ou dans l'autre (flèches noires hachurées). Un traitement conduisant à l'élimination de CS (par exemple par action de chondroïtinases ABC, ChABC) ou à l'élimination de Otx2 (par exemple par infusion d'anticorps anti-Otx2) réinitialise la période critique CP (état plastique) à pré-CP (état immature). Il en est de même pour l'état post-CP (état mature) qui peut être réinitialisé en état CP (état plastique), comme indiqué sur la **Figure 41**.

L'interaction entre CS et Otx2 apparaît donc cruciale pour la fonction des interneurons inhibiteurs GABAergiques à PV dans les deux étapes correspondant à

l'ouverture puis la fermeture de la période critique. Otx2 ouvre non seulement la période critique, mais entraîne également sa fermeture.

En modifiant la quantité d'Otx2 qui s'accumule dans les cellules PV on peut donc moduler la plasticité du cortex cérébral visuel adulte, et en particulier revenir à un état de plasticité. C'est ce qui a été réalisé avec succès, par l'équipe du Pr A. Prochiantz au Collège de France, en infusant i) le **RK-peptide**, **RKQRRERTTFTRAQL (Figure 42)**, qui va entrer en compétition avec la fixation de la protéine endogène Otx2 aux GAGs présents à la surface des neurones, ou, en collaboration avec l'équipe de J.-M. Mallet, ii) un analogue hexasaccharidique de CS-E (HexaCSE) préparé, lors de sa thèse, par Guillaume Despras à partir du Lactose, qui va bloquer la protéine endogène Otx2 à l'extérieur des cellules PV[68]. Ces approches ont été résumées de manière schématique (**Figure 43**) dans la thèse de Damien Testa, avec qui j'ai eu le plaisir de collaborer pour ma thèse.

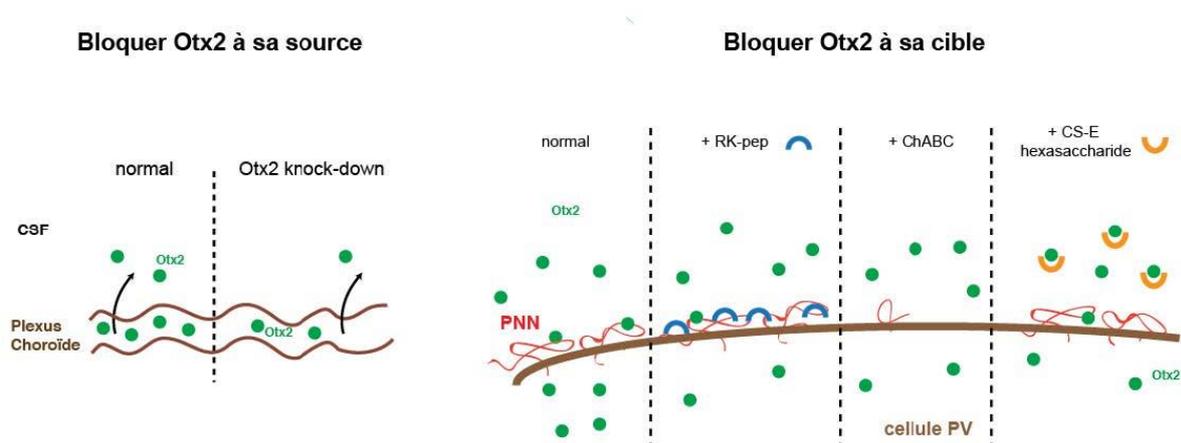


Figure 43 : Réouverture de plasticité en régulant la concentration de Otx2 dans les cellules PV. La modification de synthèse d'Otx2 par les plexus, *via* un knock-down d'Otx2, diminue la quantité d'Otx2 dans le cortex visuel. Il est aussi possible de modifier l'interaction d'Otx2 avec les PNN, soit en utilisant la séquence de reconnaissance aux GAG (RK-pep), soit en hydrolysant la matrice avec la chondroïtinase ABC (ChABC), soit en infusant un CS-E synthétique (hexasaccharide CS-E)[99], [100].

L'équipe du Collège de France a également montré que l'infusion du **RK-peptide** chez des souris adultes conduit à une amélioration d'une amblyopie induite par suture d'un œil durant la période critique de plasticité[86]. L'amblyopie est une différence d'acuité visuelle entre les yeux, qui affecte 2 à 5 % de la population. La partie du cerveau qui doit traiter l'information venant d'un œil ne fonctionne pas de manière optimale. Ce dysfonctionnement

peut être corrigé chez le jeune enfant s'il est détecté avant 6 ans mais ne peut l'être plus tard, à moins d'être en mesure de rouvrir une fenêtre de plasticité dans la zone du cortex visuel affectée, par exemple en bloquant la protéine Otx2 à l'extérieur des cellules PV.

La correction de l'amblyopie dans le cortex visuel est donc une application possible de ces travaux, mais l'implication d'Otx2 dans d'autres pathologies peut aussi être envisagée.

Dans le cadre de mon travail de thèse en collaboration avec l'équipe du Pr Alain Prochiantz au Collège de France, j'ai conçu et synthétisé des substituts peptidiques de GAGs, que nous appellerons, dans la suite de ce document, mimes de GAGs. En théorie, préparer des peptides est beaucoup plus facile et rapide que de réaliser la synthèse d'un hexasaccharide analogue de GAGs. Je décrirai les séquences choisies et je présenterai les différentes stratégies explorées et les difficultés rencontrées, auxquelles j'ai dû apporter des solutions. Ceci constituera le premier volet de mes résultats que je détaillerai dans le chapitre II de ce manuscrit.

Dans le chapitre III, je présenterai les résultats biochimiques et biologiques qui ont été obtenus avec les mimes que j'ai préparés. Ces expériences ont été intégralement réalisées par Damien Testa au cours de sa thèse au Collège de France, sous la direction du Dr Ariel Di Nardo ; je reprendrai son travail. Les deux parties de cette étude, biochimique et biologique, ont conduit à un dépôt de brevet (WO2019243272A1) et au dépôt d'une demande ANR qui a été acceptée (<https://anr.fr/Projet-ANR-18-CE16-0013>).

Dans le cadre de notre collaboration, j'ai réalisé l'étude par RMN (Résonance Magnétique Nucléaire) des interactions de ces mimes de GAGs avec la protéine Otx2, que j'ai réussi à exprimer et produire avec un double marquage ^{13}C , ^{15}N . Cette étude structurale a été réalisée sous la supervision du Dr Ludovic Carlier, maître de conférence de notre Laboratoire, le LBM. Je détaillerai les stratégies utilisées pour déduire des affinités entre les deux partenaires et identifier les acides aminés de la protéine en interaction avec les mimes. J'ai comparé les résultats obtenus avec une autre protéine à homéodomaine Engrailed 2, dont la séquence de liaison aux GAGs est différente et qui avait déjà été produite dans le laboratoire avec le double marquage *ad hoc*. En parallèle de ce travail j'ai également réalisé, avec l'aide de Tahar Bouceba, ingénieur au sein de l'Institut de Biologie Paris-Seine (IBPS) de Sorbonne

Université, une étude comparative, cette fois par SPR de type Biacore, des interactions des mimes de GAGs avec la protéine Otx2. Ces études seront détaillées dans le chapitre IV « Etudes des interactions in vitro par RMN et SPR de type Biacore ».

Dans le cinquième chapitre, je présenterai des travaux préliminaires sur la synthèse de la deuxième génération de mimes peptidiques, nos essais de mesures d'affinités des deux partenaires par polarisation de fluorescence et les résultats préliminaires obtenus.

Enfin je résumerai les faits marquants de mon travail et j'ébaucherai les perspectives ouvertes par mes travaux dans une conclusion générale.

CHAPITRE II : Synthèse des mimes de GAG

A. Introduction : des sucres aux peptides

L'étude des réseaux d'interaction, interactome, des GAGs[101], avec les nombreuses techniques développées pour les analyser[102], [103], a contribué à asseoir l'importance des GAGs dans la Chimie du Vivant[75], [104]–[106], au-delà de l'héparine et ses analogues[107], [108]. S'est posé alors le problème de l'obtention de ces GAGs, compte tenu de leurs potentiels thérapeutiques et des difficultés rencontrées pour les isoler, les purifier et les caractériser à partir de sources naturelles. Diverses approches de synthèse ont alors émergé: synthèse chimique et/ou enzymatique et/ou à partir de sources naturelles[63], [109], [110]. Mais si l'unité de base, la brique du chimiste, est relativement simple, les agencements, et SURTOUT la taille et l'inhomogénéité de sulfatations de la molécule finale, constituent les problèmes essentiels pour le chimiste impliqué dans l'étude des interactions GAGs/protéines et leur modulation. De nombreuses approches de synthèses de mimes, de plateformes de molécules organiques sulfatées ou de molécules naturelles judicieusement modifiées et de matériaux gels, nanofibres, billes etc.... sont étudiées. Nous n'en citerons que quelques-unes, cette liste n'a rien d'exhaustif car la littérature est très riche[57], [111]–[127]. Des structures peptidiques ont été modifiées pour conduire à des mimes d'oligosaccharides et également de GAGs, comme reporté dans les quelques publications[128]–[131]. Il n'existe cependant que de très rares travaux à partir de séquences minimales peptidiques incorporant des fonctions sulfate ou sulfonate conçues comme mimes de GAGs, quatre publications des années 2005-2007[132]–[135], que je détaillerai ci-dessous.

En 2005, dans la publication *GAG Mimetic Libraries: Sulphated Peptide as Heparin-like Glycosaminoglycan Mimics in Their Interaction with FGF-1*[132], M. Meldal et son équipe au Danemark ont été les premiers à proposer de développer des peptides sulfatés comme mimes de GAGs, supposant qu'il devait être possible d'obtenir, avec des peptides plus faciles à produire, une topographie de charges similaire à celle des GAGs, **Figure 44**.

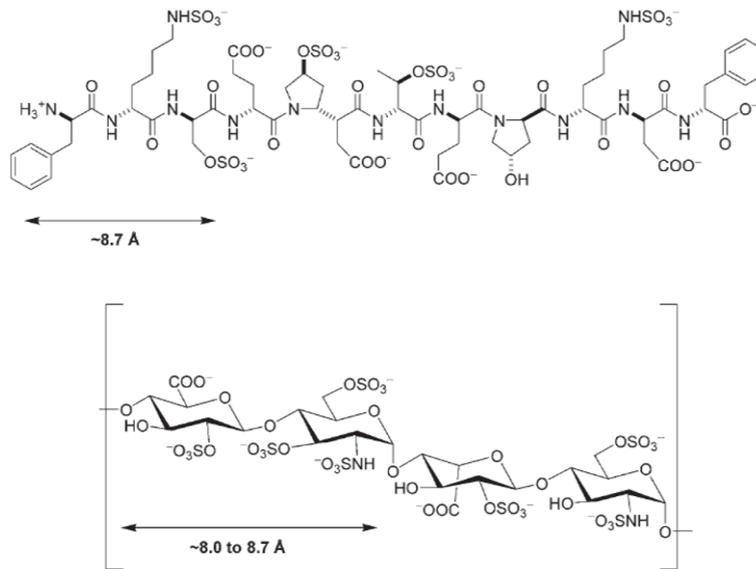


Figure 44. Structures chimiques d'un tétrasaccharide d'héparine (en bas) et d'un mime peptidique sulfaté (en haut) suggérant le degré de *O*-sulfonation nécessaire pour un mime peptidique[132].

Ils ont synthétisé, par une approche « mix-split », une petite librairie combinatoire de peptides sulfatés, par sulfatation de fonctions amine et/ou alcool. Pour cela deux stratégies ont été explorées : i) la sulfatation du peptide encore ancré sur le support solide (Méthode A), ii) l'utilisation d'acides aminés sulfatés disponibles commercialement ou préparés spécifiquement pour cette étude (Méthode B)[136], **Figure 45**.

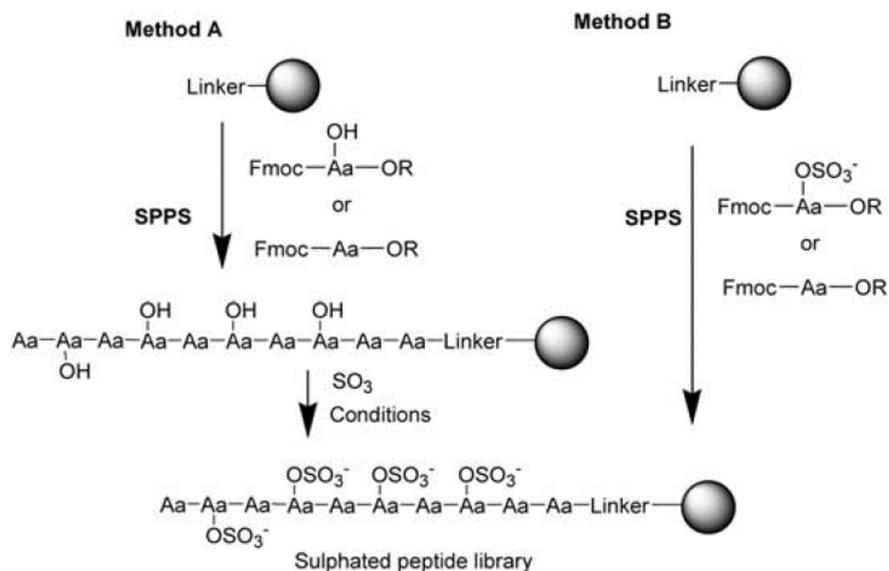


Figure 45. Deux stratégies de préparation de librairies par une approche « mix-split »[132].

La deuxième voie a été privilégiée, elle conduit à des peptides où la sulfatation et la pureté du peptide final peuvent être maîtrisées. Dans les deux stratégies le problème majeur reste la stabilité de la Sérine (Thréonine) sulfatée en position bêta, la fonction sulfate étant un excellent groupe partant, la réaction de β -élimination qui conduit à la déhydroalanine est difficilement contrôlable (cf. pages 69-70).

Deux hendecapeptides de cette librairie, acétylés en *N*-terminal, amides en *C*-terminal et porteurs de 5 ou 6 sulfates, Ac-E-T-E-T*-S*-S*-E-S*-E-S*-K-NH₂ et Ac-T*-L-D-Hyp*-T*-S-L-S*-T*-S*-K-NH₂, l'étoile indiquant les positions sulfatées, inhibent l'interaction du facteur de croissance des fibroblastes (FGF) avec l'héparine à des concentrations micromolaires, ce qui équivaut à l'affinité d'un tétrasaccharide de l'héparine contenant pareillement 6 sulfates.

La même année, dans la publication *Discovery of a sulfated tetrapeptide that binds to vascular endothelial growth factor*[133], Maynard et Hubell ont préparé, par synthèse en phase solide et par stratégie "mix-split", une librairie, d'approximativement 6600 membres, de petits peptides sulfatés sur des chaînes latérales de tyrosine. Ces peptides contiennent, comme les saccharides de l'héparine, des fonctions alcool, acide et sulfates, portées respectivement par les chaînes latérales de Sérine, acide Aspartique ou acide Glutamique et Tyrosine sulfatée, **Figure 46**.

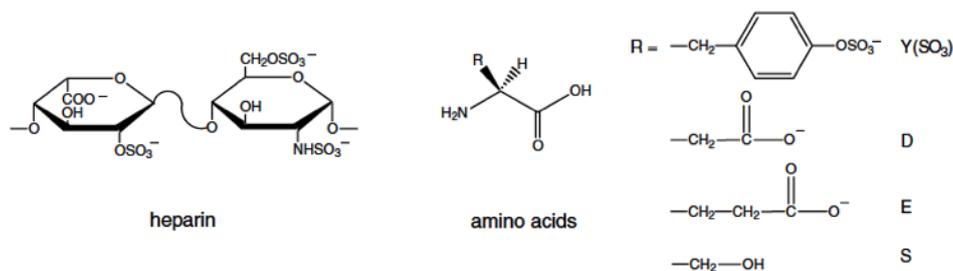


Figure 46. Structure du bloc disaccharide de l'héparine et formules des acides aminés incorporés dans la librairie de peptides[133].

Le pentapeptide disulfaté Ser-Tyr(SO₃)-Asp-Tyr(SO₃)-Gly se lie au facteur de croissance de l'endothélium vasculaire (l'isoforme VEGF₁₆₅), avec une affinité (K_d) micromolaire cent fois meilleure que celle de la suramine. Cette différence d'affinité est essentiellement apportée par un k_{off} mille fois meilleur que celui de la suramine, l'urée symétrique de l'acide

(benzamido-3-méthyl-4-benzamido)-1-naphtalène trisulfonique-4.6.8, qui contient six fonctions sulfonates, **Figure 47**.

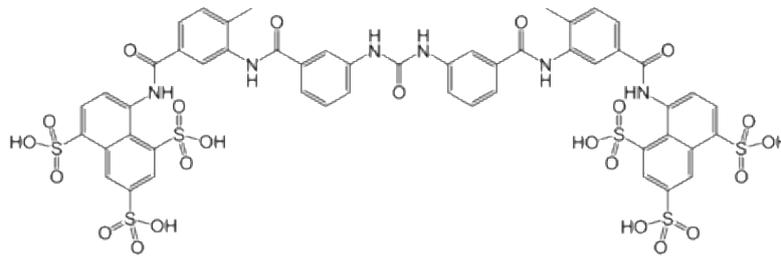


Figure 47. Formule de la suramine.

Deux ans plus tard, dans la publication *Heparin-mimetic sulfated peptides with modulated affinities for heparin-binding peptides and growth factors*[134], Kim et Kiick ont repris la séquence de base Ser-Tyr(SO₃)-Asp-Tyr(SO₃)-Gly, décrite par Maynard et Hubell[133], et ont synthétisé, de manière conventionnelle par synthèse sur support solide, un dimère de ce dipeptide, Gly-Gly-Gly-Gly-Ser-Tyr(SO₃)-Asp-Tyr(SO₃)-Ser-Tyr(SO₃)-Asp-Tyr(SO₃)-Gly, dans le but de préparer des hydrogels pour la délivrance de facteurs de croissance, comme décrit **Figure 48**. Ils ont mesuré son affinité pour différentes protéines de liaison de l'héparine et les ont comparé aux affinités du monomère de Maynard et Hubell[133] et du peptide Gly-Gly-Gly-Gly-Tyr(SO₃)-Tyr(SO₃)-Gly-Gly-Tyr(SO₃)-Asp-Tyr(SO₃)-Gly-OH avec des distributions de sulfates différentes.

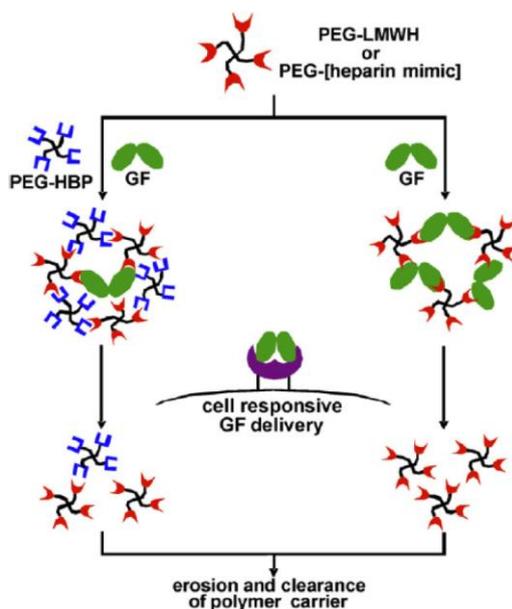


Figure 48. Hydrogels assemblés de manière non covalente pour la délivrance de facteurs de croissance. PEG : poly (éthylène glycol), LMWH: héparine de bas poids moléculaire, HBP : peptide liant l'héparine, GF : facteur de croissance[134].

Quelle que soit la protéine de liaison étudiée, c'est le peptide sulfaté le plus court, celui initialement proposé par Maynard et Hubell[133], qui présente la meilleure affinité. Kim et Kiick ont proposé qu'avec ce pentapeptide disulfaté la perte entropique devait être moins importante au moment de la liaison qu'avec un dimère[134].

En 2012, dans la publication *Growth Factor Binding on Heparin Mimetic Peptide Nanofibers*[135], Mammadov R. *et al.* se sont intéressés à l'immobilisation des facteurs de croissance dans des gels formés de peptides amphiphiles substitués d'héparine, et leurs applications en médecine régénérative pour la délivrance contrôlée de facteurs de croissance. Pour ce faire, ils ont produit des nanofibres par mélange d'un peptide amphiphile (PA) synthétique chargé positivement, K-PA : Lauryl-Val-Val-Ala-Gly-Lys-NH₂ avec un peptide amphiphile synthétique chargé négativement, HM-PA ou E-PA, acylés en N-terminal et amide en C-terminal pour HM-PA, acide en C-terminal dans le cas de E-PA, ou avec de l'héparine. Après la formation de nanofibres, l'héparine et les groupes fonctionnels des peptides amphiphiles sont supposés être présentés par les nanofibres dans les milieux aqueux environnants, comme représenté schématiquement **Figure 49**.

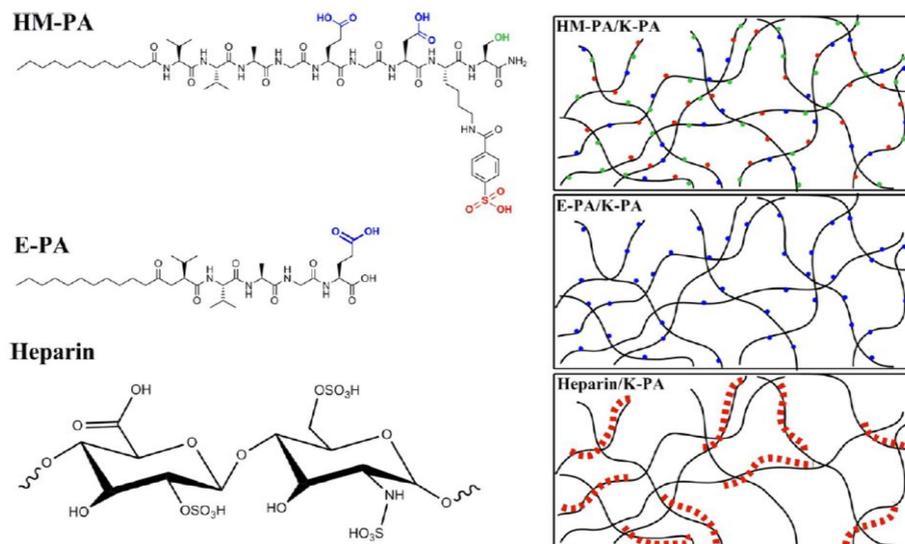


Figure 49. Formation de nanofibres avec des peptides amphiphiles[135]. La ligne rouge pointillée indique l'héparine. Les sphères sur les nanofibres représentent les groupes fonctionnels des peptides amphiphiles (rouge : sulfonate, vert : hydroxyle, bleu : carboxylate). La densité de l'héparine et des groupes fonctionnels de HM-PA sur les nanofibres de PA ne reflète pas la « réalité », ce n'est qu'une représentation schématisée.

Les résultats montrent que les nanofibres constituées d'un mélange HM-PA/K-PA, et donc avec des fonctions carboxylates (comme PE-PA/K-PA) mais aussi alcools et sulfonates, sont les meilleures pour lier par exemple le VEGF, avec un pouvoir de fixation comparable aux nanofibres produites avec de l'héparine.

Les informations que nous retiendrons de ces travaux sont qu'il est possible, avec des peptides, de préparer des mimes de GAGs, même si les affinités restent de l'ordre du micromolaire. Seule la tyrosine sulfatée a été utilisée pour « mimer » la fonction sulfate présente dans les GAGs, puisque c'était le seul acide aminé sulfaté commercialement accessible, les essais avec des Sérines (ou Thréonines) sulfatées ayant avorté pour un problème de chimie, difficile à contourner. Les motifs proposés sont relativement courts en raison du coût de l'acide aminé sulfaté protégé de manière adéquate. Depuis ces travaux, datant de dix à quinze ans, il n'y a pas eu d'autre publication sur la préparation de peptides sulfatés comme mimes de GAGs.

Avant de décrire mes travaux, je vais résumer les stratégies, avantages et problèmes de la synthèse peptidique sur support solide.

B. Principe de la synthèse peptidique sur support solide

Après les premières synthèses d'un dipeptide à la fin du XIX^{ième} siècle[22], puis, plus de 50 ans plus tard, de la vasopressine et de l'ocytocine, des nonapeptides cycliques[18], [19], [24], les avancées majeures qui ont démocratisé la synthèse peptidique sont celles proposées, moins de 10 ans plus tard, par B. Merrifield avec la synthèse peptidique sur support solide (SPPS)[25] et l'automatisation de la SPPS en collaboration avec J. Stewart[27], **Figure 50**. Le principe de la SPPS est résumé sur la **Figure 51**.



Figure 50. B. Merrifield et J. Stewart devant le premier synthétiseur automatique, (1965), Archive du « The Rockefeller Center ».

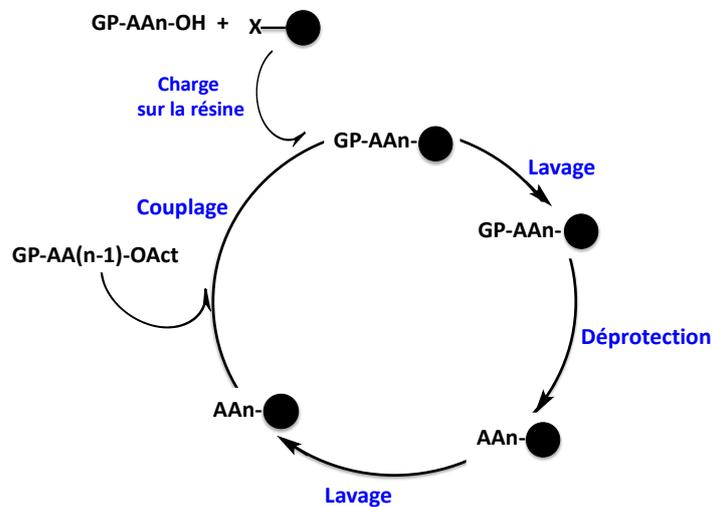


Figure 51. Principe de la synthèse sur support solide, un cycle de synthèse. GP : Groupe Protecteur ; AA : Acide Aminé.

L'acide aminé C-terminal protégé, GP-AA_n, est lié à la résine. Après déprotection du groupe protecteur temporaire GP de l'amine de l'acide aminé AA_n, la fonction acide de l'acide aminé GP-AA_(n-1) est activée puis couplée à l'amine déprotégée de AA_n, sur le support solide. Si nécessaire, un groupe protecteur orthogonal, permanent, est introduit sur la chaîne latérale de certains AA pour éviter des réactions secondaires. Entre chaque opération, des cycles de lavage/filtration de la résine sont réalisés pour éliminer les excès de réactifs. Ceci constitue un cycle qui sera répété n fois, n étant le nombre d'acides aminés présents dans la séquence du peptide ciblé. La chaîne peptidique est coupée du support en fin d'assemblage, cette réaction élimine en même temps tous les groupes protecteurs permanents présents sur les chaînes latérales. Merrifield, dans son article *princeps*, a utilisé le couple de groupes protecteurs

orthogonaux Boc/Bzl (Boc : *t*-butyloxycarbonyl et Bzl : « de type benzylique»), ce qui impose d'utiliser l'acide fluorhydrique (HF) en dernière étape[137], **Figure 52**.

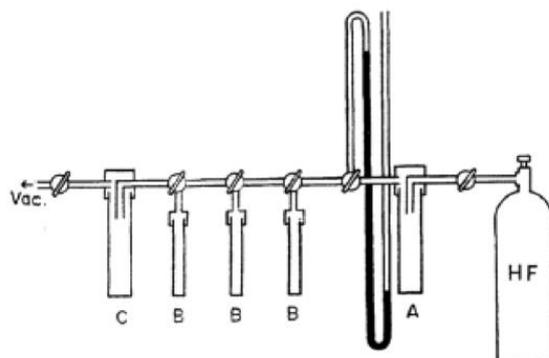


Figure 52. Premier appareil proposé par S. Sakakibara *et al.*, pour le traitement par HF de peptidyl-résine jusqu'à 1g. A : cylindre (30x200 mm) pour la purification et la mesure de HF, B : cylindre pour la réaction de coupure par HF (20x150 mm), C : cylindre pour piéger le HF en fin de réaction, Vac. : Vide[137].

Pour éviter cette étape délicate, voire « dangereuse », le couple Fmoc/tBu a été proposé par Carpino *et al.*[138], [139] (Fmoc : fluorénylméthoxyloxycarbonyl et tBu : de type « *tertio*-butylique »), pour une revue récente *Cf.* Behrendt *et al.*[140]. Il suffit alors de traiter avec de l'acide trifluoroacétique (TFA) en fin de réaction.

Quelle que soit la stratégie, Boc/Bzl ou Fmoc/tBu, en fonction de la nature du support et du lien avec le support, une fonction acide carboxylique ou amide en C-terminal est générée lors de la coupure.

La SPPS, décrite comme une « messy chemistry » par un des rapporteurs du premier article de B. Merrifield, a été très critiquée. Max Brenner, un chercheur incontournable de la chimie des peptides en solution, a résumé les critiques véhémentes lors de la parution/présentation des premiers travaux de Merrifield par cette phrase : « *The invention of the solid-phase method looked like an ingenious trick to overcome some of the unpleasant features of the classical methods. As we know today, the ingenuity of the trick remains, but only a large investment of heavy real effort will eventually, if ever, work it into a real progress over the classical approach* », ce qui peut se traduire par : L'invention de la méthode en phase solide ressemblait à une astuce ingénieuse pour surmonter certains des désagréments des méthodes classiques. Comme nous le savons aujourd'hui, l'ingéniosité de l'astuce demeure,

mais seul un réel investissement de gros efforts finira éventuellement, ou pas, par la transformer en un véritable progrès par rapport à l'approche classique[141].

Les efforts ont été fournis ! La SPPS a pu conduire à des bruts de synthèse moins complexes avant purification grâce aux évolutions des supports solides[142] mais surtout des agents de couplage permettant d'activer l'acide carboxylique des AA : des chlorures d'acide aux carbodiimides, aux sels d'uronium[143]–[148], pour des articles résumant ces évolutions, Cf. [149]–[151]. Les réactifs les plus utilisés, le couple DCC/HOBt, HATU, HBTU, COMU et Oxyma Pure, sont décrits **Figure 53**. Un exemple de mécanisme de réaction d'activation d'un AA, avec le couple DIC/HOBt, puis de couplage est donné **Figure 54**.

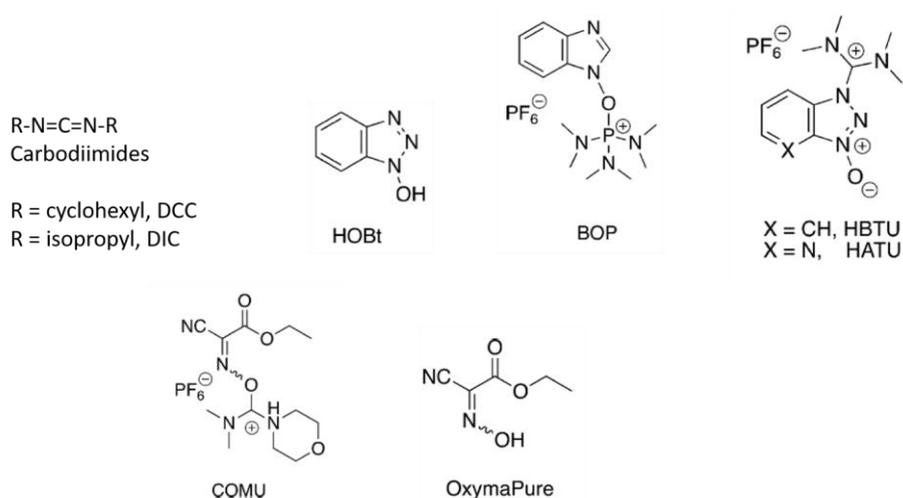


Figure 53. Structures des agents de couplage les plus utilisés.

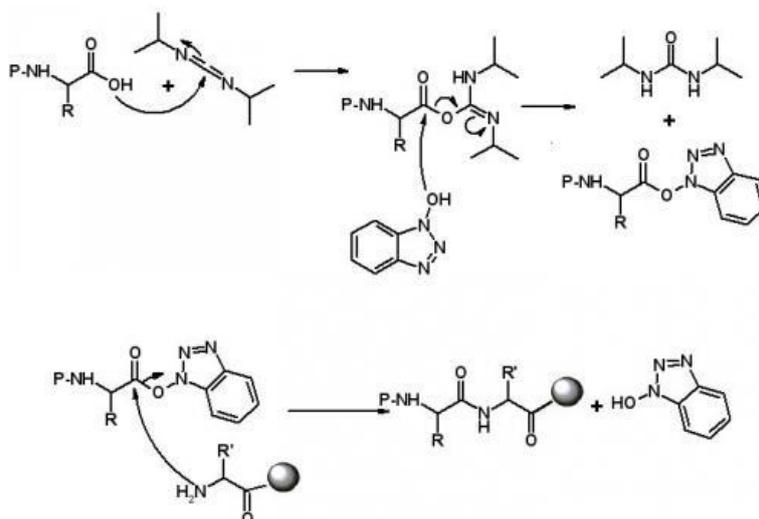


Figure 54. Couplage de deux acides aminés après activation par le couple DIC/HOBt.

Cette réaction de couplage entre deux acides aminés peut être également activée par l'utilisation de microondes, comme proposé en 1995, une approche controversée mais néanmoins implantée sur certains synthétiseurs automatiques[152]–[155].

Les systèmes de purification, tout particulièrement les colonnes HPLC (High Performance Liquid Chromatography) en phase inverse, ont aussi considérablement évolué, Cf. revue de Mant *et al.*[156].

Les schémas de synthèse et les systèmes de purification font encore l'objet de nombreuses études, car les grandes Sociétés pharmaceutiques s'intéressent de plus en plus au potentiel des peptides médicaments. Dans les 20 prochaines années, les peptides deviendraient leur mine d'or, alors que, pendant près de 50 ans, leur credo était qu'un peptide ne pourrait devenir un médicament[157], **Figure 55**. Cet intérêt est important si l'on veut faire évoluer cette chimie vers une chimie plus « verte »[158], [159].

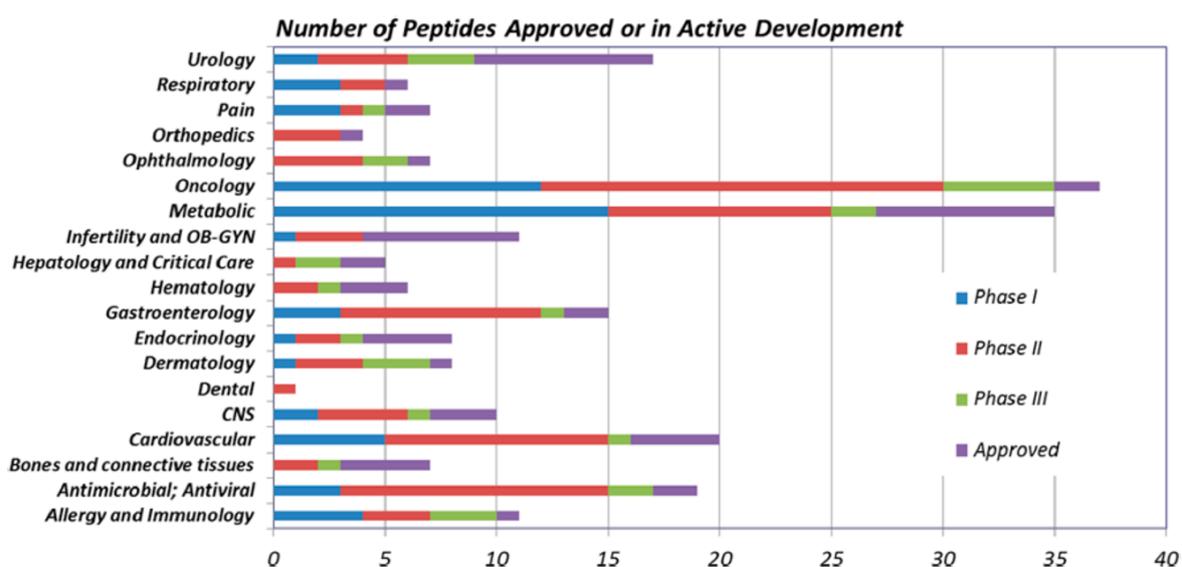


Figure 55. Peptides médicaments mis sur le marché ou en phase active d'étude (données de 2016)[157].

Les avancées de la synthèse peptidique ont permis d'obtenir des peptides totalement déprotégés relativement longs, avec une fonction électrophile en C-terminal, qui peut alors être couplée, par ligation chimique native, avec l'amine N-terminale d'un autre fragment en milieu aqueux pour conduire à des protéines, **Figure 56**.

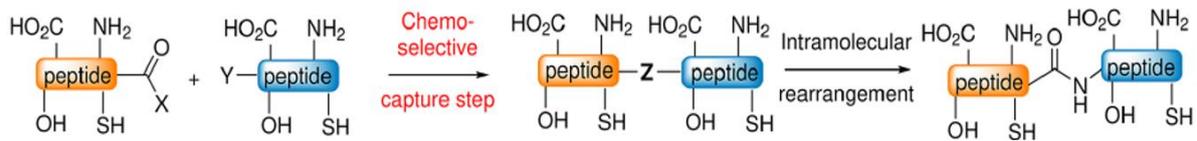


Figure 56. Principe général de la formation chémosélective de liaison amide par ligation chimique native[31].

Pour clore cette introduction à la synthèse peptidique sur support solide, je citerai quelques revues récentes, en plus de celles mentionnées dans le chapitre I, sur les avancées de la ligation chimique[31], [160] ou la ligation assistée par des enzymes[161], [162] pour ouvrir l'accès à des protéines.

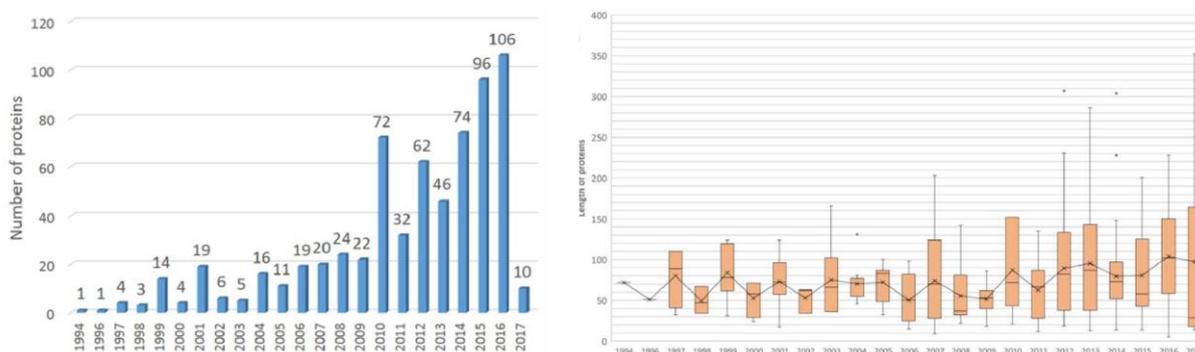


Figure 57. Nombre de protéines produites par ligation chimique (à gauche) et taille des protéines produites (à droite) par rapport à l'année de publication[160]

À ce jour, la question n'est probablement pas de savoir si cette stratégie de ligation est plus « rentable » pour accéder à des protéines (jusqu'à 350 acides aminés, le record, **Figure 57**) que le génie génétique et les approches recombinantes mais, plutôt, que peut apporter cette stratégie que ne peuvent faire d'autres approches ?

C. Stratégie de préparation des mimes peptidiques de CS-E

Nous avons vu précédemment qu'il était possible de mimer des séquences de GAGs à partir d'une séquence peptidique, **Figure 44**[132]. Nous avons également vu, avec les exemples de polymères cités dans la partie introductive de ce chapitre[135], qu'une fonction sulfonate ($R-SO_3^-$), présente dans de nombreux polymères, pouvait « mimer » une fonction sulfate ($-O-SO_3^-$) présente dans les GAGs. De plus des études, réalisées par Vlachy et *al*,

Figure 58[163] et Skvarnavičius *et al.*[164], indiquent qu'un sulfonate est un bon ligand de fonction guanidinium et/ou ammonium.

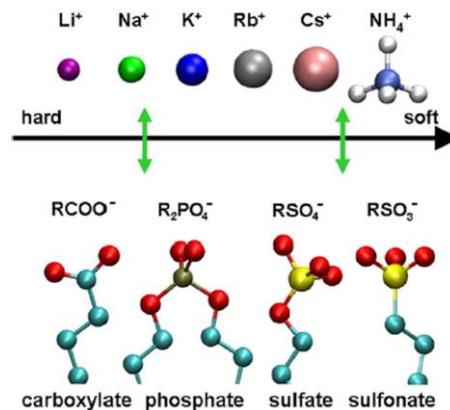


Figure 58. Têtes polaires de surfactants arrangés suivant leur capacité à former des paires d'ions intimes avec des cations[163].

G. Chassaing, dans le laboratoire, a montré, qu'en solution, le polymère PSS-Na, poly(styrène sulfonate) de sodium, piégeait de manière extrêmement efficace des peptides pénétrants, tels que Arg₉ ou la pénétratine (résultats non publiés). Rappelons que le PSS-Na est biocompatible et est utilisé pour piéger les ions K⁺ chez les personnes atteintes d'hyperkaliémie (taux de potassium sérique anormalement élevé).

1. Séquence modèle pour les essais préliminaires, un dimère de térapeptide.

Dans un premier temps, nous avons choisi une séquence peptidique modèle pour étudier les conditions expérimentales qui vont conduire, à partir d'acides aminés naturels, comme respectivement les acides Aspartique ou Glutamique, la Cystéine, la Sérine ou la Thréonine, à des peptides porteurs des fonctions trouvées majoritairement dans les GAGs : les fonctions carboxylate, sulfonate et/ou sulfate et alcools.

Les différents points que nous avons analysés avaient pour but d'établir les meilleures conditions :

- de synthèse, stratégie de synthèse Boc ou Fmoc pour le peptide initial,
- d'oxydation pour obtenir les fonctions sulfonates et/ou sulfates,
- de purification par HPLC et/ou autres techniques,
- d'analyse par spectrométrie de masse MALDI-TOF du produit final,
- et de stockage, attestant de la stabilité des peptides ciblés.

Ces différents points n'ont pas été étudiés de manière aussi linéaire et hiérarchisée, comme listé ci-dessus. C'est en faisant qu'on découvre comment faire et/ou où sont les problèmes ! Je ne ferai pas souffrir le lecteur avec tous les aléas auxquels j'ai dû faire face, ce fut un long combat de plus d'une année, je tenterai de faire didactique.

J'ai jusque-là occulté les problèmes qu'un chimiste des peptides se doit de connaître et il y en a. Les « grandes » réactions secondaires ont été listées pour chaque acide aminé, dès le début des années 1980, dans au moins deux monographies produites par M. Bodanszky et J. Martinez et qui ont fait loi dans le domaine[165], et pour des bilans plus facilement accessibles Cf. la référence[140] et, par exemple, le site de Bachem[166]. Je ne mentionnerai que les réactions secondaires susceptibles de poser problème pour notre objectif, quelques-unes rencontrées avec les acides Aspartique et Glutamique, la Cystéine et la Sérine, **Figure 59**.

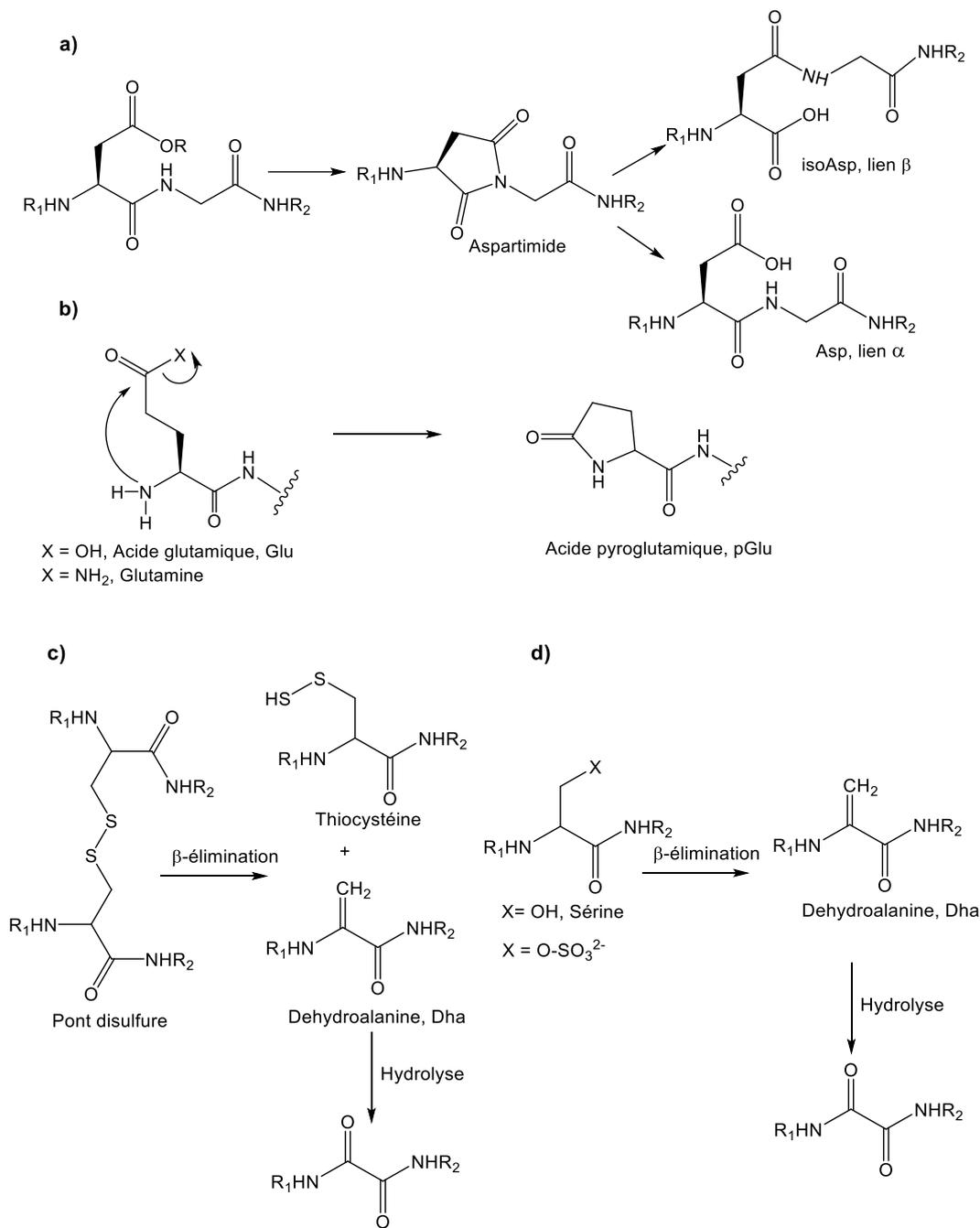


Figure 59. Réactions secondaires avec : a) l'acide Aspartique, avec formation d'aspartimide, b) l'acide Glutamique et la Glutamine, avec formation d'acide pyroglutamique, c) la Cystéine et d) la Sérine, avec formation de déhydroalanine, Dha.

* Un acide Aspartique dans une séquence peptidique peut conduire à la formation d'aspartimide, qui, par ouverture, donnera un mélange de peptide ciblé, lien α , et de peptide avec un lien β , correspondant à un acide β -aspartique (isoAsp), **Figure 59a**[167]–[169].

* De même, tout peptidiste sait que la réaction secondaire avec un acide Glutamique, ou une Glutamine, est la formation d'acide pyroglutamique (pGlu), **Figure 59b**. On ne prépare

jamais un peptide avec un Glu ou une Gln en *N*-terminal ! Le plus bel exemple d'une telle modification est illustré par les difficultés rencontrées par les équipes de R. Guillemin et A. Schally pour déterminer la séquence de la première hormone hypothalamique isolée, pGlu-His-Pro-NH₂[170]. Au milieu d'une séquence, les conditions de couplage de l'acide aminé suivant un Glu ou une Gln doivent/peuvent être maîtrisées[171]. Nous avons donc préféré insérer un acide Glutamique dans la séquence plutôt qu'un acide Aspartique, la réaction lien α , lien β avec l'acide Aspartique pouvant aussi se produire sur le peptide final.

* La fonction thiol de la Cystéine n'est pas stable en milieu aqueux, elle conduit facilement à des polymères par formation de ponts -S-S- intra- et/ou intermoléculaires, qui peuvent conduire à des agrégats insolubles. La fonction thiol en β , ou même après la formation d'un pont disulfure, peut, dans certaines conditions, conduire à la formation de déhydroalanine, Dha, **Figure 59c**[172].

* La Sérine (ou la Thréonine) modifiée avec un groupe partant, tel qu'un sulfate, peut, par β -élimination, conduire aussi à la déhydroalanine, Dha, **Figure 59d**. C'est certainement la réaction secondaire la plus difficile à éviter. Des conditions expérimentales ont même été développées pour optimiser la formation de Dha à partir de Cys et/ou Ser pour faire des modifications spécifiques de protéines[172].

L'oxydation de Cystéine (thiol, -SH degré d'oxydation du soufre : -1) en acide Cystéique (sulfonate, -SO₃H, degré d'oxydation du soufre : +4) par de l'acide performique a été décrite depuis les balbutiements du séquençage de protéines. C'est une stratégie utilisée pour déterminer la quantité exacte de Cystéine dans une protéine, les thiols et ponts disulfures étant oxydés en sulfonates. Il a été montré que, en suivant un protocole d'oxydation précis, les acides Aspartique et Glutamique restaient intacts dans le cas de la ribonucléase et que la Sérine et la Thréonine étaient peu affectées[173], [174]. Il a été établi plus récemment que la *O*-sulfonation de la fonction alcool des résidus Ser et/ou Thr en sulfates (-O-SO₃⁻) était une modification post-traductionnelle, stable, présente au cœur de certaines protéines[175].

Pour rappel, le pas de l'hélice α , d'une longueur de 0,54 nm, implique 3,6 acides aminés consécutifs par tour d'hélice, celui de l'hélice 3_{10} , d'une longueur de 0,6 nm, implique 3 acides aminés consécutifs par tour d'hélice et, dans une structure étendue en brin β , la

distance entre le C α _i de l'acide aminé *i* et le C α _(i+2) de l'acide aminé *i+2* est de 0,6 nm. Ainsi, en considérant que la longueur d'un disaccharide est de 0,7 à 0,8 nm, nous avons estimé qu'un térapeptide (4 acides aminés consécutifs), dans l'hypothèse où il serait structuré en hélice, correspondait plus ou moins, en termes de longueur, à une unité disaccharidique.

Forts de ces connaissances, nous avons choisi, comme premier peptide modèle et afin d'étudier les conditions expérimentales de synthèse, de purification et de caractérisation, le dimère du térapeptide Glu-Cys-Cys-Ser-NH₂, (ECCS)₂-NH₂, porteur des fonctions carboxylates (acide Glutamique), sulfonates (Cystéines oxydées en acides Cystéiques) et alcool (Sérine) caractéristiques des GAGs.

2. Essais préliminaires, recherche d'un protocole.

Les premiers essais ont été réalisés sur la séquence Biot(O₂)-Apa-(ECCS)₂-NH₂, **Figure 60**, obtenue par stratégie Boc sur synthétiseur automatique de peptides Applied Biosystems 433A, à partir d'une résine *p*-méthylbenzhydrylamine, MBHA, pour générer la fonction amide en C-terminal après coupure du support par l'acide fluorhydrique (HF). Les acides aminés utilisés ont été Boc-Glu(OBzl)-OH, Boc-Cys(*p*MeOBzl)-OH et Boc-Ser(OBzl)-OH, avec, comme méthode d'activation, le couple DCC/HOBt et, comme réactif de déprotection du Boc, l'acide trifluoroacétique (TFA). Nous avons introduit, en N-terminal de (Glu-Cys-Cys-Ser)₂-NH₂, un bras espaceur, apporté par le couplage de l'acide 5-aminopentanoïque, Boc-Apa-OH, puis une biotine sulfone, Biot(O₂), nécessaire pour les expériences biochimiques et biologiques ultérieures. Pour un suivi de synthèse par HPLC plus aisé, nous avons également synthétisé le peptide *p*-nitrobenzoyl-Apa-(ECCS)₃-NH₂, portant un chromophore, l'acide *p*-nitrobenzoïque, **Figure 60**.

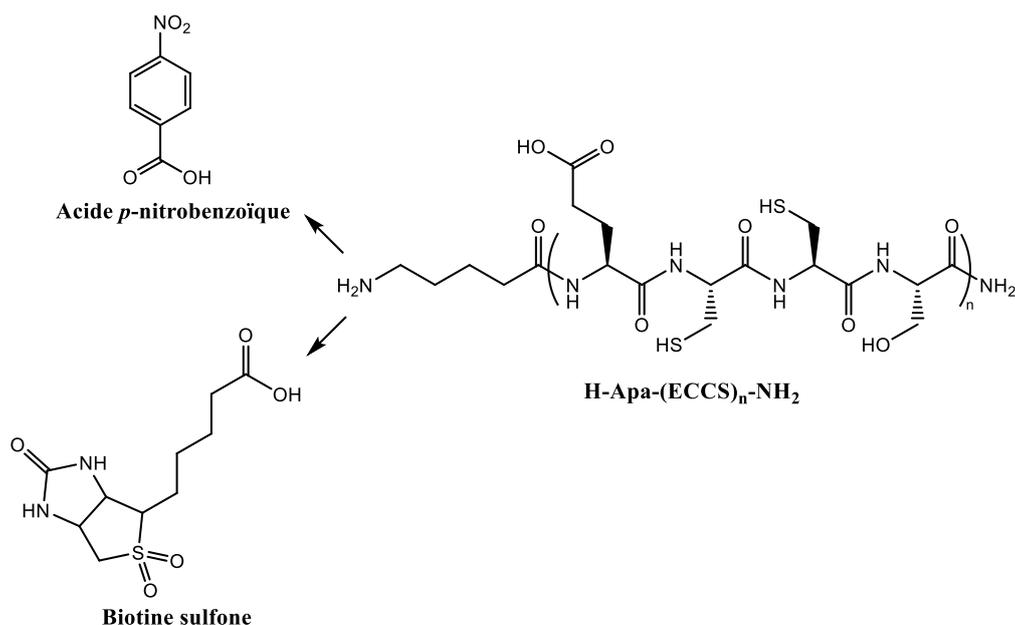
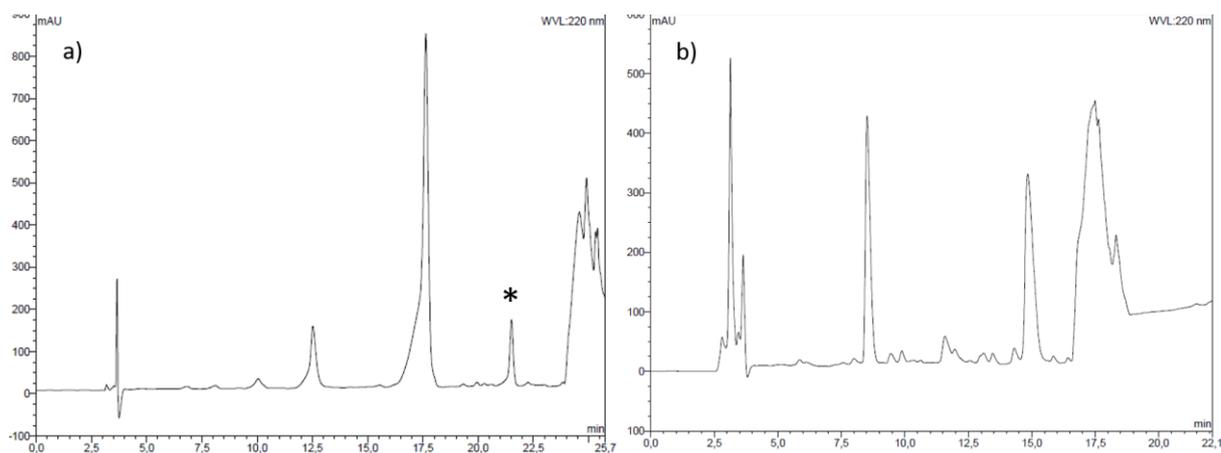


Figure 60. Séquence modèle utilisée pour les essais préliminaires.

Après coupure du support et retrait des groupements protecteurs des chaînes latérales par HF dans un mélange HF/diméthylsulfide/anisole/*p*-toluène-thiol, la solubilisation du produit brut dans des solvants compatibles avec les analyses HPLC en phase inverse (H₂O, acétonitrile...), s'est avérée difficile, nécessitant l'addition de DMSO et une sonication. Le *N*-terminal de Biot(O₂)-Apa-(ECCS)₂-NH₂ est hydrophobe, et, même avant oxydation, la région C-terminale (ECCS)₂-NH₂, est hydrophile. Le peptide final possède un caractère amphiphile susceptible de conduire à des agrégats, gel... Qui plus est, ce peptide possède quatre Cystéines dans sa séquence, susceptibles de former des ponts disulfures intra- et/ou inter-moléculaires, avant oxydation en sulfonate. Ce brut, analysé par HPLC en phase inverse (solvants : H₂O (0,1 % TFA) et CH₃CN (0,1 % TFA)), présentait 3 pics, **Figure 61a**, un des deux pics minoritaires étant le produit attendu, comme établi par spectrométrie de masse MALDI-TOF avec une matrice α -cyano-4-hydroxycinnamique, HCCA. Nous avons cependant engagé la réaction d'oxydation sur ce mélange brut, en émettant l'hypothèse que, s'il y avait des produits oxydés en disulfures, le milieu oxydant devrait les sur-oxyder en sulfonates, **Figure 62**, si tant est qu'on pouvait déplacer l'équilibre des « espèces avec ponts disulfures », moins ou peu solubles, vers le peptide linéaire avec des fonctions sulfonates et donc plus soluble.



**Figure 61. Profils HPLC de Biot(O₂)-Apa-(ECCS)₂-NH₂ brut : a) avant oxydation, b) après oxydation. Conditions : Colonne C18 Ace (AIT France), 5 µm, 300 Å, 4,6x250 mm - Gradient en 30 min de 5 à 20 % de CH₃CN (0,1 % TFA) – Débit : 1 mL/min - Lecture UV à 220 nm.
* produit attendu.**

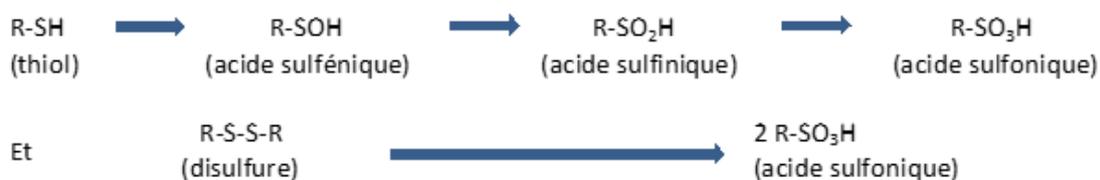


Figure 62. Oxydation par de l'acide performique de thiol ou disulfure en acide sulfonique

Les premières conditions d'oxydation étudiées ont été : 10 % H₂O₂ (30 % dans H₂O) dans de l'acide formique HCOOH durant 1 h à température ambiante, avant de faire réagir cet acide performique, préformé, sur le peptide, 4 h dans la glace, puis d'ajouter de l'eau et lyophiliser. Il s'est avéré qu'aucun des pics présents dans le brut après lyophilisation et analysés par HPLC, **Figure 61b**, et par spectrométrie de masse MALDI-TOF, ne présentait la masse (m/z) attendue. Mais ces premières analyses MALDI-TOF ont été réalisées avec une matrice HCCA, qui par la suite s'est avérée peu compatible avec ces molécules polyanioniques hydrophiles (*cf.* paragraphe 3.b ci-après).

Cette première étude concentre tous les points pour lesquels j'ai dû apporter des solutions, solutions que je résumerai ci-dessous :

- Nous avons supprimé la Sérine de la séquence, ayant de sérieux doutes sur sa stabilité totale ou partielle lors de l'oxydation.

- Nous avons introduit une Alanine dans la séquence de ce tétrapeptide mime d'un disaccharide, la L-Ala devant :
 - o apporter « un peu » d'hydrophobie, et ainsi minimiser la trop grande hydrophilie du peptide final, pour faciliter sa purification par HPLC en phase inverse,
 - o favoriser les structures en hélice du peptide. L-Ala a un bon indice d'hélicité et, dans notre séquence, les Alanines seront placées idéalement dans les positions $i, i+4$, etc...[176].
- Initialement, nous avons analysé les bruts de synthèse peptidique, puis choisi au final de conduire les réactions d'oxydation sans purification préalable.
- Nous avons fait varier les conditions d'oxydation du brut de synthèse peptidique, pour établir les meilleures conditions

3. Les premiers mimes peptidiques synthétisés.

Nous avons alors entrepris d'étudier les multimères des tétrapeptides $(EC'AC')_n$, $(EC'C'A)_n$ et $(EAC'C')_n$ qui posséderont un groupement carboxylate et, après oxydation des Cystéines, deux groupements sulfonates par bloc de tétrapeptide, tout comme l'unité disaccharidique du CS-E, et donc, *in fine*, n carboxylates et $2n$ sulfonates. L'acide Glutamique (E) mimera l'acide uronique et les acides Cystéiques (C') mimeront les fonctions sulfates du saccharide, **Figure 63**.

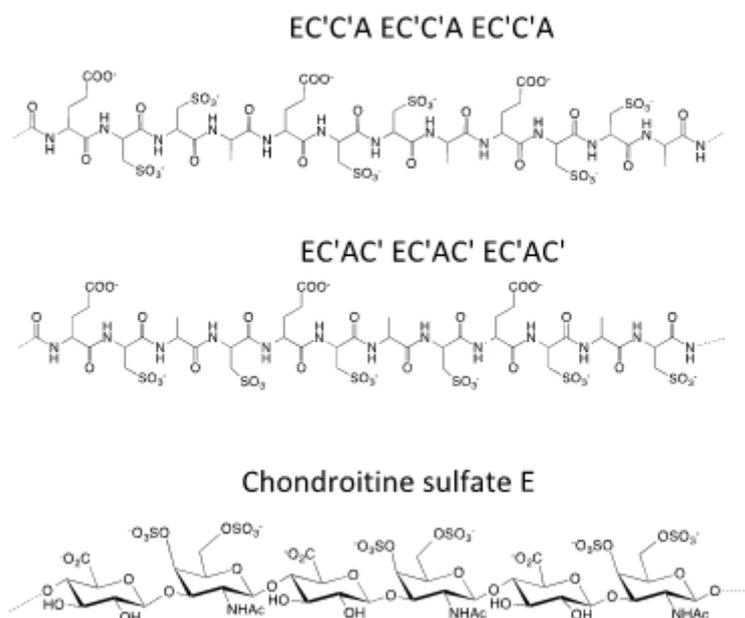


Figure 63. Tétrapeptides (EC'C'A) et (EC'AC'), mimes du bloc disaccharidique de CS-E.
C' symbolise l'acide Cystéique avec une fonction sulfonate en β .

Pour rappel nous avons choisi d'utiliser des répétitions d'unités tétrapeptidiques car, si elles se structurent en hélice α , chacune de ces unités aura une taille proche de celle d'une unité disaccharidique. Dans cette hypothèse, la position des acides aminés dans la séquence primaire aura un impact sur la répartition des fonctions carboxylate et sulfonate autour de l'hélice et donc sur leur présentation, **Figure 64**. Voilà pourquoi nous avons choisi d'étudier différentes variations de la séquence : $(EC'AC')_n$, $(EC'C'A)_n$ et $(EAC'C')_n$.

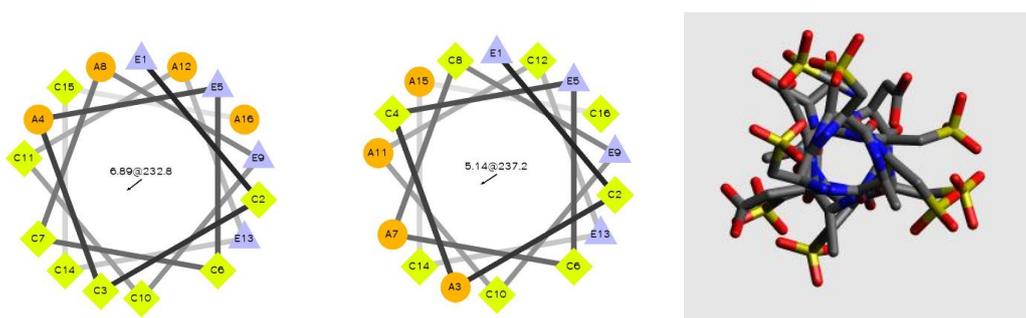


Figure 64. Représentations en roue des structures hélicoïdales théoriques de $(EC'C'A)_4$ (à gauche) et $(EC'AC')_4$ (au milieu et à droite), <http://rzlab.ucr.edu/>.

a) Protocole de synthèse des premiers mimes.

Les premières séquences synthétisées ont été : $(EC'C'A)_n$ et $(EC'AC')_n$. Les peptides listés **Tableau 4** ont été obtenus par synthèse automatique, sur un synthétiseur Applied

Biosystems 433A, par stratégie Fmoc avec une résine Fmoc-Rink-Amid MBHA (Fmoc-Rink-Amid-4-methylbenzhydrylamine) en utilisant Fmoc-Glu(OtBu)-OH, Fmoc-Ala-OH et Fmoc-Cys(Trt)-OH. Les couplages ont été réalisés avec le couple d'agents activateurs HBTU/DIEA (Diisopropyléthylamine) et la déprotection du Fmoc par action de pipéridine. En N-terminal, ont été ajoutés, manuellement, un bras espaceur, l'acide 6-aminohexanoïque (Ahx, commercial, un peu plus long que Apa et moins cher) après activation de Fmoc-Ahx-OH, puis une biotine sulfone, Biot(O₂), ou un groupement acétyl, Ac.

Motif	N-terminal	Code Personnel	Séquence
(EC'CA) ₃	Biotine(O ₂)-Ahx-	RM3-84A : Biot(O₂)-Ahx-(EC'CA)₃	Biot(O ₂)-Ahx- EC'CAEC'CAEC'CA-NH ₂
(EC'AC) ₃	Biotine(O ₂)-Ahx-	RM3-84B : Biot(O₂)-Ahx-(EC'AC)₃	Biot(O ₂)-Ahx- EC'ACE'ACE'AC'-NH ₂
	Acétyl-Ahx-	RM3-103A : Ac-Ahx-(EC'AC)₃	Ac-Ahx- EC'ACE'ACE'AC'-NH ₂
(EC'CA) ₄	Biotine(O ₂)-Ahx-	RM3-84C : Biot(O₂)-Ahx-(EC'CA)₄	Biot(O ₂)-Ahx- EC'CAEC'CAEC'CAEC'CA-NH ₂
	Acétyl-Ahx-	RM3-103B : Ac-Ahx-(EC'CA)₄	Ac-Ahx- EC'CAEC'CAEC'CAEC'CA-NH ₂
(EC'AC) ₄	Biotine(O ₂)-Ahx-	RM3-84D : Biot(O₂)-Ahx-(EC'AC)₄	Biot(O ₂)-Ahx- EC'ACE'ACE'ACE'ACE'AC'-NH ₂
	Acétyl-Ahx-	RM3-102A : Ac-Ahx-(EC'AC)₄	Ac-Ahx- EC'ACE'ACE'ACE'ACE'AC'-NH ₂
(EC'AC) ₅	Biotine(O ₂)-Ahx-	RM3-92A : Biot(O₂)-Ahx-(EC'AC)₅	Biot(O ₂)-Ahx- EC'ACE'ACE'ACE'ACE'ACE'AC'-NH ₂
	Acétyl-Ahx-	RM3-100A : Ac-Ahx-(EC'AC)₅	Ac-Ahx- EC'ACE'ACE'ACE'ACE'ACE'AC'-NH ₂
(EC'CC) ₄ *	Biotine(O ₂)-Ahx-	Biot(O₂)-Ahx-(EC'CC)₄	Biot(O ₂)-Ahx- EC'CC'EC'CC'EC'CC'EC'CC'-NH ₂
	Acétyl-Ahx-	Ac-Ahx-(EC'CC)₄	Ac-Ahx- EC'CC'EC'CC'EC'CC'EC'CC'-NH ₂

Tableau 4. Liste et nomenclature de la première série de sulfopeptides synthétisés. Synthèse sur Abi 433 par stratégie Fmoc et ajout des modifications en N-ter manuellement.

* Les bruts des synthèses de Biot(O₂)-Ahx-(EC'CC)₄-NH₂ et Ac-Ahx-(EC'CC)₄-NH₂ n'ont pas pu être purifiés par HPLC du fait de leur trop forte polarité, différentes stratégies ont été étudiées, en vain.

Les peptides ont été libérés de la résine et les groupements protecteurs des chaînes latérales ont été retirés par action d'acide trifluoroacétique (TFA) dans le mélange suivant : TFA/eau/éthanedithiol (EDT)/triisopropylsilane (TIS). Les peptides ont été précipités par addition d'éther diéthylique, préalablement refroidi à 0 °C. Ces bruts de synthèse obtenus par centrifugation, peu ou pas solubles dans des solvants compatibles avec les analyses HPLC en phase inverse (H₂O, acétonitrile...), ont été repris dans de l'eau avant d'être lyophilisés. Après lyophilisation, les peptides mis « en solution/suspension » dans la solution d'oxydation, souvent aidé par sonication, ont été oxydés.

b) Oxydation des premiers mimes.

La solution oxydante d'acide performique a été préparée en mélangeant, à température ambiante, de l'eau oxygénée avec de l'acide formique (respectivement H₂O₂ : 70 µL à 35 % dans l'eau et HCOOH : 1,93 mL, pureté 99 %). La solution résultante a été agitée 1h à température ambiante et refroidie (4 °C) avant utilisation. Les conditions d'oxydation à 4 °C ont été : peptide (10 mg) et solution d'acide performique (1 mL) avec une sonication si nécessaire pour la dispersion/solubilisation du peptide brut après lyophilisation. Après agitation pendant 15 h à 4 °C, la solution a été diluée avec de l'eau (5 volumes), et neutralisée avec de l'ammoniaque (pH = 7), puis lyophilisée.

La purification du mélange réactionnel a ensuite été réalisée par HPLC en phase inverse sur colonne C18, pour un exemple d'analyse HPLC avant et après oxydation/purification voir **Figure 65**.

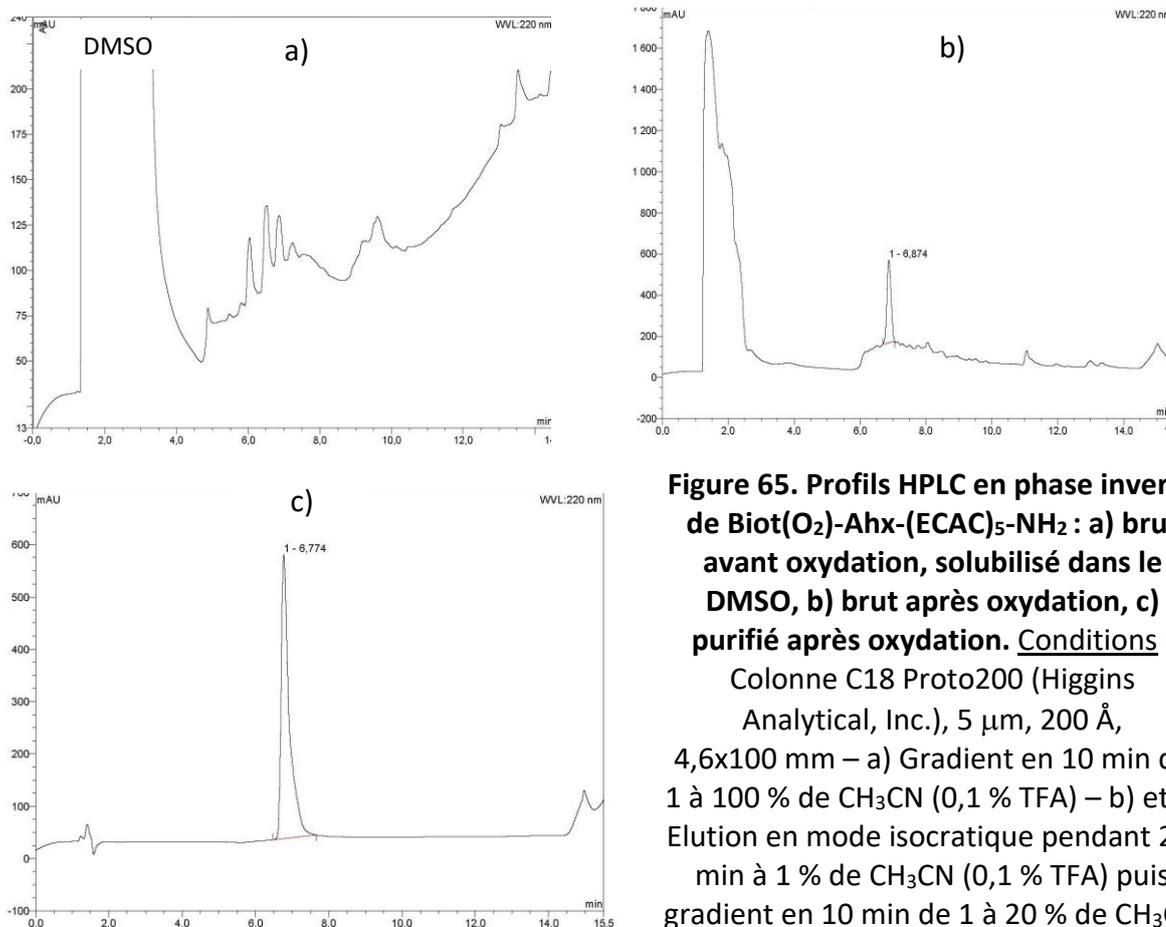


Figure 65. Profils HPLC en phase inverse de Biot(O₂)-Ahx-(ECAC)₅-NH₂ : a) brut avant oxydation, solubilisé dans le DMSO, b) brut après oxydation, c) purifié après oxydation. Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 5 µm, 200 Å, 4,6x100 mm – a) Gradient en 10 min de 1 à 100 % de CH₃CN (0,1 % TFA) – b) et c) Elution en mode isocratique pendant 2,5 min à 1 % de CH₃CN (0,1 % TFA) puis gradient en 10 min de 1 à 20 % de CH₃CN (0,1 % TFA) - Débit : 1 mL/min - lecture à 220 nm.

Les fractions pures ont été lyophilisées et analysées par spectrométrie de masse. Nous avons établi que les conditions de choix pour l'analyse par spectrométrie de masse MALDI-TOF (MALDI-TOF MS) de ces peptides porteurs de multiples fonctions sulfonates sont celles utilisées pour des dérivés de CS, c'est à dire une matrice neutre de *p*-nitroaniline en présence de pyrèneméthylguanidinium (pmg)[177]. Ohara *et al.* ont montré que l'addition de pmg (40 eq) à une solution de chondroïtine sulfate, avant ajout de matrice *p*-nitroaniline et analyse MALDI-TOF MS, conduisait à l'apparition de pics, en mode positif ou négatif, alors même qu'aucun signal n'était observable en l'absence de pmg. Ainsi, alors que les sulfonates de nos molécules rendaient impossible l'analyse directe de nos mimes par MALDI-TOF MS en utilisant la matrice HCCA classique des peptides, l'addition de pmg et l'utilisation de la matrice *p*-nitroaniline a permis de les caractériser. Les spectres obtenus, pour exemple *cf.* **Figure 66**, montrent un ensemble de pics dont le nombre est fonction du nombre de groupements sulfonates portés par la molécule. De plus, la différence de masse moléculaire entre chaque pic correspond à la masse d'un pmg, soit 273 g.mol⁻¹.

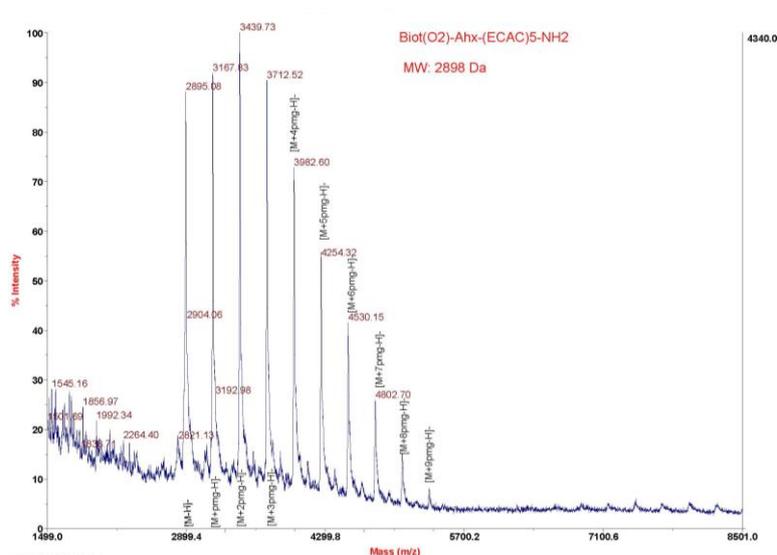


Figure 66. Profils MALDI-TOF de Biot(O₂)-Ahx-(EC'AC')₅-NH₂, avec matrice p-nitroaniline en présence de pyrèneméthylguanidinium (pmg), en mode linéaire négatif.

c) Etude des conditions de conservation/stockage des premiers mimes.

Ces peptides oxydés sont des produits hygroscopiques, ils doivent être pesés et aliquotés immédiatement en fin de lyophilisation. Sans chromophore dans leur séquence il est délicat de connaître la concentration finale pour les essais biologiques. J'ai toutefois pu estimer la concentration des peptides mimes en mesurant leur absorbance à 205 nm et en calculant leur coefficient d'extinction molaire théorique moyen à cette longueur d'onde (ϵ_{th})[178], [179]. Pour valider la méthode, j'ai ensuite comparé ces concentrations à celles obtenues en mesurant l'absorbance à 500 nm du complexe acide 4'-hydroxyazobenzène-2-carboxylique (HABA)/Avidine en absence puis en présence des analogues biotinylés de ces peptides[180], [181], **Tableau 5**, en utilisant, $A = \epsilon lc$, avec A = absorbance, c = concentration et l = longueur du trajet parcouru par la lumière dans la solution. Pour résumer le principe : l'ajout de molécules biotinylés, ayant une forte affinité pour l'avidine, va facilement libérer l'HABA du complexe et ainsi faire diminuer l'absorbance à 500 nm. Le calcul de la différence d'absorbance du complexe HABA/Avidine (dont le coefficient d'extinction molaire à 500 nm est de $34500 \text{ M}^{-1} \cdot \text{cm}^{-1}$) permet de déterminer la concentration de HABA libéré et donc la concentration de la solution de molécules biotinylés utilisée. Comme indiqué sur le tableau, les résultats obtenus par les deux approches sont cohérents à $\pm 5-10 \%$.

Peptide mime*	$A_{205\text{ nm}}$	$\epsilon_{\text{th}} (\text{M}^{-1} \text{cm}^{-1})$	Concentration 1 (μM)	Concentration 2 (μM)
<i>Biot-Ahx-(EC'CA)₃</i>	14,94	39510	378	363
<i>Biot-Ahx-(EC'AC')₃</i>	33,12	39510	838	922
<i>Biot-Ahx-(EC'CA)₄</i>	34,66	51310	674	667
<i>Biot-Ahx-(EC'AC')₄</i>	35,87	51310	699	742
<i>Biot-Ahx-(EC'AC')₅</i>	31,36	63110	497	455

* Tous les peptides possèdent en N-ter : Biotine(O₂)-Ahx- et en C-ter un amide : -NH₂

Tableau 5. Concentrations de quelques peptides biotinylés calculées en mesurant l'absorbance à 205 nm[178], [179] (Concentration 1) ou en utilisant l'absorbance à 500 nm du complexe HABA/Avidine[180], [181] (Concentration 2).

Surpris par l'évolution des profils HPLC de certains aliquots de peptides oxydés, nous avons réalisé une étude systématique, par suivi HPLC sur 5 jours, de la stabilité d'un de ces peptides, Biot(O₂)-Ahx-(EC'CA)₃-NH₂, dans différentes conditions de conservation : à température ambiante en solution dans l'eau et sous forme lyophilisée, et à -20 °C sous forme lyophilisée, les lyophilisats étant repris dans l'eau avant analyse HPLC.

Il s'est avéré qu'un deuxième pic pouvait apparaître dans certains échantillons ; or ce pic supplémentaire présentait la même masse que le peptide attendu, après analyse par spectrométrie de masse. Nous avons émis l'hypothèse que le gradient HPLC avec de l'acide trifluoroacétique conduisait à un peptide dont toutes les fonctions sulfonates n'étaient pas totalement protonées, si on considère les différences de pKa du TFA (0,3) et de l'acide sulfonique (-3). Suite à cette observation et pour ne pas risquer une dégradation du peptide liée à des conditions trop acides de stockage, il a été décidé, qu'après lyophilisation des sulfopeptides purifiés, ces mimes de GAGs seraient remis en solution dans de l'eau distillée, puis que le pH de la solution serait neutralisé (pH 7) avec une solution de NaOH à 1 mol.L⁻¹ et conservés ainsi à -20 °C. En opérant de la sorte nous avons établi que ces peptides étaient stables sur plusieurs mois, au moins.

4. Protocole de synthèse de pentamères et hexamères de mimes peptidiques.

Nous avons ensuite synthétisé des mimes de GAGs plus longs que ceux décrits ci-dessus : des pentamères et des hexamères des blocs tétrapeptidiques, **Tableau 6**. Ces peptides ont été préparés, sur résine, par Christophe Piesse, ingénieur de recherche de la plateforme de synthèse peptidique de l'Institut de Biologie Paris-Seine (IBPS) située sur le campus P. et M. Curie de Sorbonne Université, sur un synthétiseur Liberty Blue (CEM Corporation), sous micro-ondes, en stratégie Fmoc avec une résine Rink amide ProTide LL en utilisant Fmoc-Glu(OtBu)-OH, Fmoc-Ala-OH et Fmoc-Cys(Trt)-OH. Des doubles couplages systématiques ont été réalisés avec le couple d'agents activateurs DIC (*N,N'*-diisopropylcarbodiimide)/Oxyma Pur et la déprotection du Fmoc a été réalisée par action de pipérazine. En *N*-terminal, ont été ajoutés manuellement, comme pour les mimes plus courts, un bras espaceur Ahx puis une biotine sulfone, Biot(O₂) ou un groupement acétyl, Ac.

Motif	N-terminal	Code personnel	Séquence
(EC'CA) ₅	Biotine(O ₂)-Ahx	RM3-149A : Biot(O₂)-Ahx-(EC'CA)₅	Biot(O ₂)-Ahx- EC'CAEC'CAEC'CAEC'CAEC'CA-NH ₂
	Acétyl-Ahx	RM3-154A : Ac-Ahx-(EC'CA)₅	Ac-Ahx- EC'CAEC'CAEC'CAEC'CAEC'CA-NH ₂
(EAC'C') ₅	Biotine(O ₂)-Ahx	RM3-151A : Biot(O₂)-Ahx-(EAC'C')₅	Biot(O ₂)-Ahx- EAC'CEAC'CEAC'CEAC'CEAC'CE-NH ₂
	Acétyl-Ahx	RM3-155A : Ac-Ahx-(EAC'C')₅	Ac-Ahx- EAC'CEAC'CEAC'CEAC'CEAC'CE-NH ₂
(EC'CA) ₆	Biotine(O ₂)-Ahx	RM3-162 : Biot(O₂)-Ahx-(EC'CA)₆	Biot(O ₂)-Ahx- EC'CAEC'CAEC'CAEC'CAEC'CAEC'CA-NH ₂
	Acétyl-Ahx	RM3-190B : Ac-Ahx-(EC'CA)₆	Ac-Ahx- EC'CAEC'CAEC'CAEC'CAEC'CAEC'CA-NH ₂
(EC'AC') ₆	Biotine(O ₂)-Ahx	RM3-164 : Biot(O₂)-Ahx-(EC'AC')₆	Biot(O ₂)-Ahx- EC'ACE'ACE'ACE'ACE'ACE'ACE'AC'-NH ₂
	Acétyl-Ahx	RM3-190A : Ac-Ahx-(EC'AC')₆	Ac-Ahx- EC'ACE'ACE'ACE'ACE'ACE'ACE'AC'-NH ₂

Tableau 6. Liste et nomenclature de la deuxième série de sulfopeptides synthétisés.
Synthèse sur CEM micro-onde (plateforme de Jussieu) par stratégie Fmoc et ajout des modifications en N-ter manuellement.

Les peptides ont été coupés de leur support comme les peptides précédents, puis oxydés comme ci-dessus également. Le problème auquel nous avons dû faire face était celui de leur purification après oxydation. Le mélange des bruts, trop chargés en sels de formiate d'ammonium issus de la neutralisation par l'ammoniaque après oxydation, n'était pas ou très peu retenu sur les colonnes C18. Nous avons donc dû au préalable dessaler ces bruts par passage sur une colonne Séphadex G25 (colonne d'exclusion stérique permettant une séparation par taille) avant purification HPLC. La purification a pu alors être réalisée sur colonne C18, pour un exemple d'analyse HPLC avant et après purification voir **Figure 67**.

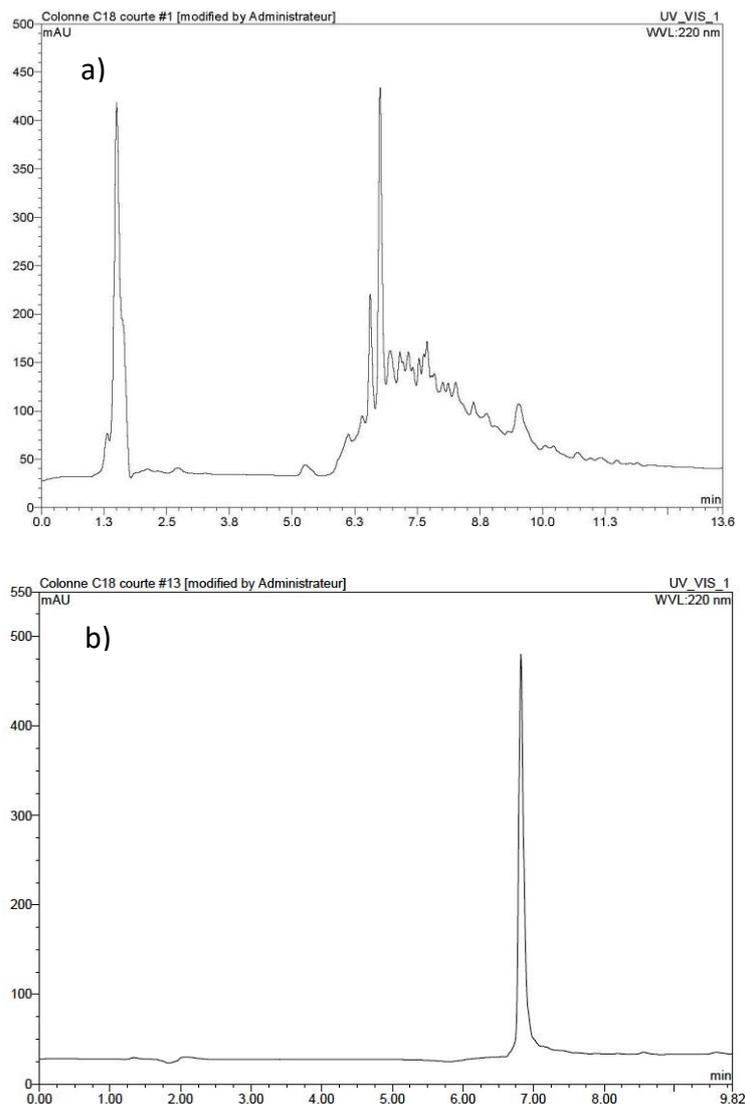


Figure 67. Profils HPLC en phase inverse de Biot(O₂)-Ahx-(EC'AC')₆-NH₂ après oxydation : a) brut , b) purifié. Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 3 μm, 200 Å, 4,6x100 mm - Elution en mode isocratique pendant 2,5 min à 1 % de CH₃CN (0,1 % TFA) puis avec un gradient en 10 min de 1 à 20 % de CH₃CN (0,1 % TFA) - Débit : 1 mL/min - Lecture à 220 nm.

Les analyses par spectrométrie de masse MALDI-TOF ont été réalisées comme précédemment avec une matrice *p*-nitroaniline et en présence de pyrène-méthylguanidinium, pour un exemple d'analyse voir **Figure 68**.

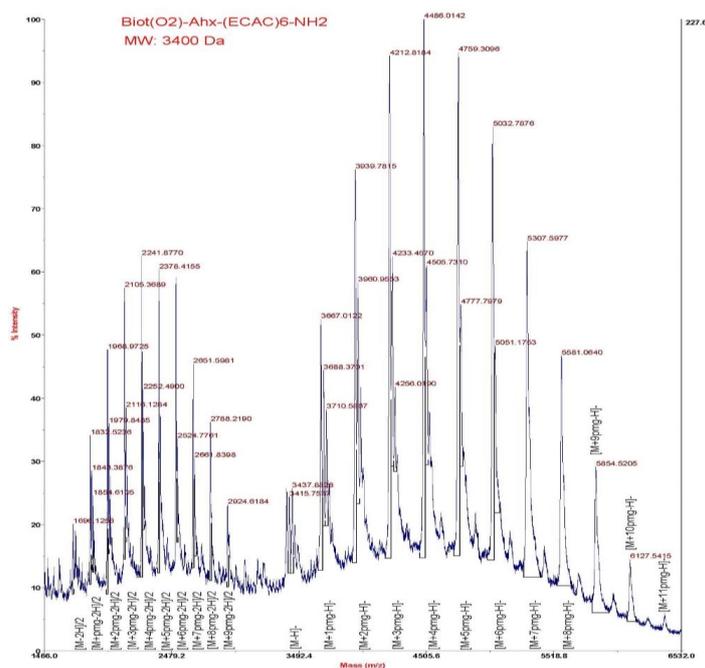


Figure 68. Profil MALDI-TOF de Biot(O₂)-Ahx-(EC'AC')₆-NH₂, avec matrice p-nitroaniline en présence de pyrèneméthylguanidinium (pmg), en mode linéaire négatif.

De la même manière que pour les premiers mimes synthétisés, j'ai pu estimer la concentration de ces nouveaux peptides mimes en mesurant leur absorbance à 205 nm et en calculant leur coefficient d'extinction molaire théorique moyen à cette longueur d'onde (ϵ_{th})[178], [179].

Ces sulfopeptides ont été utilisés par nos collègues du Collège de France avec qui nous collaborons, les résultats des études biochimiques et biologiques sont présentés dans le chapitre III.

5. Partie expérimentale

Le protocole général de synthèse, l'étape d'oxydation et les conditions de stockage des peptides ont été décrits dans les paragraphes précédents et sont repris dans le détail dans le « Supplementary Information » du manuscrit pour publication inséré dans le chapitre III. Les chromatogrammes HPLC et les spectres MALDI-TOF des peptides utilisés pour les études biochimiques et biologiques détaillées dans le chapitre III sont également fournis dans la partie « Supplementary information » du manuscrit pour publication.

Nous ne donnerons dans ce chapitre que les conditions de purification des peptides, leurs caractéristiques HPLC et spectrométrie de masse sous forme de tableau, **Tableau 7**, ainsi que les méthodes que nous avons utilisées pour déterminer leur concentration.

Conditions de purifications :

- des « premiers mimes peptidiques synthétisés » : Colonne C18 XBridge (Waters), 5 μm , 300 Å, 19x50 mm, de façon isocratique pendant 2,5 min à 0 % de CH_3CN (0,1 % TFA) puis avec un gradient en 10 min de 0 à 10 % de CH_3CN (0,1 % TFA), avec un débit de 10 mL/min et une lecture à 220 nm.
- des « pentamères et hexamères de mimes peptidiques » : Colonne C18 Ace (AIT France), 5 μm , 300 Å, 10x250 mm, de façon isocratique pendant 2,5 min à 0 % de CH_3CN (0,1 % TFA) puis avec un gradient en 30 min de 0 à 10 % de CH_3CN (0,1 % TFA), avec un débit de 6 mL/min et une lecture à 220 nm.

Caractéristiques HPLC et spectrométrie de masse

Code personnel	Gradient analytique	R_T (min)	m/z calculé [M-H]⁻	m/z mesuré [M-H]⁻
RM3-84A : Biot(O₂)- Ahx-(EC'C'A)₃	Isocratique à 0 % B pendant 2,5 min puis gradient de 0 à 20 % B en 10 min	7,1	1893	1889
RM3-84B : Biot(O₂)- Ahx-(EC'AC')₃	Isocratique à 0 % B pendant 2,5 min puis gradient de 0 à 20 % B en 10 min	7,3	1893	1890
RM3-103A : Ac- Ahx - (EC'AC')₃	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,9	1677	1678
RM3-84C : Biot(O₂)- Ahx-(EC'C'A)₄	Isocratique à 0 % B pendant 2,5 min puis gradient de 0 à 20 % B en 10 min	6,9	2395	2395
RM3-103B : Ac- Ahx - (EC'C'A)₄	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,6	2179	2178
RM3-84D : Biot(O₂)- Ahx-(EC'AC')₄	Isocratique à 0 % B pendant 2,5 min puis gradient de 0 à 20 % B en 10 min	7,1	2395	2395
RM3-102A : Ac- Ahx - (EC'AC')₄	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,6	2179	2179
RM3-149A : Biot(O₂)- Ahx-(EC'C'A)₅	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,9	2897	2895
RM3-154A : Ac- Ahx - (EC'C'A)₅	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,7	2681	2675
RM3-92A : Biot(O₂)- Ahx-(EC'AC')₅	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,8	2897	2895
RM3-100A : Ac- Ahx - (EC'AC')₅	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,4	2681	2680
RM3-151A : Biot(O₂)- Ahx-(EAC'C')₅	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	7,1	2897	2893
RM3-155A : Ac- Ahx - (EAC'C')₅	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,7	2681	2675
RM3-162 : Biot(O₂)- Ahx-(EC'C'A)₆	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,7	3399	3396
RM3-190B : Ac- Ahx - (EC'C'A)₆	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,9	3183	3190
RM3-164 : Biot(O₂)- Ahx-(EC'AC')₆	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,8	3399	3394
RM3-190A : Ac- Ahx - (EC'AC')₆	Isocratique à 1 % B pendant 2,5 min puis gradient de 1 à 20 % B en 10 min	6,9	3183	3190

Tableau 7. Caractéristiques HPLC et spectrométrie de masse des différents peptides synthésisés (Eluants HPLC : A = H₂O (0,1 % TFA) et B = CH₃CN (0,1 % TFA))

Méthodes utilisées pour déterminer la concentration des peptides

L'estimation de la concentration des sulfopeptides est réalisée par mesure au Nanodrop (Thermo Scientific) de l'absorbance des solutions à 205 nm, en utilisant la formule, $c = A/(\epsilon_{th} \times l)$, avec A = absorbance, c = concentration, et l = longueur du trajet parcouru par la lumière dans la solution. Le coefficient d'extinction molaire théorique moyen à 205 nm (ϵ_{th}) de chaque sulfopeptide est calculé en additionnant, pour chaque résidu, les valeurs de coefficients d'extinction molaire suivantes[178], [179] :

- Liaison peptidique (amide) : $2780 \text{ M}^{-1} \cdot \text{cm}^{-1}$
- Alanine : $80 \text{ M}^{-1} \cdot \text{cm}^{-1}$
- Acide Glutamique : $200 \text{ M}^{-1} \cdot \text{cm}^{-1}$
- Acide Cystéique : $200 \text{ M}^{-1} \cdot \text{cm}^{-1}$

Pour obtenir la valeur du coefficient d'extinction molaire à 205 nm de la biotine sulfone (ϵ), une mesure de l'absorbance à 205 nm est réalisée sur une solution aqueuse de biotine sulfone préparée, de manière la plus précise possible, à $800 \mu\text{M}$. L'utilisation de la formule $\epsilon = A/(c \times l)$ permet d'obtenir la valeur $\epsilon = 1330 \text{ M}^{-1} \cdot \text{cm}^{-1}$.

Un deuxième protocole permettant d'estimer, par spectrophotométrie (colorimétrie), la concentration d'une solution contenant des molécules biotinylées, a aussi été développé[180], [181]. Il repose sur le fait que le colorant HABA (acide 4'-hydroxyazobenzène-2-carboxylique) se lie à l'avidine pour produire un complexe coloré jaune-orange qui absorbe à 500 nm. Une solution de HABA est préparée en solubilisant 2,7 mg de HABA (Interchim - Ref 05361E) dans 1,114 mL d'eau distillée et 11,16 μL de NaOH 1N. À 2,9 mg d'avidine (Avidine de blanc d'œuf - pureté > 98 % - Sigma-Aldrich - Ref A9275) sont ajoutés 58 μL de cette solution puis le mélange est complété à 645 μL par du tampon PBS (Tampon Phosphate Salin, comprimé dissout dans 200 mL d'eau distillée, pH 7,4 à 25 °C - Sigma-Aldrich - Ref P4417). L'absorbance DO1 à 500 nm du complexe HABA/Avidine ainsi formé, mesurée au Nanodrop, est de 0,73. La solution de sulfopeptide biotinylé (10 μL) est alors ajoutée à 90 μL de solution de complexe HABA/Avidine et une nouvelle absorbance DO2 à 500 nm est mesurée. La concentration (c) de la solution de sulfopeptide utilisée est déterminée en utilisant la formule suivante : $c = ((0,9 \times \text{DO1}) - \text{DO2})/(34500 \times l)$, avec 0,9 = facteur de dilution du complexe

HABA/Avidine avec la solution de sulfopeptide biotinylé et l = longueur du trajet parcouru par la lumière dans la solution.

CHAPITRE III : Etudes biochimiques et biologiques des mimes peptidiques de GAGs

Remarque liminaire : Ces analyses biochimiques et biologiques ont été réalisées par Damien Testa, lors de sa thèse au Collège de France, sous la direction du Dr Ariel Di Nardo. Ces travaux sont décrits dans son manuscrit de thèse intitulé : « *Contrôler la plasticité du cortex cérébral adulte à travers l'action non-autonome de l'homéoprotéine Otx2* »[100].

Damien a fait le choix de présenter nos travaux :

- 1) au travers d'une courte introduction sur la partie biologique, qui, par ailleurs, est au cœur de son manuscrit,
- 2) puis d'insérer le manuscrit de nos travaux à soumettre pour publication, dont nous sommes les deux premiers co-auteurs avec « contributions à parts égales »,
- 3) pour tirer les conclusions majeures de cette étude.

Je ferai de même, si ce n'est que je présenterai plus dans le détail les méthodes d'études biochimiques et biologiques que Damien a utilisées. Sans y avoir participé, j'ai beaucoup échangé et discuté avec lui pour comprendre, je vais donc tenter, avec ma compréhension du domaine, de faire passer les messages essentiels.

A. Introduction

Nous possédons des peptides de formules $\text{Biot}(\text{O}_2)/\text{Ac-Ahx}-(\text{EC}'\text{CA}')_n\text{-NH}_2$ et $\text{Biot}(\text{O}_2)/\text{Ac-Ahx}-(\text{EC}'\text{AC}')_n\text{-NH}_2$ avec $n = 3$ à 6 . À présent, nous souhaiterions savoir :

- 1) si ces peptides sont des mimes de GAGs, tout comme l'hexaCSE, analogue saccharidique de chondroïtine sulfate-E[68]
- 2) s'ils se lient à la protéine Otx2
- 3) s'ils peuvent réguler *in vivo* les fonctions de Otx2 dans le cortex visuel ?

Pour rappel : Les molécules, CS-E, hexaCSE et mimes de GAGs, sont représentées

Figure 69.

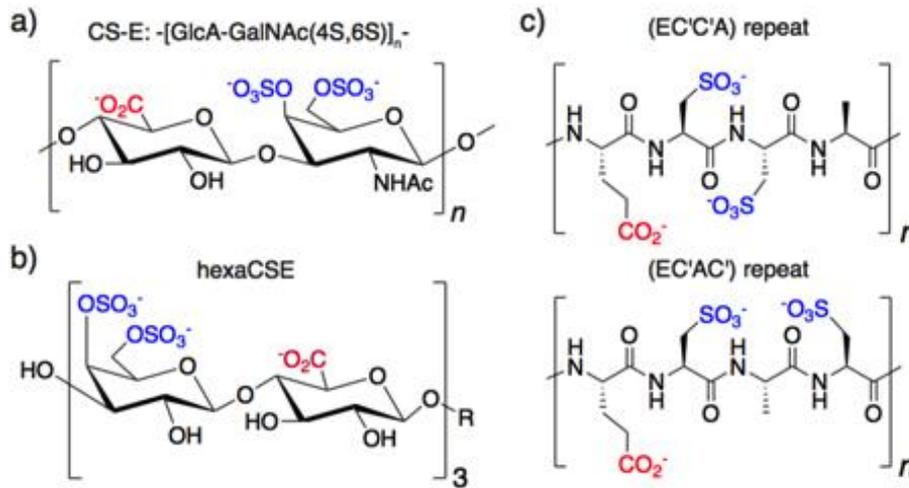


Figure 69. Structure chimique des unités répétitives. a) Chondroïtine-4,6-sulfate, CS-E, b) plateforme de poly-lactose, hexaCSE, c) mimes peptidiques de GAG, EC'C'A et EC'AC' (C' = acide Cystéique).

Pour rappel : La protéine Otx2 présente un site de liaison pour les GAGs, en particulier les CS-D et CS-E, de séquence **RK-QRRERTTFTRAQL**, **Figure 70**, à partir de laquelle un peptide, le **RK-peptide**, a été synthétisé. Cette séquence se situe du côté *N*-ter de l'homéodomaine de Otx2. L'étape de liaison de Otx2 aux CS-D et CS-E de la matrice extracellulaire est nécessaire à l'internalisation de Otx2 dans les neurones GABAergiques à parvalbumine (PV) du cortex visuel. Cette internalisation induit une maturation du réseau périneuronal (PNN) de ces cellules et une modification des positions de sulfatation des CS du PNN.

MMSYLKQPPY¹⁰ AVNGLSLTTS²⁰ GMDLLHPSVG³⁰ YPATP⁴⁰ **RKQRR**⁴¹ **ERTTFTRAQL**⁵⁰ DVLEALFAKT⁵⁰ RY⁵⁰
 PDIFMREE⁷⁰ VALKINLPES⁸⁰ **RVQVWFKNRR**⁸⁰ AKCRQ⁹⁰ QQQQQ¹⁰⁰ QNGGQNKVVRP¹¹⁰ AKKKTSPARE¹²⁰
 VSSESGTSGQ¹³⁰ FPPSSTSV¹⁴⁰ TIASSAPVS¹⁵⁰ WSPASISPL¹⁶⁰ SDPLSTSSC¹⁷⁰ MQRSYPMITY¹⁸⁰ QAS¹⁸⁰
 GYSQGYA¹⁹⁰ GSTSYFGGD²⁰⁰ CGSYLTPMH²¹⁰ QLPGPGATLS²²⁰ PMGTNAVTS²³⁰ LNQPASLST²⁴⁰ Q²⁴⁰
 GYGASSLGF²⁵⁰ NSTTDCLDYK²⁶⁰ DQTASWKLNF²⁷⁰ NADCLDYKDQ²⁸⁰ TSSWKFQVL²⁸⁹



Figure 70. Séquence de Otx2 chez l'Homme[98]

La double mutation du doublet basique **RK** par deux Alanines **AA**, dans la séquence de reconnaissance des GAGs d'Otx2, conduit à une perte d'affinité pour les CS-D, démontrant l'importance de ce doublet RK pour la reconnaissance des CS doublement sulfatés. Enfin, un peptide dit « scrambled », possédant les mêmes acides aminés que le **RK-peptide** mais aléatoirement placés dans la séquence primaire, montre une moins bonne affinité (facteur 5)

pour les CS-D que le **RK**-peptide[86], ce qui atteste de l'importance de la distribution des acides aminés dans le domaine de reconnaissance des GAGs d'Otx2.

B. Les expériences

Les études mises en jeu impliquent :

- 1) Des expériences de « dot blot » visant à comparer les affinités relatives des mimes biotinylés $\text{Biot}(\text{O}_2)\text{-Ahx-(EC'CA')}_n\text{-NH}_2$ et $\text{Biot}(\text{O}_2)\text{-Ahx-(EC'AC')}_n\text{-NH}_2$ vis-à-vis de la protéine Otx2 ou du **RK**-peptide, avec une comparaison avec l'hexaCSE biotinylé.
- 2) Des expériences de « pull-down » de lysats de cortex visuel de souris adultes avec ces mimes peptidiques biotinylés.
- 3) Des expériences d'infusion de ces mimes de GAGs dans le cortex visuel de souris adultes avec l'analyse de l'évolution de l'environnement des cellules PV.

C. Les résultats

1. Expérience de dot blots.

Les mimes biotinylés $\text{Biot}(\text{O}_2)\text{-Ahx-(EC'CA')}_n\text{-NH}_2$ et $\text{Biot}(\text{O}_2)\text{-Ahx-(EC'AC')}_n\text{-NH}_2$ ou l'hexaCSE biotinylé sont déposés sur une membrane de nitrocellulose, seuls ou après incubation avec la protéine Otx2 ou les peptides **RK**-peptide, **AA**-peptide ou le contrôle SC-peptide (scrambled peptide). La révélation se fait ensuite *via* la biotine des mimes, par incubation avec de la streptavidine couplée à la peroxydase HRP (SA-HRP), suivie d'une réaction de chimioluminescence. Les mimes biotinylés n'étant pas retenus sur la membrane après lavages, ne sont révélés que les complexes formés par l'interaction des mimes avec un partenaire retenu par la membrane (Otx2, **RK**-peptide, **AA**-peptide ou SC-peptide). Les membranes sont ensuite numérisées et quantifiées par densitométrie, **Figure 71**.

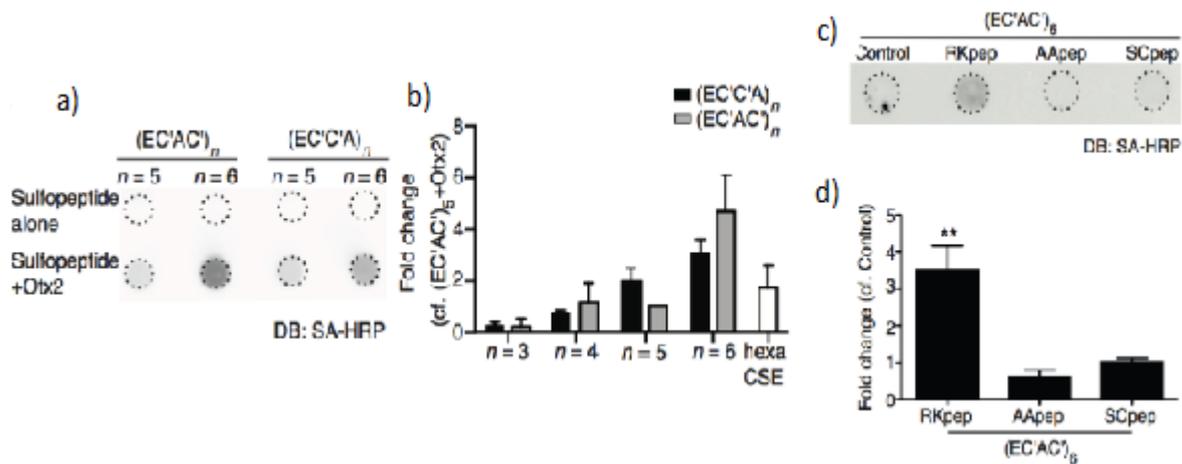


Figure 71. Expériences de dot-blot et quantification de la chimioluminescence. a) et b) Exemple de dot blot (DB) de sulfopeptides biotinylés, seuls ou incubés avec la protéine Otx2 et quantification de la chimioluminescence. c) et d) DB de Biot(O₂)-Ahx-(EC'AC')₆-NH₂ seul (contrôle) ou incubé avec le RK-peptide ou les peptides contrôles (AA-peptide, SC-peptide) et quantification de la chimioluminescence.

Conclusions :

** Les séquences de tétrapeptides de formule Biot(O₂)-Ahx-(EC'CA')_n-NH₂ et Biot(O₂)-Ahx-(EC'AC')_n-NH₂, avec n = 4 ou 5, se lient à la protéine Otx2 aussi bien que l'hexaCSE. Les séquences plus longues, avec n = 6, montrent une liaison 2 à 3 fois meilleure que celle de l'hexaCSE pour Otx2.*

** La liaison des sulfopeptides Biot(O₂)-Ahx-(EC'CA')_n-NH₂ et Biot(O₂)-Ahx-(EC'AC')_n-NH₂ à la protéine Otx2 se ferait via le site de fixation aux GAGs de la protéine Otx2, site qui correspond au RK-peptide.*

** Comme le CS-D (cf. introduction de ce chapitre), les sulfopeptides étudiés ne se lient que faiblement aux peptides contrôle AA-peptide et SC-peptide.*

2. Expérience de pull-down.

Remarque : La concentration de Otx2 dans le cortex visuel est extrêmement faible. La protéine Sémaphorine-3A (Sema-3A) est une protéine présente dans les PNNs, à une concentration plus élevée, et, comme Otx2, Sema-3A joue un rôle dans la plasticité des PNNs, **Figure 72**[100]. Elle se lie aux CS *via* une séquence de liaison similaire au RK-peptide, **Figure 73a**. Ainsi, nous avons utilisé nos sulfopeptides biotinylés pour isoler/extraire Sema-3A d'un lysat de cortex visuel de souris adultes, à défaut de pouvoir le faire avec Otx2.

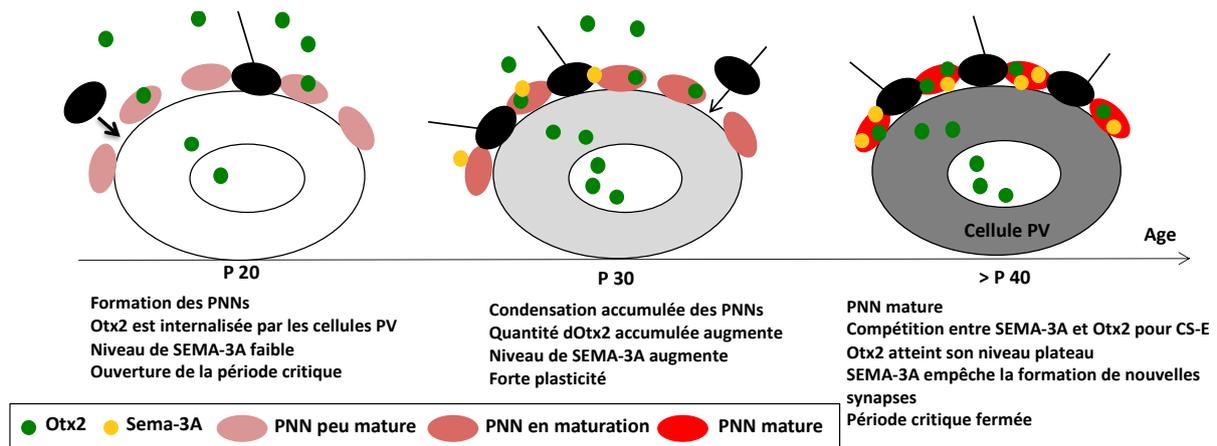


Figure 72. Modèle de compétition dans les PNNs entre Otx2 et Sema-3A. L'accumulation d'Otx2 dans les cellules PV engendre une condensation en PNNs. Les modifications synaptiques sont possibles pendant la plasticité. Après cette fenêtre de plasticité, Otx2 atteint un plateau et Sema-3A empêche la formation de nouvelles synapses (adaptée de ref[100]).

Après sacrifice de souris adultes, dissection puis lyse de leur cortex visuel, les échantillons sont centrifugés et le surnageant est incubé avec les sulfopeptides biotinylés. Pour les tests de compétition, les mimes de GAG sont pré-incubés avec le **RK**-peptide, le **AA**-peptide ou le **SC**-peptide. Puis ces solutions sont incubées avec des billes Dynabeads recouvertes de streptavidine. Les billes chargées sont alors lavées puis chauffées dans du tampon de Laemmli pour décrocher/dénaturer les protéines qui sont alors analysées par Western Blot (WB) : les protéines sont séparées en fonction de leur taille/poids moléculaire sur gel d'électrophorèse puis transférées sur des membranes PVDF (poly(fluorure de vinylidène)). Ces membranes sont ensuite incubées avec un anticorps primaire de lapin anti-Sema-3A, puis lavées et incubées avec un anticorps secondaire anti-lapin lié à la peroxydase HRP. Les membranes sont numérisées et quantifiées par densitométrie, **Figure 73 b-e**.

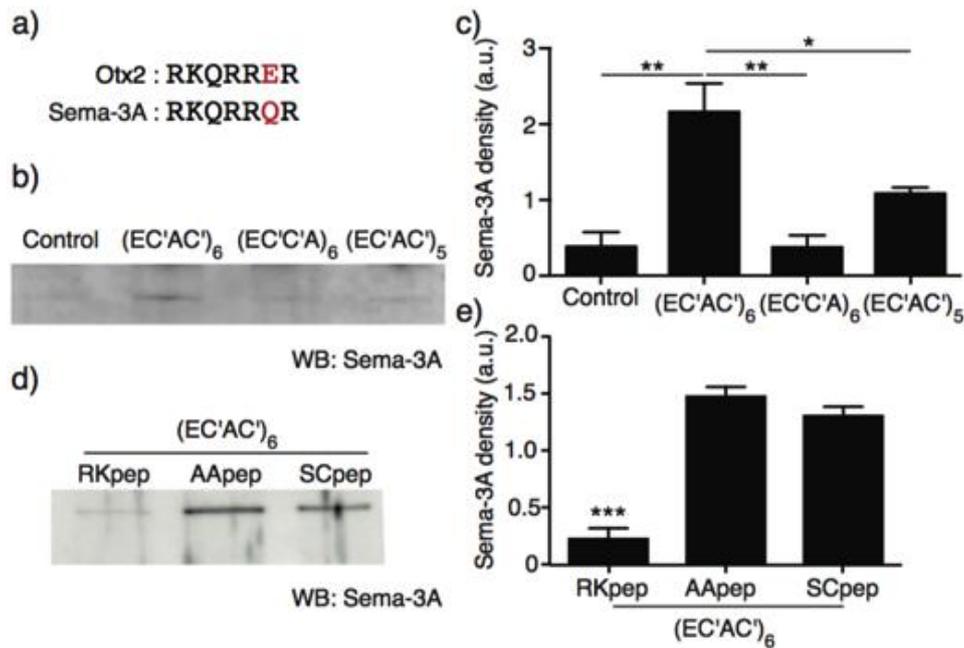


Figure 73. Expériences de pull-down de lysats de cortex visuel de souris adulte avec les mimes peptidiques de GAG. a) Comparaison des motifs de liaison aux GAG de Otx2 et Sema-3A. b) et c) Western blot (WB) pour Sema-3A, après un pull-down avec les sulfopeptides biotinylés retenus sur des billes de streptavidine (le contrôle correspond à des billes seules) et la quantification correspondante. d) et e) WB pour Sema-3A après un pull-down avec Biot(O₂)-Ahx-(EC'AC')₆-NH₂ pré-incubé avec les peptides **RK**-peptide, **AA**-peptide ou **SC**-peptide et la quantification correspondante.

Conclusions :

- * *Biot(O₂)-Ahx-(EC'AC')₅-NH₂ et Biot(O₂)-Ahx-(EC'AC')₆-NH₂ interagissent avec Sema-3A.*
- * *Sema-3A reconnaît préférentiellement le motif Biot(O₂)-Ahx-(EC'AC')₆-NH₂ par rapport au motif Biot(O₂)-Ahx-(EC'C'A)₆-NH₂.*
- * *Le RK-peptide affecte spécifiquement l'interaction de Sema-3A avec Biot(O₂)-Ahx-(EC'AC')₆-NH₂ dans ces lysats.*

3. Expérience d'infusion de sulfopeptides *in vivo*.

Pour rappel : Comme écrit précédemment dans le chapitre I, rouvrir la plasticité du cortex cérébral visuel adulte en ciblant Otx2 peut se faire en modifiant son interaction avec la matrice extracellulaire. Ceci peut être réalisé par les deux approches représentées **Figure 74**. On peut bloquer l'interaction de Otx2 avec les CS-D et CS-E de la matrice par infusion du **RK**-peptide. Une telle infusion a permis de guérir l'amblyopie chez des souris adultes[86]. On peut

aussi séquestrer la protéine Otx2 par infusion d'un mime de GAGs tel que l'analogue synthétique du CS-E (hexaCSE)[68].

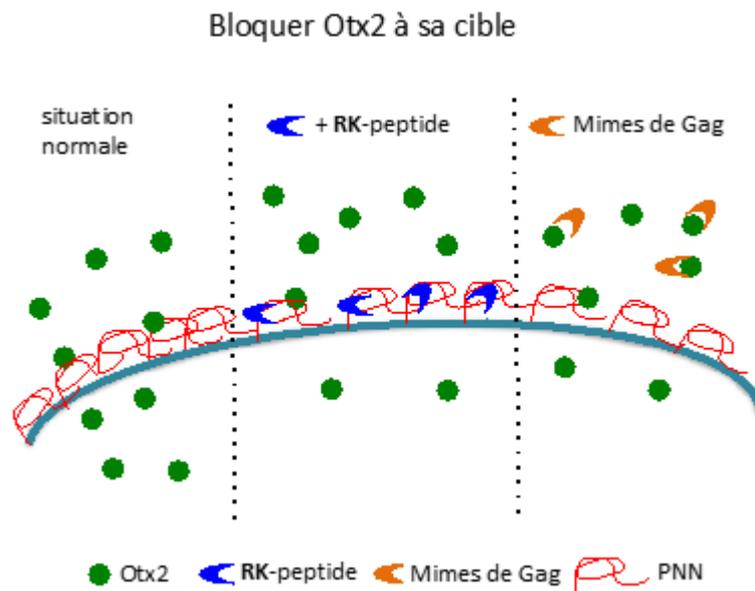


Figure 74. Méthodes de réouverture de la plasticité en interférant avec Otx2. Modification de l'interaction d'Otx2 avec les PNNs : soit en utilisant la séquence de reconnaissance aux GAGs (RK-peptide), soit en infusant un mime de GAGs tel qu'un CS-E de synthèse (hexaCSE).

Il existe une boucle de régulation positive entre l'accumulation d'Otx2 dans les cellules PV et l'assemblage des PNNs, **Figure 75**. Les PNNs attirent la protéine Otx2 vers les cellules PV et l'internalisation d'Otx2 permet d'augmenter l'expression du PNN. À l'inverse, si Otx2 est séquestré à l'extérieur des cellules PV, nous devrions observer une diminution de l'assemblage des PNNs et un retour à un état pré-mature des cellules PV, permettant une réouverture de la plasticité neuronale.

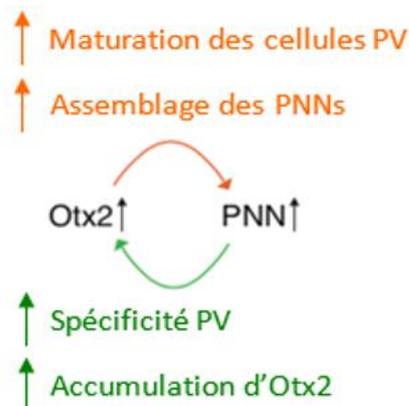


Figure 75. Boucle de rétroaction positive entre Otx2 et les PNNs.

Des souris C57BL/6J de trois mois sont perfusées dans le cortex pendant 7 jours avec diverses concentrations de sulfopeptides (4 pM, 400 pM ou 4 μ M), en utilisant des pompes micro-osmotiques (0,5 μ L/h). Les animaux sacrifiés sont ensuite perfusés avec du PBS et du paraformaldéhyde à 4 %. Des coupes de cortex (20 μ m) de leurs cerveaux sont alors incubées, durant toute une nuit, avec du WFA-FITC (Agglutinine de Wisteria Floribunda couplée à la fluorescéine- qui marque les PNNs) et un anticorps primaire de lapin anti-parvalbumine, suivi par l'incubation avec un anticorps secondaire anti-lapin Alexa Fluor-546 pendant 1h. Les images sont acquises avec un microscope et quantifiées, **Figure 76**.

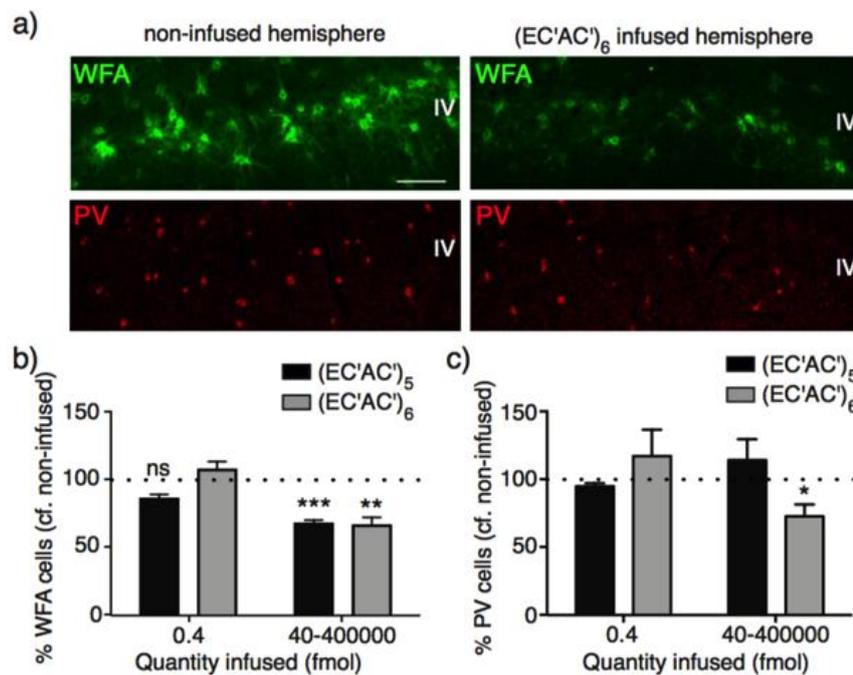


Figure 76. Infusion, *in vivo*, des sulfopeptides Biot(O₂)-Ahx-(EC'AC')_n-NH₂, n = 5 et n = 6.

a) Images représentatives de la coloration de WFA-FITC, qui marque les PNNs, et de la parvalbumine (PV), dans la couche IV de cortex visuel primaire de souris adulte perfusée ou non avec 40 fmol de Biot(O₂)-Ahx-(EC'AC')₆-NH₂. b) Quantification du nombre de cellules WFA+. c) Quantification du nombre de cellules PV+.

L'infusion de Biot(O₂)-Ahx-(EC'AC')₅-NH₂ et Biot(O₂)-Ahx-(EC'AC')₆-NH₂ (> 40 fmol) entraîne une diminution de l'assemblage du PNN mais seule l'infusion de Biot(O₂)-Ahx-(EC'AC')₆-NH₂ (> 40 fmol) conduit à une réduction significative de l'expression de PV, témoin de l'état de maturation des cellules. Cette réduction peut apparaître modeste (≤ 25 %), mais elle devrait être suffisante, comme elle s'est avérée suffisante pour rouvrir la plasticité cérébrale dans le cas de l'hexaCSE[68].

Conclusion :

- * Le motif (EC'AC') interfère *in vivo* avec le transfert d'Otx2.

D. Conclusion générale de ce chapitre

Nous retiendrons de ce travail que l'hypothèse de départ est vérifiée :

- * Il est possible de partir d'une séquence peptidique plus facile d'accès, quand les difficultés sont maîtrisées, pour préparer des mimes de GAGs capables d'interagir avec la protéine Otx2.
- * Il est possible d'utiliser Biot(O₂)-Ahx-(EC'AC')₆-NH₂ *in vivo* pour moduler l'activité du facteur de transcription Otx2.

Il doit être possible de trouver une sélectivité permettant aux mimes peptidiques de GAGs d'inhiber préférentiellement une protéine liant les GAGs par rapport à une autre. Otx2 et Sema-3A ont des domaines de liaison aux GAGs quasi identiques RK-QRRER vs RK-QRRQR, un acide Glutamique pour Otx2 et une Glutamine pour Sema-3A et, néanmoins, Sema-3A reconnaît préférentiellement le motif (EC'AC')₆ par rapport au motif (EC'C'A)₆.

Forts de cette information et sans *a priori* dans un premier temps sur le mode de reconnaissance entre les deux partenaires, nous pouvions proposer des variations autour du résidu hydrophobe dans la séquence du térapeptide, (EC'XC')_n, pour identifier des mimes de GAGs avec une meilleure affinité et une sélectivité de reconnaissance pour un facteur de transcription donné.

Cependant, à ce stade, nous avons décidé d'obtenir des informations structurales sur les domaines/mode de reconnaissance par les sulfopeptides de deux facteurs de transcription étudiés depuis des années par nos collègues du Collège de France : Otx2 et Engrailed 2.

Cette étude fera l'objet du chapitre suivant. Elle implique différentes étapes allant de la production/purification des protéines jusqu'aux études d'interaction par RMN et SPR de type Biacore.

Cysteic acid-rich peptides as bioactive glycosaminoglycan mimics

Rodrigue Marquant^{§,‡}, Damien Testa^{§,†}, Solange Lavielle[‡], Alain Prochiantz[†], Ariel A. Di Nardo^{*,†}, Jean-Maurice Mallet^{*,‡}

[†]Centre for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS UMR 7241 / INSERM U1050, PSL University, Labex Memolife, 11 place Marcelin Berthelot, 75005 Paris, France.

[‡]Laboratoire des BioMolécules (LBM), Département de Chimie, École Normale Supérieure, PSL University, Sorbonne Université, CNRS UMR 7203, 24 rue Lhomond, 75005 Paris, France.
12/3/2018

[§]These authors contributed equally.

ABSTRACT: *Glycosaminoglycans (GAGs) form important extracellular interactions that impact brain physiology. However, there is a lack of synthetic GAG mimics with high biological specificity to probe GAG physiological functions. Here, we report the synthesis of peptides with oxidized cysteines leading to sulfonated analogues of chondroitin sulfates. These GAG mimics interact with Otx2 homeoprotein transcription factor and Sema-3A specifically through their GAG-binding motifs. These proteins are key actors of visual cortex critical period plasticity that function through extracellular GAG interactions. Indeed, GAG mimic infusion in adult mouse visual cortex simultaneously reduced perineuronal net assembly and parvalbumin expression, which are hallmarks of reopened cortical plasticity.*

Glycosaminoglycans (GAGs) are sulfated polysaccharides found on cell surfaces and in the extracellular matrix (ECM) where they interact with a large number of proteins (cytokines, growth factors, serpins, etc). The biologically important chondroitin sulfate (CS) proteoglycan consists of a protein backbone to which is linked chains

containing hundreds of disaccharide repeats. A fascinating characteristic of GAGs is that their polysaccharide units are modified through epimerisation, *N*- and *O*-sulfation, and deacetylation, and thus can be finely tuned for interactions. Consequently, they present non-regular sequences with specific affinity.

In postnatal brain development, specific sulfation patterns are crucial for the timing of critical periods, which are short windows in time during which a given neural circuit is highly plastic.¹ These periods are necessary for proper environmental adaptation and arise from the maturation of cortical inhibitory GABAergic parvalbumin neurons (PV cells). This maturation is driven by non-cell autonomous activity of the Otx2 homeoprotein transcription factor: Otx2 is synthesized outside the cerebral cortex, transported in the extracellular milieu, and internalized specifically by PV cells.^{2,3} This internalization specificity is mediated by disulfated CS (type D (CS-D) or type E (CS-E, Figure 1a) contained within a specialized ECM called perineuronal net (PNN) enwrapping these cells.^{4,5} Thus, plasticity timing is controlled in part by the interaction of GAGs with Otx2. Interfering with Otx2-

GAG interactions provides an approach to disrupt critical periods for fundamental research and therapeutic applications.⁶

To unravel the extraordinary complexity of GAG-protein interactions, defined compounds are necessary. Unfortunately, both extraction of pure compounds from natural sources and chemical or enzymatic synthesis of GAG fragments remain difficult and inefficient. However, interactions between proteins and GAGs are electrostatic with the most critical aspect being the correct space distribution of charges.⁷ Therefore, an effective strategy could involve a scaffold approach allowing for spatial presentation of sulfates and carboxylates. Modified natural and synthetic polymers have been proposed as GAG surrogates, but they lack sequence modulations and show only non-specific electrostatic interactions with proteins.⁸⁻¹⁰

Previously, we used an approach based on a poly-lactose scaffold whereby the D-GalNAc in natural CS was replaced by D-Gal and the number of synthetic steps was reduced by half by starting from a lactose disaccharide.¹¹ While a hexasaccharide (hexaCSE, Figure 1b) showed promising mimicry, longer and more diverse oligosaccharides mimics are now required and require alternative scaffolds. One promising approach that permits structure modulation involves peptides with either sulfated tyrosines¹² or sulfated serines¹³, even though they show limited stability. Given that GAGs have an extended helical structure (e.g. heparin, 1hpn) with sulfates and carboxylates aligned along the main axis, we hypothesized that peptide analogues are ideal candidates for GAGs mimics as they have the potential for high helical propensity and are relatively easy to synthesize by solid phase strategy.

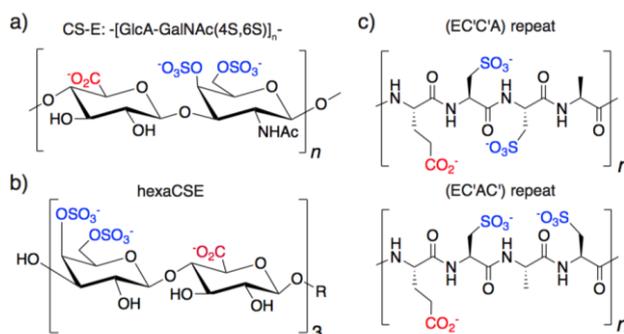


Figure 1. Chemical structure of repeat units. a) Chondroitin-4,6-sulfate, CS-E. b) poly-lactose scaffold, hexaCSE. c) GAG mimics peptide scaffold: EC'C'A and EC'AC', where C' represents cysteic acid.

A tetrapeptide in helix conformation is similar in length (0.8 nm) to a CS disaccharide. We chose to mimic disulfated CS, such as CS-E (Figure 1a), by replacing each disaccharide unit with a tetrapeptide while keeping the chemistry as simple as possible: glutamate to mimic uronic carboxylate; two cysteic acids to mimic sulfates; and alanine to favor helicity and introduce some hydrophobicity. Thus, repeating units each present a carboxylate group and two sulfonate groups (same as 2:1 sulfate:carboxylate ratio for CS-E).

The sequences (ECCA)_n and (ECAC)_n ($n = 3 - 6$) were synthesized by Fmoc strategy on a Rink amide MBHA resin. A 6-aminohexanoic acid (Ahx) spacer was introduced at the N-terminus and the peptidyl resins were either acetylated (Ac) or acylated with a biotin sulfone (Biot(SO₂)). The peptides were cleaved from the resin by TFA and precipitated in diethyl ether to give either Ac-Ahx-(ECAC)_n-NH₂, Ac-Ahx-(ECCA)_n-NH₂, Biot(SO₂)-Ahx-(ECAC)_n-NH₂, or Biot(SO₂)-Ahx-(ECCA)_n-NH₂. These precursors were oxidized with performic acid^{14,15} to give cysteic acid (C') peptides, then neutralized with aqueous ammonia resulting in GAG mimics (EC'AC')_n and (EC'C'A)_n (Figure 1c), with either an acetyl or a biotin sulfone at the N-terminus and an amide at the C-terminus. Crude sulfopeptides were

desalted on Sephadex G25, purified by reverse phase HPLC and characterized by mass spectroscopy. Attempts to purify the precursors prior to oxidation were unsuccessful due to insolubility. These GAG mimic peptides can be stored for months at -20°C as sodium salts.

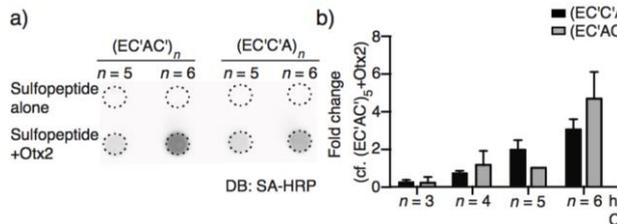


Figure 2. GAG mimics interact with Otx2. (a) Example dot blot (DB) of biotinylated sulfopeptides incubated or not with Otx2 protein. (b) Quantification of DB chemiluminescence with at least 3 duplicate experiments per data point.

To evaluate affinity of the biotinylated (EC'C'A)_n and (EC'AC')_n libraries to Otx2 protein, we performed dot blots with nitrocellulose membranes for which sulfopeptide retention requires interaction with protein (Figure 2a). While Otx2 binding did not favor one motif over the other, there was a clear increase in affinity as a function of *n* repeats (Figure 2b). For comparison, dot blots were performed with hexaCSE, which was previously shown to bind Otx2 and interfere with its *in vivo* activity in the mouse brain.¹¹ Sequences with *n* = 4 or 5 repeats were found to bind to Otx2 equally as well as hexaCSE, while those with *n* = 6 were found to bind 2- to 3-fold better.

To confirm the biotinylated sulfopeptides interact with Otx2 through its previously identified GAG-binding site, we performed DNA chase and peptide binding experiments (Figure 3). In Otx2, a GAG-binding motif is located in the first helix of its DNA-binding domain,⁵ thus specific binding of GAG molecules may interfere with DNA binding. Assays show the (EC'AC')₅ mimic is able to chase the *IRBP1* DNA probe from Otx2 to the same extent as biotinylated hexaCSE positive

control (Figure 3a). The GAG-binding site in Otx2 (RKQRRER) contains an arginine-lysine doublet (RK) which when mutated to an alanine doublet (AA), no longer binds ECM⁵ and causes critical period defects in the *Otx2*^{+/AA} mouse model.¹⁶ Dot blot assays with peptides (15-mer) based on this motif were used to assess whether (EC'AC')₆ binds specifically (Figure 3b). While the wild type peptide (RKpеп: RKQRRERTTFTRAQL) retained the GAG mimic, the mutated peptide (AApеп: AAQRRERTTFTRAQL) did not; neither did a scrambled peptide (SCpеп: RTQTRFRTRARLEQK) containing the same residues as the RKpеп but in random order (Figure 3c). These assays confirm the interaction is not simply electrostatic but requires a specific residue sequence.

To measure the *in vivo* activity and specificity of biotinylated GAG mimics, we performed biochemical and immunohistochemical analyses. Otx2 and Semaphorin-3A (Sema-3A) are both key actors of visual cortex plasticity and share a similar motif for binding CS-E (Figure 4a).^{5,17} We first focused on Sema-3A, as cortical Otx2 levels are too low to be reliably detected biochemically. By using GAG mimics in pull-down experiments with lysates of adult mouse visual cortex, we found that (EC'AC')₅ and (EC'AC')₆ motifs but not (EC'C'A)₆ motif interact with Sema-3A (Figure 4b-c). The RKpеп specifically disrupts Sema-3A interaction with Biot(O₂)-Ahx-(EC'AC')₆-NH₂ in these lysates (Figure 4d-e), suggesting involvement of Sema-3A GAG-binding motif.

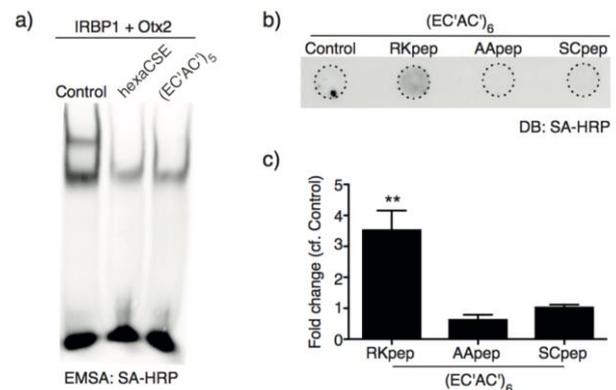


Figure 3. GAG mimics interact with the GAG-binding site of Otx2. (a) Electrophoretic mobility shift assay (EMSA) with biotinylated *IRBP1* DNA probe and Otx2 protein shows loss of shift when incubated with hexaCSE and Biot(O₂)-Ahx-(EC'AC')₅-NH₂. (b) Dot blot (DB) of Biot(O₂)-Ahx-(EC'AC')₆-NH₂ incubated with GAG motif peptide (RKpep) or control peptides (AApep, SCpep). (c) Quantification of DBs. All values: *N* = 3; mean ± SEM; one-way ANOVA with Bonferonni posthoc test; ***P* < 0.01.

After 7-day infusion of either Biot(O₂)-Ahx-(EC'AC')₅-NH₂ or Biot(O₂)-Ahx-(EC'AC')₆-NH₂ in adult mouse visual cortex, 40 fmol caused a reduction in PNN assembly (Figure 5a-b). Infusion with up to 10⁵-fold more of either mimic provided the same reduction. A feedback loop exists between Otx2 accumulation in PV cells and assembly of surrounding PNNs; PNNs attract extracellular Otx2 while Otx2 activity in PV cells increases PNN expression.^{18,19} These results suggest that these GAG mimics interfere with Otx2 signaling enough to break this feedback loop and diminish PNN assembly. Furthermore, interfering with Otx2 signaling in the adult visual cortex has been shown to reverse PV cell maturation state and induce plasticity.^{2,5,20} Here, only infusion of Biot(O₂)-Ahx-(EC'AC')₆-NH₂ (> 40 fmol) resulted in significantly reduced PV expression (Figure 5c). This modest reduction (> 25 %) has been previously shown to be sufficient for re-opening brain plasticity.²⁰

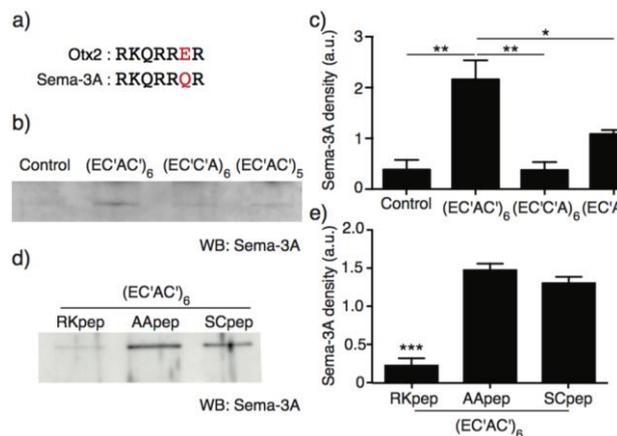


Figure 4. GAG mimic pull-down experiments with lysates of adult mouse visual cortex. (a) Comparison of GAG-binding motifs. (b) Western blot (WB) for Sema-3A after pull-down with GAG mimics retained on streptavidin beads (Control is beads alone). (c) Quantification of WBs. (d) WB for Sema-3A after pull-down with (EC'AC')₆ incubated with GAG motif peptide (RKpep) or control peptides (AApep, SCpep). (e) Quantification of WBs. All values: *N* = 3; mean ± SEM; one-way ANOVA with Bonferonni posthoc test; **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

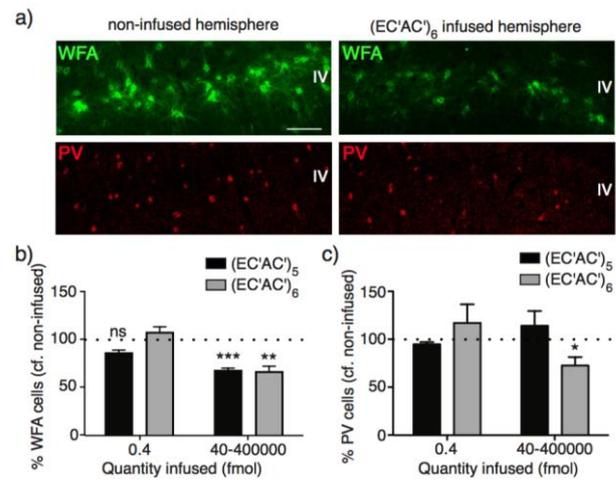


Figure 5. GAG mimics have *in vivo* activity. (a) Representative images of staining for *Wisteria floribunda* agglutinin (WFA), which labels PNNs, and parvalbumin (PV) in adult mouse primary visual cortex layer IV infused or not with 40 fmol Bio(O₂)-Ahx-(EC'AC')₆-NH₂. (b) Quantification of WFA+ cell numbers. (c) Quantification of PV cell numbers. Scale bar = 100 μm. All values: *N* = 3-6; mean ± SEM; t-test; **P* < 0.05, ***P* < 0.01, ****P* < 0.001.

These pull-down and infusion experiments confirm *in vitro* findings that the longer GAG mimics have higher affinity for GAG-binding proteins; only when *n* = 6 for the Biot(O₂)-Ahx-(EC'AC')_{*n*}-NH₂ mimic do we observe full effects on PV cell maturation (Figure 5c). They also suggest that Biot(O₂)-Ahx-(EC'AC')_{*n*}-NH₂ and Biot(O₂)-Ahx-(EC'C'A)_{*n*}-NH₂ mimics can provide specificity and selectivity. While Otx2 shows no preference for either motif *in vitro* (Figure 2b), Sema-3A interacts with the (EC'AC')_{*n*} motif but not the (EC'C'A)_{*n*} motif in brain lysates (Figure 4c).

Thus, these GAG mimics contain valid electrostatic patterns to recapitulate specific sulfation patterns present in natural GAGs. They could be used to construct more complete libraries and screen a set of proteins to generate an interaction matrix with biological relevance. For example, given that the 300-strong homeoprotein family contains a large variety of GAG-binding motifs, such an approach could unlock the hypothesized sugar code for homeoprotein *in vivo* internalization specificity.⁶

The problem of selectivity between GAG and protein remains paramount. CS chains are known to interact with many ECM proteins mainly through electrostatic interactions such that weak or non-productive bindings may interfere with selectivity. Our GAG mimics have a putative alpha-helical form with different carboxylate and sulfonate charge patterns, which allow exploring the charge pattern recognition. Extended structure or induced conformations may also be present, though the final (or major) conformation (in solution or induced through protein interaction) is difficult to predict at this stage. To improve GAG mimic selectivity, different charge-patterned motifs could be assembled along with neutral spacer motifs, while also modifying the repeating motif and length.

Homeoproteins have a high therapeutic potential and their intercellular transfer properties make them particularly interesting.⁶ Otx2 has been shown to act in a non-cell autonomous manner in various brain areas,^{2,16} making it theoretically possible to design various brain plasticity therapies through direct delivery of GAG mimics that alter PV cell maturation state. More generally, GAGs have important and specific physiological functions but are not developed for therapeutic use owing to the difficulty of chemical synthesis. Alternative efforts have focused on peptides and peptidomimetics,^{5,21} while others have

developed small molecules that disrupt sulfation enzymes.²² The development of easy to synthesize GAG mimics with high biological specificity may lead to a new and important class of pharmacological tools.

AUTHOR INFORMATION

Corresponding Authors

*E-mail: ariel.dinardo@college-de-france.fr

*E-mail: Jean-Maurice.Mallet@ens.fr

Author Contributions

§These authors contributed equally.

ORCID

AA Di Nardo: [0000-0003-2944-488X](https://orcid.org/0000-0003-2944-488X)

JM Mallet: 0000-0001-8570-9542

REFERENCES

- (1) Hensch, T. K. *Nat Rev Neurosci* **2005**, *6* (11), 877.
- (2) Spatazza, J.; Lee, H. H. C.; Di Nardo, A. A.; Tibaldi, L.; Joliot, A.; Hensch, T. K.; Prochiantz, A. *Cell Rep* **2013**, *3* (6), 1815.
- (3) Sugiyama, S.; Di Nardo, A. A.; Aizawa, S.; Matsuo, I.; Volovitch, M.; Prochiantz, A.; Hensch, T. K. *Cell* **2008**, *134* (3), 508.
- (4) Miyata, S.; Komatsu, Y.; Yoshimura, Y.; Taya, C.; Kitagawa, H. *Nat Neurosci* **2012**, *15* (3), 414.
- (5) Beurdeley, M.; Spatazza, J.; Lee, H. H. C.; Sugiyama, S.; Bernard, C.; Di Nardo, A. A.; Hensch, T. K.; Prochiantz, A. *Journal of Neuroscience* **2012**, *32* (27), 9429.
- (6) Prochiantz, A.; Di Nardo, A. A. *Neuron* **2015**, *85* (5), 911.
- (7) Johnson, D. J. D.; Li, W.; Adams, T. E.; Huntington, J. A. *EMBO J* **2006**, *25* (9), 2029.
- (8) Henry, B. L.; Desai, U. R. *Thromb. Res.* **2014**, *134* (5), 1123.
- (9) Malineni, J.; Singh, S.; Tillmann, S.; Keul, H.; Möller, M. *Macromol Biosci* **2017**, *17* (5), 1117.
- (10) Sangaj, N.; Kyriakakis, P.; Yang, D.; Chang, C.-W.; Arya, G.; Varghese, S. *Biomacromolecules* **2010**, *11* (12), 3294.
- (11) Despras, G.; Bernard, C.; Perrot, A.; Cattiaux, L.; Prochiantz, A.; Lortat-Jacob, H.; Mallet, J.-M. *Chemistry* **2013**, *19* (2), 531.

- (12) Kim, S. H.; Kiick, K. L. *Peptides* **2007**, *28* (11), 2125.
- (13) Vázquez-Campos, S.; St Hilaire, P. M.; Damgaard, D.; Meldal, M. *QSAR Comb. Sci.* **2005**, *24* (8), 923.
- (14) Hirs, C. H. W. *J Biol Chem* **1956**, *219* (2), 611.
- (15) Moore, S. *Journal of Biological Chemistry* **1963**, *238* (1), 235.
- (16) Lee, H. H. C.; Bernard, C.; Ye, Z.; Acampora, D.; Simeone, A.; Prochiantz, A.; Di Nardo, A. A.; Hensch, T. K. *Mol Psychiatry* **2017**, *22* (5), 680.
- (17) Dick, G.; Tan, C. L.; Alves, J. N.; Ehlert, E. M. E.; Miller, G. M.; Hsieh-Wilson, L. C.; Sugahara, K.; Oosterhof, A.; van Kuppevelt, T. H.; Verhaagen, J.; Fawcett, J. W.; Kwok, J. C. F. *J Biol Chem* **2013**, *288* (38), 27384.
- (18) Bernard, C.; Prochiantz, A. *Neural Plast.* **2016**, *2016* (22, article rc40), 1.
- (19) Hou, X.; Yoshioka, N.; Tsukano, H.; Sakai, A.; Miyata, S.; Watanabe, Y.; Yanagawa, Y.; Sakimura, K.; Takeuchi, K.; Kitagawa, H.; Hensch, T. K.; Shibuki, K.; Igarashi, M.; Sugiyama, S. *Sci Rep* **2017**, *7* (1), 12646.
- (20) Bernard, C.; Vincent, C.; Testa, D.; Bertini, E.; Ribot, J.; Di Nardo, A. A.; Volovitch, M.; Prochiantz, A. *PLoS Genet* **2016**, *12* (5), e1006035.
- (21) Corredor, M.; Bonet, R.; Moure, A.; Domingo, C.; Bujons, J.; Alfonso, I.; Pérez, Y.; Messeguer, À. *Biophys J* **2016**, *110* (6), 1291.
- (22) Cheung, S. T.; Miller, M. S.; Pacoma, R.; Roland, J.; Liu, J.; Schumacher, A. M.; Hsieh-Wilson, L. C. *ACS Chem. Biol.* **2017**, acschembio.7b0088

Supplementary information

Peptides syntheses – general procedure.

Peptides were synthesized by solid-phase on an Applied Biosystems 433 using a Rink amide (*p*-methylbenzhydrylamine)-resin (MBHA-resin, 100–200 mesh, 0.62 mmol/g) by Fmoc strategy on a 0.1 mmol scale. For the coupling reactions HBTU/DIEA was used with 10 equiv of each amino acid (1 mmol). Fmoc-Glu(OtBu)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Ala-OH and Fmoc-6-amino-hexanoic acid, as a spacer at the *N*-terminus, were selected. The spacer and the biotiny sulfone were coupled manually on the peptidyl-resin as specifically described on 0.05 mmol scale. After removal of the Fmoc group with 20 % piperidine in NMP (3 min, 5 min then 10 min) followed by washing steps (NMP), Fmoc-6-amino-hexanoic acid (5 equiv, 0.25 mmol) was activated by HBTU (4.5 equiv) in the presence of DIEA (10 equiv) in NMP for 3 min. The activated species was then added on the resin and the reaction was left for 3 h at room temperature, then the peptidyl-resin was washed (NMP). The Fmoc group was removed and, after washings (NMP), the biotiny sulfone, activated as above, was coupled overnight. After extensive washings (NMP, CH₂Cl₂, MeOH) the peptidyl-resin was dried.

For the cleavage reactions: The resin was stirred for 2 h 30 at room temperature with a mixture of (TFA/water/ethanedithiol/Triisopropylsilane: 94/2.5/2.5/1, v/v). The resin was filtered off and the peptide was precipitated by cold diethyl ether, the precipitate was collected after centrifugation and the pellet was taken with water (insoluble) and lyophilized.

Oxidation steps.

A performic acid solution (2 mL) was prepared by mixing at room temperature: H₂O₂ (70 μL, 35 % in water) with formic acid (1.93 mL, purity 99 %). The resulting solution was stirred 1 h at room temperature then cooled at 0 °C before use. Peptide (10 mg) was dissolved in the performic solution (1 mL) at 0 °C (sonication needed for dispersion). After stirring for 15 h at 0 °C, the solution was diluted with water (6 mL) and neutralized (pH = 7) with aqueous ammonia and lyophilized.

Purification steps/Storage.

Peptides solubilized in acetic acid (10 % in water) were desalted on Sephadex G25 Fine (Pharmacia), with 10 % acetic acid, and the pooled fractions were lyophilized. Peptide purifications were done by preparative HPLC (Waters 600 pump and controller + 2487 Dual absorbance detector) with a C18 column (19x50 mm, XBridge Waters) 5 microns at 10 mL/min, with eluent A: (H₂O + 0.1 % TFA) and eluent B: (CH₃CN + 0.1 % TFA). General elution conditions were: for 2.5 min isocratic conditions with 0 % B, then a gradient from 0 to 10 % B in 10 min. Pure fractions were collected and lyophilized and stored at -20 °C. For biological evaluation, concentrated aqueous solution (ca 3 mg/mL, ca 1 mM) were prepared by dissolution of the peptide in water, neutralization to pH 7 using NaOH (1 mol/L⁻¹) and kept at -20 °C.

MALDI-ToF characterization.

MALDI-ToF mass spectra were acquired on a 4700 Proteomics Analyzer MALDI-TOF-TOF (AB Sciex), equipped with a 355-nm Nd-YAG laser (200 Hz), in a linear negative mode using a guanidinium derivative, pyrenemethylguanidine (pmg), and a *p*-nitroaniline in EtOH matrix[177].

Dot blots.

For competition assays, 400 pmol of biotinylated GAG mimics or hexaCSE were incubated 30 min at 37 °C in 100 mM ammonium acetate with 1 μg of Otx2 protein (in-house)

and with 3 μg of RK-, AA- or SC-peptide. Each solution was then spotted on a nitrocellulose membrane and biotin was detected by 30-min incubation with streptavidin-HRP (ThermoFisher Scientific) followed by chemiluminescence (#34580, ThermoFisher Scientific) reaction. Membranes were digitized with an LAS-4000 (Fujifilm) and quantified by densitometry with ImageJ.

Gel shift.

Otx2 protein (0.1 μg) was incubated at room temperature for 30 min with 40 fmol of biotinylated *IRBP1* oligonucleotide and 4 pmol of (EC'AC')₅ or hexaCSE in 50 ng/ μL dIdC, PBS. Samples were separated on 6 % native polyacrylamide gels at 100 V in TBE then transferred onto a nylon membrane at 380 mA for 45 min, crosslinked with UV (120,000 $\mu\text{J}/\text{cm}^2$, Amersham). The LightShift Chemiluminescent EMSA Kit (#89880, ThermoFisher Scientific) was used for detection and membranes were digitized with an LAS-4000 (Fujifilm) and quantified by densitometry with ImageJ.

Immunoprecipitation.

Visual cortex of adult mice were dissected and lysed in homogenization buffer (0.32 M sucrose, 5 mM HEPES, 10 mM MgCl_2 and protease inhibitors). Samples were centrifuged (8 min, 1700 g) at 4 °C and the supernatant was incubated 2 h at 37 °C with 5 nmol of GAG mimics. For the competition assay, GAG mimics were pre-incubated 30 min with 50 nmol of RK-, AA- or SC-peptide at 37 °C prior to incubation overnight at 4 °C with streptavidin-coupled Dynabeads (Life technologies). The loaded beads were washed with homogenization buffer and heated 10 min at 95 °C in Laemmli buffer (with DTT) to detach proteins for western blot analysis.

Western blot.

Immunoprecipitated proteins were separated on NuPAGE 4-12 % Bis-Tris pre-cast gels (Invitrogen) for 1 h at 200 V and transferred onto a methanol-activated PVDF membrane at 400 mA for 1 h. Membranes were blocked with 5 % non-fat dry milk for 1 h before incubation with primary antibody anti-Sema3A (rabbit, 1/1000, Millipore) overnight at 4 °C. Membranes were washed and incubated 1 h with secondary antibody anti-rabbit HRP-linked (Cell Signaling). Membranes were digitized with an LAS-4000 (Fujifilm) and quantified by densitometry with ImageJ.

Ethics statement.

All animal procedures were carried out in accordance with the guidelines of the European Economic Community (2010/63/UE) and the French National Committee (2013/118). For surgical procedures, animals were anesthetized with Xylazine (Rompun 2 %, 5 mg/kg) and Ketamine (Imalgene 1000, 80 mg/kg). This project (no. 00704.02) obtained approval from Ethics committee n° 59 of the French Ministry for Research and Higher Education.

Brain infusions and immunohistochemistry

Three-month-old C57BL/6J mice (Janvier) were infused for 7 days into V1 (lambda: x = 1.7 mm, y = 0 mm, z = 0.5 mm) with various concentrations of GAG mimics (4 pM, 400 pM or 4 μM), using Alzet micro-osmotic pumps (0.5 $\mu\text{L}/\text{h}$). Animals were then perfused with PBS and 4 % paraformaldehyde. Cryostat sections (20 μm) were incubated overnight with primary antibody anti-parvalbumin (rabbit, 1/500, Swant) and WFA-FITC (1/100, Vector), followed by secondary antibody anti-rabbit Alexa Fluor-546 (1/2000, Molecular Probes) for 1 h. Images were acquired with a Leica SP5 confocal microscope and quantified by analysis with ImageJ.

Statistical analysis

Analysis was performed with Prism 6 (GraphPad). Single comparisons were made by t-test, whereas multiple group analyses were made by ANOVA followed by Bonferonni's posthoc test.

Annexes

HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'AC')₃-CONH₂ (RM3-84B)

HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'AC')₄-CONH₂ (RM3-84D)

HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'AC')₅-CONH₂ (RM3-92A)

HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'AC')₆-CONH₂ (RM3-164)

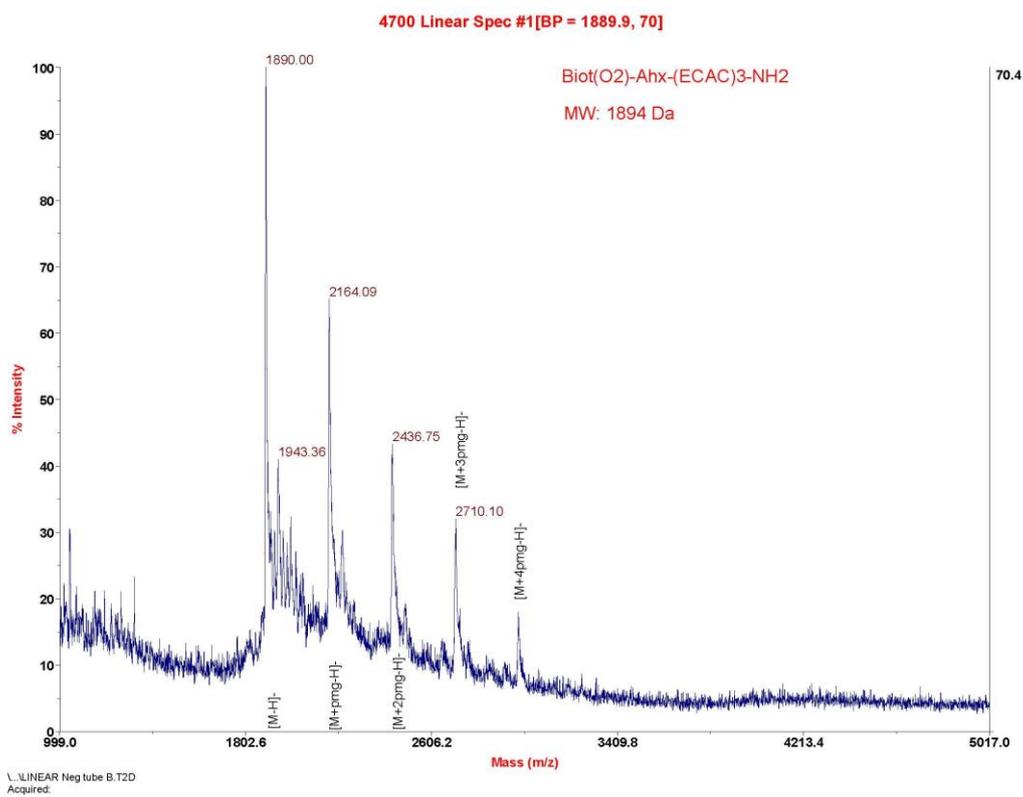
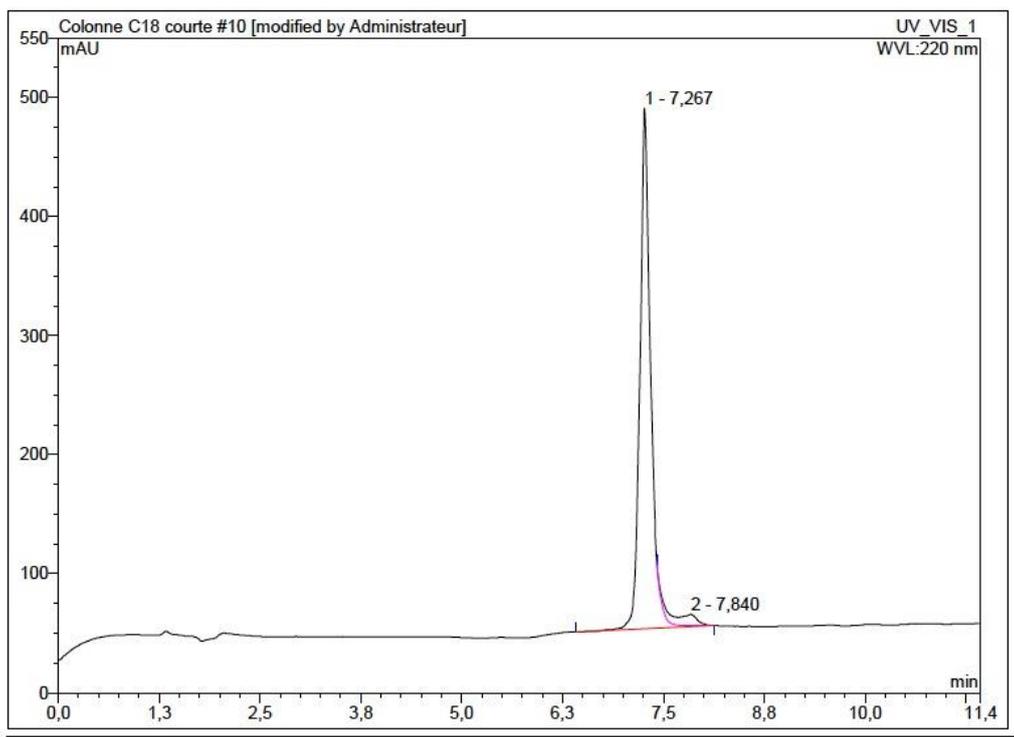
HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'C'A)₃-CONH₂ (RM3-84A)

HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'C'A)₄-CONH₂ (RM3-84C)

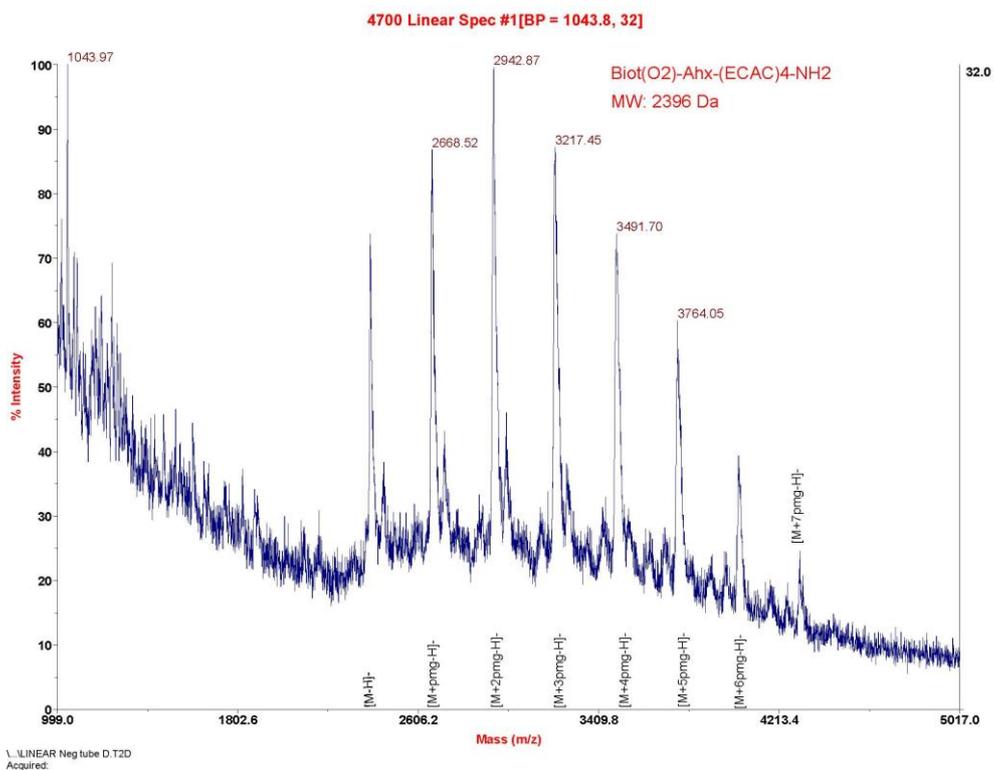
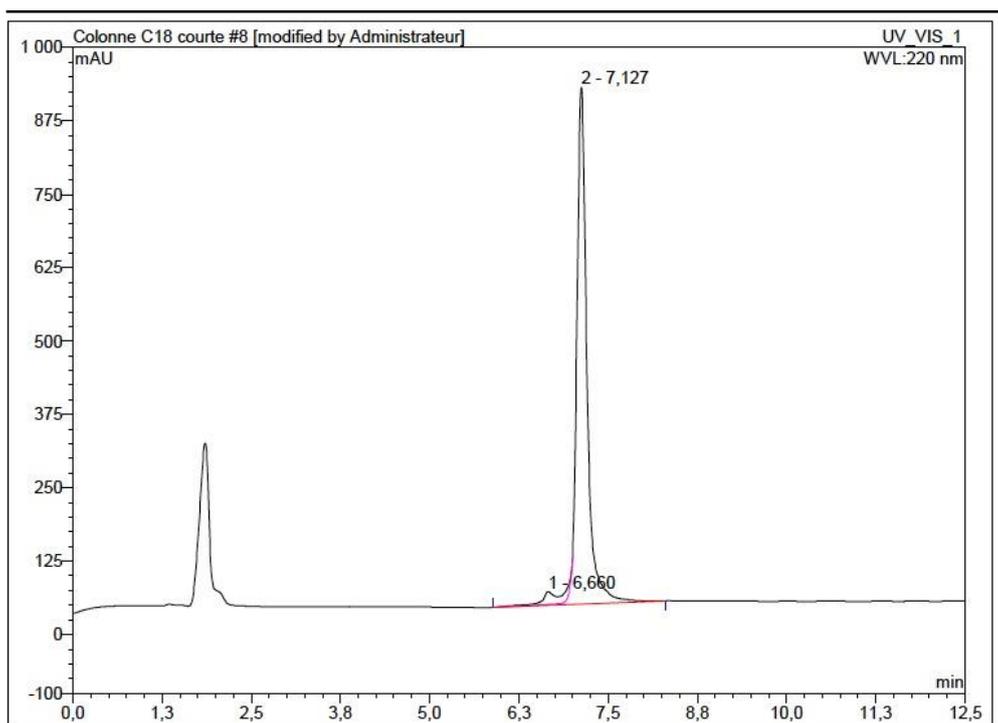
HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'C'A)₅-CONH₂ (RM3-149A)

HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'C'A)₆-CONH₂ (RM3-162)

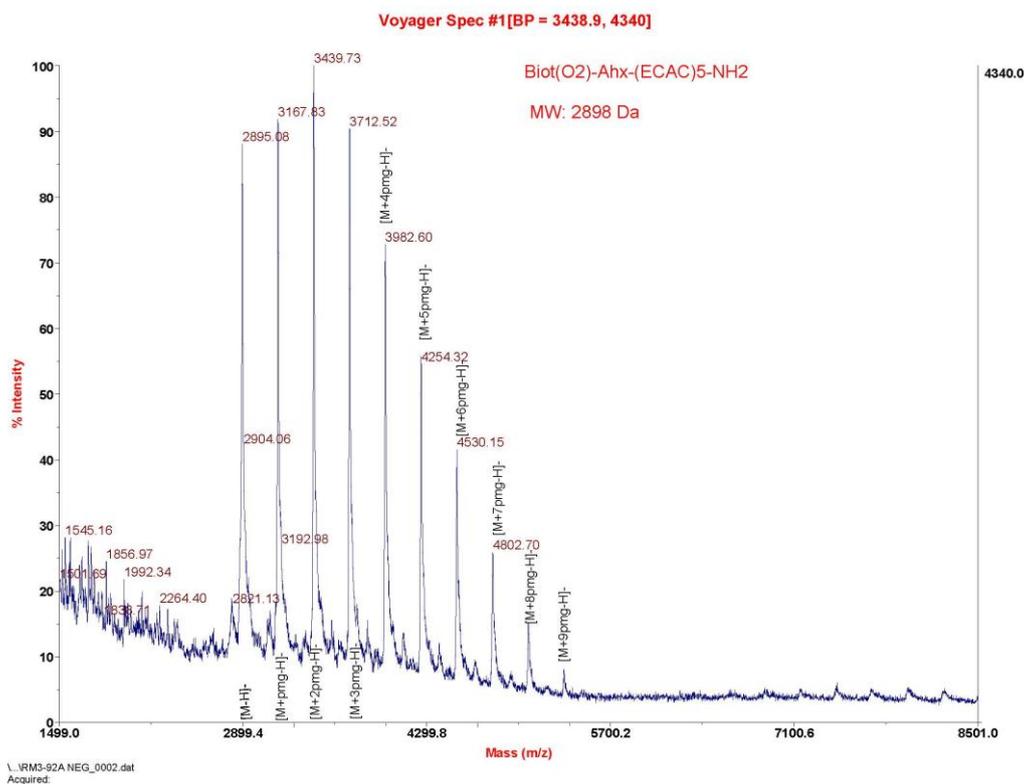
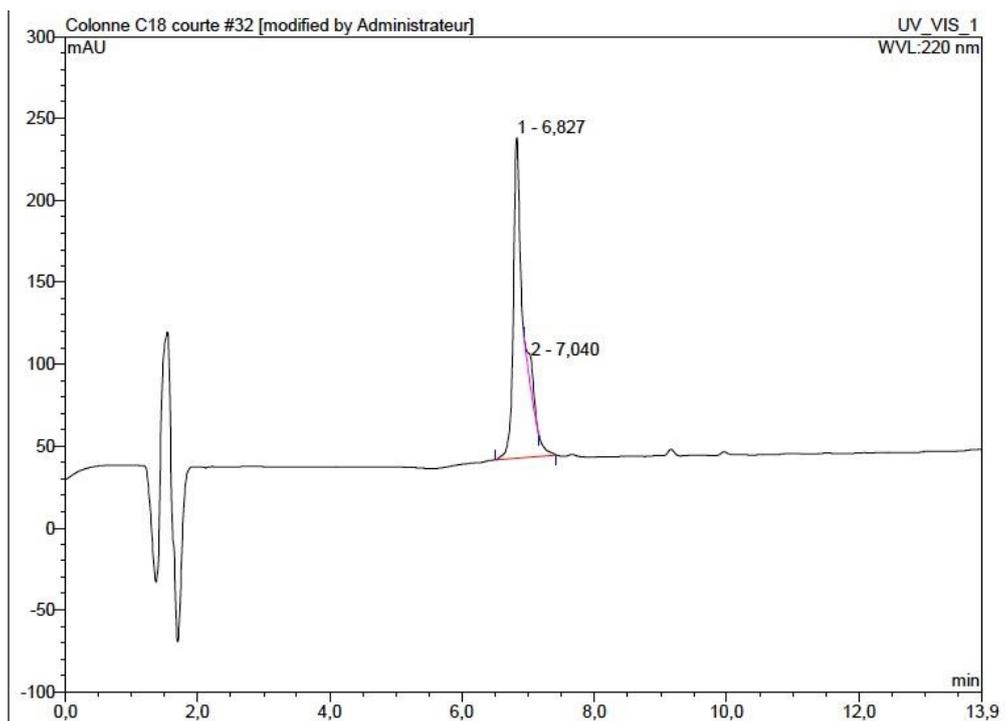
HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'AC')₃-CONH₂ (RM3-84B)



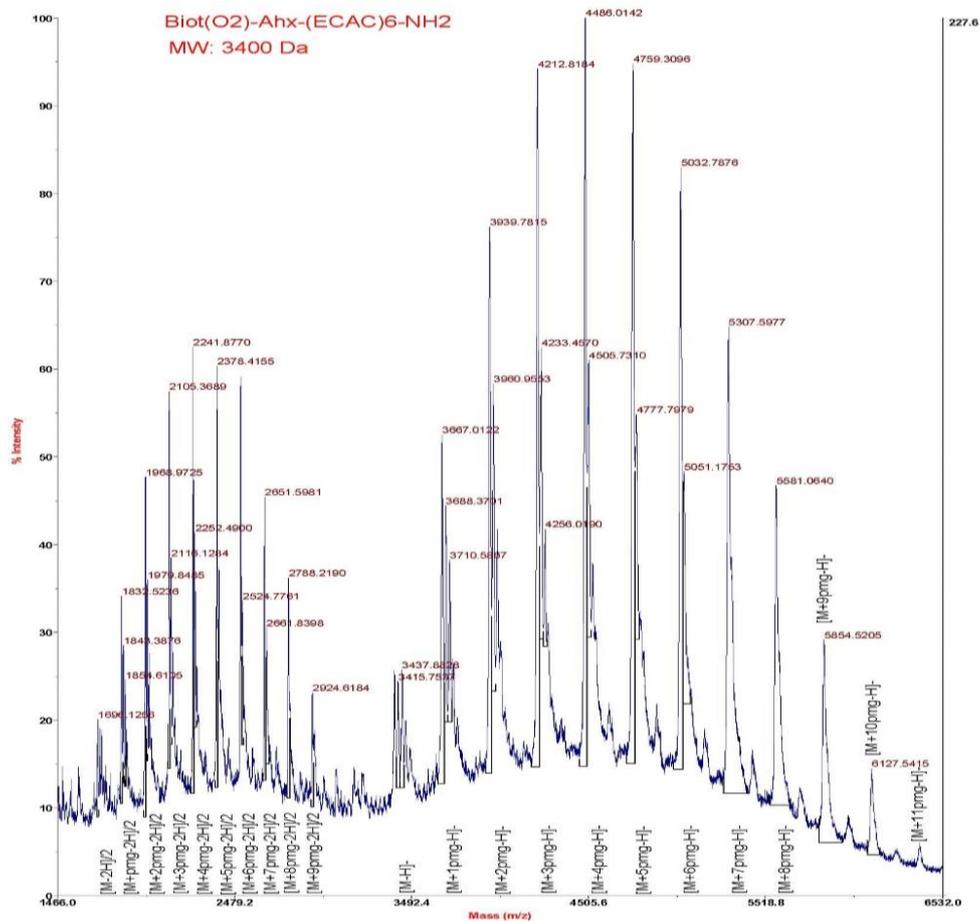
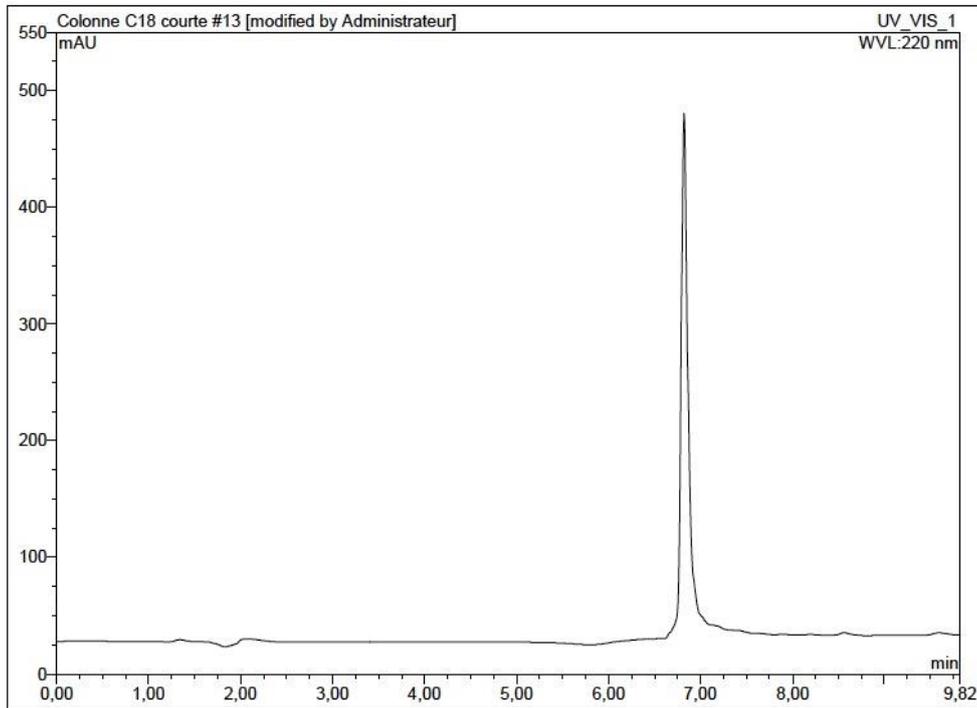
HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'AC')₄-CONH₂ (RM3-84D)



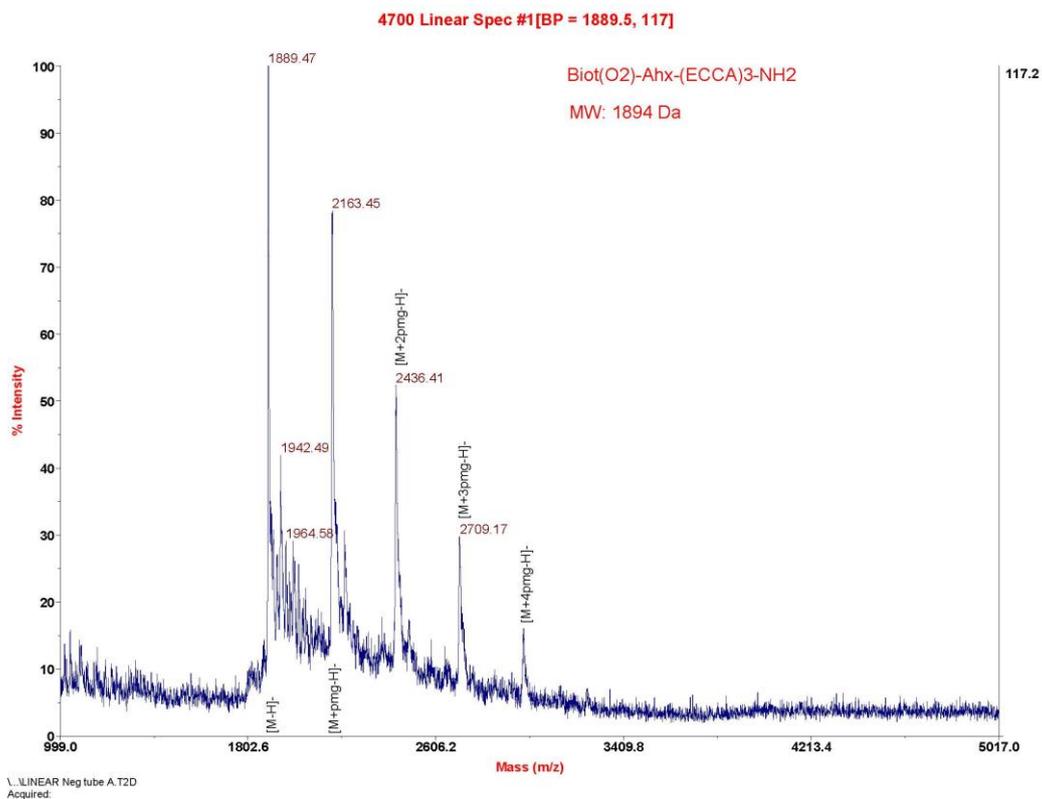
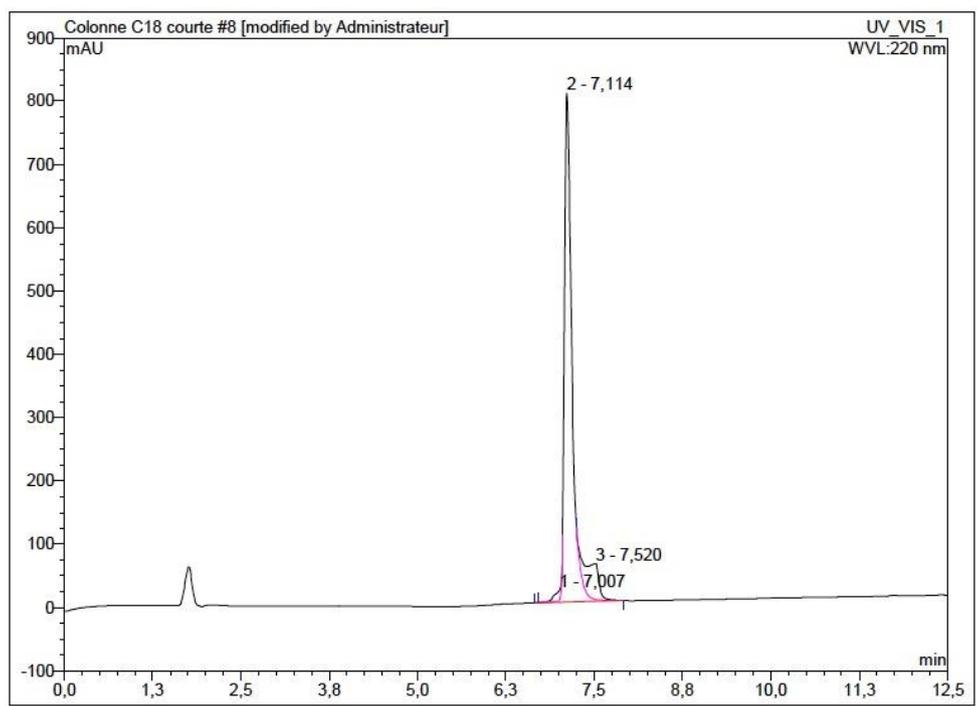
HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'AC')₅-CONH₂ (RM3-92A)



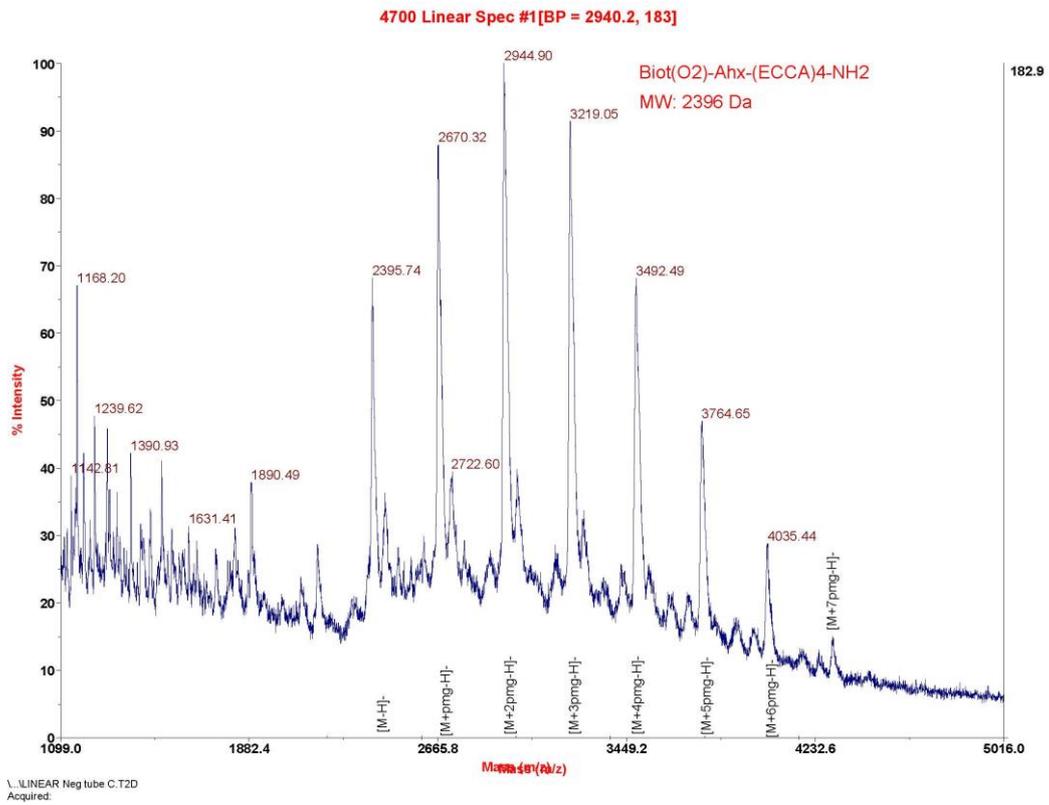
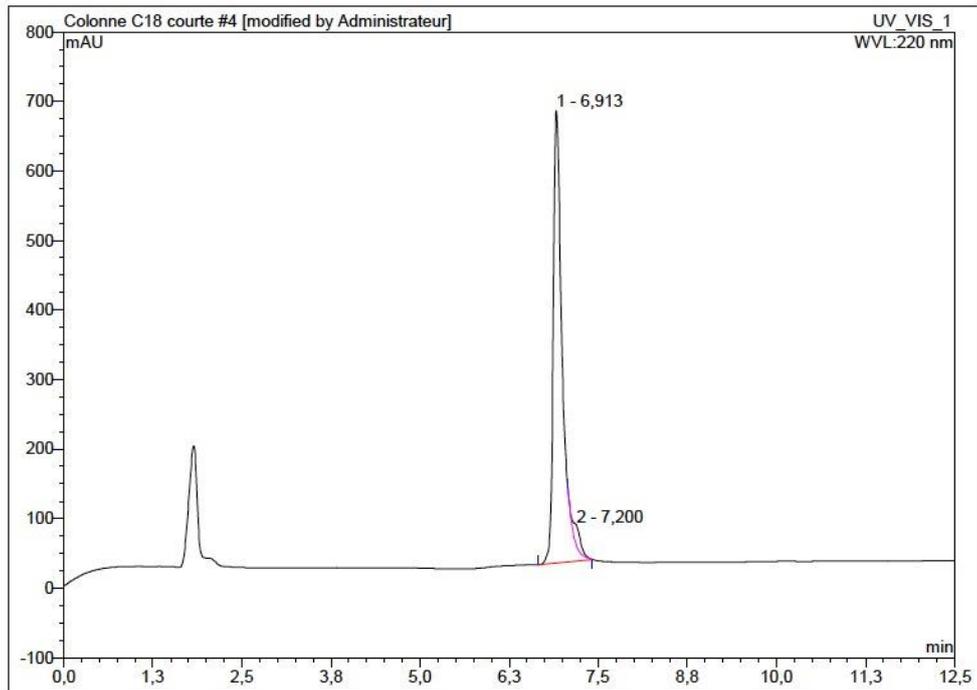
HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'AC')₆-CONH₂ (RM3-164)



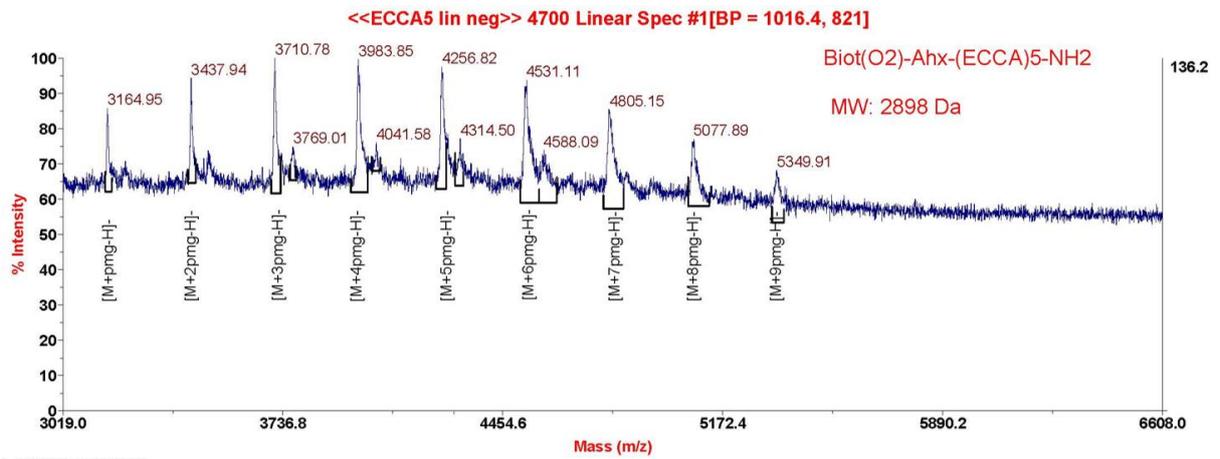
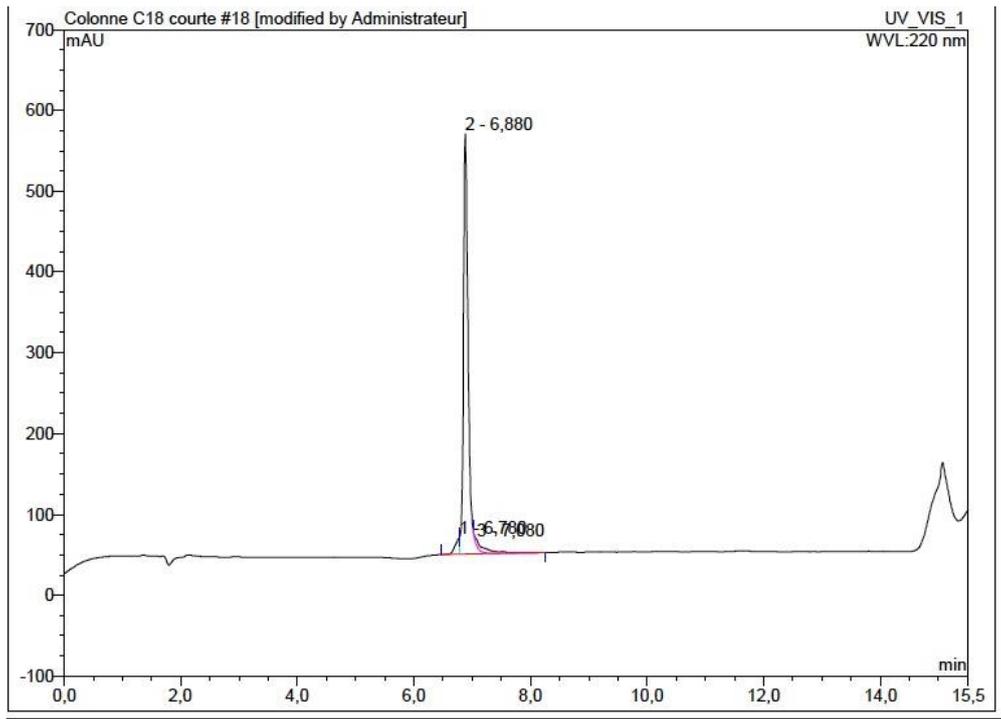
HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'C'A)₃-CONH₂ (RM3-84A)



HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'C'A)₄-CONH₂ (RM3-84C)

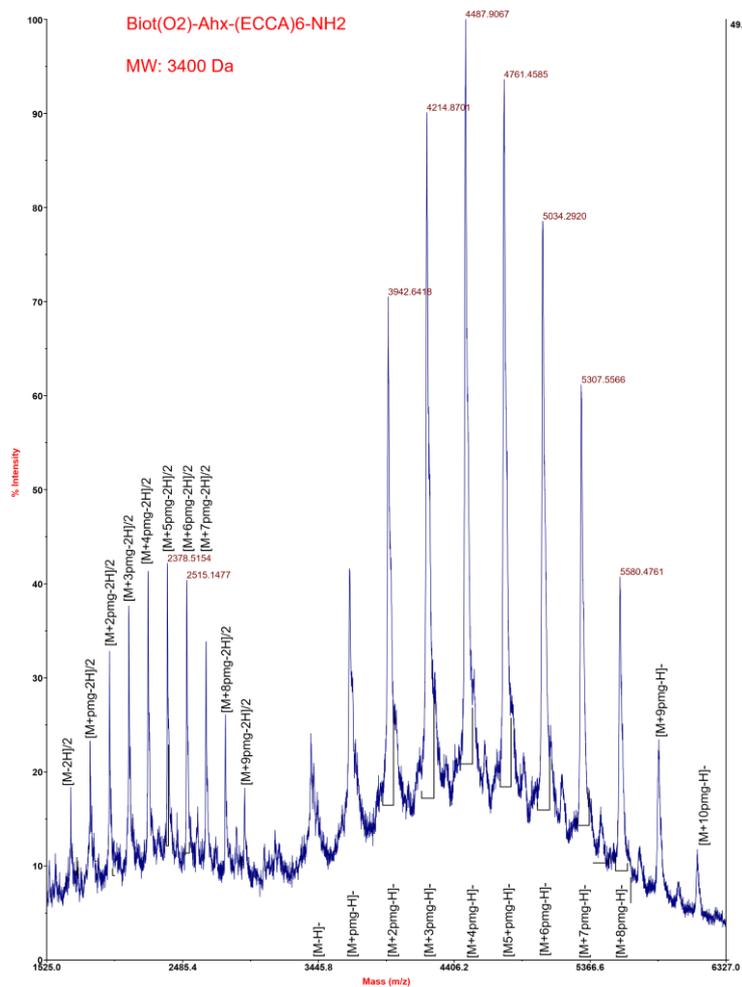
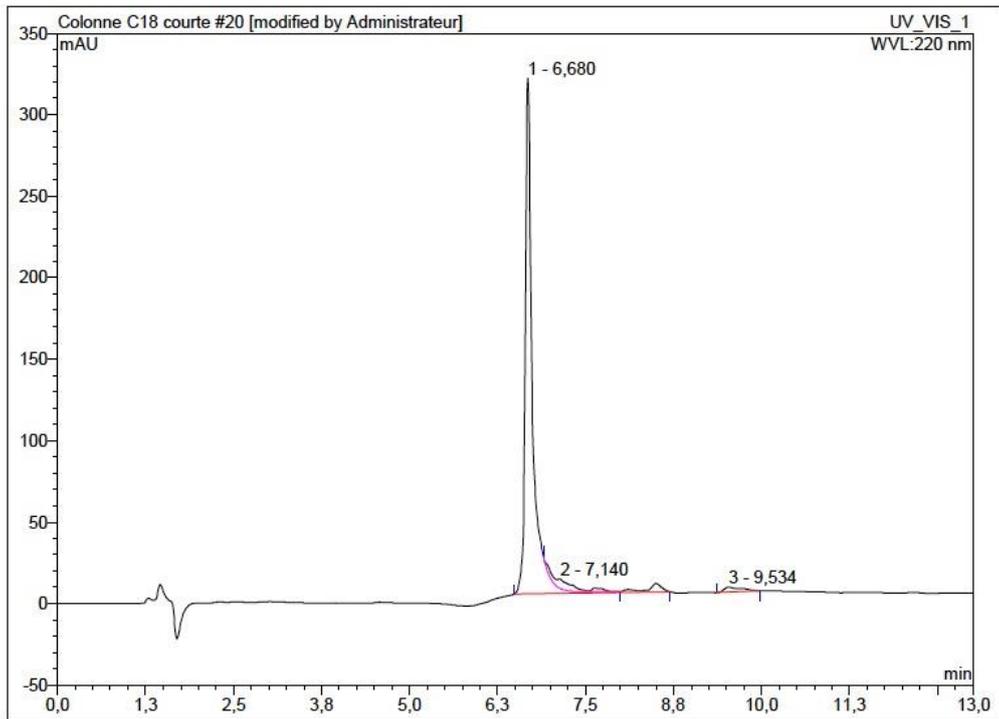


HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'C'A)₅-CONH₂ (RM3-149A)



\\.\ECCA5 lin postT2D.T2D
Acquired:

HPLC trace and negative ion mass spectra of Biot(O₂)-Ahx-(EC'C'A)₆-CONH₂ (RM3-162)



CHAPITRE IV : Etudes des interactions *in vitro* par RMN et SPR de type Biacore.

A. Introduction

En conclusion du travail décrit dans le chapitre III, j'insistai sur deux points qui sont la pierre angulaire de toute recherche de nouvelle molécule ciblant une activité biologique, à savoir l'affinité et la sélectivité envers la biomolécule cible. Modifier l'une ou l'autre de ces propriétés nécessite une meilleure connaissance au niveau moléculaire de l'interaction des partenaires mis en jeu. La question de l'orientation à donner à mes travaux s'est donc posée à ce stade, concevoir et synthétiser des molécules plus affines, plus sélectives étant bien évidemment l'objectif final. Mais comment s'y prendre ?

Sans *a priori* dans un premier temps sur le mode de reconnaissance entre les deux partenaires, nous pouvions proposer des variations dans la séquence du térapeptide, $(EC'XC')_n$, autour du résidu X : avec X plus ou moins hydrophobe, inducteur ou interrupteur de structure hélicoïdale. Je présenterai les résultats préliminaires de cette approche dans le dernier chapitre de ce manuscrit, « *Travaux en cours et perspectives* ».

Nous pouvions aussi tenter une approche aléatoire, justifiée par la littérature qui tend à montrer que les interactions carbohydrates/protéines englobent très souvent des domaines espacés dans la séquence primaire de la protéine et que des substrats plurifonctionnels conduisent à de meilleures affinités avec la protéine cible. Une hypothèse « évidente » (sur le papier !) était de préparer des polymères ou dendrimères, avec plusieurs séquences de carbohydrates identiques (homopolymères ou homodendrimères) ou avec des séquences différentes (hétéropolymères ou hétérodendrimères). Des constructions pouvaient aussi être obtenues par conjugaison du (ou des) monomère(s) carbohydrates sur un polymère synthétique ou naturel, comme décrit dans la littérature et résumé dans différentes revues, nous n'en citerons que deux[182], [183], et réalisé dans le laboratoire par Antonio Mazzoleni durant sa thèse[184]. Sur ce sujet, je présenterai également des résultats préliminaires dans le dernier chapitre de ce manuscrit, « *Travaux en cours et perspectives* ».

Au-delà de ces démarches n'impliquant que des compétences de synthèse, il nous fallait également obtenir des informations sur les interactions de la protéine Otx2 avec nos sulfopeptides et ainsi baser nos nouvelles stratégies de synthèses sur des études de relations

structure-activité/affinité. Nous avons alors analysé ce qui était du domaine du faisable en associant des compétences présentes au sein du laboratoire.

Dans le laboratoire, l'équipe du Pr Olivier Lequin étudie depuis plusieurs années les protéines à homéodomaine d'un point de vue structural. Le Dr Ludovic Carlier, au sein de cette équipe, a tout particulièrement étudié par RMN (Résonance Magnétique Nucléaire) la structure du facteur de transcription Engrailed 2 (En2)[185] et plus récemment son interaction avec des fragments de chondroïtines sulfates et d'héparine dans le cadre de la thèse de Sébastien Cardon[186]. En nous appuyant sur ce travail, nous avons entrepris de caractériser les bases moléculaires de la reconnaissance des sulfopeptides par le domaine d'intérêt d'Otx2 ainsi que d'En2, la cible étudiée par Sébastien Cardon durant ses travaux de thèse.

Dans la première partie de ce chapitre, je décrirai le travail réalisé en RMN et notamment :

- la production recombinante du domaine d'intérêt de la protéine Otx2 chez la bactérie *E. coli*, ainsi que sa purification par chromatographie d'affinité et d'exclusion stérique[185],
- les méthodes utilisées au cours de mon travail de thèse pour caractériser l'interaction d'une protéine avec un ligand par spectroscopie RMN,
- les résultats des études RMN d'interactions des sulfopeptides avec les domaines d'intérêt des protéines Otx2 et En2.

Puis, dans une deuxième partie, je présenterai les études d'interaction réalisées par la technique SPR de type Biacore, sur les mimes de première génération décrits précédemment mais également sur de nouveaux analogues, Ac-Ahx-(EC'PC')₆-NH₂ et Ac-Ahx-(EC'LC')₆-NH₂.

B. Etude des interactions par RMN

1. Surexpression et purification des protéines pour l'étude structurale

a) Surexpression et purification du domaine NLS-HD de la protéine Otx2 humaine

La protéine Otx2 humaine est un polypeptide basique de 289 résidus (**Figure 77**) qui contient de nombreuses Arginines et Lysines. Son point isoélectrique (théorique) est de 8,78 (<http://isoelectric.org/calculate.php>).

La production de la protéine Otx2 humaine dans son intégralité n'a pu être réalisée par nos collègues du Collège de France, néanmoins le Dr Alain Joliot (Collège de France - Centre de recherche de l'Institut Curie) a préparé un plasmide qui permet de produire la région étendue de l'homéodomaine correspondant aux résidus 22-97. Ce fragment de protéine, appelé NLS-HD, englobe les 3 hélices de l'homéodomaine et notamment le motif de pénétration cellulaire, ainsi que le site de liaison aux GAGs en amont de l'homéodomaine. Le Dr Joliot nous a gracieusement fourni un lot de bactéries *E. Coli* BL21 DE3 préalablement transformées par le plasmide.

MMSYLKQPPY¹⁰AVNGLSLTTS²⁰GMDLLHPSVG³⁰YPATP **PKQRR⁴⁰ERTTFTRAQL⁵⁰DVLEALFAKT⁶⁰RY**
PDIFMREE⁷⁰VALKINLPES⁸⁰RVQVWFKNRR⁹⁰AKCRQQQQQ¹⁰⁰QNGGQNKVRP¹¹⁰AKKKTSPARE¹²⁰
VSSEGTSGQ¹³⁰FTPPSSTSVP¹⁴⁰TIASSAPVS¹⁵⁰IWSPASISPL¹⁶⁰SDPLSTSSSC¹⁷⁰MQRSYPMTYT¹⁸⁰Q
ASGSQGYA¹⁹⁰GSTSYFGGMD²⁰⁰CGSYLTPMH²¹⁰QLPGPGATLS²²⁰PMGTNAVTS²³⁰LNQSPASLST
²⁴⁰QGYGASSLGF²⁵⁰NSTTDCLDYK²⁶⁰DQTASWKLNF²⁷⁰NADCLDYKDQ²⁸⁰TSSWKFQVL²⁸⁹

Figure 77. Séquence de la protéine Otx2 humaine (Uniprot P32243). Le domaine de liaison aux GAGs est surligné en cyan, les hélices H1, H2 et H3 de l'homéodomaine sont en rouge.

Comme le montre la **Figure 78a**, le vecteur d'expression de NLS-HD encode une protéine de fusion de 24 kDa comprenant une étiquette 6XHis *N*-terminale, la séquence de la protéine Cherry, le site de clivage LEVLFQGP spécifiquement reconnu par l'enzyme PreScission, et le domaine NLS-HD d'Otx2 en *C*-terminal. L'étiquette 6XHis (6 résidus histidine consécutifs) a été introduite en *N*-terminal pour que la protéine de fusion puisse être purifiée par chromatographie d'affinité sur colonne de nickel. La présence du partenaire Cherry en *N*-terminal a pour objectif d'augmenter la solubilité de la protéine de fusion. L'originalité du partenaire Cherry réside également dans sa capacité à fixer l'hème bactérien, ce qui induit une coloration rosée des fractions protéiques et facilite la détection de la protéine de fusion lors des étapes de purification. L'introduction du site de reconnaissance LEVLFQGP de l'enzyme PreScission en amont de la séquence de NLS-HD va permettre de cliver la protéine

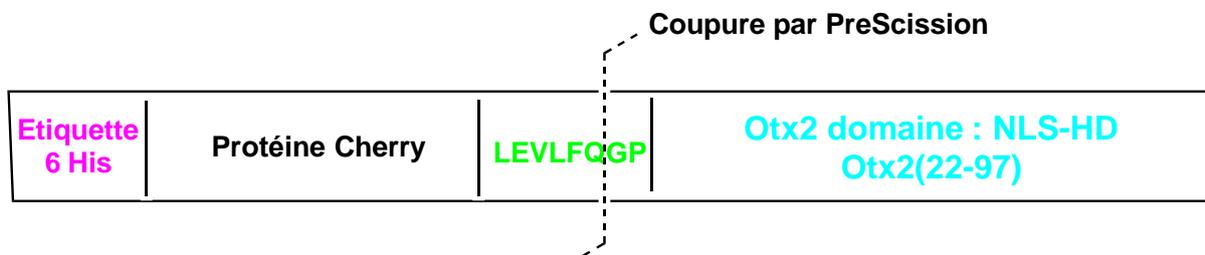
de fusion spécifiquement entre les résidus Gln (Q) et Gly (G), et ainsi de séparer l'étiquette 6XHis et le partenaire de fusion de la protéine d'intérêt, Cf. **Figure 78b**.

Les séquences détaillées ainsi qu'une version schématisée sont représentées avec un code couleur sur cette même figure, **Figures 78a et 78b**.

a) Protéine exprimée par le plasmide : 206 résidus / 24 kDa, Point isoélectrique théorique : 5,68, (<http://isoelectric.org/calculate.php>)

MHHHHHHTSEFAEQSDKDVKYITLLEIQKHKDSKSTWVILHHKVYDLTKFLEEHPGGEEVLGEQAGGDA
TENFEDVGHSTDARELSKTYIIGELHPDDRSKIAPSETLDDDDKHMGESEFLEVLFGGPASM²²DLLHPSVG
YPATPRKQRRERTTFTRAQLDVLEALFAKTRYPDIFMREEVALKINLPESRVQVWFKNRRAKCRQQ⁹⁷

Etiquette 6xHis ; coupure Q/G par PreScission ; Otx2(22-97), domaine de liaison aux GAGs



b) Séquence après coupure par l'enzyme PreScission : 80 acides aminés / 9444Da, Point isoélectrique théorique : 10,60, (<http://isoelectric.org/calculate.php>)

GPASM²²DLLHPSVGYPATPRKQRRERTTFTRAQLDVLEALFAKTRYPDIFMREEVALKINLPESRVQVWFKNRRAKCRQQ⁹⁷

Composition en acides aminés : Ala (A) : **7**, 8,8 % ; Arg(R) : **11**, 13,8 % ; Asn (N) : **2**, 2,5 % ; Asp (D) : **3**, 3,8 % ; Cys (C) : **1**, 1,2 % ; Gln (Q) : **6**, 7,5 % ; Glu (E) : **5**, 6,2 % ; Gly (G) : **2**, 2,5 % ; His (H) : **1**, 1,2 % ; Ile (I) : **2**, 2,5 % ; Leu (L) : **7**, 8,8 % ; Lys (K) : **5**, 6,2 % ; Met (M) : **2**, 2,5 % ; Phe (F) : **4**, 5,0 % ; Pro (P) : **6**, 7,5 % ; Ser (S) : **3**, 3,8 % ; Thr (T) : **5**, 6,2 % ; Trp (W) : **1**, 1,2 % ; Tyr (Y) : **2**, 2,5 % ; Val (V) : **5**, 6,2 %.

Composition atomique : Carbone : 418 ; Hydrogène : 675 ; Azote : 129 ; Oxygène : 115 ; Soufre : 3. Coefficient d'extinction : 8480 M⁻¹.cm⁻¹ ; Abs 0.1 % (= 1 g/l) 0.898, si Cys sous forme réduite.

Figure 78. a) Séquence de la protéine exprimée et b) séquence de la protéine d'intérêt incluant le domaine 22-97 de Otx2 humaine. Les hélices H1, H2 et H3 de l'homéodomaine sont en rouge.

La RMN liquide est une technique de choix pour étudier la structure et l'interaction d'une protéine de 10 kDa à l'échelle de l'acide aminé. Elle nécessite cependant la préparation d'échantillons de protéine uniformément enrichie en isotopes ¹⁵N et/ou ¹⁵N/¹³C. Les protocoles d'expression, marquage isotopique, et purification des protéines à

homéodomaine, entre autres Antennapedia et Engrailed 2, ont été mis au point dans le laboratoire au fil des années par l'équipe du Pr Olivier Lequin à partir de productions recombinantes chez la bactérie *E. coli*. Ils ont été décrits dans le détail dans les travaux du Dr Ludovic Carlier[185] et dans le manuscrit de thèse de Sébastien Cardon[186]. Nous avons reproduit les grandes lignes de ces protocoles pour l'expression et la purification de la construction NLS-HD d'Otx2 non marqué (utilisée pour les études d'interaction hors RMN – se reporter au chapitre V) et doublement marqué $^{15}\text{N}/^{13}\text{C}$ (pour les études d'interaction par RMN).

Les bactéries *E. Coli* BL21 DE3 transformées par le plasmide ont été cultivées à 37 °C sous agitation. La présence du plasmide confère aux bactéries une résistance à la kanamycine, un antibiotique bactériolytique de la famille des aminosides, ce qui permet de ne cultiver que les bactéries portant le plasmide en présence de kanamycine dans le milieu de culture. Ainsi, toutes les bactéries non transformées seront lysées par la kanamycine. La surexpression de la protéine de fusion est alors réalisée de façon conventionnelle avec une induction par de l'IPTG, (isopropyl- β -D-1-thiogalactopyranoside), comme détaillé dans la partie expérimentale. Les suivis de l'extraction du milieu bactérien et des étapes de purification ont été réalisés par électrophorèse sur gel SDS-PAGE. Comme mentionné plus haut, la fixation de l'hème bactérien par la protéine Cherry permet d'identifier facilement les fractions contenant la protéine de fusion à partir de leur coloration rosée, au-delà du suivi par UV à 280 nm. Les étapes de purification correspondent essentiellement à des chromatographies d'affinité, *via* des colonnes fixant du Ni^{2+} pour chélater les histidines de l'étiquette *N*-terminale en amont de la protéine Cherry (**Figure 79**). L'éluion est réalisée par un gradient d'imidazole.

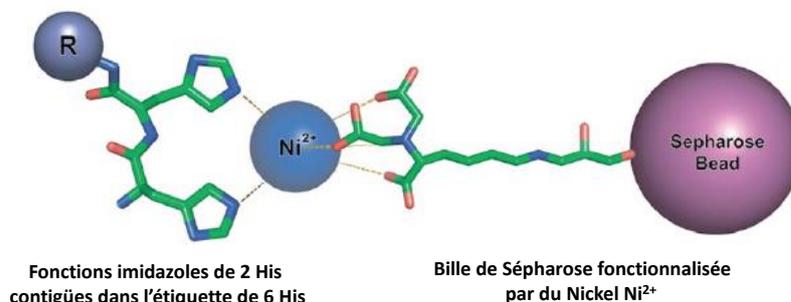


Figure 79. Principe de la chromatographie d'affinité entre Histidines et colonne de Nickel[187].

Après une première étape de purification sur colonne de Ni²⁺, l'enzyme PreScission est introduite dans les fractions d'élution afin de cliver la protéine de fusion entre la Glutamine et la Glycine au niveau de la séquence LEVLFQ/GP introduite entre Cherry et NLS-HD d'Otx2. Une deuxième étape de purification est réalisée sur colonne de Ni²⁺ afin de séparer le fragment N-terminal 6XHis-Cherry de la protéine d'intérêt NLS-HD d'Otx2. Le fragment 6XHis-Cherry sera en effet retenu sur la colonne contrairement à NLS-HD qui ne porte pas d'étiquette 6XHis. Il est à noter que l'enzyme PreScission utilisée a été produite au laboratoire et porte également une étiquette 6XHis, elle sera donc également séparée de la protéine d'intérêt lors de la deuxième étape de purification.

Les fractions contenant le domaine NLS-HD d'Otx2 sont concentrées par ultrafiltration sur cellule Amicon. Avant une dernière étape de purification par chromatographie d'exclusion stérique (également appelée « filtration sur gel »), la concentration de la protéine, dosée par mesure de l'absorbance à 280 nm, est de l'ordre de 180 µM. Le domaine NLS-HD d'Otx2 est alors purifié sur colonne de filtration sur gel Hiload 16/60 - Superdex 75 - pepgrade (GE Healthcare). Le rendement de production de protéine purifiée est de l'ordre de 2 mg par litre de milieu de culture pour la protéine non marquée, et de 4 mg pour la protéine doublement marquée ¹⁵N/¹³C.

La caractérisation finale de la protéine a été réalisée par HPLC et spectrométrie de masse MALDI-TOF, ce qui nous a permis de déterminer le pourcentage d'incorporation des isotopes ¹⁵N et ¹³C dans les échantillons doublement marqués pour les études RMN.

A titre d'exemples, nous présentons ci-dessous (**Figures 80 à 82**)

- **Figure 80** : Un gel SDS-Page des fractions d'élution du domaine NLS-HD d'Otx2 doublement marqué ¹⁵N/¹³C après les deux premières étapes de purification sur colonne Ni²⁺,
- **Figure 81** : Le profil HPLC en phase inverse du domaine NLS-HD d'Otx2 complètement purifié (non marqué) qui montre un seul pic d'élution,
- **Figure 82** : Le spectre MALDI-TOF du domaine NLS-HD d'Otx2 complètement purifié **a)** marqué et **b)** non marqué.

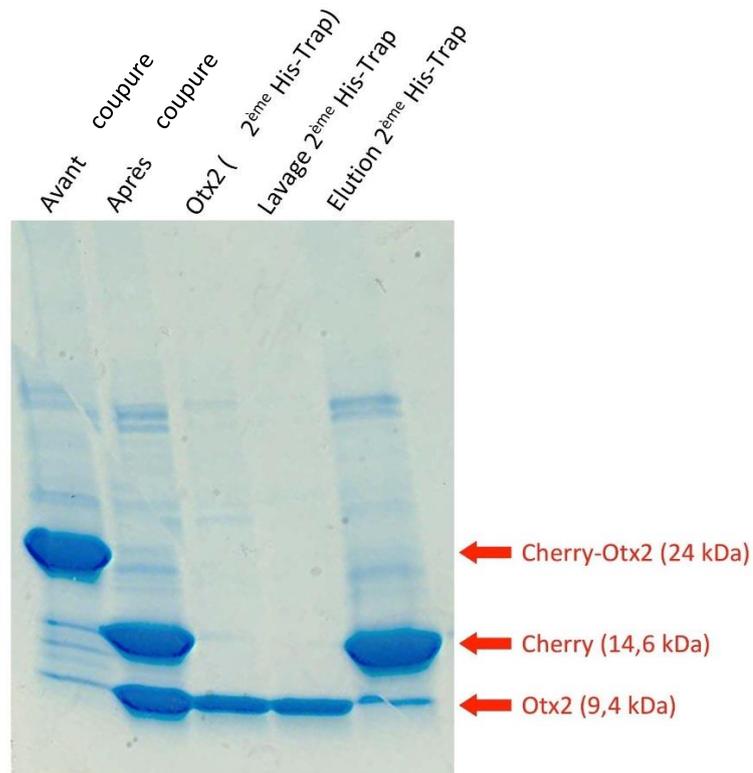


Figure 80. Gel SDS-PAGE (12,5 % de polyacrylamide) des différentes étapes de purification du domaine NLS-HD d'Otx2, avant et après coupure par l'enzyme PreScission, et les différentes fractions analysées après cette coupure.

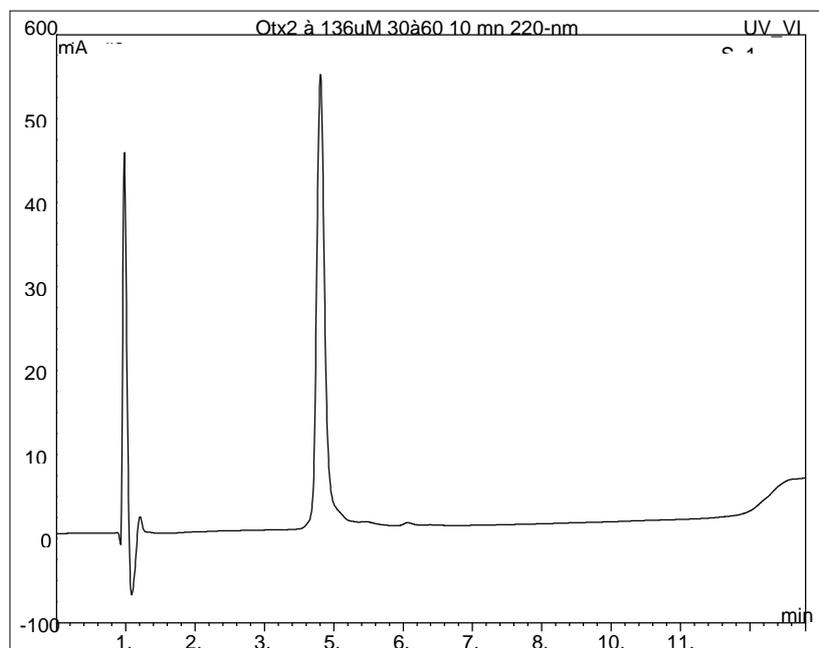


Figure 81. Profil HPLC en phase inverse du domaine NLS-HD d'Otx2 non marqué.
Conditions : Colonne C18 Proto200, 3 µm, 200 Å, 4,6x100 mm - Gradient en 10 min de 30 à 60 % de CH₃CN (0,1 % TFA) - Débit : 1 mL/min - Lecture à 220 nm

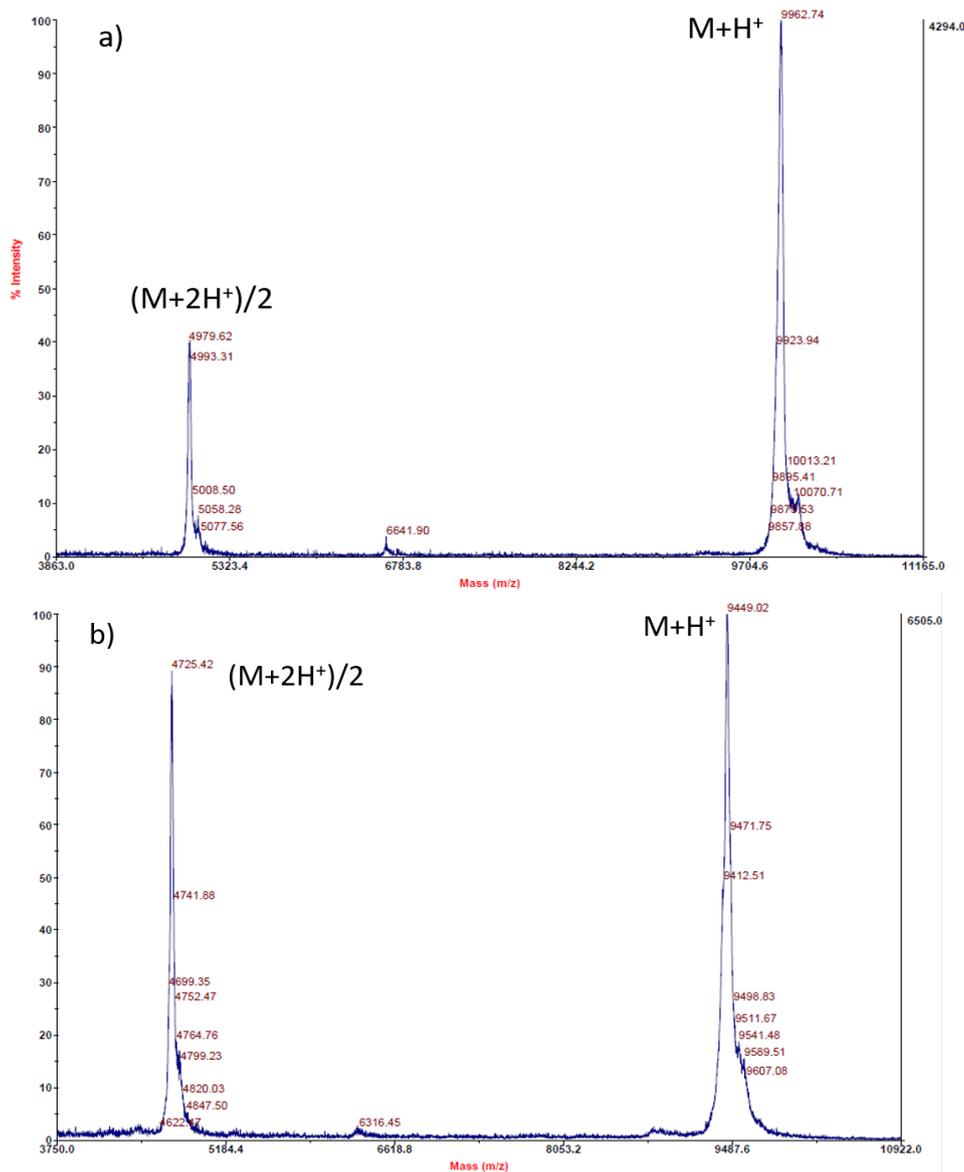


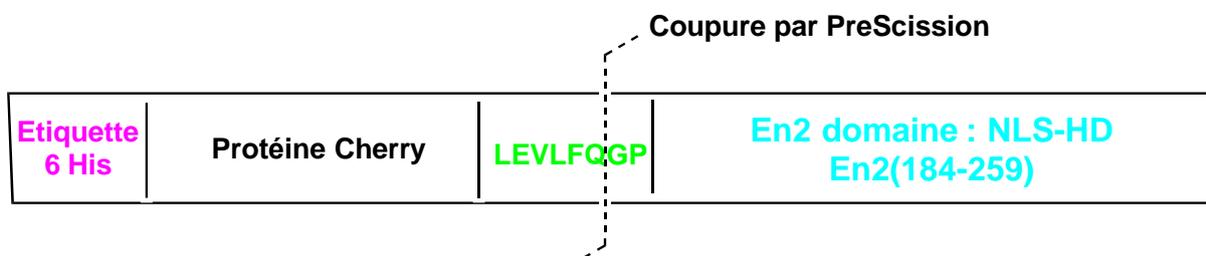
Figure 82. Profil MALDI-TOF du domaine NLS-HD d'Otx2 a) marqué et b) non marqué, avec matrice HCCA, en mode linéaire positif.

Le pourcentage d'incorporation des isotopes ^{15}N et ^{13}C dans le domaine NLS-HD d'Otx2 doublement marqué est obtenu en calculant le rapport de l'augmentation de la masse moléculaire observée entre la protéine marquée et celle non marquée (ici $9962 - 9449 = 513$) sur l'augmentation de masse moléculaire théorique à 100 % d'incorporation (correspondant ici à 129 ^{15}N et 418 ^{13}C , soit 547). Ainsi le pourcentage d'incorporation est ici de $(513 / 547) \times 100 = 93,8 \%$.

b) Surexpression et purification du domaine NLS-HD de la protéine Engrailed 2 (En2) de poulet

Les échantillons contenant le domaine NLS-HD de la protéine En2 (non marqué, simplement marqué ¹⁵N, et doublement marqué ¹⁵N/¹³C) ont été préparés au laboratoire par Sébastien Cardon[186] au cours de son doctorat avec la même stratégie de production et purification (**Figure 83a**). La séquence obtenue après coupure par l'enzyme PreScission correspond aux résidus 184-259 de la protéine En2 (**Figure 83b**). A l'instar d'Otx2, le domaine NLS-HD d'En2 est particulièrement basique avec 9 Arginines et 12 Lysines qui conduisent à un point isoélectrique théorique de 11,09, soit environ 0,5 unité de pH de plus que celui d'Otx2.

a) Protéine de fusion et principe de la stratégie de production de NLS-HD En2



b) Séquence après coupure par l'enzyme PreScission de NLS-HD de En2 : 80 acides aminés / 9,35 kDa, Point isoélectrique théorique : 11,09, (<http://isoelectric.org/calculate.php>)

GPAS¹⁸⁴GPRSRK PKKKNPVKED KRPR²⁰⁴TAFTAE QLQRLKAEFQ TNRYLTEQRR QSLAQELGLN
ESQIKIWFQN KRAKIKKAT²⁵⁹

En2(184-259) : domaine de liaison putatif aux GAGs, puis séquences des hélices H1, H2 et H3.

Composition en acides aminés : Ala (A) : 7, 8.8 % ; Arg(R) : 9, 11.2 % ; Asn (N) : 5, 6,2 % ; Asp (D) : 1, 1,2 % ; Cys (C) : 0, 0 % ; Gln (Q) : 8, 10,0 % ; Glu (E) : 6, 7,5 % ; Gly (G) : 3, 3,7 % ; His (H) : 0, 0 % ; Ile (I) : 3, 3,8 % ; Leu (L) : 6, 7,5 % ; Lys (K) : 12, 15,0 % ; Met (M) : 0, 0 % ; Phe (F) : 3, 3,8 % ; Pro (P) : 5, 6,2 % ; Ser (S) : 5, 6,2 % ; Thr (T) : 5, 6,2 % ; Trp (W) : 1, 1,2 % ; Tyr (Y) : 1, 1,2 % ; Val (V) : 0, 0 %.

Composition atomique : Carbone : 408 ; Hydrogène : 679 ; Azote : 133 ; Oxygène : 119 ; Soufre : 0. Coefficient d'extinction : 6990 M⁻¹.cm⁻¹ à 280 nm dans H₂O ; Abs 0.1 % (= 1 g/l) 0.747.

Figure 83. Expression et purification de NLS-HD En2. a) Construction de la protéine de fusion et principe de la stratégie et b) séquence de la protéine d'intérêt incluant le domaine 184-259 de En2 de poulet

2. RMN et études structurales

Les interactions mises en jeu entre une protéine et un ligand peuvent être étudiées par différentes techniques biophysiques dont la plupart sont accessibles au sein du laboratoire ou *via* des plateformes du campus. Je citerai le dichroïsme circulaire, la spectroscopie de fluorescence, la résonance plasmonique de surface (SPR), la calorimétrie de titrage isotherme (ITC), la spectrométrie de masse, la diffusion des rayons X aux petits angles et la RMN. Si certaines de ces techniques fournissent des données au niveau moléculaire, d'autres peuvent même nous procurer des informations au niveau atomique ou de l'acide aminé, informations susceptibles d'aider dans la conception de ligands de deuxième génération.

Les techniques les plus couramment utilisées pour appréhender les interactions mises en jeu à l'échelle de l'atome sont la radiocristallographie (RX) et la résonance magnétique nucléaire (RMN) de l'état liquide ou solide. La première nécessite l'obtention d'un cristal de la protéine étudiée. La RMN quant à elle nécessite que la protéine soit uniformément marquée avec différents isotopes (^{15}N , ^{13}C et/ou ^2D) en fonction de son poids moléculaire. Ces étapes sont généralement considérées à la fois comme limitantes et déterminantes.

La RMN liquide est certainement l'outil privilégié pour étudier la dynamique des interactions protéine-ligand, probablement car c'est la seule des deux méthodes qui fournit des informations sur les molécules en solution. L'analyse par RMN liquide étant limitée par la taille de la biomolécule étudiée, elle est généralement réservée à des protéines dont le poids moléculaire n'excède pas 30 à 40 kDa. Cependant, en quelques décennies les progrès instrumentaux et le développement de techniques de marquage isotopique sélectif ont permis de repousser cette limite de poids moléculaire, et il est aujourd'hui possible d'obtenir des informations structurales sur des protéines ou complexes de plusieurs mégadaltons. La première structure *de novo* d'une protéine globulaire, un inhibiteur de protéase (BPTI) a été déterminée par RMN liquide par le groupe de K. Wuthrich en 1985[188], [189]. À ce jour, plus de 13000 structures de protéines ont été déterminées par RMN et déposées dans la PDB (Protein Data Bank), entre 400 et 500 par an ces quinze dernières années[190]. Bien que le nombre de structures résolues par RMN ne représente à peine que 10 % des structures déposées dans la PDB, la RMN en solution reste cependant une technique de choix pour l'étude de la structure des protéines et de leurs interactions. Elle demeure la seule méthode

qui apporte des informations structurales et dynamiques haute résolution à l'état liquide, en l'occurrence la forme native d'un grand nombre de protéines *in vivo*. Elle est également la seule technique qui permette d'étudier des complexes de protéine flexibles ou semi-flexibles à l'échelle de l'atome, qui peuvent être rencontrés dans le cas d'interactions protéines-GAGs. En parallèle des informations structurales apportées à l'échelle de l'atome, la RMN permet l'obtention des paramètres cinétiques et thermodynamiques de l'interaction, avec en particulier la mesure des constantes de dissociation d'un complexe de protéine.

a) La Résonance Magnétique Nucléaire, les séquences et stratégies utilisées

Il est hors de propos et surtout hors de mon champ de compétences de présenter dans ce manuscrit un « cours de RMN » ou une introduction détaillée sur cette technique spectroscopique. Je ne présenterai donc que les grands principes des expériences de RMN biomoléculaire que j'ai utilisées, aidé par les lectures de ces dernières années et surtout par les nombreuses questions/discussions avec mes collègues, qui m'ont fait évoluer dans mes connaissances.

Attribution des signaux dans le spectre de la protéine

Dans toute étude structurale de protéine par RMN, la première étape consiste à attribuer les résonances ^1H , ^{15}N et ^{13}C de la chaîne principale de chaque acide aminé. Cette étape préliminaire est indispensable pour pouvoir ensuite obtenir des informations structurales ou identifier les acides aminés ou surfaces de la protéine impliqués dans une interaction avec un ligand.

Les protéines NLS-HD d'Otx2 et En2 contenant près de 700 atomes d'hydrogène, les expériences RMN classiques homonucléaires à 1 ou 2 dimensions (1D et 2D) ne sont pas adaptées pour étudier des molécules de cette taille et conduiraient à de nombreuses zones de recouvrement sur les spectres. Des expériences RMN triple résonance ($^1\text{H}/^{15}\text{N}/^{13}\text{C}$) à 3 dimensions (3D) ont été développées afin de limiter la quantité de signal sur les spectres et ainsi limiter les zones de recouvrement. Ces expériences 3D triple résonance fournissent des informations sur la structure (connectivité, distances interprotons, angles dièdres...) et la dynamique de la protéine.

Les expériences de RMN hétéronucléaire classiquement utilisées pour l'attribution de la chaîne principale d'une protéine sont basées sur le couplage entre les noyaux ^1H et ^{15}N de la liaison peptidique. Ce couplage peut être observé pour chaque acide aminé d'une protéine, à l'exception des Prolines qui sont dépourvues de proton amide. Il s'agit de faire corrélérer le proton amide ^1H d'un résidu i avec l'azote ^{15}N qui le porte. L'expérience de spectroscopie de corrélation utilisée pour enregistrer ce couplage est appelée ^1H - ^{15}N HSQC (heteronuclear single quantum correlation). Dans cette expérience 2D HSQC apparaîtront des pics de corrélation qui correspondent aux groupements amides du squelette de la protéine, ainsi qu'aux groupements NH_2 des chaînes latérales des résidus asparagine et glutamine. Le spectre 2D ^1H - ^{15}N HSQC présente donc l'avantage de ne faire apparaître qu'un seul pic de corrélation amide par acide aminé, et le nombre de pics sur le spectre doit correspondre globalement au nombre d'acides aminés de la séquence. C'est la raison pour laquelle cette expérience est très largement utilisée pour identifier les acides aminés impliqués dans une interaction avec un ligand. Le spectre ^1H - ^{15}N HSQC représente également une signature du repliement de la protéine car la distribution des pics sur le spectre (leur dispersion) dépend fortement du type de repliement.

Afin de pouvoir attribuer chaque pic du spectre HSQC à l'acide aminé correspondant, plusieurs expériences 3D de triple résonance ont été développées afin de faire corrélérer le proton amide ^1H d'un résidu i avec l'azote ^{15}N qui le porte et un carbone $^{13}\text{C}_i$ du même résidu et/ou $^{13}\text{C}_{i-1}$ du résidu précédent dans la séquence de la protéine. Les expériences 3D ainsi réalisées sont nommées HNCO, HN(CA)CO, HNCA, CBCA(CO)NH et HNCACB[191]–[193].

Les objectifs des différentes expériences sont résumés ci-dessous :

- HNCO : corrélation du groupement amide du résidu i avec le carbonyle $^{13}\text{C}_{i-1}$ du résidu précédent $i-1$,
- HN(CA)CO : corrélation du groupement amide du résidu i avec les carbonyles $^{13}\text{C}_i$ du résidu i et du résidu précédent $i-1$,
- HNCA : corrélation du groupement amide du résidu i avec les $^{13}\text{C}_\alpha$ du résidu i et du résidu précédent $i-1$,
- CBCA(CO)NH : corrélation du groupement amide du résidu i avec le $^{13}\text{C}_{\alpha-1}$ et le $^{13}\text{C}_{\beta-1}$ du résidu précédent $i-1$,

- HNCACB : corrélation du groupement amide du résidu i avec les $^{13}\text{C}_\alpha$ et $^{13}\text{C}_\beta$ du résidu i et du résidu précédent $i-1$.

Les noyaux impliqués dans les transferts d'aimantation mis en jeu dans les expériences CBCA(CO)NH et HNCACB sont indiqués dans la **Figure 84**.

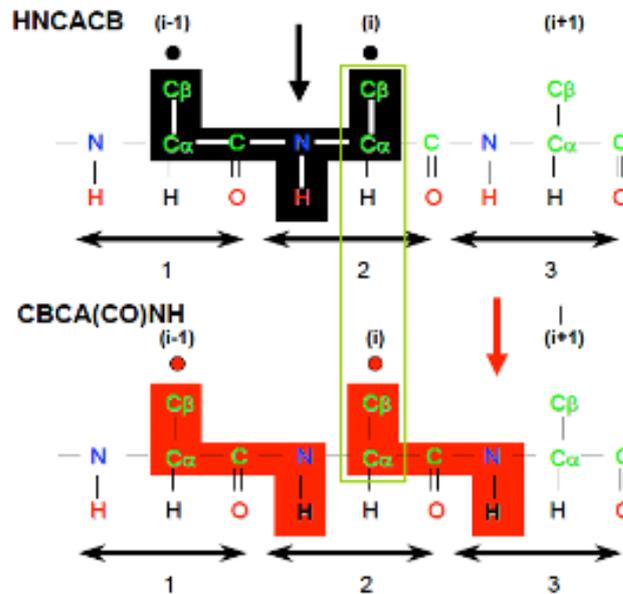


Figure 84. Principe des expériences HNCACB et CBCA(CO)NH. Les noyaux impliqués dans les transferts d'aimantation sont indiqués par un fond noir pour l'HNCACB et un fond rouge pour la CBCA(CO)NH. Le groupement amide ^1H - ^{15}N du résidu (i) est corrélé à la paire $^{13}\text{C}_\alpha$ - $^{13}\text{C}_\beta$ des résidus (i) et ($i-1$) dans l'expérience HNCACB alors qu'il est uniquement corrélé à la paire $^{13}\text{C}_\alpha$ - $^{13}\text{C}_\beta$ du résidu ($i-1$) dans l'expérience CBCA(CO)NH. Si des valeurs identiques de déplacements chimiques $^{13}\text{C}_\alpha$ - $^{13}\text{C}_\beta$ sont retrouvées pour 2 groupements amides différents respectivement en i et $i-1$ (comme indiqué dans le cadre vert), alors ces 2 groupements amides indiqués par des flèches noire et rouge sont voisins dans la séquence peptidique[194].

Les informations recueillies avec ces expériences 3D permettent d'identifier les déplacements chimiques des noyaux ^{13}CO , $^{13}\text{C}_\alpha$ et $^{13}\text{C}_\beta$ des résidus i et $i-1$ relatifs à chaque groupement amide. Ainsi, en comparant les valeurs de déplacements chimiques $^{13}\text{C}_i$ aux valeurs de déplacements chimiques $^{13}\text{C}_{i-1}$, il est possible de relier les systèmes de corrélation amides deux à deux, c'est ce qu'on appelle l'attribution séquentielle. Celle-ci est alors positionnée dans la séquence primaire de la protéine, ici information connue, en s'appuyant sur les acides aminés dont les déplacements chimiques sont caractéristiques (par exemple la Glycine qui ne possède pas de C_β ou les Thréonine, Sérine et Alanine dont les valeurs de C_α et

C_β sont spécifiques). Cette méthode permet d'attribuer à chaque pic de corrélation amide du spectre HSQC l'acide aminé correspondant.

Mesure par RMN de l'affinité entre ligand et protéine

Pour déterminer l'affinité de nos ligands envers les protéines Otx2 et En2, nous avons enregistré des expériences de perturbation de déplacement chimique. La méthode consiste à suivre les variations de déplacement chimique de chaque pic sur un spectre ¹H-¹⁵N HSQC lors du titrage d'une protéine par un ligand. Il est à noter que cette méthode ne permet d'obtenir des informations que du point de vue de la protéine ; les ligands ayant été synthétisés en abondance naturelle, ils seront invisibles sur les spectres de RMN hétéronucléaire.

Sur un spectre RMN, l'interaction d'une protéine avec un ligand peut prendre plusieurs formes en fonction du régime d'échange et notamment de la vitesse d'interconversion entre la forme libre et la forme liée. Dans le cas d'une affinité faible ($K_d > 1-10 \mu\text{M}$), le système est dit **en échange rapide** et l'interaction aura principalement pour conséquence des variations de déplacement chimique et éventuellement quelques élargissements de raies. Dans le cas d'un équilibre à deux états, chaque noyau est caractérisé par 2 déplacements chimiques : celui de la forme libre δ_f (free) et de la forme liée δ_b (bound). Cependant, comme le montre la **Figure 85a**, en échange rapide la valeur mesurée de déplacement chimique pourra varier entre la forme libre et la forme liée en fonction de la proportion de forme liée en solution. En ajoutant des quantités croissantes de ligand, on observera alors des déplacements de pics sur le spectre HSQC avec des trajectoires linéaires de la position δ_f (free) à δ_b (bound) s'il s'agit d'un équilibre à deux états (**Figure 85b**). Le déplacement chimique d'un noyau étant fonction de son environnement chimique et électronique, les pics seront plus ou moins affectés par la présence du ligand. Les résidus directement impliqués dans l'interaction avec le ligand ou subissant une modification conformationnelle induite par l'interaction présenteront les perturbations de déplacement chimique (CSP) ¹H et ¹⁵N les plus importantes. Pour chaque groupement amide, la variation moyenne de déplacements chimiques est quantifiée en utilisant la formule $\text{CSP} = \sqrt{(\Delta\delta^1\text{H}_\text{N})^2 + (0,1\Delta\delta^{15}\text{N}_\text{H})^2}$, où $\Delta\delta^1\text{H}_\text{N}$ est la variation du déplacement chimique du proton amide et $\Delta\delta^{15}\text{N}_\text{H}$ celle de l'azote[195]. Ainsi les acides aminés les plus affectés par l'interaction avec le ligand auront un CSP élevé alors que ceux les moins affectés auront un CSP proche de zéro.

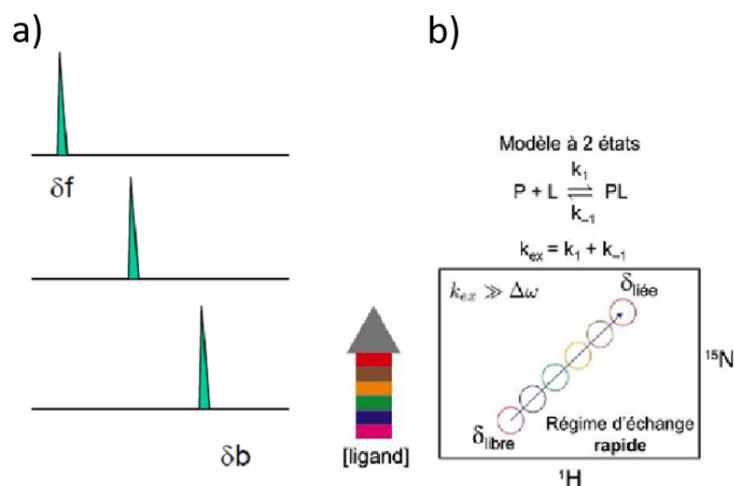


Figure 85. a) Représentation schématique d'un titrage par RMN d'un ligand avec une protéine en échange rapide. En haut à gauche, le spectre représente la situation de départ où aucun ligand n'a été ajouté. Le signal RMN est à δ_f . La ligne médiane est la situation où suffisamment de ligand a été ajouté pour saturer la moitié des sites de liaison dans la protéine. Ligne d'en bas : le ligand ajouté occupe tous les sites de liaison et le signal RMN est maintenant à δ_b [194]. **b) Evolution des pics de corrélation sur un spectre HSQC ^1H - ^{15}N lors d'une expérience de titrage en échange rapide, dans un modèle à 2 états** (Figures adaptées de la thèse de Cyril Charlier soutenue en 2015).

D'autre part, la valeur de déplacement chimique mesurée lors du titrage étant directement liée à la quantité de complexe en solution, il est possible de calculer des constantes de dissociation K_d dites apparentes pour chaque acide aminé suffisamment affecté par la présence du ligand, à partir d'une fonction mathématique du second degré[196] (**Figure 86**) en ajustant les courbes des CSP en fonction du nombre d'équivalents de ligand introduit[197], comme représenté **Figure 87**.

$$\text{CSP} = \text{CSPmax} \left\{ \frac{([P]_t + [L]_t + K_d) - \sqrt{([P]_t + [L]_t + K_d)^2 - 4[P]_t[L]_t}}{2} \right\} / [P]_t$$

Avec : $[P]_t$ = concentration totale de la protéine

$[L]_t$ = concentration totale du ligand

CSPmax = CSP maximale obtenue pour 100 % de protéine liée

Figure 86. Equation permettant d'obtenir les K_d apparents des acides aminés impliqués dans l'interaction entre une protéine et un unique ligand (ratio 1 :1), en échange rapide[196].

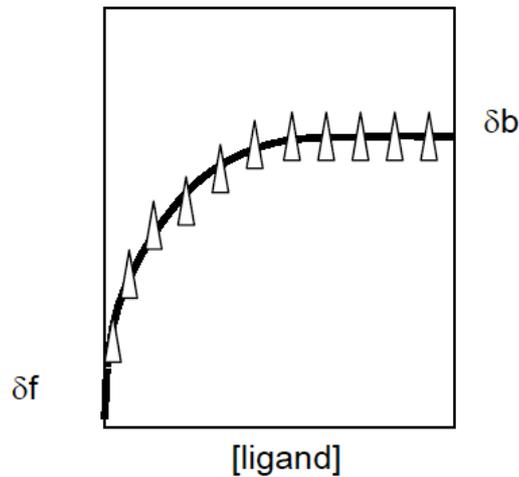


Figure 87. Courbe de liaison d'un titrage. La constante de dissociation, K_d , peut être déterminée à partir de la courbe de liaison[194]

b) Résultats des études structurales par RMN

Interactions avec Otx2

Attribution des signaux du spectre HSQC du domaine NLS-HD d'Otx2 doublement marqué $^{15}\text{N}/^{13}\text{C}$:

L'attribution des noyaux de la chaîne principale des 80 résidus du domaine NLS-HD d'Otx2 a été obtenue selon la méthode présentée plus haut. Des expériences 2D HSQC ^1H - ^{15}N (**Figure 88**) et 3D triple résonance (HNCO, HN(CA)CO, HNCA, CBCA(CO)NH et HNCACB) ont été enregistrées à 25 °C et à pH 6,3 sur un spectromètre 500 MHz équipé d'une cryosonde $^1\text{H}/^{15}\text{N}/^{13}\text{C}$. Un échantillon unique de protéine doublement marquée à 500 μM a été suffisant pour obtenir l'attribution séquentielle.

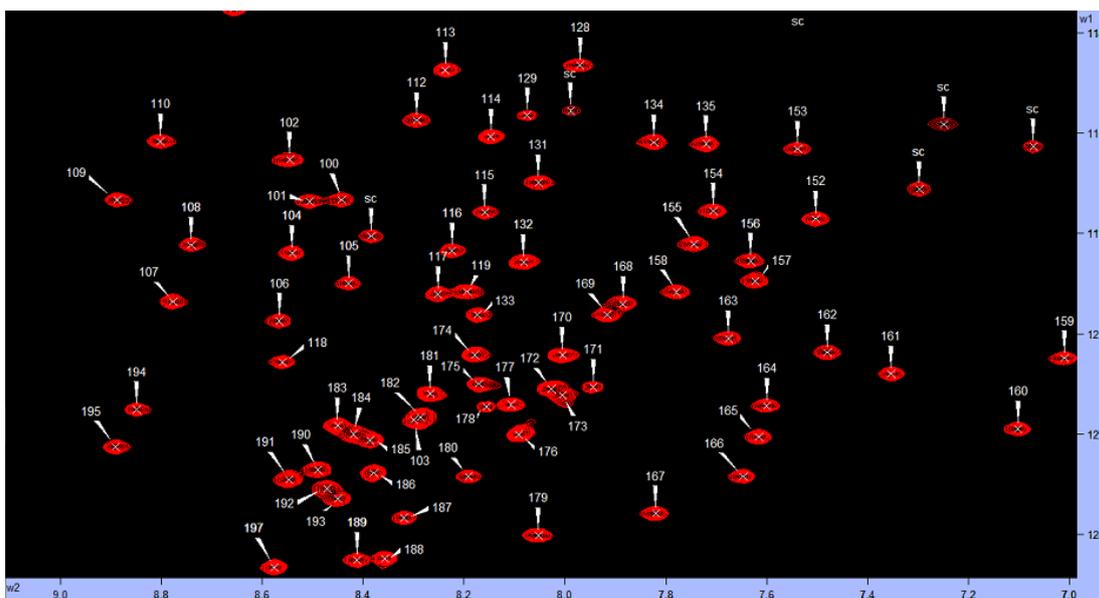


Figure 88. Spectre ^1H - ^{15}N HSQC numéroté et non attribué de NLS-HD Otx2 doublement marqué. Les numéros ont été ajoutés arbitrairement.

Les expériences 3D triple résonance ont permis d'obtenir, pour chaque corrélation amide ^1H - ^{15}N du spectre HSQC, les déplacements chimiques des noyaux ^{13}CO , $^{13}\text{C}_\alpha$ et $^{13}\text{C}_\beta$ du même résidu (intra-résiduelles) et du résidu le précédent dans la séquence peptidique. Toutes ces valeurs de déplacements chimiques ont été collectées sous forme de tableau. Une partie des résultats, pour exemple, est reportée dans le **Tableau 8**.

	15N	1H		C α	C β	CO	C α -1	C β -1	CO-1
100	117,331	8,442	100	63,642	37,891	175,446	52,867	41,452	176,334
101	117,344	8,506	101	60,723	31,57	178,28	58,838	33,873	179,838
102	116,52	8,545	102	58,436	63,937	174,712	63,444	32,232	177,083
104	118,387	8,541	104	57,654	39,991	179,751	57,681	41,594	178,583
105	118,987	8,429	105	57,722	41,565	178,59	58,79	28,571	177,943
106	119,743	8,565	106	60,205	28,38	178,707	67,11	31,949	177,677
107	119,362	8,779	107	59,611	29,345	178,958	60,762	31,627	178,277
108	118,214	8,741	108	64,082	39,577	177,781	63,235	28,469	179,489
109	117,338	8,888	109	59,21	29,046	180,103	58,554	38,416	178,222
110	116,156	8,804	110	59,847	68,978	181,861	58,065	33,366	177,307
111	113,494	8,655	111	61,463	???	178,27	60,976	28,533	178,261
112	115,73	8,295	112	61,893	69,787	174,369	56,138	30,964	176,421
113	114,734	8,238	113	58,449	63,705	174,898	52,622	19,159	177,927
114	116,052	8,149	114	59,955	69,829	172,904	52,333	19,3	177,732
115	117,582	8,16	115	60,372	31,742	178,722	56,712	28,303	179,541
116	118,335	8,226	116	55,485	37,971	177,307	60,358	32,292	179,905
117	119,204	8,253	117	54,968	18,337	181,347	58,964	30,129	177,634
118	120,57	8,559	118	58,844	33,723	179,846	60,696	37,757	179,542

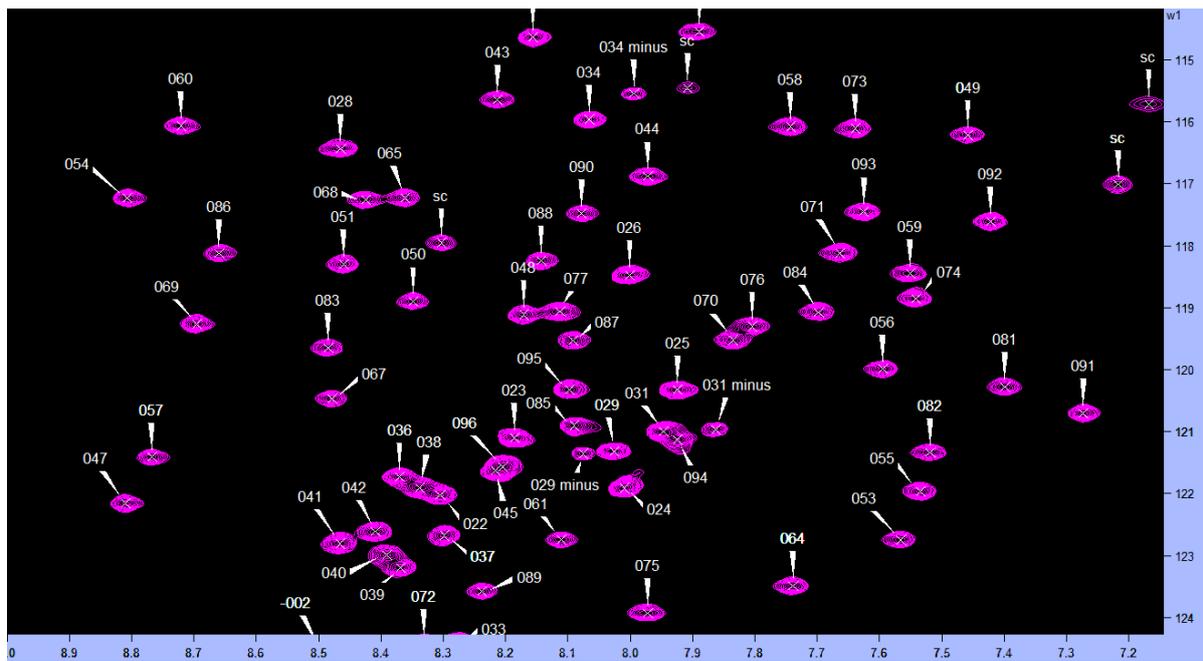
Tableau 8. Quelques valeurs, pour exemple, de déplacements chimiques des noyaux ^{13}CO , $^{13}\text{C}_\alpha$ et $^{13}\text{C}_\beta$ associés à un couple ^1H - ^{15}N , obtenues grâce aux expériences 3D triple résonance.

La comparaison des valeurs de déplacements chimiques $^{13}\text{C}_i$ aux valeurs de déplacements chimiques $^{13}\text{C}_{i-1}$ a permis de connecter les systèmes de corrélation deux à deux et de les situer dans la séquence de la protéine à partir des valeurs attendues pour chaque type d'acide aminé. L'attribution de la région 22-39 d'Otx2 avec cette approche numérique est montrée pour exemple dans le **Tableau 9**.

Numérotation arbitraire	15N	1H	C α	C β	CO	C α -1	C β -1	CO-1	Séquence primaire	Numérotation de la séquence
185	122,123	8,387	55,784	32,972	175,933	58,535	63,697	174,89	M	22
181	121,203	8,268	54,393	41,014	176,332	55,793	33,015	175,918	D	23
176	121,997	8,092	55,57	42,224	177,497	54,363	41,021	176,339	L	24
170	120,417	8,007	55,336	42,251	176,858	55,333	42,274	177,476	L	25
132	118,564	8,084	53,36	29,092	172,332	55,251	42,311	176,849	H	26
									P	27
102	116,52	8,545	58,436	63,937	174,712	63,444	32,232	177,083	S	28
177	121,404	8,11	62,448	32,839	176,406	58,442	63,827	174,721	V	29
124	111,789	8,335	44,926		173,276	62,446	32,856	176,393	G	30
172	121,087	8,029	55,841	38,363	174,071	44,922		173,255	Y	31
									P	32
188	124,484	8,356	52,376	19,326	177,744	63,133	31,999	176,441	A	33
114	116,052	8,149	59,955	69,829	172,904	52,333	19,3	177,732	T	34
									P	35
183	121,838	8,45	56,29	30,881	176,551	63,291	32,241	176,935	R	36
186	122,785	8,379	56,384	33,087	176,53	56,389	30,903	176,543	K	37
184	121,997	8,419	55,782	29,682	175,898	56,449	33,111	176,508	Q	38
193	123,297	8,45	56,159	30,82	176,237	56,045	29,734	175,881	R	39

Tableau 9. Attribution de la région 22-39 d'Otx2 à partir des déplacements chimiques C α , C β et CO relevés sur les spectres RMN 3D HNCA, HNCOC, HN(CA)CO, CBCA(CO)NH et HNCACB.

Tous les résidus de la séquence d'intérêt 22-97 d'Otx2, à l'exception des Prolines, ont ainsi été identifiés sur le spectre HSQC ^1H - ^{15}N . L'attribution finale des pics de corrélation amide sur le spectre HSQC ^1H - ^{15}N du domaine NLS-HD d'Otx2 est présentée dans la **Figure 89**.



**NLS-HD Otx2 : ASM²²DLLHPSVG³⁰YPATPRKQRR⁴⁰ERTTFTRAQL⁵⁰DVLEALFAKT⁶⁰
RYPDIFMREE⁷⁰VALKINLPES⁸⁰RVQVWFKNRR⁹⁰AKCRQQQ⁹⁷**

Figure 89. En haut : Spectre ^1H - ^{15}N HSQC attribué de NLS-HD Otx2 doublement marqué. Le numéro attribué à chaque pic se rapporte à celui de l'acide aminé correspondant dans la séquence de la protéine. En bas : Séquence du domaine NLS-HD d'Otx2.

Détermination par RMN de la structure secondaire du domaine NLS-HD d'Otx2 :

Wishart et Sykes sont à l'origine de la technique du « Chemical-Shift Deviation » (CSD ou déplacement chimique secondaire) qui permet de localiser et d'identifier les structures secondaires de protéines à partir des valeurs de déplacements chimiques des noyaux $^1\text{H}_\alpha$, $^{13}\text{C}_\alpha$, $^{13}\text{C}_\beta$ et ^{13}CO des acides aminés qui les composent [198], [199]. La méthode CSD consiste à comparer la valeur expérimentale des déplacements chimiques de ces noyaux, pour chaque acide aminé, avec des valeurs théoriques, pour les mêmes acides aminés, obtenues par l'analyse RMN de peptides non structurés. Ainsi une différence significative (en valeur absolue) entre la valeur théorique et la valeur expérimentale, pour une série d'acides aminés adjacents, sera caractéristique d'une structure secondaire stable. Concernant plus particulièrement les $^{13}\text{C}_\alpha$, une structure en hélice α est observée lorsque les valeurs expérimentales sont nettement supérieures aux valeurs théoriques ($\text{CSD} > 1$) et une structure en brin β est observée lorsqu'elles sont nettement inférieures ($\text{CSD} < -1$).

Les structures secondaires de NLS-HD Otx2 ont pu être localisées et identifiées en comparant les valeurs des déplacements chimiques des $^{13}\text{C}_\alpha$ de chaque résidu de la protéine avec les valeurs expérimentales pour les $^{13}\text{C}_\alpha$ des mêmes acides aminés dans un peptide non structuré[200] (**Figure 90**).

Représentées dans la **Figure 90** par des successions de « barres » d'amplitude supérieure à 1, les 3 hélices α caractéristiques de l'homéodomaine de la protéine Otx2 sont bien présentes, ce qui tend à démontrer que NLS-HD Otx2 produit est correctement structuré en solution. Des valeurs de CSD de faible amplitude (CSD ≈ 0) sont observées dans la région N-terminale précédant l'homéodomaine, ce qui indique que cette région de la protéine ne présente pas de structure secondaire stable.

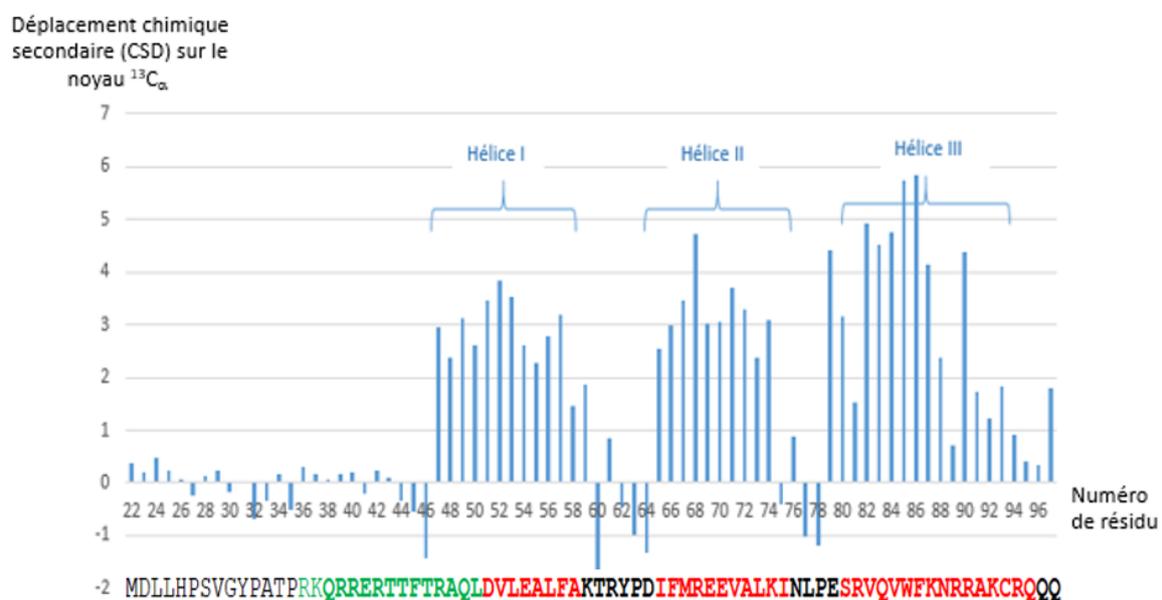


Figure 90. Histogramme représentant les valeurs de déplacements chimiques secondaires (CSD) des noyaux $^{13}\text{C}_\alpha$ des acides aminés du domaine NLS-HD d'Otx2 (région 22-97)[200].

Dans la séquence de la protéine (en bas de l'histogramme) sont représentés en vert le domaine de liaison aux GAGs, en gras l'homéodomaine et en rouge les hélices α .

Mesure par RMN de l'affinité entre un ligand et NLS-HD Otx2 marqué en ^{15}N :

Des échantillons du domaine NLS-HD d'Otx2 uniformément enrichi en isotope ^{15}N et ^{13}C ont été étudiés à la concentration de 150 μM afin d'évaluer l'affinité de la protéine vis-à-vis de différents ligands, par la méthode RMN de perturbation de déplacement chimique. Pour cela, des spectres 2D HSQC $^1\text{H}/^{15}\text{N}$ ont été enregistrés, avant et après ajouts successifs et

modérés (entre 0,12 et 0,4 équivalent molaire) de solutions concentrées de ligand, jusqu'à ne plus observer de modification sur le spectre HSQC.

Les ligands utilisés ont été :

- Ac-Ahx-(EC'AC')₆-NH₂, Ac-Ahx-(EC'A)₆-NH₂ et Ac-Ahx-(EC'AC')₃-NH₂ : 3 sulfopeptides, mimes de GAGs.
- HI dp8 : un octasaccharide commercial (Iduron) dérivé d'héparine, dont 75 % approximativement se trouve sous la forme UA,2S-GlcNS,6S-(IdoA,2S - GlcNS,6S)₃, **Figure 91a**, le reste étant sous des formes moins sulfatées. Ce fragment de GAG sulfaté a servi de référence.
- PolyGlu : un peptide commercial (Iris Biotech) constitué de 20 résidus d'acide glutamique, n-Butyl-poly(L-glutamic acid)₂₀ sodium salt, **Figure 91b**. Ce ligand, ne présentant que des charges négatives provenant de fonctions carboxylates uniquement, a été utilisé afin de déterminer si la reconnaissance des GAGs et mimes de GAGs était seulement basée sur la densité de charges négatives exposées aux protéines ou si la nature des fonctions et leur répartition pouvaient jouer un rôle dans cette reconnaissance.

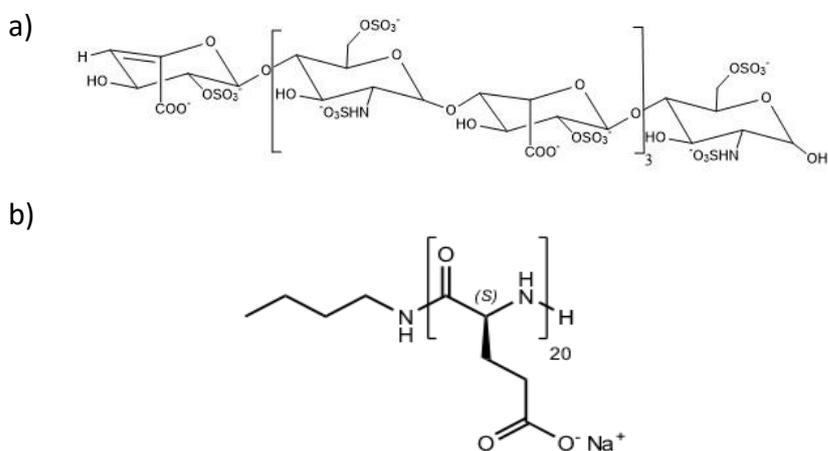


Figure 91. Formules de a) HI dp8 et b) PolyGlu

Pour chaque ligand étudié, des variations de déplacement chimique ont été observées dans les spectres 2D HSQC au fur et à mesure des ajouts (voir pour exemple la **Figure 92**). Cela atteste d'une interaction entre le domaine NLS-HD d'Otx2 et le ligand. Cette interaction perturbe l'environnement électronique et donc modifie le déplacement chimique des noyaux ¹H et ¹⁵N des acides aminés de la protéine impliqués dans l'interaction. De plus, les trajectoires

linéaires observées confirment qu'il s'agit d'un modèle à 2 états du régime d'échange rapide pour lequel un seul pic par résidu est observé, à une position intermédiaire entre celle du résidu dans la protéine libre (en rouge dans la **Figure 92**) et celle du résidu dans la protéine liée au ligand (en rose dans la **Figure 92**).

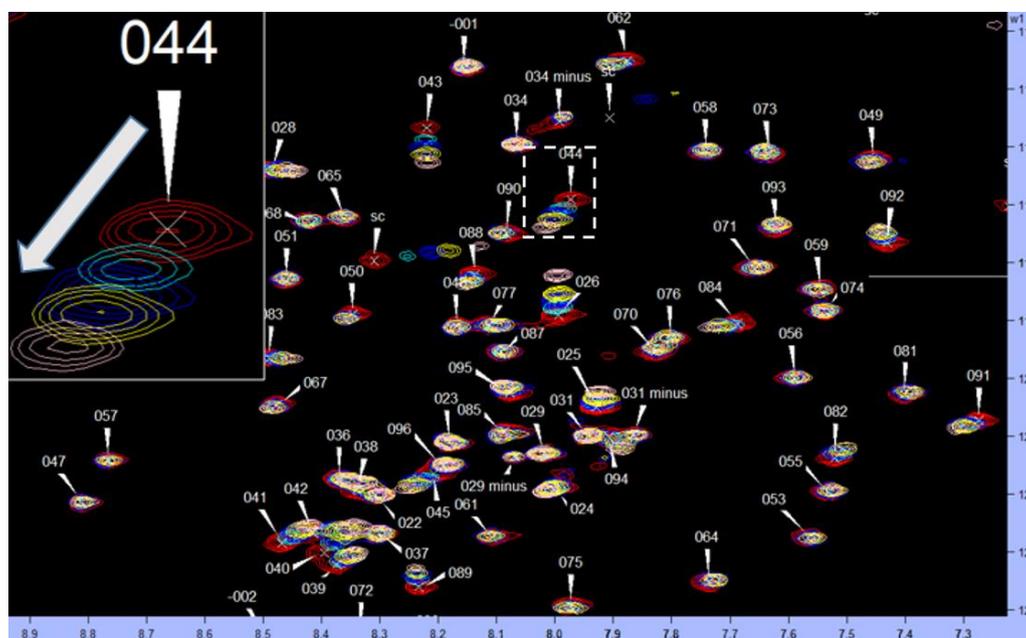


Figure 92. Superposition des spectres 2D HSQC ^1H - ^{15}N du domaine NLS-HD d'Otx2, avant (en rouge) et après ajouts successifs de Ac-Ahx-(EC'AC') $_6$ -NH $_2$ (cyan : 0,26 eq / bleu : 0,52 eq / jaune : 0,78 eq et rose : 1,9 eq).

A partir de ces expériences, les valeurs des déplacements chimiques des noyaux ^1H ($\delta^1\text{H}_\text{N}$) et ^{15}N ($\delta^{15}\text{N}_\text{H}$) des acides aminés de la protéine, sous sa forme libre et liée, ont été extraites afin de pouvoir calculer les perturbations de déplacement chimique (CSP). Les valeurs de CSP ont ensuite été représentées sous forme d'histogrammes pour chaque ligand ce qui permet de visualiser les zones de NLS-HD Otx2 perturbées par l'ajout de ligand (voir pour exemple la **Figure 93**).

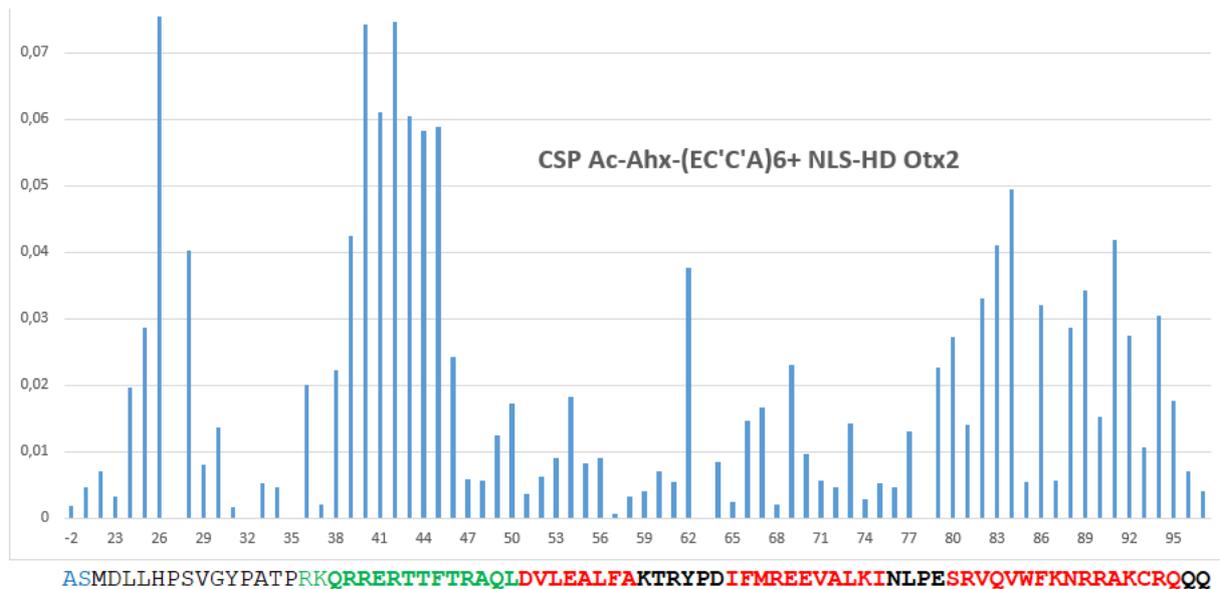


Figure 93. CSP amides (en ppm) de NLS-HD Otx2 en interaction avec 2,6 eq de Ac-Ahx-(EC'C'A)₆-NH₂. Dans la séquence de la protéine, en bas de l’histogramme, sont représentés en vert le domaine de liaison aux GAGs, en gras l’homéodomaine et en rouge les hélices α .

Il est important de noter que les CSP mesurés dans les expériences RMN sont ceux des groupements amides du squelette peptidique qui sont plus ou moins éloignés des groupements fonctionnels en bout de chaîne latérale selon le type d’acides aminés. Par conséquent, une valeur élevée de CSP amide pour un résidu donné peut être la conséquence de l’interaction directe du ligand avec la chaîne latérale de ce résidu ou celle d’un résidu adjacent qui modifie l’environnement électronique des acides aminés voisins. Je me garderai donc de tirer des conclusions hâtives sur des CSP d’acides aminés isolés et raisonnerai en termes de région ou surface d’interaction à partir des régions qui présentent les valeurs de CSP les plus élevées.

Il est également important de noter que les valeurs de CSP ne peuvent être directement utilisées pour tirer des conclusions sur les affinités des ligands envers la protéine. Il n’y a en effet pas de relation directe entre l’intensité des CSP et l’affinité d’interaction, comme nous pourrions le constater par exemple pour His²⁶, et seul le calcul des K_d apparents permet d’apporter des informations raisonnables et comparables concernant cette affinité. C’est pourquoi l’analyse des résultats se fera en considérant parallèlement les régions présentant des valeurs élevées de CSP et les K_d apparents correspondants.

A partir des spectres 2D HSQC $^1\text{H}/^{15}\text{N}$ obtenus après chaque ajout de ligand, les valeurs des déplacements chimiques des noyaux ^1H ($\delta^1\text{H}_\text{N}$) et ^{15}N ($\delta^{15}\text{N}_\text{H}$) des acides aminés de la protéine présentant des valeurs de CSP élevées ont été extraites et utilisées pour tracer les courbes des CSP en fonction du nombre d'équivalents de ligand introduit. L'ajustement de ces courbes a permis d'obtenir des constantes de dissociation apparentes (K_d) pour les résidus d'intérêt. Des exemples de courbes de liaison, avec les résidus 26 (His de la région N-terminale), 42 (Arg de la région de liaison aux GAGs) et 92 (Lys de la région de l'hélice 3), pour l'interaction de NLS-HD Otx2 avec Ac-Ahx-(EC'C'A) $_6$ -NH $_2$, sont reportés dans la **Figure 94**.

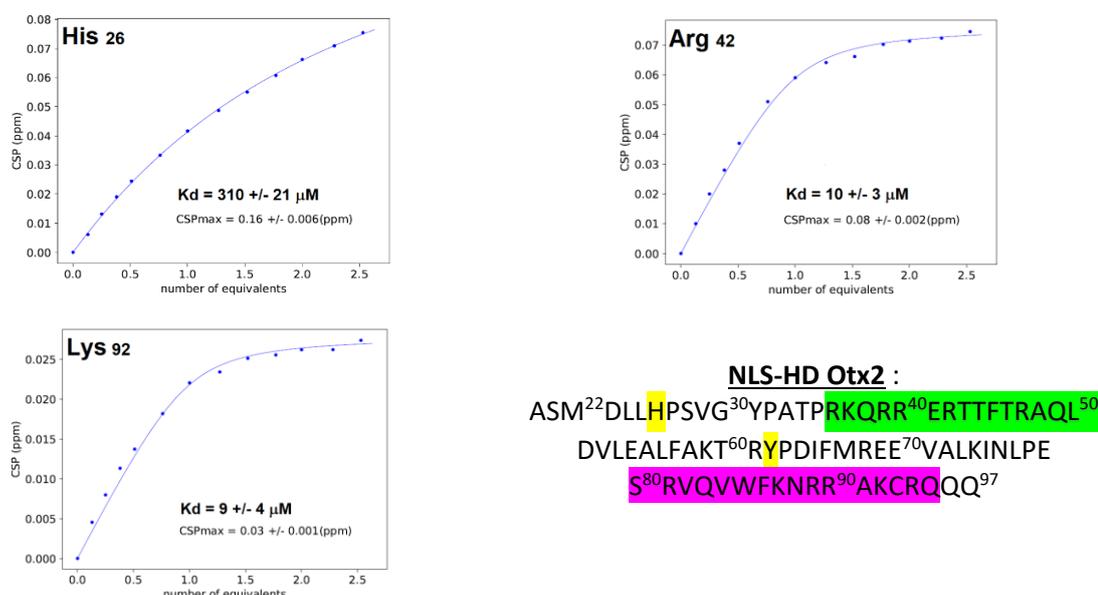
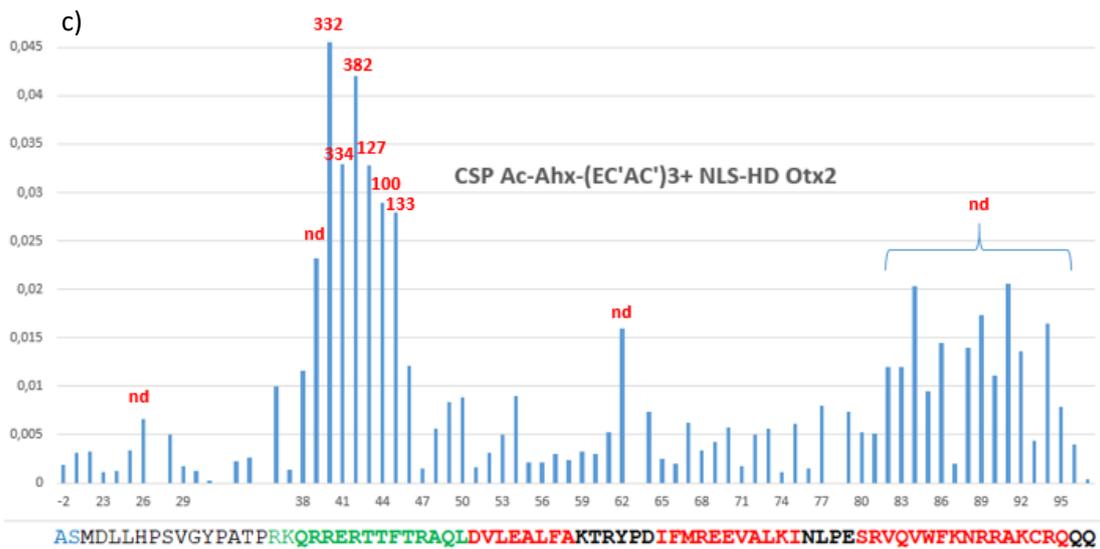
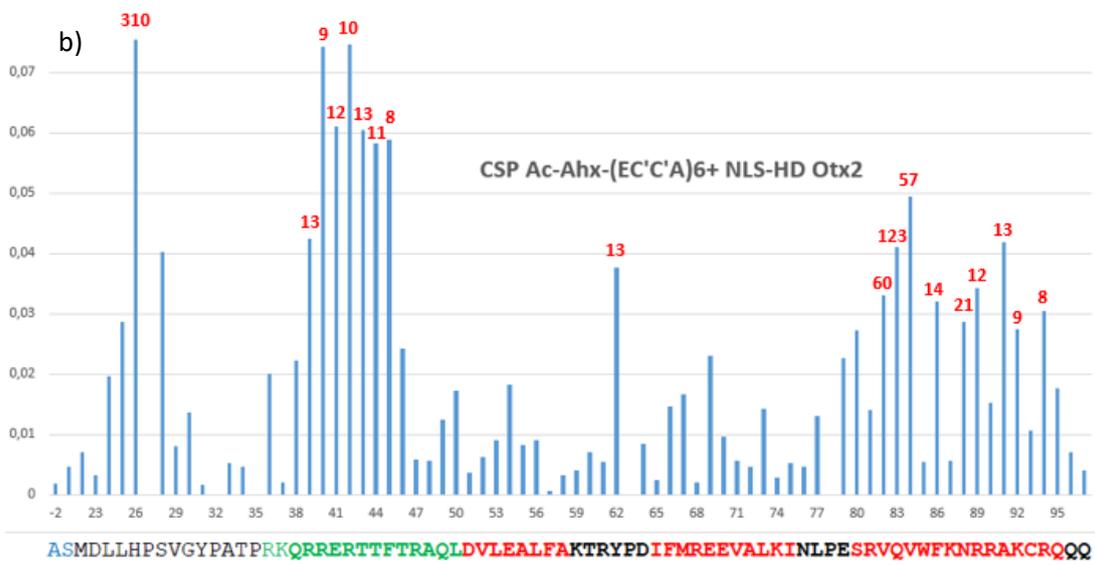
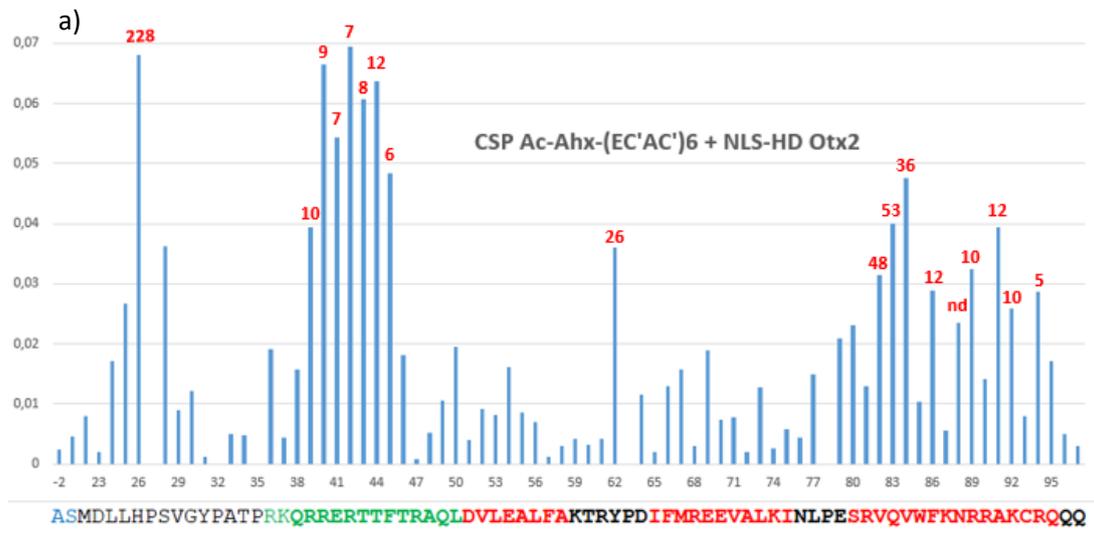


Figure 94. Exemples de courbes de liaison permettant la détermination des K_d apparents des résidus His²⁶, Arg⁴² et Lys⁹² du domaine NLS-HD d'Otx2 en interaction avec Ac-Ahx-(EC'C'A) $_6$ -NH $_2$.

La valeur des K_d apparents (en μM) des résidus les plus affectés par l'interaction protéine/ligand est directement reportée, en rouge, sur les histogrammes de la **Figure 95**. Pour davantage de lisibilité, les erreurs standards n'ont pas été indiquées sur ces histogrammes, mais reportées dans la partie expérimentale de ce chapitre (**Tableau 13**).



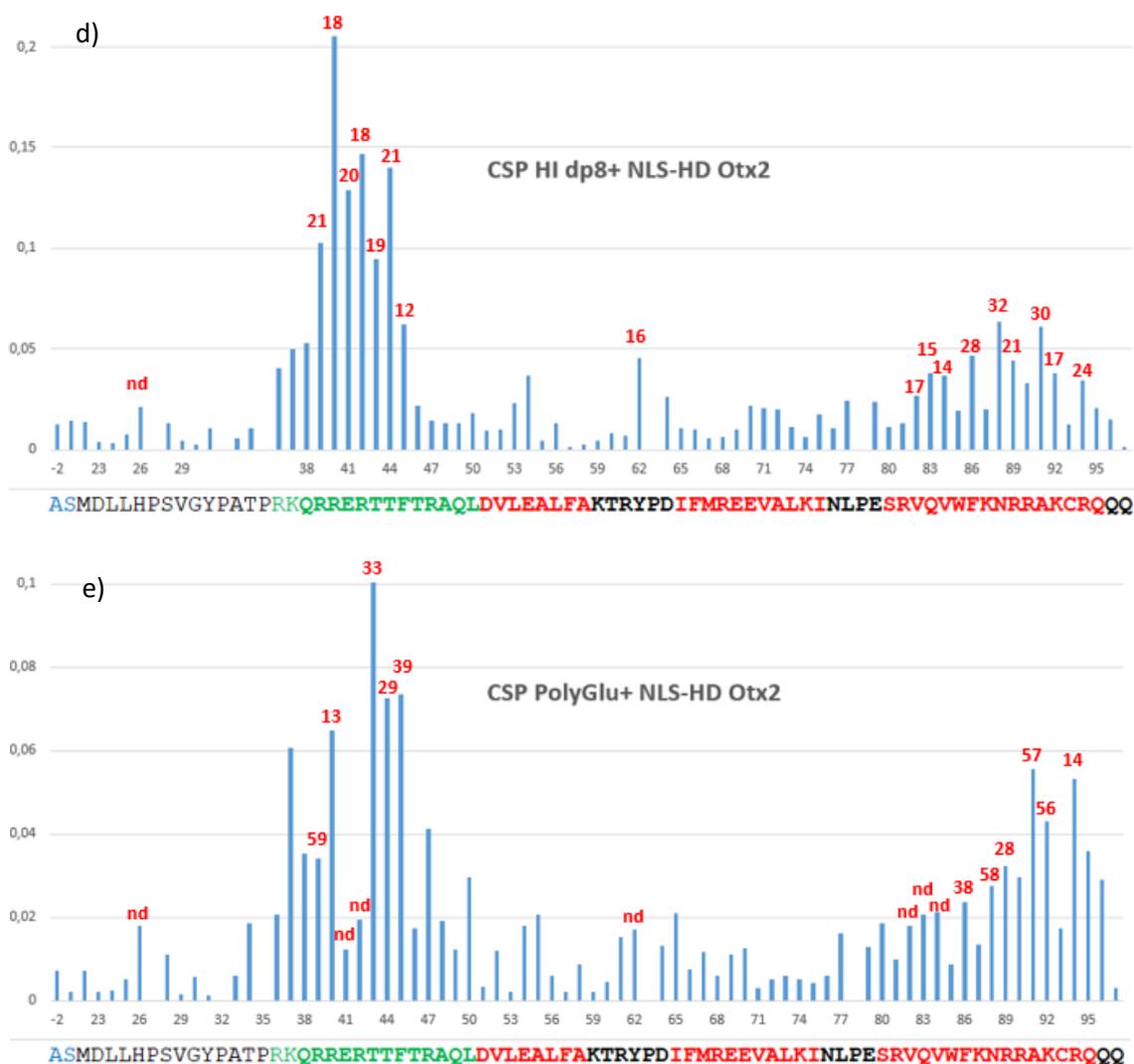


Figure 95. CSP amides (en ppm) de NLS-HD Otx2 en interaction avec a) 1,9 eq de Ac-Ahx-(EC'AC')₆-NH₂, b) 2,6 eq de Ac-Ahx-(EC'C'A)₆-NH₂, c) 8,4 eq de Ac-Ahx-(EC'AC')₃-NH₂, d) 2,8 eq de HI dp8 et e) 4 eq de PolyGlu. La valeur des K_d apparents (en μM) des résidus d'intérêt est indiquée en rouge (« nd » signifie non déterminé). Dans la séquence de la protéine (en bas de chaque histogramme) sont représentés en vert le domaine de liaison aux GAGs, en gras l'homéodomaine et en rouge les hélices α. Rq : il est à noter que les échelles pour les CSP ne sont pas les mêmes pour les différents ligands étudiés.

Les valeurs de CSP, représentées ci-dessus sous forme d'histogramme, ont également été reportées dans un tableau (**Tableau 10**), en les présentant en valeur relative, à savoir en normalisant à 100 % l'acide aminé avec le CSP le plus élevé (CSP+) pour chaque ligand, les autres CSP étant exprimés en % de cette valeur maximale. Les CSP compris entre 40 % et 100 % sont indiqués en vert. Ce mode de représentation permet de mieux « visualiser » quelques points que je présenterai ci-dessous.

Ratio CSP du domaine de liaison aux GAGs (100 % pour le CSP le plus important, CSP+)

*	CSP+	n eq	36 R	37 K	38 Q	39 R	40 R	41 E	42 R	43 T	44 T	45 F	46 T	47 R	48 A	49 Q	50 L
(EC'AC') ₆	0,07	1,9	27	6	22	57	95	74	100	86	90	70	27	0	7	14	27
(EC'C'A) ₆	0,075	2,6	27	3	30	57	100	82	100	81	78	79	33	8	8	17	23
(EC'AC') ₃	0,045	8,4	21	2	24	51	100	73	93	72	64	62	26	3	11	18	19
HI dp8	0,2	2,8	20	24	26	50	100	63	72	46	68	30	11	7	7	6	9
PolyGlu	0,1	4	20	60	34	33	64	11	19	100	73	74	18	41	19	12	29

Ratio CSP de l'Hélice 3 (valeur relative)

*	CSP+	n eq	80 S	81 R	82 V	83 Q	84 V	85 W	86 F	87 K	88 N	89 R	90 R	91 A	92 K	93 C	94 R	95 Q
(EC'AC') ₆	0,07	1,9	33	17	44	57	69	16	43	8	20	46	19	57	37	11	41	24
(EC'C'A) ₆	0,075	2,6	36	19	44	55	66	7	43	7	38	46	20	56	37	14	41	24
(EC'AC') ₃	0,045	8,4	12	12	26	26	45	21	32	4	31	38	24	44	31	10	36	17
HI dp8	0,2	2,8	6	6	13	18	18	9	23	10	31	22	16	30	18	6	17	10
PolyGlu	0,1	4	19	10	18	21	21	9	23	14	28	32	30	56	43	17	53	36

Ratio CSP de His²⁶ et Tyr⁶² (valeur relative)

*	CSP+	n eq	26 H		62 Y
(EC'AC') ₆	0,07	1,9	98		51
(EC'C'A) ₆	0,075	2,6	100		51
(EC'AC') ₃	0,045	8,4	15		36
HI dp8	0,2	2,8	10		22
PolyGlu	0,1	4	19		18

* : Tous les sulfopeptides ont pour formule générale Ac-Ahx-(XXXX)_n-NH₂, avec Ahx = acide 6-amino-hexanoïque.

NLS-HD Otx2 : ASM²²DLLHPSVG³⁰YPATPRKQRR⁴⁰ERTTFTRAQL⁵⁰DVLEALFAKT⁶⁰RYPDIFMREE⁷⁰
VALKINLPES⁸⁰RVQVWFKNRR⁹⁰AKCRQQ⁹⁷

Tableau 10. CSP du domaine NLS-HD d'Otx2, obtenus après l'ajout de « n » équivalents de Ac-Ahx-(EC'AC')₆-NH₂, Ac-Ahx-(EC'C'A)₆-NH₂, Ac-Ahx-(EC'AC')₃-NH₂, HI dp8 et PolyGlu et reportés en valeur relative, le 100 % étant attribué à l'acide aminé dont le CSP est maximum pour chaque ligand (CSP+). Dans la séquence de NLS-HD Otx2 sont représentés en vert le domaine de liaison aux GAGs, en violet l'hélice 3 et en jaune His¹⁶ et Tyr⁶².

Il est attendu que les interactions entre NLS-HD Otx2 et les ligands soient de nature électrostatique et puissent également engager des liaisons hydrogène, pouvant impliquer d'une part les sulfonates (sulfopeptides, HI dp8), sulfates (HI dp8) et/ou carboxylates (sulfopeptides, HI dp8 et PolyGlu) des ligands, et d'autre part les chaînes latérales chargées positivement des Arginines et Lysines (acides aminés indiqués en bleu dans le **Tableau 10**) et

les acides aminés polaires non chargés (tels que Glutamine, Asparagine, Tyrosine, Thréonine et/ou Sérine) de la protéine. De manière générale, les valeurs de CSP les plus élevées sont bien observées dans les régions qui concentrent le plus de résidus cationiques, à savoir le motif N-terminal de liaison aux GAGs et l'hélice 3 de l'homéodomaine (**Figure 93**), confirmant ainsi l'importance des interactions électrostatiques. Je discuterai donc de ces deux régions d'interaction mais également de deux acides aminés qui, bien qu'isolés, montrent, dans certains cas, des valeurs de CSP suffisamment élevées pour retenir notre attention : His²⁶ et Tyr⁶².

Tous les ligands **affectent la région 39 – 45 du domaine de liaison aux GAGs d'Otx2**. Les acides aminés affectés sont polaires chargés (Arg, Lys et même Glu) mais aussi polaires neutres, comme Thr⁴³ et Thr⁴⁴. De plus, pour tous les ligands, à l'exception de PolyGlu, nous pouvons observer une zone homogène de CSP où tous les acides aminés de la région 39-45 sont affectés. Les mimes peptidiques de GAGs ont donc un comportement comparable à celui de HI dp8 contrairement à PolyGlu qui interagit différemment avec les acides aminés de la région 39-45, n'affectant que très peu les CSP de certains résidus pourtant cationiques comme Arg³⁹ ou Arg⁴². Concernant l'affinité, le plus petit des ligands peptidiques, Ac-Ahx-(EC'AC')₃-NH₂, présente la moins bonne affinité (entre 100 et 300 μM). Les autres ligands, Ac-Ahx(EC'AC')₆-NH₂, Ac-Ahx-(EC'C'A)₆-NH₂ et HI dp8, ont une affinité comparable pour ce domaine, globalement autour de 6 à 20 μM. Pour PolyGlu, les affinités observées sont un peu moins bonnes, se situant plutôt autour de 40 μM.

Il est à noter que les acides aminés Arg³⁶ et Lys³⁷, doublet RK du domaine de liaison aux GAGs de la protéine Otx2, dont l'importance dans l'interaction avec les CS-E a été proposée en utilisant des mutants[86], présentent des valeurs de CSP assez faibles, ce qui suggère qu'ils ne joueraient pas de rôle clé dans l'interaction des ligands étudiés avec le domaine NLS-HD d'Otx2.

Des valeurs élevées de CSP sont également observées pour des résidus de l'hélice 3 de l'homéodomaine. Pour rappel cette hélice 3 joue un rôle primordial dans l'internalisation cellulaire des protéines à homéodomaine ainsi que dans leur fixation dans le grand sillon de l'ADN. Ici encore, PolyGlu a un profil de CSP sensiblement différent de celui des autres ligands, seuls certains acides aminés à l'extrémité C-terminale de cette hélice 3 étant fortement

affectés par l'interaction. Pour les autres ligands, sulfopeptides et HI dp8, les acides aminés V⁸² à Q⁹⁵ de l'hélice 3, qu'ils soient polaire chargé comme R⁸⁹ ou apolaires comme Q⁸³, V⁸⁴, F⁸⁶ ou A⁹¹ sont affectés de manière significative, ce qui, une fois encore, tend à prouver que les sulfopeptides ont un comportement comparable à celui de HI dp8. Mais, contrairement à ce qui a été observé dans la région 39 – 45 du domaine de liaison aux GAGs, on ne retrouve pas ici de zone homogène de CSP et tous les acides aminés cationiques ne sont pas systématiquement affectés par le ligand, pour exemple K⁸⁷. De plus, si les valeurs de K_d apparents obtenues avec HI dp8 sont relativement homogènes pour l'ensemble des acides aminés les plus affectés de la région de l'hélice 3 (de 20 à 30 μM), deux zones se distinguent avec les sulfopeptides (EC'AC')₆ et (EC'C'A)₆ : le N-ter (zone V⁸²Q⁸³V⁸⁴) avec des affinités entre 40 et 120 μM *versus* le C-ter (zone F⁸⁶–R⁹⁴) avec de meilleures valeurs de K_d, autour de 10 μM. En ce qui concerne le sulfopeptide Ac-Ahx-(EC'AC')₃-NH₂, les valeurs faibles de CSP dans la région de l'hélice 3 n'ont pas permis d'obtenir de K_d, ce qui suggère une affinité très faible. PolyGlu possède, quant à lui, une affinité de l'ordre de 40 μM pour les acides aminés de la région C-terminale de cette hélice.

De manière intéressante, les valeurs de K_d apparents dans la région 39–45 du domaine de liaison aux GAGs sont du même ordre de grandeur dans la partie C-ter de l'hélice 3, pour tous les ligands testés à l'exception de Ac-Ahx-(EC'AC')₃-NH₂ pour lequel les K_d de la région de l'hélice 3 n'ont pu être déterminés. Cette observation suggère la possibilité d'une synergie d'interaction impliquant ces deux régions d'Otx2, à savoir que les motifs cationiques 39–45 et 86–94 interagiraient de manière concomitante avec une seule molécule de ligand. La structure en solution de l'homéodomaine d'Otx2 déterminée par RMN[201] confirme qu'il est tout à fait possible qu'un ligand suffisamment long puisse à la fois interagir avec le domaine N-terminal flexible de liaison aux GAGs et avec la partie C-terminale de l'hélice 3, puisque ces deux régions cationiques se trouvent du même côté de la molécule (**Figure 96**). Des valeurs similaires de K_d dans ces deux régions d'Otx2 pourraient également s'expliquer par la liaison exclusive du motif cationique N-terminal au ligand qui provoquerait de manière concomitante un changement conformationnel et/ou d'orientation de la partie C-terminale de l'hélice 3. Dans le cas de Ac-Ahx-(EC'AC')₃-NH₂, il est difficile, sans connaissance des K_d de la région de l'hélice 3, d'en avoir la certitude, mais il semble raisonnable d'imaginer que ce peptide, plus court que les autres mimes, ne puisse interagir simultanément avec les deux régions, cette

absence de synergie pouvant expliquer, en partie, la grande différence d'affinité observée par rapport aux autres ligands étudiés.

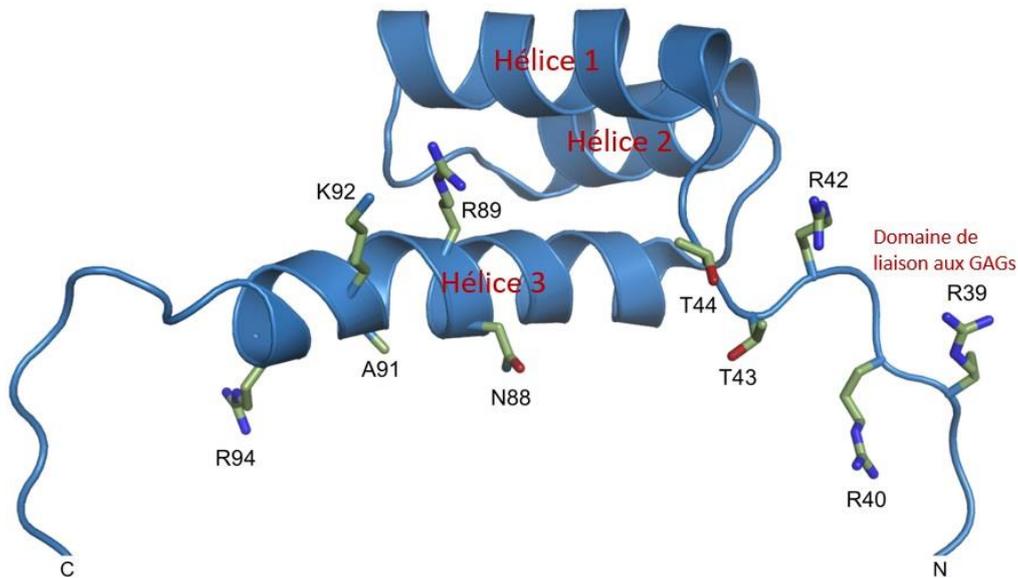


Figure 96. Structure en solution de l'homéodomaine de la protéine Otx2 (code PDB 2DMS)[201].

Deux autres CSP significatifs à noter sont celui de Tyr⁶² et de His²⁶. Tyr⁶² est un acide aminé polaire situé dans la boucle entre les hélices 1 et 2 de l'homéodomaine et qui, si on se réfère à la structure en solution de l'homéodomaine d'Otx2 (**Figure 97**), est engagé dans une liaison hydrogène avec Arg⁹⁰ de l'hélice 3. Le résidu Tyr⁶² présente des valeurs élevées de CSP amide en présence des ligands Ac-Ahx-(EC'AC')₆-NH₂, Ac-Ahx-(EC'CA)₆-NH₂ et HI dp8, avec des K_d apparents de 20 à 30 μM. Ces valeurs de K_d sont du même ordre de grandeur que celles mesurées dans le domaine N-terminal de liaison aux GAGs et la zone F⁸⁶-R⁹⁴ de l'hélice 3 contenant Arg⁹⁰. Le résidu Tyr⁶² est situé sur une surface de la protéine opposée à la surface large formée par le domaine N-terminal de liaison aux GAGs et la partie C-terminale de l'hélice 3 (**Figure 97**). Il est donc peu probable que la perturbation de ce résidu, seul résidu significativement perturbé dans la boucle entre les hélices 1 et 2, soit due à une interaction directe avec les ligands. Par contre, une réorganisation conformationnelle de l'hélice 3 contenant Arg⁹⁰ ou une interaction directe de cette hélice avec les ligands pourrait avoir pour conséquence une rupture de la liaison hydrogène entre Tyr⁶² et Arg⁹⁰, ce qui conduirait à une modification de l'environnement électronique de Tyr⁶² et une perturbation de ce résidu. Il est à noter que PolyGlu se différencie une fois de plus des 4 autres ligands puisque Tyr⁶² n'est pas

particulièrement affecté par l'interaction d'Otx2 avec ce peptide de 20 résidus chargés négativement.

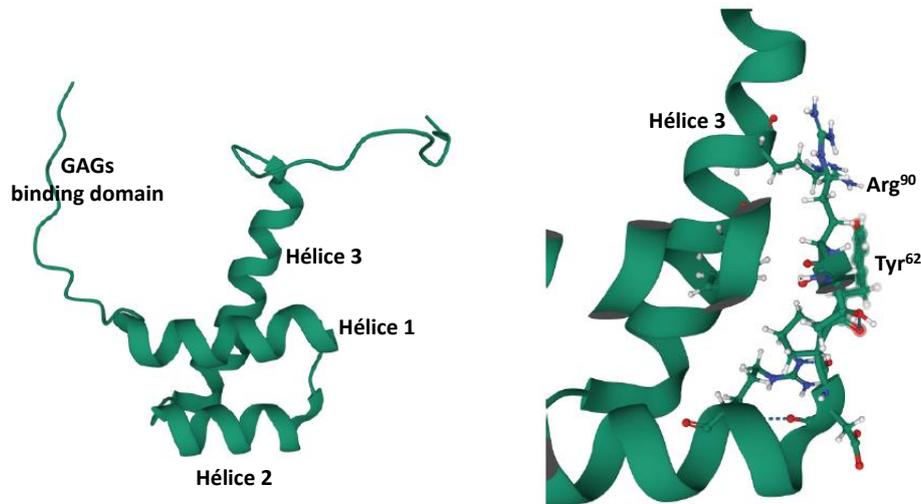


Figure 97. Structure en solution de l'homéodomaine de la protéine Otx2 (Code PDB 2DMS)[201].

L'Histidine est un acide aminé qui contient un noyau imidazole ionisable dont le pKa est proche de 6,50. Le pH du tampon utilisé pour les études d'interaction par RMN étant proche de cette valeur, les perturbations importantes observées pour His²⁶ après ajout de Ac-Ahx-(EC'AC')₆-NH₂ et Ac-Ahx-(EC'C'A)₆-NH₂ pourraient s'expliquer par l'addition de ces deux ligands peptidiques acides (12 charges de sulfonates et 6 charges de carboxylates) qui, en modifiant légèrement le pH de la solution de protéine, pourrait conduire à un changement d'état de protonation du noyau imidazole. Cependant le ligand HI dp8, malgré ses sulfates, sulfonates et carboxylates, n'a pas induit une telle perturbation de His²⁶, et tous les ligands ont été préalablement dissous dans le tampon RMN avant chaque titrage. Ceci n'est donc qu'une hypothèse, d'autant qu'en toute rigueur, nous n'avons pas mesuré le pH final des solutions tamponnées après addition des ligands. Quoiqu'il en soit, même si les CSP importants observés pour His²⁶ après ajout des deux peptides longs ne sont pas liés à un changement d'acidité du milieu, les K_d apparents obtenus pour cet acide aminé, de l'ordre de 200 à 300 μM (valeurs 10 à 14 fois supérieures aux meilleurs K_d observés), indiquent qu'il ne s'agirait que d'une interaction de faible affinité par rapport à ce qui est observé pour les autres domaines.

Une dernière observation au sujet des CSP : Comme évoqué précédemment, les valeurs de CSP ne peuvent être directement utilisées pour tirer des conclusions concernant

l'affinité de la protéine envers les ligands. Toutefois sachant que Ac-Ahx-(EC'AC')₆-NH₂, Ac-Ahx-(EC'C'A)₆-NH₂ et HI dp8 sont des molécules comparables, de par la nature et le nombre de charges qu'elles portent, nous pouvons nous interroger sur le fait que les amplitudes de CSP mesurées dans le domaine de liaison aux GAGs sont globalement deux fois plus élevées avec HI dp8 qu'avec Ac-Ahx-(EC'AC')₆-NH₂ et Ac-Ahx-(EC'C'A)₆-NH₂, alors que les valeurs de CSP sont équivalentes dans l'hélice 3 pour ces 3 molécules. En d'autres termes, les amplitudes des CSP induites par HI dp8 pour les acides aminés de l'hélice 3 ne représentent que 10 à 30 % de celles mesurées pour les acides aminés du domaine de liaison aux GAGs alors qu'avec les sulfopeptides contenant 6 motifs répétés, les amplitudes des CSP mesurées dans l'hélice 3 représentent 50% de celles mesurées dans le domaine de liaison aux GAGs.

Mais avant de discuter ces résultats obtenus avec le domaine NLS-HD de la protéine Otx2 et établir un certain nombre d'hypothèses concernant les modes d'interaction des ligands étudiés, je vais décrire ci-après les expériences préliminaires réalisées sur le domaine NLS-HD de la protéine Engrailed 2.

Interactions avec Engrailed 2.

En parallèle des études RMN menées sur NLS-HD Otx2, une étude similaire a été réalisée sur le domaine NLS-HD de l'homéoprotéine Engrailed 2 (En2) en utilisant la version marquée ¹⁵N de cette protéine, produite par mes collègues. J'ai pu me servir de l'attribution, déjà réalisée, des signaux du spectre HSQC de la protéine doublement marquée ¹⁵N/¹³C pour effectuer mon expérience de titrage par le sulfopeptide le plus prometteur : Ac-Ahx-(EC'AC')₆-NH₂.

Mesure par RMN de l'affinité du sulfopeptide Ac-(EC'AC')₆-NH₂ pour NLS-HD En2 marqué en ¹⁵N :

Des échantillons de domaine NLS-HD d'En2 uniformément enrichi en isotope ¹⁵N ont été préparés à la concentration de 150 μM afin d'évaluer l'affinité de la protéine vis-à-vis de Ac-Ahx-(EC'AC')₆-NH₂, par la méthode RMN de perturbation de déplacement chimique. Des variations de déplacement chimique avec des trajectoires linéaires ont été observées sur les spectres 2D HSQC ¹H-¹⁵N au fur et à mesure des ajouts (voir pour exemple la **Figure 98**), ce qui atteste d'une interaction entre le domaine NLS-HD d'En2 et le ligand Ac-Ahx-(EC'AC')₆-NH₂ selon le régime d'échange rapide à deux états.

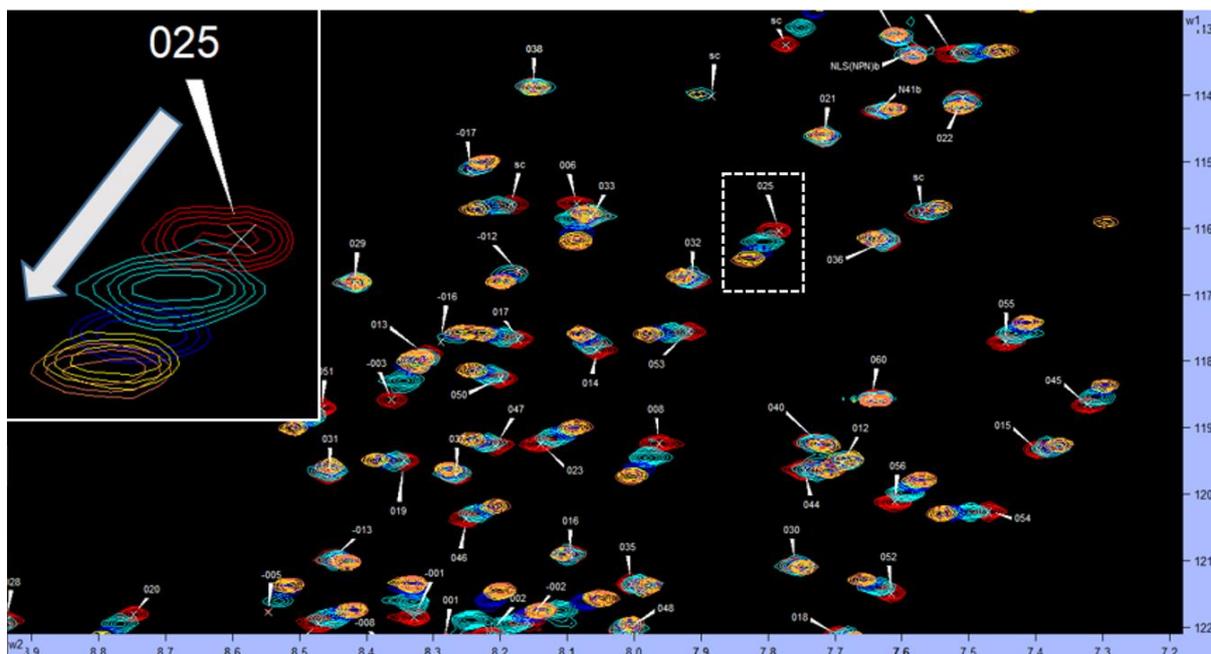


Figure 98. Superposition des spectres 2D HSQC ^1H - ^{15}N du domaine NLS-HD d'En2, avant (en rouge) et après ajouts successifs de Ac-Ahx-(EC'AC') $_6$ -NH $_2$ (cyan : 0,24 eq / bleu : 0,48 eq / jaune : 0,72 eq et rose : 0,96 eq).

Comme décrit précédemment, les valeurs de déplacements chimiques amides ont été mesurées à chaque point du titrage afin d'extraire les constantes de dissociation K_d apparentes des résidus les plus perturbés. Les valeurs de CSP amides sont représentées dans l'histogramme de la **Figure 99b**, ainsi que les valeurs de K_d des acides aminés les plus affectés par l'interaction. Pour davantage de lisibilité, les erreurs standards n'ont pas été indiquées sur ces histogrammes, mais reportées dans la partie expérimentale de ce chapitre (**Tableau 14**). Pour faciliter la comparaison j'ai reporté sur la même figure (**Figure 99a**) les CSP du domaine NLS-HD d'Otx2 induits par l'ajout de Ac-Ahx-(EC'AC') $_6$ -NH $_2$ et déjà présentés **Figure 95a**. En complément, une étude réalisée par Sébastien Cardon au cours de sa thèse[186], utilisant la méthode RMN de perturbation des déplacements chimiques et portant sur la mesure de l'affinité du domaine NLS-HD d'En2 vis-à-vis de HI dp8, permet, dans la **Figure 99c**, de comparer les CSP de la protéine induits par HI dp8 à ceux induits par Ac-Ahx-(EC'AC') $_6$ -NH $_2$, ainsi que les K_d des acides aminés les plus affectés par l'ajout des ligands.

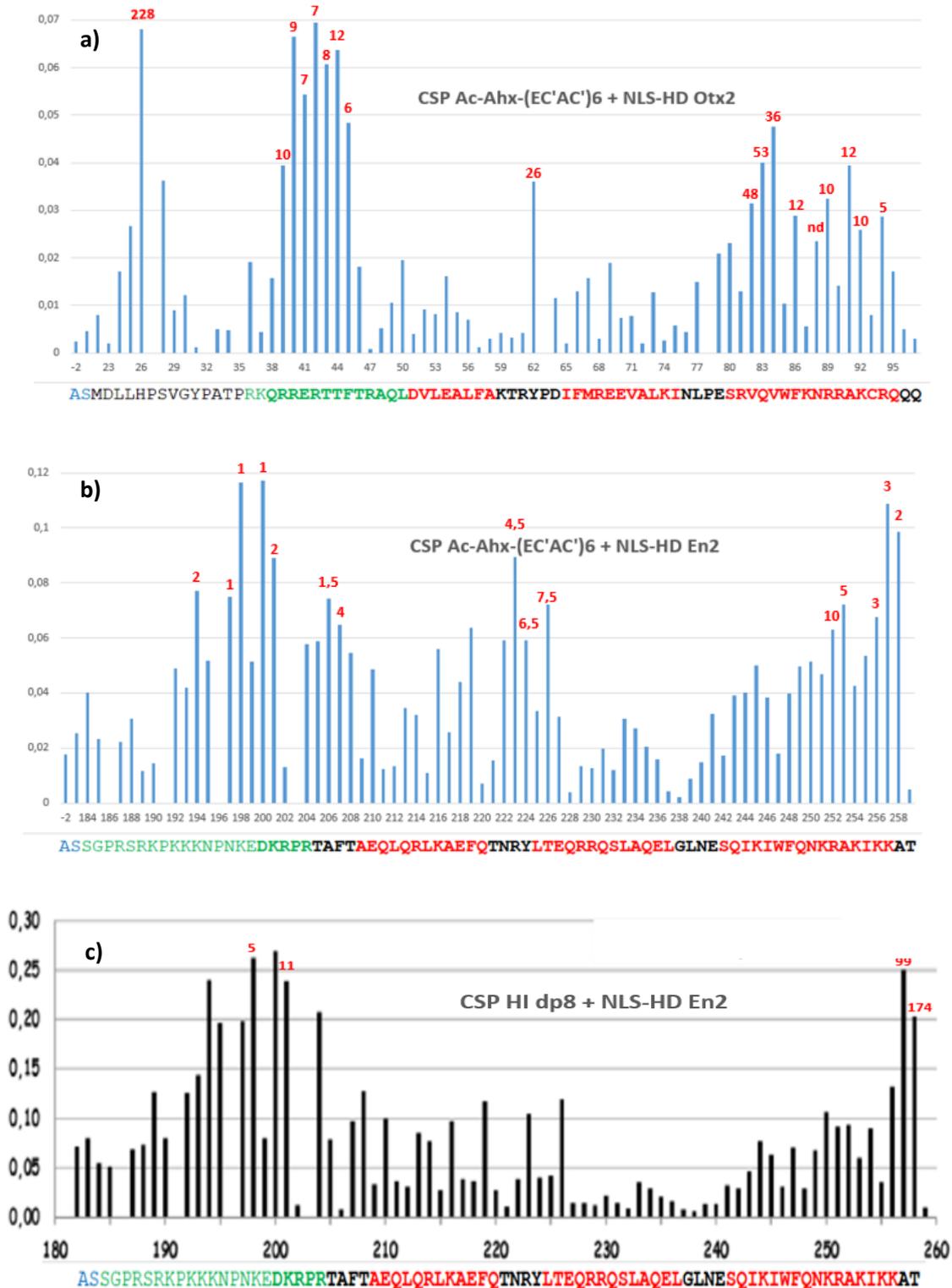


Figure 99. CSP amides (en ppm) de : a) NLS-HD Otx2 et b) NLS-HD En2 en interaction avec respectivement 1,9 et 0,96 eq de Ac-Ahx-(EC'AC')₆-NH₂ – c) CSP amides (en ppm) de NLS-HD En2 en interaction 1,4 eq de HI dp8 (figure extraite de la thèse de S. Cardon[186]). La valeur des K_d apparents (en μM) des résidus d'intérêt est indiquée en rouge (« nd » signifie non déterminé). Dans la séquence des protéines (en bas des histogrammes) sont représentés en vert le domaine de liaison putatif aux GAGs[186], en gras l'homéodomaine et en rouge les hélices α.

Comme pour les expériences réalisées sur NLS-HD Otx2, les valeurs de CSP obtenues avec NLS-HD En2 sont également présentées en valeurs relatives dans le **Tableau 11**, c'est-à-dire en normalisant à 100 % l'acide aminé avec le CSP le plus élevé (CSP+), les autres CSP étant exprimés en % de cette valeur maximale. Les CSP compris entre 40 % et 100 % sont indiqués en vert.

Ratio CSP du domaine de liaison aux GAGs (100 % pour le CSP le plus important, CSP+)

NLS-HD Otx2	CSP +	n eq	36	R	K	Q	R	R	E	R	T	T	F	T	R	A	Q	L
	0,07	1,9	27	6	22	57	95	74	100	86	90	70	27	0	7	14	27	
NLS-HD En2	CSP +	n eq	190	K	P	K	K	K	N	P	N	K	E	D	K	R	P	R
	0,12	0,96	12	-	42	36	66	44	-	64	99	44	100	76	11	-	49	

Ratio CSP de l'Hélice 3 (valeur relative)

NLS-HD Otx2	CSP +	n eq	80	S	R	V	Q	V	W	F	K	N	R	R	A	K	C	R	Q
	0,07	1,9	33	17	44	57	69	16	43	8	20	46	19	57	37	11	41	24	
NLS-HD En2	CSP +	n eq	242	S	Q	I	K	I	W	F	Q	N	K	R	A	K	I	K	K
	0,12	0,96	14	33	34	43	33	15	34	42	44	40	54	61	37	46	58	93	

Ratio CSP de Tyr⁶² et Tyr²²⁴ (valeur relative)

NLS-HD Otx2	CSP +	n eq	62	Y
	0,07	1,9	51	
NLS-HD En2	CSP +	n eq	224	Y
	0,12	0,96	50	

NLS-HD Otx2 : ASM²²DLLHPSVG³⁰YPATPRKQRR⁴⁰ERTTFTRAQL⁵⁰DVLEALFAKT⁶⁰RYPDIFMREE⁷⁰
VALKINLPES⁸⁰RVQVWFKNRR⁹⁰AKCRQQQ⁹⁷
NLS-HD En2 : ASS¹⁸⁴GPRSRK¹⁹⁰PKKKNPNKED²⁰⁰KRPRTAFTAET²¹⁰QLQRLKAEFQ²²⁰
TNR²²⁴LTEQRR²³⁰QSLAQELGLN²⁴⁰ESQIKIWFQN²⁵⁰KRAKIKKAT²⁵⁹

Tableau 11. CSP des domaines NLS-HD Otx2 et En2, obtenus après l'ajout de « n » équivalents de Ac-Ahx-(EC'AC')₆-NH₂ et reportés en valeur relative, le 100 % étant attribué à l'acide aminé dont le CSP est maximum (CSP+). Dans la séquence de ces fragments de protéine à homéodomaine NLS-HD Otx2 et NLS-HD En2 sont représentés en vert le domaine de liaison aux GAGs, en violet l'hélice 3 et en jaune Tyr⁶² de Otx2 et Tyr²²⁴ de En2.

Alors que le titrage du domaine NLS-HD d'Otx2 par le ligand Ac-Ahx-(EC'AC')₆-NH₂ montrait sur l'histogramme correspondant (Figure 99a) des zones d'interaction bien définies, ce même ligand adopte un profil d'interaction sensiblement différent avec le domaine NLS-HD d'En2 (Figure 99b) puisque plus de 50 % des acides aminés de la protéine voient leur déplacement chimique affecté par l'addition du sulfopeptide, avec des CSP supérieurs à 0,04 ppm. Les perturbations les plus marquées sont toutefois observées dans les parties C-terminales du domaine de liaison aux GAGs et de l'hélice 3 de l'homéodomaine qui concentrent le plus de résidus cationiques. Les autres régions d'En2 affectées par la présence du ligand Ac-Ahx-(EC'AC')₆-NH₂ correspondent aux résidus se situant entre le domaine de liaison aux GAGs et l'hélice 1 (région 204-208) et les résidus de la boucle entre les hélices 1 et 2 (région 222-226), dont Tyr²²⁴ qui est idéalement situé dans la structure pour établir une liaison hydrogène avec l'Arg²⁵² de l'hélice 3 (**Figure 100**). Les K_d apparents déterminés pour toutes ces zones d'interaction d'En2 (1 à 10 μM) sont globalement du même ordre de grandeur que ceux observés pour l'interaction entre le domaine NLS-HD d'Otx2 et le ligand Ac-Ahx-(EC'AC')₆-NH₂ (1 à 20 μM). De plus, les valeurs homogènes de K_d dans les régions 194-201, 222-226 et 252-258 d'En2 suggèrent, comme pour Otx2, une synergie, c'est-à-dire qu'une même molécule de ligand Ac-Ahx-(EC'AC')₆-NH₂ pourrait se lier de manière concomitante aux parties C-terminales du domaine de liaison aux GAGs et de l'hélice 3 de l'homéodomaine. L'interaction directe de l'hélice 3 avec le ligand ou un changement conformationnel impliquant cette hélice aurait pour conséquence la rupture de la liaison hydrogène entre Arg²⁵² et Tyr²²⁴, modifiant ainsi l'environnement électronique de Tyr²²⁴ et des acides aminés de la boucle entre les hélices 1 et 2 (**Figure 100**).

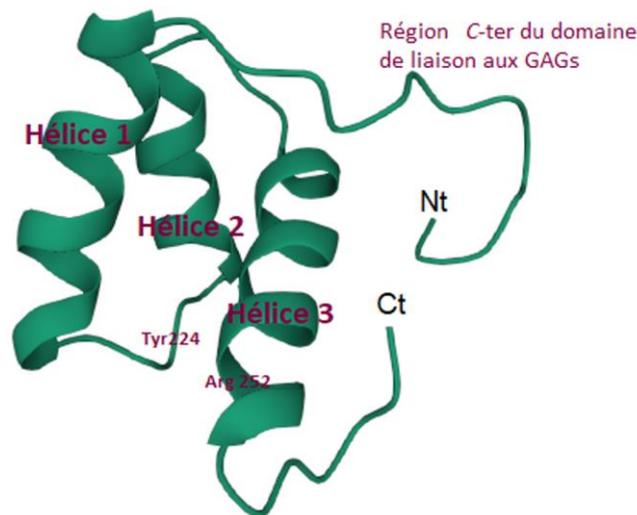


Figure 100. Structure en solution de l'homéodomaine de la protéine En2 (code PDB 3ZOB) [185].

La comparaison des profils de CSP du domaine NLS-HD d'En2 en interaction avec Ac-Ahx-(EC'AC')₆-NH₂ et HI dp8 permet de constater que, pour les deux ligands, les régions de NLS-HD En2 les plus affectées sont à nouveau les parties C-terminales du domaine de liaison aux GAGs et de l'hélice 3 de l'homéodomaine. Par contre, les régions 204-208 et 222-226 d'En2 sont sensiblement moins perturbées en présence de HI dp8. On observe plutôt des variations de déplacements chimiques d'acides aminés isolés dans ces 2 régions de la protéine avec HI dp8 (**Figure 99c**). D'autre part, alors que les valeurs de K_d mesurées dans les zones perturbées d'En2 sont toutes du même ordre de grandeur (1 à 10 μ M) avec le ligand Ac-Ahx-(EC'AC')₆-NH₂, ce n'est pas le cas avec le ligand HI dp8. HI dp8 présente une affinité beaucoup plus faible pour l'extrémité C-terminale de l'hélice 3 de NLS-HD En2, avec des valeurs de K_d entre 99 et 174 μ M, contre 5-11 μ M pour les résidus les plus perturbés du motif N-terminal de liaison aux GAGs. Il n'y a donc pas de synergie dans l'interaction des motifs cationiques en N- et C-terminal d'En2 pour le ligand HI dp8, ce qui a été démontré dans les travaux de thèse de Sébastien Cardon à l'aide d'expériences complémentaires.

3. Discussion

Avant de proposer une/des hypothèse(s) de travail, je vais au préalable récapituler de manière schématique les résultats qui m'apparaissent les plus importants dans l'interaction des ligands étudiés avec le domaine NLS-HD des protéines Otx2 et En2 :

* Tous les ligands étudiés ont montré une affinité, plus ou moins forte, pour le domaine NLS-HD d'Otx2, plus particulièrement pour le domaine *N*-terminal de liaison aux GAGs (région 39-45) et l'hélice 3 de l'homéodomaine (région 82-95). C'est également le cas du domaine NLS-HD de la protéine En2 qui est reconnu par le sulfopeptide Ac-Ahx-(EC'AC')₆-NH₂.

* PolyGlu, dont l'affinité est un peu moins bonne que celle de HI dp8 ou des sulfopeptides à 6 motifs répétés, montre un profil de CSP différent pour la protéine Otx2 en comparaison avec les autres ligands.

* Avec uniquement 3 motifs EC'AC' répétés, Ac-Ahx-(EC'AC')₃-NH₂ est le sulfopeptide qui présente la moins bonne affinité pour le domaine NLS-HD d'Otx2.

* Tous les acides aminés du domaine *N*-terminal de liaison aux GAGs d'Otx2 sont affectés de manière homogène par l'interaction avec les ligands HI dp8, Ac-Ahx-(EC'AC')₆-NH₂ et Ac-Ahx-(EC'C'A)₆-NH₂, ce qui n'est pas le cas pour les acides aminés de l'hélice 3 (région 82-95).

* HI dp8 et les deux sulfopeptides à 6 motifs répétés ont globalement une affinité équivalente pour les deux régions d'interaction identifiées dans le domaine NLS-HD d'Otx2. Par contre, si le profil de CSP est assez similaire entre ces 3 ligands pour la région de l'hélice 3 de l'homéodomaine, les amplitudes de CSP des acides aminés de la région du domaine *N*-terminal de liaison aux GAGs sont trois fois plus élevées avec HI dp8, par rapport aux peptides mimes, alors que la taille et le nombre de charges de ces ligands sont comparables.

* Dans le cas de NLS-HD d'En2, le sulfopeptide Ac-Ahx-(EC'AC')₆-NH₂ a une affinité similaire à celle de HI dp8 pour le domaine *N*-terminal de liaison aux GAGs, mais les deux ligands ne présentent pas le même comportement vis-à-vis de l'hélice 3.

De ces résultats, nous pouvons conclure que, pour qu'un ligand possède une affinité optimale pour NLD-HD Otx2 et un comportement proche de celui d'un fragment de GAG sulfaté (HI dp8), la présence de sulfonates, ou du moins une certaine répartition des charges dans la molécule, est nécessaire, ainsi qu'un nombre minimal de motifs EC'AC' ou EC'C'A, au-delà de 3 en tous les cas. Les sulfopeptides possédant 6 motifs répétés semblent bien mimer l'affinité et le mode d'action du fragment de GAG sulfaté HI dp8 dans le cas du domaine NLS-HD d'Otx2, à savoir qu'ils induisent des perturbations conformationnelles concomitantes du domaine de liaison aux GAGs et de l'hélice 3. Nous constatons toutefois que, contrairement à nos attentes, la position des acides aminés dans la séquence primaire des sulfopeptides, (EC'C'A)₆ ou (EC'AC')₆, n'a pas eu d'impact significatif sur l'affinité des molécules ou leur mode d'interaction vis à vis du domaine NLS-HD d'Otx2.

Considérant ces données RMN, nous serions tentés de conclure que nous n'avons pas synthétisé de mime interagissant sélectivement avec notre protéine cible Otx2, puisque, contrairement à HI dp8, Ac-Ahx-(EC'AC')₆-NH₂ interagit de la même manière avec les domaines NLS-HD d'Otx2 et d'En2. Il faut noter cependant que les profils d'interaction varient d'une homéoprotéine à l'autre et il est, en l'état, difficile de conclure si ces variations observées se traduisent d'un point de vue biochimique/biologique. Mais, quoi qu'il en soit, nous avons bel et bien produit des mimes de GAGs.

A partir des résultats ainsi résumés que pouvons-nous proposer quant au mode d'interaction des ligands étudiés avec la protéine NLS-HD Otx2 ?

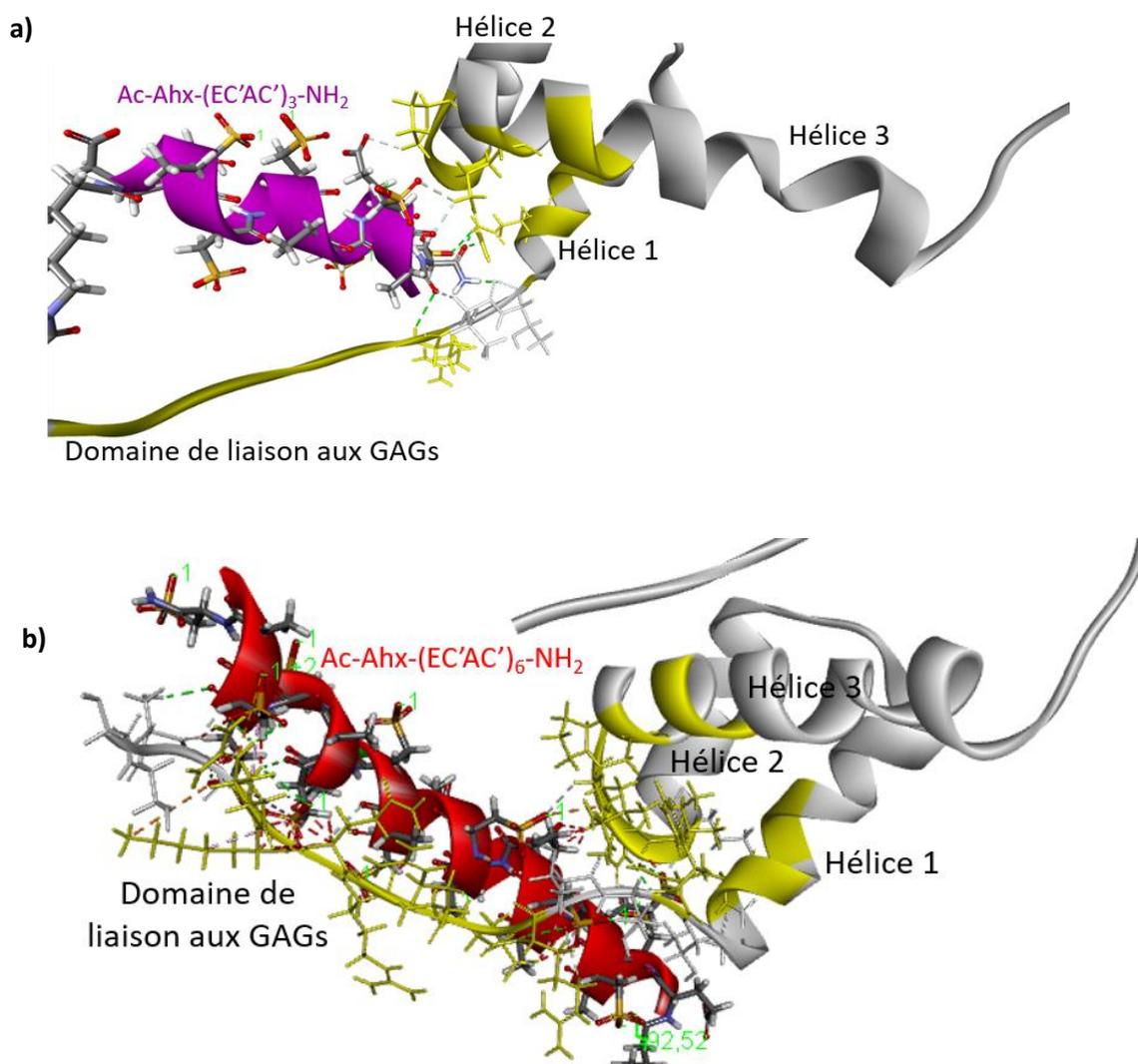
Nous avancerons deux hypothèses pour expliquer les interactions des mimes synthétisés et de HI dp8 avec le domaine de liaison aux GAGs et l'hélice 3 de l'homéodomaine de la protéine NLS-HD Otx2 :

- Les ligands se lieraient à la fois au domaine de liaison aux GAGs et à l'hélice 3. Nous avons pu observer, dans la **Figure 96** - structure RX de l'homéodomaine de la protéine Otx2 en solution - que cette hypothèse était envisageable puisque ces deux régions se trouvent du même côté de la protéine. Une molécule suffisamment longue pourrait donc interagir simultanément avec les deux régions et affecter également la Tyrosine en position 62 par la rupture de sa liaison hydrogène avec l'Arginine 90 située dans l'hélice 3.

- Les ligands, en se liant au domaine *N*-terminal de liaison aux GAGs, pourraient engendrer un changement de conformation et/ou d'orientation des hélices de la protéine, affectant l'environnement électronique de l'autre région.

Afin de tenter d'apporter des éléments de réponses en faveur de l'une ou l'autre de ces hypothèses, des expériences de modélisation ont été réalisées au laboratoire par le Dr Nicolas Pietrancosta, selon un protocole déjà décrit[202]. Sur la base de la structure en solution de l'homéodomaine de la protéine Otx2[201], Ac-Ahx-(EC'AC')₆-NH₂ et Ac-Ahx-(EC'AC')₃-NH₂ ont été proposés comme ligands potentiels dans l'application de modélisation et de simulation BIOVIA Discovery Studio. Les images obtenues sont reportées dans la **Figure 101**.

Ces expériences de modélisation de l'interaction entre l'homéodomaine d'Otx2 et les ligands Ac-Ahx-(EC'AC')₃-NH₂ et Ac-Ahx-(EC'AC')₆-NH₂ nous apportent plusieurs informations intéressantes. Tout d'abord, pour obtenir une interaction entre les mimes et la protéine, il a fallu structurer les ligands sous forme d'hélice. Par eux-mêmes, les mimes forment une « boule », avec un cœur hydrophobe, et n'interagissent pas avec la protéine. Mais, ainsi pré-structurées, cette structure hélicoïdale en interaction avec l'homéodomaine d'Otx2 reste stable, ceci tendant à prouver que, si les mimes ne sont pas structurés en solution, ils pourraient se structurer en hélice au contact de la protéine. Nous constatons alors que l'interaction entre le mime et la protéine a lieu essentiellement avec les acides aminés du domaine de liaison aux GAGs, les sulfopeptides se plaçant le long de ce domaine. Certains acides aminés isolés des hélices 1 et 3 ainsi que de la boucle entre les hélices 2 et 3 pourraient créer des liaisons, *via* des liaisons hydrogène ou de van der Waals, avec les sulfopeptides. Mais nous ne trouvons pas d'interaction directe avec la partie cationique de l'hélice 3 (incluant R⁹⁰), interaction suggérée par les résultats des expériences de RMN.



Séquence de l'homéodomaine de la protéine Otx2 :

RKQRR⁴⁰ERTTFTRAQL⁵⁰DVLEALFAKT⁶⁰RYPDIFMREE⁷⁰VALKINLPES⁸⁰RVQVWFKNRR⁹⁰AKCRQQ⁹⁷

Figure 101. Modélisation de l'interaction de l'homéodomaine de la protéine Otx2 avec a) Ac-Ahx-(EC'AC')₃-NH₂ et b) Ac-Ahx-(EC'AC')₆-NH₂. Sur les figures sont indiquées en jaune les zones d'interaction putatives de la protéine avec les ligands. Dans la séquence de l'homéodomaine d'Otx2 sont représentés en jaune les acides aminés interagissant potentiellement avec le ligand, en gras le domaine de liaison aux GAGs et surlignés en vert l'hélice 1, en gris l'hélice 2 et en bleu l'hélice 3.

Le domaine de liaison aux GAGs serait bel et bien l'unique domaine d'interaction directe avec les sulfopeptides (et aussi très probablement avec HI dp8). Cette liaison engendrerait un changement de conformation et/ou d'orientation de l'hélice 3 de l'homéodomaine, affectant ainsi l'environnement électronique des acides aminés la composant. Cela permettrait d'expliquer que les CSP du domaine de liaison aux GAGs montrent une zone homogène, l'interaction directe affectant **tous** les acides aminés se

trouvant à proximité du ligand, alors que les CSP de l'hélice 3 sont hétérogènes pour un même ligand (tous les résidus basiques ne sont pas affectés), les variations d'environnement étant liées à un « évènement indirect », ou allostérie.

Les résultats des expériences de RMN présentées ci-dessus sont bien en adéquation avec les résultats de l'étude de biochimie décrite dans le chapitre III puisque, pour rappel, dans ces expériences de biochimie, les peptides Ac-Ahx-(EC'AC')₆-NH₂ et Ac-Ahx-(EC'CA)₆-NH₂ étaient bien plus affins pour la protéine Otx2 que Ac-Ahx-(EC'AC')₃-NH₂. Le peptide Ac-Ahx-(EC'AC')₆-NH₂ que nous avons choisi, au vu de ces expériences de biochimie, comme étant le meilleur parmi les molécules testées, inhibait même *in vivo* l'internalisation de Otx2 dans les neurones à parvalbumine.

À ce stade de notre réflexion, nous souhaitons émettre deux hypothèses, concernant, cette fois, l'inhibition de l'internalisation de Otx2 dans les cellules PV :

i) La première hypothèse, sans doute la plus évidente, serait de dire que l'infusion d'une grande quantité de mimes assez affins pour Otx2, de type Ac-Ahx-(EC'AC')₆-NH₂, dans le cerveau de souris crée une compétition avec les GAGs naturellement présents autour des cellules PV, vis-à-vis de l'interaction avec la protéine Otx2. L'idée est donc ici que l'interaction entre le mime et Otx2 est de même nature que celle existant entre Otx2 et les GAGs naturels et que nos mimes, s'ils ont une affinité suffisante pour Otx2, permettent de conserver l'homéoprotéine éloigné du réseau périneuronal.

Cette hypothèse ne tient toutefois pas compte de ce qui a été observé lors des expériences de RMN, à savoir les perturbations de déplacement chimique des acides aminés de l'hélice 3.

ii) Nous proposons donc un deuxième modèle, plus spéculatif. **Pour qu'Otx2 ne pénètre plus dans les cellules, ce qui est notre objectif, il faudrait que le ligand conçu se lie au domaine de liaison aux GAGs et induise un changement conformationnel autour de l'hélice 3**, domaine indispensable à la pénétration d'Otx2 dans les cellules. Cette deuxième hypothèse tendrait à indiquer que l'interaction entre le mime et Otx2 ne serait pas de même nature que celle qui existe entre Otx2 et les GAGs naturels. En effet les CS ou HS présents en surface des cellules n'empêchent pas l'internalisation de Otx2, au contraire ils se fixent de

manière transitoire pour « guider » la protéine ou « aider » sa diffusion vers la membrane cellulaire et conduire à son internalisation, médiée par l'hélice 3. Mais les CS et HS naturels n'affectent peut-être pas la structure de l'hélice 3, comme nous l'observons avec les sulfopeptides à 6 motifs et HI dp8. Il serait, à ce sujet, intéressant de réaliser de nouvelles expériences d'infusion dans le cerveau de souris, mais cette fois avec HI dp8 afin de savoir si ce ligand, qui semble avoir le même mode d'action que les sulfopeptides à 6 motifs, peut également affecter l'internalisation de Otx2 dans les cellules.

Tout cela reste bien entendu à l'état d'hypothèse et nécessite d'être confronté à des expériences complémentaires.

4. Conclusion - Perspectives

Ces études RMN, en permettant de localiser à l'échelle de l'acide aminé les zones du domaine NLS-HD Otx2 affectées par l'ajout de ligand et de déterminer des affinités apparentes pour ces zones, nous ont permis de mieux comprendre le mode d'interaction de nos sulfopeptides mimes de GAGs avec l'homéodomaine d'Otx2. Des expériences complémentaires de modélisation ont permis de confirmer le rôle clef du domaine *N*-terminal de liaison aux GAGs de la protéine dans l'interaction avec nos mimes.

Toutefois, le temps nécessaire à la réalisation d'une expérience et la lourdeur du traitement des données ne font pas de la méthode RMN de perturbation de déplacement chimique la méthode de choix pour une étude comparative rapide de l'affinité des mimes de GAGs pour NLS-HD Otx2. De plus, les affinités apparentes obtenues à l'échelle du résidu par RMN donnent une indication approximative de la gamme d'affinité, mais seule la mesure de la constante thermodynamique moléculaire est suffisamment précise pour comparer l'affinité de nos ligands envers la protéine Otx2. D'autre part, en avançant dans cette étude, une interrogation s'est posée : la RMN en solution est-elle la bonne approche pour analyser des interactions de molécules dans un milieu riche en GAGs, comme c'est le cas à la surface des cellules ? Nous avons analysé ici le phénomène en employant des stratégies habituelles d'étude de protéine en solution à laquelle sont progressivement ajoutées des fractions de mimes de GAGs. Ne faut-il pas analyser le phénomène différemment, par exemple en ajoutant de faibles fractions de protéine à des mimes de GAGs concentrés ? Ou, en poussant plus loin le questionnement, l'environnement d'une cellule peut-il être assimilé à une solution comme

dans le tube RMN tel que nous l'avons analysé dans ce chapitre ou plutôt à un « gel dense » ? Pour « créer » un environnement riche en mimes de GAGs sur lequel on ajoute une protéine, une approche SPR de type Biacore nous est apparue plus appropriée.

C. Etude des interactions par Biacore

Dans le but de trouver une approche permettant de déterminer « rapidement » les affinités de nos peptides pour la protéine NLS-HS Otx2, nous nous sommes donc intéressés à la technologie du Biacore. Le Laboratoire des BioMolécules ne possédant pas cet appareil dont l'utilisation nécessite le savoir-faire d'une personne qualifiée, les expériences de mesures d'affinité par Biacore ont été réalisées par Tahar Bouceba, ingénieur au sein de l'Institut de Biologie Paris-Seine (IBPS) de Sorbonne Université.

1. Synthèse de nouveaux analogues

Remarque liminaire : Deux nouveaux sulfopeptides, Ac-Ahx-(EC'PC')₆-NH₂ et Ac-Ahx-(EC'LC')₆-NH₂, ont été étudiés par Biacore et le résultat de cette étude sera exposé plus loin dans ce paragraphe. Le choix de ces analogues et leur synthèse sont donc décrits ci-après.

Notre hypothèse de départ était que les sulfopeptides, mimes peptidiques de GAGs, pouvaient adopter une structuration en hélice. Or, en solution, des peptides de cette taille présentent généralement un mélange de conformères, comme le suggère l'expérience acquise par notre laboratoire dans l'analyse structurale de peptides de tailles variées (± 10 à 30 acides aminés) et le confirme une rapide analyse par dichroïsme circulaire (DC) réalisées sur les sulfopeptides Ac-Ahx-(EC'AC')₃-NH₂ et Ac-Ahx-(EC'AC')₆-NH₂. Le dichroïsme circulaire est basé sur les propriétés d'absorption d'une lumière polarisée de molécules chirales en solution. Avec un peptide, molécule chirale par essence, le DC permet d'évaluer, après déconvolution du spectre, le pourcentage de structure(s) secondaire(s) d'un peptide en solution (hélice versus feuillet et structure aléatoire).

Quoi qu'il en soit, peu ou pas structurés en solution, ces peptides semblent devoir adopter une structure privilégiée pour se lier à la protéine, si on se réfère aux études de modélisation réalisées par le Dr Nicolas Pietrancosta. Dans ces conditions, l'affinité des sulfopeptides pour la protéine cible NLS-HD Otx2 ne pourrait-elle pas être améliorée si, avant

même d'interagir avec la protéine, le mime présentait déjà une structuration en hélice ou, tout au moins, un pourcentage non négligeable de conformères en structure hélicoïdale ? C'est pourquoi, dans le but d'obtenir des sulfopeptides plus structurés en solution, nous avons envisagé de modifier la séquence de base de notre meilleur candidat, Ac-Ahx-(EC'AC')₆-NH₂, *via* une approche justifiée par nos lectures.

Ainsi, selon le même protocole de synthèse que celui employé pour le peptide Ac-Ahx-(EC'AC')₆-NH₂ (pour les détails, se référer au chapitre II-C.4), j'ai synthétisé deux peptides de séquence Ac-Ahx-(EC'XC')₆-NH₂ dans lesquels j'ai remplacé les acides aminés Fmoc-Ala-OH par des Fmoc-Leu-OH (X = L) ou des Fmoc-Pro-OH (X = P). Les Leucines devraient stabiliser une structure hélicoïdale dans le peptide final, avec un coefficient stabilisateur d'hélice du même ordre de grandeur que l'Alanine[203]–[205]. De plus, la présence de Leucines, avec les chaînes isopropyles plus hydrophobes, pourrait également stabiliser, par des interactions hydrophobes, le complexe peptide-protéine Otx2. A l'inverse, la Proline est considérée comme un interrupteur d'hélice α , mais pourrait aussi favoriser des conformations privilégiées de type polyprolines.

Après synthèse, la coupure du support et l'oxydation des mélanges réactionnels ont été réalisées selon le mode opératoire décrit pour les précédents peptides (*Cf.* chapitre II-C.3.a et b), à l'exception d'une simple modification de traitement final. En effet, pour éviter les sels dans le mélange final après oxydation, j'ai supprimé la neutralisation par l'ammoniaque pour n'ajouter qu'un grand volume d'eau froide (x 5) avant lyophilisation[206]. Les mélanges finaux ont ensuite été purifiés par HPLC préparative puis les peptides purifiés ont été analysés par HPLC analytique, **Figure 102** et par spectrométrie de masse MALDI-TOF avec une matrice *p*-nitroaniline, en présence de pyrène-méthylguanidinium, **Figure 103**, avant d'être lyophilisés.

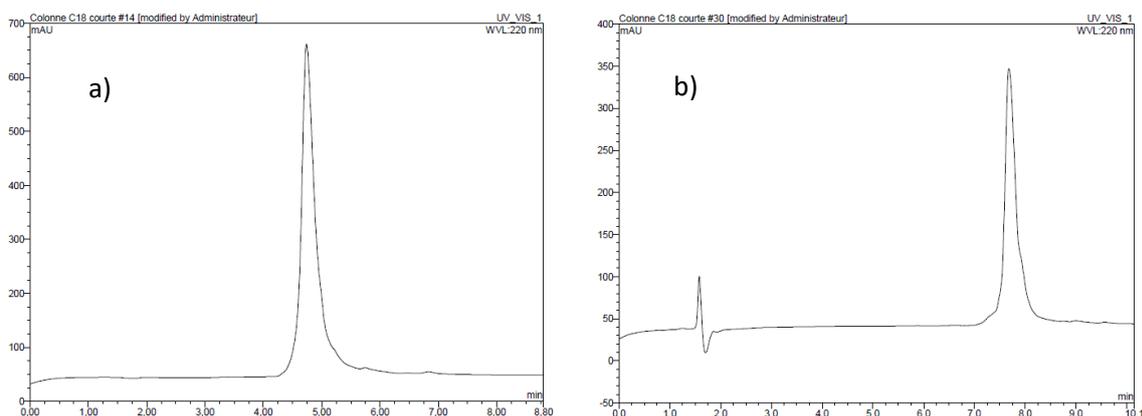


Figure 102. Profils HPLC en phase inverse de : a) Ac-Ahx-(EC'PC')₆-NH₂ purifié et b) Ac-Ahx-(EC'LC')₆-NH₂ purifié. Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 5 µm, 200 Å, 4,6x100 mm – a) Elution avec un gradient en 10 min de 2 à 15% de CH₃CN (0,1% TFA) – b) Elution avec un gradient en 10 min de 5 à 20% de CH₃CN (0,1% TFA) - Débit : 1 mL/min - Lecture à 220 nm.

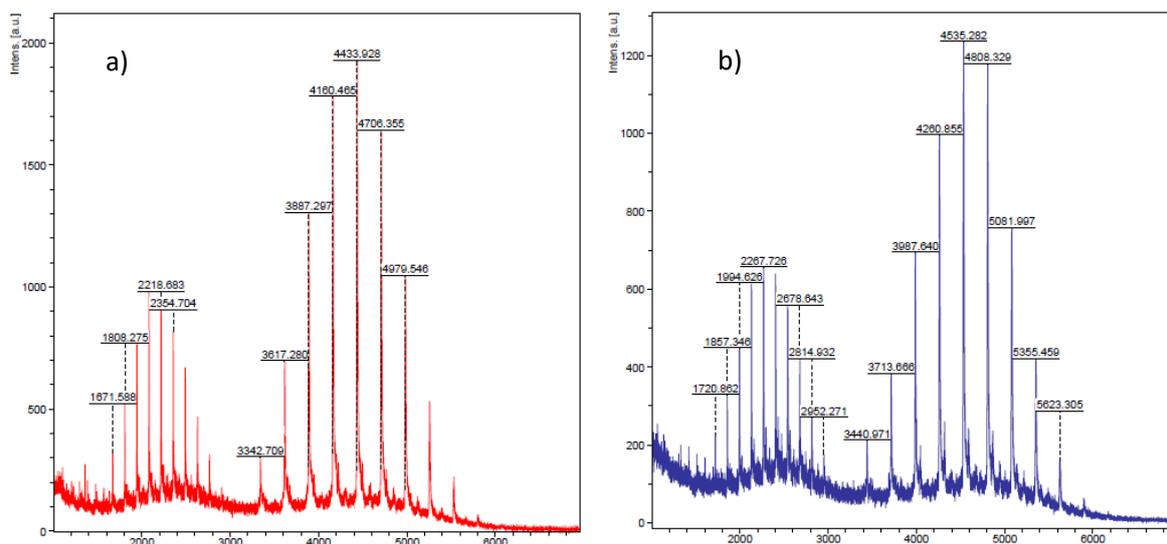


Figure 103. Profil MALDI-TOF de : a) Ac-Ahx-(EC'PC')₆-NH₂ et b) Ac-Ahx-(EC'LC')₆-NH₂, avec matrice *p*-nitroaniline en présence de pyrèneméthylguanidinium (pmg), en mode linéaire négatif.

Les peptides ont ensuite été mis en solution dans de l'eau distillée et leur pH a été amené autour de 7 par ajout d'une solution de NaOH à 1M. La mesure de leur absorbance à 205 nm a alors permis de déterminer la concentration des solutions, en utilisant la loi de Lambert-Beer (*cf.* chapitre II-C.3.c).

2. Le Biacore, principe et résultats

a) Principe de la technologie Biacore

« Le Biacore » (Bia : pour "biomolecular interaction analysis" et core : pour "au cœur de") est le nom commercial d'un appareil mis au point par la Société Pharmacia (rachetée par GE Healthcare). Cet appareil a été mis au point pour analyser les interactions protéine/protéine ou protéine/ligand. La technologie mise en jeu repose sur les principes de la résonance plasmonique de surface (SPR pour *Surface Plasmon Resonance*) qui est une technique optique. Comme très fréquemment, la technique a pris le nom de l'objet et on parle de Biacore et non de SPR !

Ce phénomène de résonance plasmonique de surface se produit lorsqu'une lumière polarisée rencontre un film métallique à l'interface de supports avec des indices de réfraction différents. Sous certaines conditions (longueur d'onde, polarisation et angle d'incidence), les électrons libres à la surface du prisme recouvert d'une fine couche métallique, en général de l'or, absorbent les photons incidents et les convertissent en ondes de plasmons de surface. À un certain angle d'incidence (ou angle de résonance), les plasmons vont résonner avec la lumière, entraînant l'absorption de la lumière à cette valeur d'angle et créant une ligne sombre dans le faisceau réfléchi, comme illustré **Figure 104**. Ces conditions de SPR étant très sensibles, lorsqu'une interaction moléculaire se produit entre un partenaire immobilisé sur la puce et un deuxième partenaire circulant à la surface de la puce, l'indice de réfraction change et les conditions de SPR sont modifiées (**Figure 104**).

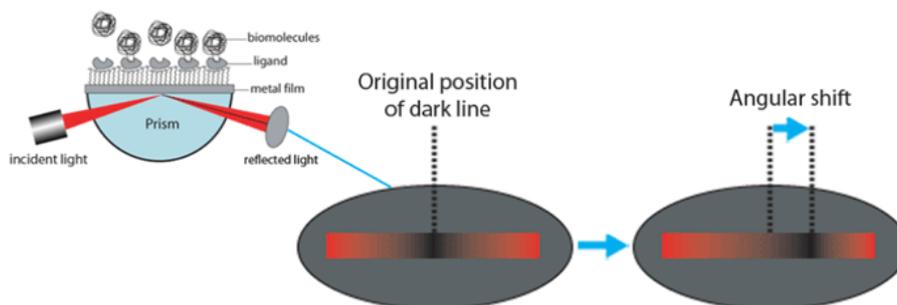


Figure 104. L'excitation des plasmons de surface se traduit par une ligne sombre dans le faisceau réfléchi, et la position angulaire de la ligne sombre se déplace lors de la liaison de molécule. Figure extraite de[207].

La SPR permet donc la détection des variations de masse à la surface d'une puce (sensorchip) sur lequel un des deux partenaires est immobilisé alors que l'autre partenaire est injecté dans un flux continu de tampon à la surface de la puce. La liaison du partenaire injecté, **I**, avec le partenaire fixé sur la puce, **F**, induit une réponse proportionnelle à la quantité (masse) de **I** lié. Le décalage de la courbe de réflectivité (**Figure 105a**) est donc lié à l'interaction de deux partenaires sur ou à proximité du film métallique. En analysant ce décalage en fonction du temps, on peut déduire la cinétique de cette liaison (**Figure 105b**).

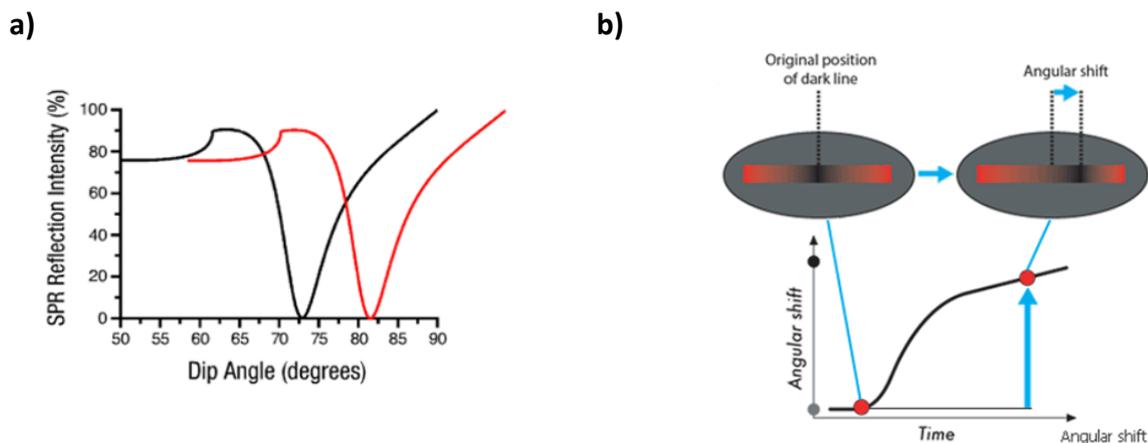


Figure 105. La résonance plasmonique de surface a) La SPR provoque une baisse d'intensité de la lumière réfléchiée à la surface du capteur, permettant de déterminer l'angle de résonance. Le décalage angulaire observé correspond à l'interaction entre les deux partenaires. **b)** La mesure du décalage angulaire en fonction du temps permet d'étudier la cinétique de liaison. Figures extraites de[207].

La réponse est quantifiée en « resonance units » (RU). Elle est directement proportionnelle à la concentration de complexes formés à la surface de la puce. Un sensorgramme reflète la réponse (RU) en fonction du temps et permet de visualiser l'évolution de l'interaction entre les deux partenaires, **Figure 106**.

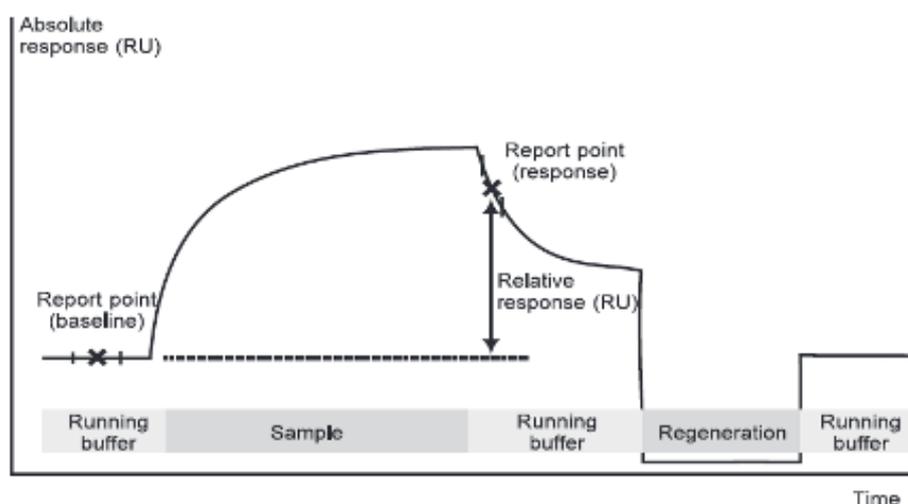


Figure 106. Sensorgramme illustrant l'interaction en fonction du temps de deux partenaires, l'un fixé sur une puce (sensorchip) et l'autre injecté en continu à la surface de cette puce. Figure extraite de[208].

Les avantages de la résonance plasmonique de surface sont de deux ordres. Elle ne nécessite pas de marquage préalable des partenaires étudiés et elle autorise une détection en temps réel de l'interaction entre les deux partenaires. Ce suivi en temps réel permet de déterminer les constantes cinétiques d'association (k_{on}) et de dissociation (k_{off}), et d'en déduire la valeur de la constante d'affinité (K_d). Qui plus est, les puces fonctionnalisées (sensorchip) peuvent être réutilisées après avoir été régénérées par injection d'une solution dite de régénération.

Les puces que nous avons utilisées pour étudier l'interaction entre NLS-HD Otx2 et les mimes de GAGs ont été fonctionnalisées à partir de puce de type CM5 de GE Healthcare, **Figure 107a**. Le film d'or déposé sur ces puces est recouvert par une matrice de dextran carboxyméthylé sur laquelle nous avons fixé de façon covalente de la streptavidine (SA), en suivant les modes opératoires proposés par GE Healthcare, **Figure 108**. Nous avons alors complexé sur les sites de la streptavidine les mimes biotinylés Biot(O₂)-Ahx-(EC'AC')_n-NH₂ avec $n = 3$ ou $n = 6$, **Figure 107b**. L'interaction de la biotine avec la streptavidine est quasi covalente puisque le K_d de l'interaction est de l'ordre de 10^{-14} M, même si une diminution d'affinité est envisageable pour la biotine fixée sur une chaîne peptidique[181], ce que nous avons tenté de compenser en introduisant un bras espaceur de type acide 6-aminohexanoïque, Ahx.

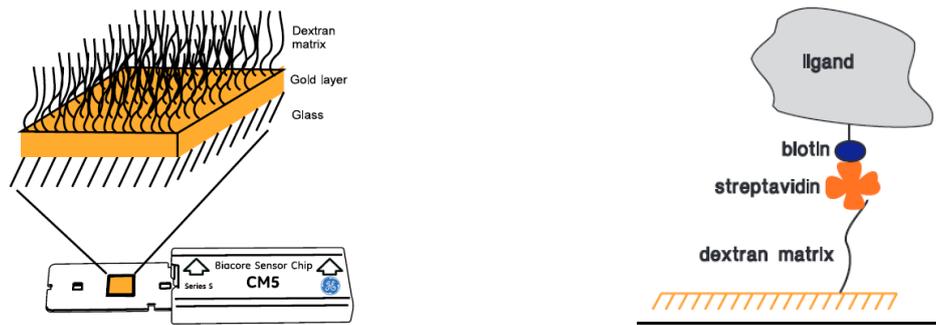


Figure 107. a) Illustration d'une puce de type CM5 (GE Healthcare) que nous avons utilisée pour nos expériences après b) fixation covalente de streptavidine. Figures extraites de[208].

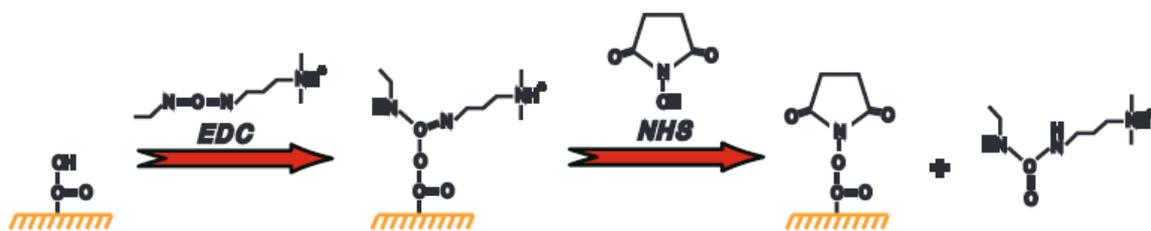


Figure 108. Activation de la matrice carboxyméthylé par le couple EDC/NHS (1-éthyl-3-(3-diméthylaminopropyl)carbodiimide et N-hydroxysuccinimide), pour fixer la streptavidine. Figure extraite de[208].

b) Résultats des expériences réalisées sur Biacore

Mesure des interactions directes

La protéine NLS-HD-Otx2 non marquée, que j'ai produite et purifiée, a été injectée, en flux continu et à différentes concentrations, sur la puce fonctionnalisée par de la streptavidine (CM5-SA), elle-même complexée à un mime biotinylé, Biot(O₂)-Ahx-(EC'AC')_n-NH₂ avec n = 3 ou n = 6. Une augmentation du signal correspondant à l'interaction mime/NLS-HD Otx2 (permettant de déterminer un k_{on}) puis, après interruption du flux, une diminution de ce signal (permettant de déterminer un k_{off}) ont été observées. Des affinités pour les complexes formés, K_d = k_{off}/k_{on}, de l'ordre de quelques dizaines de nM, ont ainsi été calculées, quelle que soit la sonde fixée, Biot(O₂)-Ahx-(EC'AC')₃-NH₂ ou Biot(O₂)-Ahx-(EC'AC')₆-NH₂, un exemple de sensorgramme est donné, pour chacune des sondes, **Figure 109**.

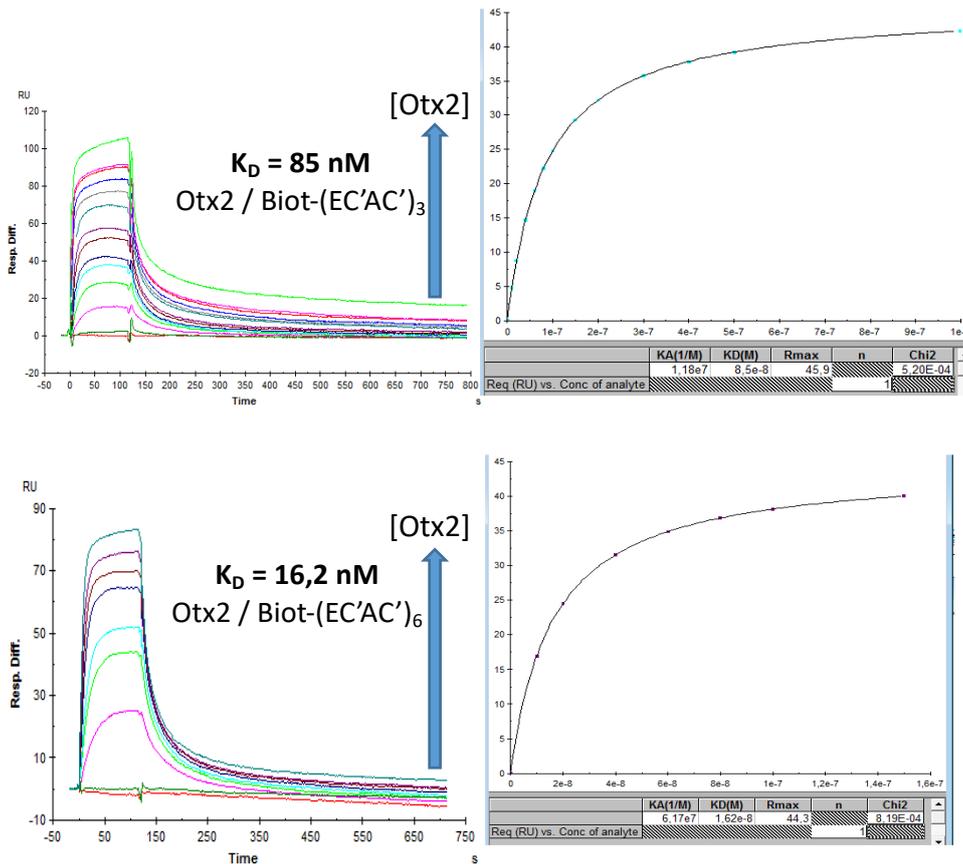


Figure 109. Mesure de l'affinité par Biacore de l'interaction des mimes avec NLS-HD d'Otx2, sensogramme. Figures en haut : Un exemple de la cinétique de fixation de NLS-HD Otx2 (injecté à 0, 10, 20, 40... 1000 nM) sur une puce CM5-SA après fixation de Biot(O₂)-Ahx-(EC'AC')₃-NH₂ sur la streptavidine (SA). **Figures en bas :** Un exemple de la cinétique de fixation de NLS-HD d'Otx2 (injecté à 0, 10, 20, 40... 1000 nM) sur une puce CM5-SA après fixation de Biot(O₂)-Ahx-(EC'AC')₆-NH₂. **Figures de gauche :** Les réponses (RU) en fonction du temps pour chaque concentration de protéine injectée. **Figures de droite :** Réponses à l'équilibre en fonction de la concentration en protéine injectée.

Les affinités, K_d , des sulfopeptides Biot(O₂)-Ahx-(EC'AC')₃-NH₂ et Biot(O₂)-Ahx-(EC'AC')₆-NH₂ pour NLS-HD Otx2 présentent peu de différence (facteur 4) entre la version courte et la version longue, étant respectivement de 86 nM et 23 ± 7 nM (moyennes obtenues à partir des résultats de 3 expériences). Ainsi, ces affinités nanomolaires sont autour de 500 fois meilleures que celles trouvées par RMN. La BSA (pour Albumine de Serum Bovin), utilisée comme témoin de la spécificité de reconnaissance de NLS-HD Otx2 pour les mimes de GAGs, n'a montré aucune affinité pour la surface de la puce. Toutefois, à ce stade, il convient de rappeler que la matrice déposée sur la puce est du dextran carboxyméthylé, que la streptavidine fixée sur cette matrice a un point isoélectrique (pI) autour de 5 et que le pI de NLS-HD-Otx2 est autour de 10,5 alors que celui de la BSA est autour de 5. La BSA n'est donc probablement pas la meilleure des protéines pour évaluer la composante de liaison non-

spécifique. Cependant, l'injection de NLS-HD Otx2 sur une puce CM5-SA ne portant pas de mime biotinylé n'a montré aucune interaction de la protéine avec la surface de la puce, ce qui prouverait que les résultats observés avec les puces CM5-SA porteuses de mimes biotinylés sont bien dus à des interactions entre les mimes fixés et la protéine NLS-HD Otx2 circulante.

J'ai écrit dans le paragraphe « 4. Conclusion – Perspectives » de l'étude des interactions par RMN de ce chapitre : « *Nous avons analysé ici le phénomène en employant des stratégies habituelles de la protéine en solution à laquelle on ajoute des fractions de mimes de GAGs. Ne faut-il pas analyser le phénomène différemment, par exemple en ajoutant de faibles fractions de protéine à des mimes de GAGs concentrés ? Ou, en poussant plus loin le questionnement, l'environnement d'une cellule peut-il être assimilé à une solution comme dans le tube RMN tel que nous l'avons analysé dans ce chapitre ou plutôt à un « gel dense » ? Pour « créer » un environnement riche en mimes de GAGs sur lequel on ajoute une protéine, une approche de type Biacore nous est apparue probablement plus appropriée.* ». Mais entre des affinités nanomolaires obtenues par Biacore et des affinités micromolaires par RMN : que « croire » ?

La science, ne relevant pas du domaine de la croyance, est une affaire d'hypothèses : la fixation des mimes sur la puce de streptavidine doit conduire à une densité locale élevée, induisant une synergie de reconnaissance. Dans la technique Biacore, c'est une masse fixée qui « est mesurée », donc ces affinités de 15 à 100 nM reflètent les molécules de Otx2 qui, pour une concentration en protéine donnée, sont retenues par le mime fixé sur la puce et ceci en plus grand nombre qu'escompté par rapport aux K_d estimés par RMN. Synergie signifierait qu'ici, grâce à une présentation en deux dimensions des mimes (sur support vs. en solution), NLS-HD Otx2 pourrait trouver plus « facilement » un partenaire. On pourrait envisager une fixation concomitante de la protéine à différentes molécules de mimes fixés sur la puce ou, du moins, après libération de la protéine d'avec un premier mime, une cinétique d'interaction avec un nouveau partenaire rendue plus rapide par la proximité spatiale des mimes fixés.

Tout ceci nous rappelle la présentation en réseau des GAGs dans la matrice extracellulaire. L'affinité des unités disaccharidiques prises individuellement n'est pas élevée, c'est la multitude et la proximité spatiale de ces sites d'interaction faible qui permet aux

protéines cibles d'être captées dans le milieu puis transportées de proche en proche le long du réseau de GAGs pour être acheminées vers la cellule et y être internalisées.

Mais, dans ces conditions, si les mimes fixés à la puce n'agissent pas comme des molécules individuelles mais comme un réseau de molécules, peut-on encore déterminer une affinité par un simple calcul à deux états, libre vs. lié ?

Mesure des interactions par compétition

Nous nous sommes alors engagés dans des analyses par compétition, pour nous permettre de comparer les analogues synthétisés et avoir une stratégie de comparaison pour des études ultérieures de structure-affinité. Le mime Biot(O₂)-Ahx-(EC'AC')₆-NH₂ a été fixé sur la puce, comme détaillé ci-dessus, et nous avons fait passer un flux de protéine, soit NLS-HD Otx2 seul, soit NLS-HD Otx2 incubé au préalable avec des concentrations variables de ligands compétiteurs sous leur forme acétylée en *N*-terminal. L'idée était que l'utilisation d'un mime « efficace », c'est-à-dire capable d'entrer en compétition avec Biot(O₂)-Ahx-(EC'AC')₆-NH₂ fixé sur la puce, entraînerait une diminution du signal par rapport à la réponse maximale (RU_{max}) obtenue par la fixation de NLS-HD Otx2 injecté seul sur Biot(O₂)-Ahx-(EC'AC')₆-NH₂. Cette diminution reflèterait la capacité du mime compétiteur à se fixer, à une concentration plus ou moins forte, à NLS-HD Otx2 en solution, rendant ce dernier moins disponible pour une fixation aux mimes de la puce.

Nous avons ainsi testé Ac-Ahx-(EC'AC')_n-NH₂ avec n = 6, 5 4 ou 3, Ac-Ahx-(EAC'C')₅-NH₂, Ac-Ahx-(EC'C'A)_n-NH₂ avec n = 6 ou 5, HI dp8, PolyGlu et les nouveaux analogues décrits dans le paragraphe précédent, Ac-Ahx-(EC'PC')₆-NH₂ et Ac-Ahx-(EC'LC')₆-NH₂. Quelques sensorgrammes sont donnés, à titre d'exemples, dans les figures suivantes **Figures 110, 111 et 112.**

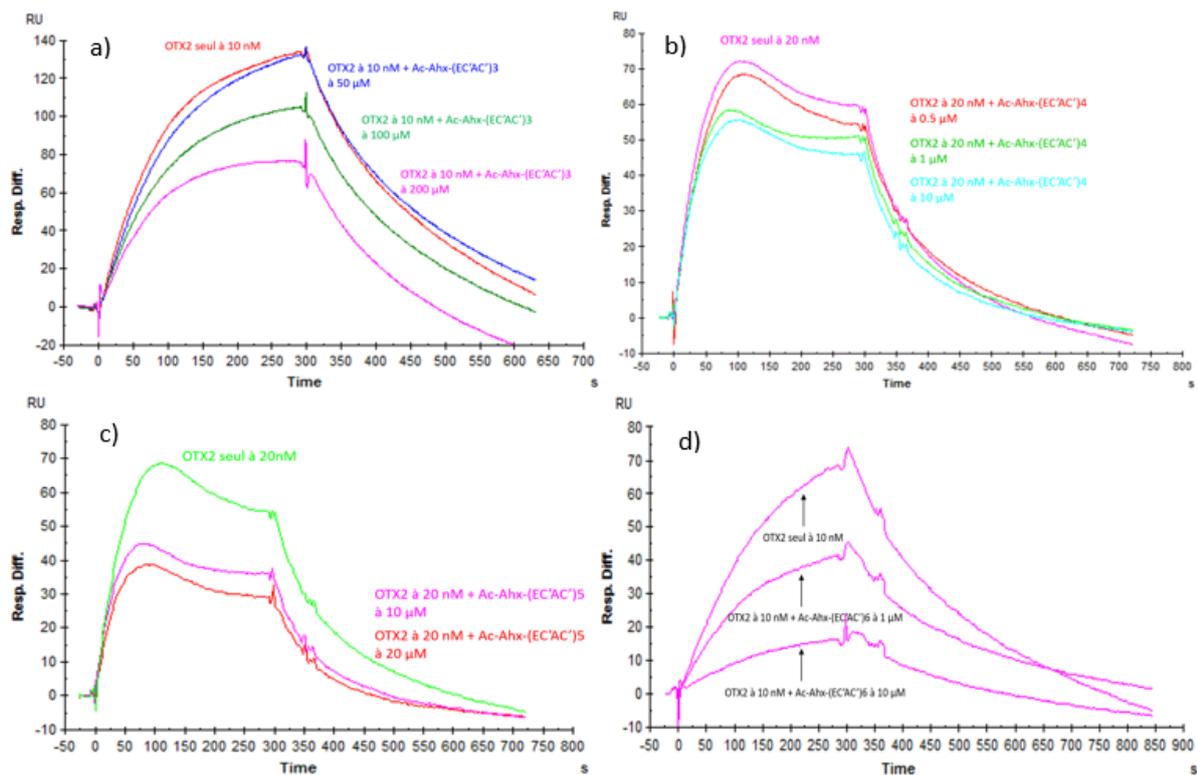


Figure 110. Analyse par Biacore de l'interaction de Ac-Ahx-(EC'AC')₃-NH₂, Ac-Ahx-(EC'AC')₄-NH₂, Ac-Ahx-(EC'AC')₅-NH₂ ou Ac-Ahx-(EC'AC')₆-NH₂ avec NLS-HD Otx2 (10 ou 20 nM), exemples d'expériences de compétition avec Biot(O₂)-Ahx-(EC'AC')₆-NH₂ fixé sur puce « coâtée » par de la streptavidine : a) compétition avec Ac-Ahx-(EC'AC')₃-NH₂ incubé à 50, 100 ou 200 μM avec NLS-HD Otx2 (10 nM), b) compétition avec Ac-Ahx-(EC'AC')₄-NH₂ incubé à 0,5, 1 ou 10 μM avec NLS-HD Otx2 (20 nM), c) Ac-Ahx-(EC'AC')₅-NH₂ incubé à 10 ou 20 μM avec NLS-HD Otx2 (20 nM), d) Ac-Ahx-(EC'AC')₆-NH₂ incubé à 1 ou 10 μM avec NLS-HD Otx2 (10 nM).

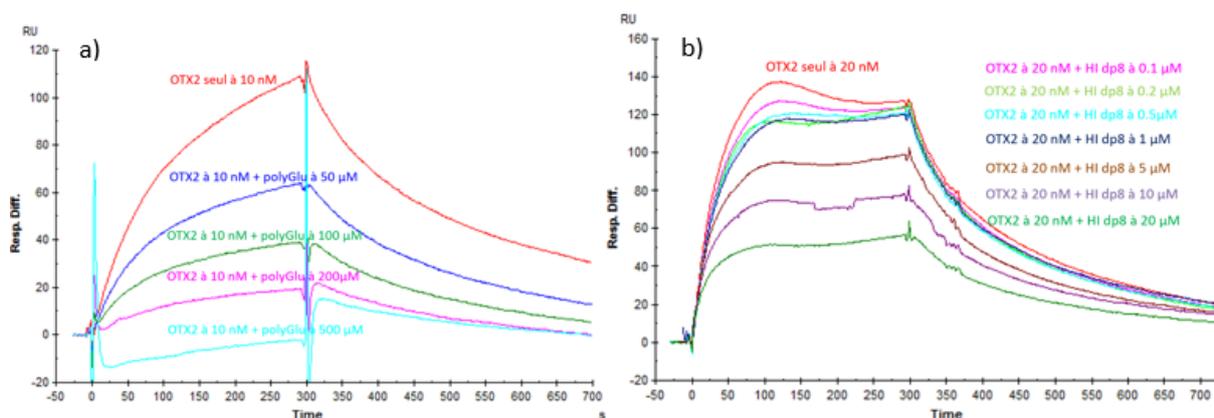


Figure 111. Analyse par Biacore de l'interaction de polyGlu ou Hipd8 avec NLS-HD Otx2 (10 ou 20 nM), exemples d'expériences de compétition avec Biot(O₂)-Ahx-(EC'AC')₆-NH₂ fixé sur puce « coâtée » par de la streptavidine : a) compétition avec polyGlu incubé de 50 à 500 μM avec NLS-HD Otx2 (10 nM) et b) compétition avec HI dp8 incubé de 0,1 à 20 μM avec NLS-HD Otx2 (20 nM).

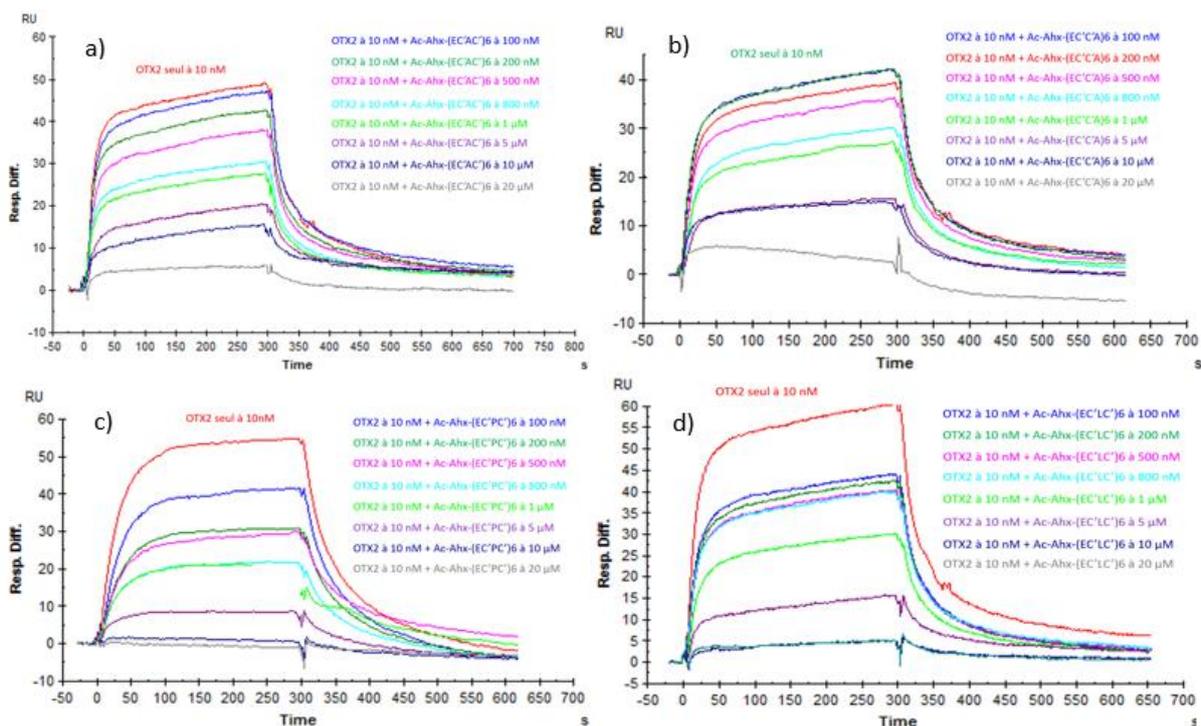


Figure 112. Analyse par Biacore de l'interaction de Ac-Ahx-(EC'AC')₆-NH₂, Ac-Ahx-(EC'C'A)₆-NH₂, Ac-Ahx-(EC'PC')₆-NH₂ ou Ac-Ahx-(EC'LC')₆-NH₂ avec NLS-HD Otx2 (10 nM), exemples d'expériences de compétition avec Biot(O₂)-Ahx-(EC'AC')₆-NH₂ fixé sur puce « coatée » par de la streptavidine : a) compétition avec Ac-Ahx-(EC'AC')₆-NH₂, b) compétition avec Ac-Ahx-(EC'C'A)₆-NH₂, c) compétition avec Ac-Ahx-(EC'PC')₆-NH₂, d) compétition avec Ac-Ahx-(EC'LC')₆-NH₂. Tous les peptides ont été incubés avec NLS-HD Otx2 (10 nM) avec des concentrations de peptides variant de 100 nM à 20 μM.

Certaines de ces expériences sont encore préliminaires et demanderont à être reproduites afin de pouvoir calculer des valeurs d'affinité relative (CI_{50}) fiables. Je ne donnerai donc que des tendances, avec des calculs de « CI_{50} estimée » sans erreur standard. La « CI_{50} estimée » correspond à la concentration du compétiteur qui, après incubation avec NLS-HD Otx2, donnera une concentration de NLS-HD Otx2 libre conduisant à une diminution de 50 % du signal RU_{max} . Les IC_{50} sont calculées en utilisant l'équation de Hill-Langmuir pour un modèle 1:1[209]: $RU = RU_{max} / (1 + CI_{50} / [C])$, [C] représentant la concentration en compétiteur. En traçant RU_{max} / RU en fonction de [C], alors, à $RU_{max} / RU = 2$, on obtient $[C] = CI_{50}$. Un exemple représentant $RU_{max} / RU = fn([Ac-Ahx-(EC'PC')_6-NH_2])$ est donné **Figure 113**.

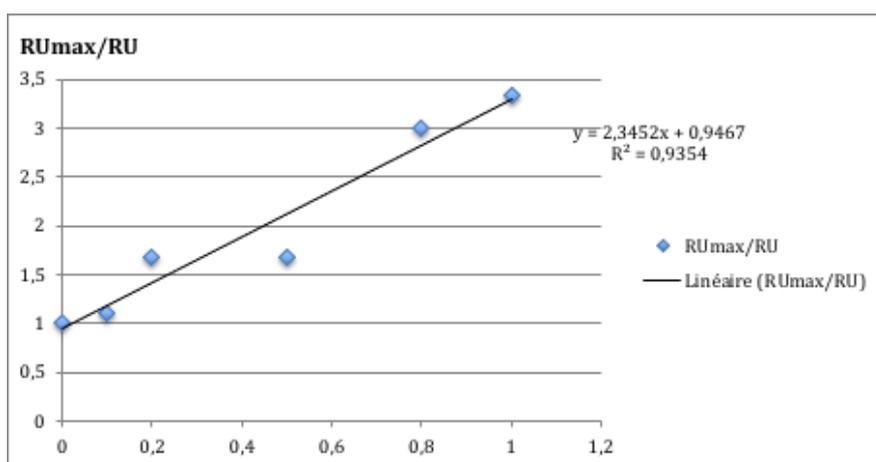


Figure 113. Exemple de détermination de CI_{50} par Biacore avec une expérience de compétition conduite avec NLS-HD Otx2 (10 nM) et des concentrations variables du compétiteur Ac-Ahx-(EC'PC')₆-NH₂. Le mime fixé sur la puce est Biot(O₂)-Ahx-(EC'AC')₆-NH₂. Tracé de $RU_{max}/RU = fn([Ac-Ahx-(EC'PC')_6-NH_2])$.

Les affinités « CI_{50} estimée » de huit molécules sont présentées dans le **Tableau 12**.

Compétiteur	Ac-Ahx-(EC'PC') ₆ -NH ₂	Ac-Ahx-(EC'LC') ₆ -NH ₂	Ac-Ahx-(EC'C'A) ₆ -NH ₂	Ac-Ahx-(EC'AC') ₆ -NH ₂
CI_{50} (μM)	0,5	1	1,5	2,5
Compétiteur	HI dp8	Ac-Ahx-(EC'AC') ₅ -NH ₂	PolyGlu	Ac-Ahx-(EC'AC') ₃ -NH ₂
CI_{50} (μM)	17	22	30	250

Tableau 12. « CI_{50} estimée » : Expériences de compétition par Biacore pour les différents compétiteurs étudiés. Biot-Ahx-(EC'AC')₆-NH₂ est fixé sur la puce. NLS-HD Otx2 (à 10 ou 20 nM) est injecté après préincubation avec différentes concentrations de compétiteur.

Dans ce tableau, j'ai classé les huit compétiteurs du meilleur au moins bon : Ac-Ahx-(ECP'C')₆-NH₂, Ac-Ahx-(EC'LC')₆-NH₂, Ac-Ahx-(EC'C'A)₆-NH₂, Ac-Ahx-(EC'AC')₆-NH₂, HI dp8, Ac-Ahx-(EC'AC')₅, polyGlu et Ac-Ahx-(EC'AC')₃-NH₂. Pour les compétiteurs de première génération, les CI_{50} , de 1,5 μM à 250 μM, présentent donc un facteur 170 entre le meilleur et le moins bon des compétiteurs. Plus le nombre de motif (EC'AC')_n augmente (n = 3 à 6) meilleure est l'affinité.

Concernant Ac-Ahx-(EC'C'A)₆-NH₂, Ac-Ahx-(EC'AC')₆-NH₂, HI dp8, polyGlu et Ac-Ahx-(EC'AC')₃-NH₂, nous retrouvons sensiblement le même classement que celui obtenu avec les K_d apparents des expériences de RMN (cf. paragraphe B.2.b) de ce chapitre)

Les derniers analogues que j'ai synthétisés, avec une Proline ou une Leucine à la place de l'Alanine, sont meilleurs que Ac-Ahx-(EC'AC')₆-NH₂, avec respectivement une CI₅₀ de 0,5 µM et 1 µM. Un facteur 2,5 à 5 a été gagné par rapport à l'analogue portant une Alanine.

3. Discussion

Les analogues de deuxième génération portant une Proline ou une Leucine à la place de l'Alanine ont des affinités meilleures pour NLS-HD Otx2 que les mimes de première génération, cependant le gain n'est au maximum que d'un facteur 5. Le résultat obtenu avec Ac-Ahx-(EC'PC')₆-NH₂ suggère que, fixé à NLS-HD Otx2, le mime pourrait adopter une conformation privilégiée induite par les Prolines en i, i+4, mais laquelle ? Ceci restera à analyser. L'analogue portant une Leucine en i, i+4 peut aussi adopter une meilleure structure hélicoïdale et/ou favoriser des interactions hydrophobes stabilisantes entre la chaîne isopropyle de la Leucine et NLS-HD Otx2. Pour savoir s'il y a additivité des phénomènes, structure particulière et interaction hydrophobe stabilisante, nous pourrions introduire un analogue contraint de la Leucine, « ProLeu », dont le laboratoire s'est fait une spécialité, tout particulièrement le Pr Philippe Karoyan et ses étudiants[210]–[213]. Quoi qu'il en soit, ces résultats nous ont d'ores et déjà convaincus d'explorer la piste des peptides à agrafes, peptides « préstructurés », dont la synthèse sera évoquée dans le chapitre suivant.

Les affinités, K_d, des sulfopeptides Biot-Ahx-(EC'AC')₃-NH₂ et Biot-Ahx-(EC'AC')₆-NH₂ pour NLS-HD Otx2 par mesure directe au Biacore sont respectivement de 86 nM et 23 ± 7 nM, soit environ 1000 fois meilleure que ce qui est obtenu par RMN, ce qui est très intrigant ! J'ai invoqué la concentration élevée fixée sur la puce, qui pourrait conduire à cet artefact et il me faudrait donc faire de nouvelles déterminations de K_d avec des taux de fonctionnalisation moindre de la puce, c'est-à-dire avec des concentrations de Biot-Ahx-(EC'AC')₆-NH₂ moindre. Ceci relèverait toutefois encore un peu de « la boîte noire », puisqu'en théorie la streptavidine, avec 4 sites de fixation pour la biotine, a une affinité qui relève de la liaison « quasi-covalente », c'est à dire qu'il n'y a pas de réarrangement possible, la vitesse de dissociation étant très lente. Donc, même en utilisant une concentration moindre de Biot-Ahx-(EC'AC')₆-NH₂, nous pourrions toujours avoir une hétérogénéité de fonctionnalisation en surface de la couche de streptavidine, avec des zones localement concentrées et d'autres vides de mime. Une autre idée serait de diminuer la concentration de streptavidine utilisée

pour la fonctionnalisation de la puce. Ainsi, avec moins de streptavidine fixée, il serait possible d'obtenir une quantité moindre de Biot-Ahx-(EC'AC')₆-NH₂ fixé.

D. Conclusion

Les affinités relatives, CI₅₀, trouvées par les expériences de compétition par Biacore sont du même ordre de grandeur que les K_d apparents déterminés par RMN, ce qui est très satisfaisant !

- Les affinités relatives, CI₅₀, de Ac-Ahx-(EC'AC')₃ et Ac-Ahx-(EC'AC')₆ pour NLS-HD Otx2 trouvées par les expériences de compétition par Biacore sont respectivement 250 μM et 2,5 μM.
- Les affinités apparentes, K_d, pour les résidus de NLS-HD Otx2 dont le déplacement chimique du NH est le plus affecté par l'interaction avec Ac-Ahx-(EC'AC')₃-NH₂, sont de l'ordre de 300 à 350 μM pour les 3 premiers acides aminés du domaine de liaison aux GAGs et de 100 à 120 μM pour les 4 acides aminés suivants de ce domaine. Concernant le peptide Ac-Ahx-(EC'AC')₆-NH₂, l'affinité pour le domaine de liaison aux GAGs et la région C-terminale de l'hélice 3 est autour de 8 μM.

Nous pouvons donc nous réjouir d'avoir mis au point une stratégie de criblage (« screening ») par Biacore facile à mettre en œuvre pour des études structure/affinité. Il restera bien évidemment à reproduire ces expériences, ne serait-ce que pour obtenir des CI₅₀ avec des erreurs standards.

Nous ne reprendrons des études structurales par RMN, telles que décrites dans ce chapitre, qu'avec le meilleur des analogues que nous aurons synthétisés, sachant que si un facteur 10 en affinité pouvait être gagné, soit un CI₅₀ de l'ordre d'une centaine de nM à quelques dizaines de nM, le régime d'échange pourrait changer dans le complexe ligand-protéine pour passer de rapide à plus lent, voire lent.

E. Partie expérimentale

1. Production de protéines recombinantes et RMN

a) Production des protéines recombinantes

Généralités

Le milieu LB, utilisé pour la culture des bactéries *E. Coli*, est constitué d'extrait de levure (5 g/L), de peptone de caséine (10 g/L), de NaCl (10 g/L), mis en solution dans de l'eau distillée dont le pH est ensuite ajusté à 7,2. Ce milieu de culture est stérilisé par passage en autoclave (121 °C, 20 min, 2 bars).

Les solutions stock de bactéries *E. Coli* BL21 DE3 transformées par le plasmide d'intérêt sont conservées à -80 °C dans des tubes type « eppendorf » contenant du milieu LB, de la Kanamycine (30 µg/mL) et 25 % de glycérol, pour un volume total de solution stock de 400 µL par tube.

Le milieu minimum, utilisé pour l'incorporation de ¹⁵N et ¹³C dans les acides aminés de la protéine marquée, est composé de Na₂HPO₄ anhydre (6,78 g/L), de KH₂PO₄ (3 g/L), de NaCl (0,5 g/L), de MgSO₄ (1 mM), de CaCl₂ (0,1 mM), de Thiamine (10 mg/L) et de Yeast nitrogen Base (1,7 g/L), mis en solution dans de l'eau distillée dont le pH est ensuite ajusté à 7,2.

Le tampon de lyse bactérienne est composé de 20 mM de NaH₂PO₄, de 500 mM de NaCl, de 1 mM de DTT et de 25 mM d'imidazole, mis en solution dans de l'eau distillée dont le pH est ensuite ajusté à 8.

Les tampons utilisés en FPLC pour la purification par chromatographie d'affinité sur colonne de Nickel HisTrap sont :

- Le tampon de liaison dont la composition est identique à celle du tampon de lyse décrit précédemment, pH 8. Cependant l'eau utilisée pour ce tampon est de l'eau ultrapure MilliQ.
- Le tampon d'élution qui diffère du tampon de liaison uniquement par sa concentration en imidazole qui est ici de 500 mM et non de 25 mM. Cette forte concentration en imidazole permet de rompre les liaisons entre les fonctions

imidazoles des six His introduites dans l'étiquette en *N*-terminal de la protéine de fusion et les Ni²⁺ de la colonne HisTrap, libérant ainsi la protéine de fusion.

Le tampon de dialyse est composé de 150 mM de NaCl, de 20 mM de Tris, de 1 mM de DTT, de 20 mM d'imidazole et de 1 mM d'EDTA, mis en solution dans de l'eau ultrapure MilliQ dont le pH est ensuite ajusté à 7,5.

Les électrophorèses ont été réalisées à 180 V dans un tampon de migration Tris-Glycine (Tris-HCl (25 mM), Glycine (200 mM), SDS (0,1 %) à pH 8,3). Le gel d'électrophorèse utilisé est composé de 12,5 % d'acrylamine, de 0,1 % de SDS et de 25 % de Tris-HCl à 3 M – pH 8,8. Il est surmonté d'un gel dit « stacking » (servant à concentrer les échantillons) composé de 4,5 % d'acrylamine, de 0,1 % de SDS et de 12,6 % de Tris-HCl à 0,5 M – pH 6,8. La réaction de polymérisation de l'acrylamine est initiée par l'APS (persulfate d'ammonium) et catalysée par le TEMED (*N,N,N',N'*-tétraméthyl-1-,2-diaminométhane). Avant de les déposer sur ce gel SDS-PAGE 12,5 %, les échantillons sont dilués par deux dans un tampon de charge composé à 80 % de bleu de charge 2X et à 20 % d'une solution de β-mercaptoéthanol (20 % dans l'éthanol), puis sont portés à ébullition durant 10 minutes pour dénaturer la protéine et permettre au bleu de charge d'interagir avec les protéines. En fin de migration, le gel est plongé quelques instants dans une solution d'acide acétique à 10 % dans l'eau, puis les protéines sont révélées pendant 2 h sous agitation par une solution 0,1 % de bleu de Coomassie dans EtOH/H₂O/CH₃COOH (45/45/10). L'excès de bleu de Coomassie est éliminé sous agitation du gel dans de l'acide acétique 10 % dans l'eau.

Le tampon de Filtration sur Gel (FG) est composé de 100 mM de NaCl, de 40 mM de Na₂HPO₂ et de 2 mM de DTT, mis en solution dans de l'eau ultrapure MilliQ dont le pH est ensuite ajusté à 6,3.

Surexpression de NLS-HD Otx2 non marqué.

Afin d'obtenir toutes les bactéries dans un même état de croissance et avant de procéder à la culture à plus grande échelle, les bactéries doivent être mises en préculture. Pour cela le contenu d'un tube de solution stock, soit 400 µL, est incorporé dans un erlen contenant 100 mL de milieu LB stérile en présence de kanamycine (50 µg/mL). L'erlen est incubé une nuit à 30 °C sous agitation (180 rpm). Après mesure de la DO à 600 nm, cette

préculture est diluée dans 8 erlens de 2 litres contenant chacun 475 mL de LB afin d'obtenir une DO finale de 0,05 environ dans chaque erlen.

Après addition de kanamycine (50 µg/mL), les erlens sont incubés à 37 °C sous agitation (180 rpm) en suivant régulièrement la DO, jusqu'à obtention d'une DO de 0,6 environ, ce qui correspond à la phase exponentielle de croissance des bactéries. L'expression de la protéine est alors induite par addition d'IPTG à 500 µM en concentration finale. L'incubation est poursuivie pendant 5 h à 27 °C et à 180 rpm.

Rq : À cette température les bactéries arrêtent de se multiplier. L'IPTG induit l'expression de la protéine de fusion, qui dans le plasmide a été placée sous le contrôle du promoteur lac, en déplaçant la protéine LacI de son site de liaison à LacO. En conséquence, l'ARN polymérase T7 peut se lier au promoteur et initier la transcription génique conduisant à la production de la protéine d'intérêt.

Surexpression de NLS-HD Otx2 doublement marqué ¹⁵N/¹³C pour les études RMN.

Le protocole de la production de protéine marquée diffère de celui que nous venons de décrire : il faut changer de milieu de culture avant l'induction des bactéries par l'IPTG pour permettre l'incorporation de ¹⁵N et ¹³C dans les acides aminés de la protéine.

Par centrifugation on élimine le milieu d'incubation LB dit riche pour le remplacer par un milieu d'incubation dit minimum suivant le protocole de Marley[214]. Pour ce faire, les cultures des 8 erlens, correspondant à 4 litres de milieu de culture, sont centrifugées à 3000 g pendant 15 min à 20 °C pour retirer le milieu riche. Les culots sont ensuite repris dans 1 L de milieu minimum, en présence de kanamycine (50 µg/mL). Les bactéries sont alors réparties dans 3 erlens de 2 litres puis incubées à 37 °C pendant 40 min sous agitation (180 rpm) pour leur permettre de consommer le LB résiduel (sources d'énergie non marquées). Les sources de ¹³C et ¹⁵N sont ajoutées au milieu sous forme de ¹³C-glucose (180 mg/L) et ¹⁵NH₄Cl (9 mg/L). Les bactéries sont préincubées pendant 40 min à 37 °C pour permettre d'incorporer les sources d'énergie isotopiques puis la température est baissée à 27 °C avant addition d'IPTG (500 µM final) afin d'induire la production d'Otx2. Les bactéries sont alors incubées 5 h à 27 °C, sous agitation à 180 rpm.

Extraction/purification de NLS-HD Otx2 (marqué ou non marqué ¹⁵N/¹³C)

Après cette phase de surexpression, les milieux de culture contenant les bactéries sont centrifugés à 3000 g pendant 20 min à 4 °C pour retirer le milieu de culture. Les culots cellulaires, repris dans 80 mL de tampon de lyse et mis sur la glace pour éviter l'action des protéases, sont lysés par l'utilisation du sonificateur à ultrasons, Vibra-Cell™ 75115, de la société Bioblock Scientific. Le protocole que nous avons adopté est le suivant : 10 cycles de 15 secondes de sonication (à 300 W) suivis par 60 secondes de refroidissement dans l'eau glacée puis 10 cycles de 20 secondes de sonication / 60 secondes dans la glace. À ce stade des inhibiteurs de protéases sont ajoutés : du PMSF (fluorure de phénylméthylsulfonyle) (à 1 mM final), de la pepstatine (à 1 µg/mL final) et de la leupeptine (à 1 µg/mL final). Puis, ces lysats sont centrifugés à 30000 g pendant 40 min à 4 °C afin de récupérer le surnageant contenant notamment la protéine de fusion à purifier.

Le surnageant est filtré sur filtre seringue 0,45 µm avant purification par FPLC (Fast Performance Liquid Chromatography). La protéine de fusion est purifiée en chambre froide (à 6 °C environ) en utilisant un appareil « ÄKTA™ start » équipé d'une colonne HisTrap HP / GE Healthcare / 5 mL, de 16 mm × 25 mm, avec des particules de 34 µm en moyenne. Cette étape correspond à une purification par chromatographie d'affinité, la protéine de fusion étant fixée sur la colonne constituée de billes d'agarose fonctionnalisées par du Nickel (Ni²⁺) *via* les fonctions imidazoles des six His introduites dans l'étiquette en *N*-terminal.

La colonne et l'ensemble du système sont équilibrés dans le tampon de liaison grâce à une pompe principale fonctionnant à un débit de 1 mL/min. Puis, le lysat bactérien est injecté *via* « une pompe échantillon » à un débit de 1 mL/min tandis que le débit de la pompe principale est abaissé à 0,2 mL/min. Une fois l'ensemble de l'échantillon injecté, la pompe échantillon est stoppée et la pompe principale remise à un débit de 1 mL/min de tampon de liaison pour permettre à la protéine de fusion d'interagir avec la phase de la colonne (étape de lavage de 45 min environ). Tout ce qui n'est pas retenu par la colonne de Ni²⁺ est directement élué, la présence de 500 mM de NaCl et de 25 mM d'imidazole dans le tampon de liaison permet d'éviter les liaisons non spécifiques sur la colonne. Le suivi de ce qui sort du système, et qui est systématiquement collecté pour analyse, se fait par conductimétrie et lecture de l'absorption UV à 280 nm. Puis, on applique un gradient progressif d'imidazole dans le système en augmentant la proportion de tampon d'éluion tout en diminuant celle du

tampon de liaison. Lorsque la concentration d'imidazole est suffisante dans la colonne (ici autour de 40 % de tampon d'élution/60 % de tampon de liaison), la protéine de fusion se décroche et est alors collectée en sortie du système. Le système est ensuite lavé par 100 % de tampon d'élution afin de décrocher de la colonne ce qui pourrait rester.

L'enzyme « PreScission » est ajoutée à la protéine de fusion récoltée en sortie de FPLC. L'ensemble (30 mL environ) est séparé dans deux boudins de dialyse ayant une membrane seuil de 3500 Da. La dialyse permet de diminuer la concentration en imidazole, pour permettre la prochaine étape de purification sur colonne de Ni²⁺, et en NaCl, pour une meilleure action de la protéase PreScission. Les boudins sont agités une nuit en chambre froide (à 6 °C environ) dans un récipient contenant le tampon de dialyse.

Le contenu des boudins de dialyse est filtré sur filtre seringue 0,45 µm puis dilué par deux dans le tampon de liaison afin de faire remonter la concentration en NaCl du mélange et ainsi éviter les interactions non spécifiques sur la colonne. Le mélange est ensuite injecté sur le système « ÄKTA™ start » équipé de la colonne Ni²⁺ équilibrée dans le tampon de liaison, comme expliqué précédemment. Seules la protéine Cherry, libérée de NLS-HD Otx2 et toujours munie du tag His, ainsi que l'enzyme « PreScission », elle-même munie d'un tag His, seront retenues par la colonne. NLS-HD Otx2, libre du tag His, sera directement élué dans le tampon de liaison et récolté en sortie d'appareil. Le système est ensuite nettoyé par 100 % de tampon d'élution, pour décrocher Cherry et PreScission.

Après chaque étape de purification, la présence de la protéine d'intérêt dans les fractions collectées est confirmée par le dépôt d'échantillons sur un gel d'électrophorèse SDS-Page. L'électrophorèse permet de séparer les éléments d'un échantillon en fonction de leur masse moléculaire et une révélation au bleu de Coomassie permet de les visualiser.

Les fractions contenant le domaine NLS-HD d'Otx2 (environ 100 mL de solution) sont concentrées, par centrifugations successives de 15 mL durant 1 h environ, sur une unité d'ultrafiltration à centrifuger Amicon de 3000 Da (nommé par la suite « Amicon 3000 »), à 4000 rpm à 4 °C. Après plusieurs rinçages du filtre de l'unité de filtration par le tampon de filtration sur gel (FG), la concentration de la protéine, dosée par mesure de l'absorbance à 280 nm selon la loi de Beer-Lambert, est de l'ordre de 180 µM pour un volume de 14 mL environ. Cette solution, trouble, est centrifugée à 30000 g durant 15 min à 4 °C.

Le surnageant est ensuite injecté (2 mL par injection) sur le système « ÄKTA™ start », mais cette fois équipé d'une colonne Hiloal 16/60 - Superdex 75 – pegrade pour une séparation par chromatographie d'exclusion stérique ou filtration sur gel, à un débit de 1 mL/min de tampon de FG. La séparation des macromolécules ne se fait alors plus en fonction de leur affinité vis-à-vis de la phase contenue dans la colonne mais en fonction de leur volume hydrodynamique. La détection de la protéine se fait par lecture de l'absorption UV à 280 nm. Le domaine NLS-HD d'Otx2 est récupéré en sortie du système environ 80 min après chaque injection. Ainsi 50 mL du domaine NLS-HD d'Otx2 pur en solution dans le tampon de FG sont récupérés et stockés à -80 °C en attendant d'être concentrés sur une unité « Amicon 3000 » et repris dans le tampon et à la concentration adéquats aux expériences à mener.

b) Etudes RMN

Généralités

Les expériences de RMN ont été réalisées, à 25 °C (298 K), sur un spectromètre Bruker Avance III 500 MHz équipé d'une cryosonde triple résonance TCI 5-mm. L'appareil a été paramétré et les données traitées à l'aide du logiciel Topspin (Bruker). Puis les spectres RMN obtenus ont été analysés à l'aide du logiciel SPARKY (T. D. Goddard and D. G. Kneller, SPARKY 3, University of California, San Francisco).

Le tampon phosphate 2X utilisé est composé de 200 mM de NaCl, 4 mM de DTT et de 80 mM de Na₂HPO₄, mis en solution dans de l'eau ultrapure milliQ dont le pH est ensuite ajusté à 6,3. Le tout est filtré sur filtre seringue de 0,2 µm. Le tampon phosphate 1X est obtenu en diluant par deux, avec de l'eau milliQ, le tampon 2X.

Attribution par RMN des raies de résonance de NLS-HD Otx2 doublement marqué ¹⁵N/¹³C

Concentration de la protéine : NLS-HD Otx2 doublement marqué ¹⁵N/¹³C (environ 50 mL dans le tampon de FG, stockés à -80 °C), est décongelé lentement dans la glace puis est concentré, par centrifugations successives de 15 mL durant 1 h environ, sur une unité « Amicon 3000 », à 4000 rpm à 4 °C. La concentration de la solution de protéine ainsi obtenue (environ 5 mL), dosée par mesure de l'absorbance à 280 nm, est de l'ordre de 182 µM. Une nouvelle concentration de 15 min à 4000 rpm et à 4 °C sur l'unité « Amicon 3000 », suivie d'un

lavage du filtre par 90 μL de tampon de FG, permet d'obtenir environ 1,5 mL de solution de protéine Otx2 doublement marquée à la concentration désirée de 542 μM . Le tube type « eppendorf » contenant la protéine est conservé dans la glace en attendant la préparation de l'échantillon à étudier en RMN.

Préparation de l'échantillon : Dans un tube type « eppendorf » sont ajoutés 481 μL de NLS-HD Otx2 doublement marqué $^{15}\text{N}/^{13}\text{C}$ (à 500 μM final), 39 μL de D_2O (soit 7,5 % en volume final), 2 μL de DSS (à 470 μM final), 0,5 μL de pepstatine (à 1 $\mu\text{g}/\text{mL}$ final) et 0,5 μL de leupeptine (à 1 $\mu\text{g}/\text{mL}$ final). Après agitation la solution est transférée dans un tube RMN de 5 mm de diamètre.

Rq : Le DSS ou acide 4,4-diméthyl-4-silapentane-1-sulfonique est utilisé pour l'étalonnage interne des déplacements chimiques.

Titration par RMN de NLS-HD Otx2 et NLS-HD En2 marqués en ^{15}N par divers ligands : sulfopeptides, HI dp8 et PolyGlu

Rq : Du fait du très bon rendement de production de NLS-HD Otx2 doublement marqué ^{15}N et ^{13}C , nous avons pu utiliser cette protéine pour réaliser les expériences de titrage alors qu'une protéine simplement marquée ^{15}N aurait été suffisante.

Dosage des divers éléments : Avant chaque expérience, les protéines marquées (stockée à $-80\text{ }^\circ\text{C}$) ont été dosées par mesure de l'absorbance à 280 nm. L'héparine HI dp8 (Iduron, ref HO08, UA,2S-GlcNS,6S-(IdoA,2S-GlcNS,6S)₃, MW moyenne : 2400 Da) a été mise en solution dans l'eau distillée à la concentration de 6 mM en se basant sur la quantité totale de produit dans le flacon, indiquée par le fournisseur. Le PolyGlu (Iris Biotech, ref PGA 1005, n-Butyl-poly(L-glutamic acid)₂₀ sodium salt ou nBu-PGA(20), MW moyenne : 3492 Da) et les sulfopeptides ont été dosés par mesure de l'absorbance à 205 nm en se basant sur leur coefficient d'extinction molaire théorique moyen (se reporter à la partie expérimentale du chapitre II-C pour davantage de précision).

Préparation des échantillons pour les expériences RMN : Dans un tube type « eppendorf » sont ajoutés : NLS-HD Otx2 marqué ^{15}N pour une concentration finale de 150 μM , 39 μL de D_2O (soit 7,5 % en volume final) et du tampon phosphate 1X pour un volume total de 520 μL . Après agitation la solution est transférée dans un tube RMN de 5 mm de

diamètre. Les sulfopeptides (dosés par mesure de l'absorbance à 205 nm: Ac-Ahx-(EC'AC')₆-NH₂ et Ac-Ahx-(EC'C'A)₆-NH₂ à 4 mM dans l'eau et Ac-Ahx-(EC'AC')₃-NH₂ à 6 mM dans l'eau), polyGlu (dosé par mesure de l'absorbance à 205 nm à 6 mM dans l'eau) et HI dp8 (à 6 mM dans l'eau) sont dilués par deux dans du tampon phosphate 2X avant d'être utilisés pour les expériences de titrage.

Titration : Le tube est introduit dans le spectromètre RMN avant de lancer une expérience 2D HSQC ¹H ¹⁵N qui servira de référence. Une comparaison avec le spectre HSQC obtenu lors de l'attribution permet de s'assurer que la protéine ne s'est pas dégradée. Le tube est alors retiré du spectromètre, son contenu est transféré dans un tube type « eppendorf » afin de pouvoir y ajouter une quantité définie et relativement faible de ligand (de l'ordre de 0,12 à 0,4 équivalent par ajout par rapport à la quantité de protéine). Après mélange, l'ensemble est remis dans le tube RMN qui est ensuite placé dans le spectromètre pour réaliser une nouvelle expérience HSQC. Ce protocole est répété jusqu'à ce que l'on n'observe plus de modification dans les déplacements chimiques des protons amides et/ou des azotes marqués qui les portent (ajout maximal de ligand).

Calcul des K_d apparents : Il est à noter que les valeurs de K_d indiquées dans les **Tableaux 13** et **14** ci-dessous sont issues d'un seul titrage par ligand testé, les erreurs standards indiquées proviennent uniquement de la méthode de calcul des K_d apparents. Le script utilisé pour ce calcul peut parfois proposer un K_d avec une barre d'erreur importante (voire supérieure à la valeur du K_d indiquée) lorsque la courbe de liaison est mal ajustée aux points expérimentaux et qu'elle ne passe pas par suffisamment de points. De plus, ce script est incapable de déterminer un K_d apparent (nd) lorsque la valeur de CSPmax correspondante (CSP obtenu pour 100 % de protéine liée) est trop faible (inférieure à 0,02 ppm environ) ou lorsque la courbe obtenue ne permet pas de déterminer un CSPmax. Malgré ces imprécisions, les K_d apparents obtenus permettent de comparer globalement les zones d'interaction entre elles.

K_d apparentes (en μM) des résidus les plus affectés du domaine de liaison aux GAGs

*	39 R	40 R	41 E	42 R	43 T	44 T	45 F
(EC'AC') ₆	10 ± 9	9 ± 9	7 ± 7	7 ± 6	8 ± 7	12 ± 12	6 ± 6
(EC'C'A) ₆	13 ± 9	9 ± 7	12 ± 6	10 ± 3	13 ± 5	11 ± 7	8 ± 6
(EC'AC') ₃	nd	332 ± 37	334 ± 30	382 ± 47	127 ± 16	100 ± 18	133 ± 19
HI dp8	21 ± 1	18 ± 4	20 ± 2	18 ± 3	19 ± 2	21 ± 1	12 ± 3
PolyGlu	59 ± 11	13 ± 5	nd	nd	33 ± 4	29 ± 3	39 ± 7

K_d apparentes (en μM) des résidus les plus affectés de l'Hélice 3

*	82 V	83 Q	84 V	86 F	88 N	89 R	91 A	92 K	94 R
(EC'AC') ₆	48 ± 29	53 ± 19	36 ± 20	12 ± 11	-	10 ± 10	12 ± 12	10 ± 10	5 ± 5
(EC'C'A) ₆	60 ± 24	123 ± 32	57 ± 15	14 ± 8	21 ± 8	12 ± 7	13 ± 6	9 ± 4	8 ± 5
(EC'AC') ₃	nd	nd	nd	nd	nd	nd	nd	nd	nd
HI dp8	17 ± 5	15 ± 6	14 ± 3	28 ± 7	32 ± 4	21 ± 5	30 ± 3	17 ± 3	24 ± 8
PolyGlu	nd	nd	nd	38 ± 15	58 ± 8	28 ± 4	57 ± 6	56 ± 7	14 ± 5

K_d apparentes (en μM) des résidus His₂₆ et Tyr₆₂

*	26 H	62 Y
(EC'AC') ₆	228 ± 13	26 ± 37
(EC'C'A) ₆	310 ± 21	13 ± 13
(EC'AC') ₃	nd	nd
HI dp8	nd	16 ± 12
PolyGlu	nd	nd

nd : Non déterminé

* : Tous les sulfopeptides ont pour formule générale Ac-Ahx-(XXXX)_n-NH₂, avec Ahx = acide 6-amino-hexanoïque.

NLS-HD Otx2 : ASM²²DLL²³H²⁴PSVG³⁰YPATPRKQRR⁴⁰ERTTFTRAQL⁵⁰DVLEALFAKT⁶⁰RY⁶¹PDIFMREE⁷⁰
VALKINLPES⁸⁰RVQVWFKNRR⁹⁰AKCRQQQ⁹⁷

Tableau 13. Valeurs des K_d apparentes des résidus les plus affectés par l'interaction entre NLS-HD Otx2 et les ligands Ac-Ahx-(EC'AC')₆-NH₂, Ac-Ahx-(EC'C'A)₆-NH₂, Ac-Ahx-(EC'AC')₃-NH₂, HI dp8 et PolyGlu.

K_d apparents (en μM) des résidus les plus affectés de la partie C-ter du domaine de liaison putatif aux GAGs

	194	197	198	200	201
	K	N	K	D	K
$(\text{EC}'\text{AC}')_6^*$	2 ± 1	1 ± 1	1 ± 1	1 ± 1	2 ± 1
HI dp8**	-	-	5 ± 3	-	11 ± 2

K_d apparents (en μM) des résidus les plus affectés de la région C-ter de l'Hélice 3

	252	253	256	257	258
	R	A	K	K	A
$(\text{EC}'\text{AC}')_6^*$	10 ± 9	5 ± 2	3 ± 1	3 ± 1	2 ± 1
HI dp8**	-	-	-	99 ± 7	174 ± 18

K_d apparents (en μM) des résidus les plus affectés des régions « 204-208 » et « 222-226 »

	206	207		223	224	226
	A	F		R	Y	T
$(\text{EC}'\text{AC}')_6^*$	$1,5 \pm 0,5$	4 ± 1		$4,5 \pm 1$	$6,5 \pm 3$	$7,5 \pm 3,5$
HI dp8**	-	-		-	-	-

* : Tous les sulfopeptides ont pour formule générale $\text{Ac-Ahx}-(\text{XXXX})_n-\text{NH}_2$, avec Ahx = acide 6-amino-hexanoïque.

** valeurs de K_d extraites de la thèse de S. Cardon[186].

NLS-HD En2 : ASS¹⁸⁴GPRSRK¹⁹⁰PKKKNPNKED²⁰⁰KRPR²⁰⁴TAF²⁰⁷TAE²¹⁰QLQRLKAEFQ²²⁰TNRY²²⁴LTEQ²²⁶RR²³⁰QSLAQELGLN²⁴⁰ESQIKIWFQN²⁵⁰KRAKIKKAT²⁵⁹

Tableau 14. Valeurs des K_d apparents des résidus les plus affectés par l'interaction entre NLS-HD En2 et les ligands $\text{Ac-Ahx}-(\text{EC}'\text{AC}')_6-\text{NH}_2$ et HI dp8.

2. Synthèse de nouveaux analogues et Biacore

a) Synthèse de $\text{Ac-Ahx}-(\text{EC}'\text{PC}')_6-\text{NH}_2$ et $\text{Ac-Ahx}-(\text{EC}'\text{LC}')_6-\text{NH}_2$

Le protocole général de synthèse, l'étape d'oxydation, les étapes de purification et les conditions de stockage des peptides ayant été décrits dans le chapitre II-C3 et C4 et dans le « Supplementary Information » du manuscrit pour publication inséré chapitre III, je ne détaillerai, dans cette partie expérimentale, que les éléments présentant un caractère nouveau.

Les peptides bruts ont été purifiés par HPLC préparative sur une colonne C18 Ace (AIT France), 5 μm , 300 Å, 10x250 mm, à un débit de 6 mL/min :

- pour Ac-Ahx-(EC'PC')₆-NH₂ de façon isocratique pendant 5 min à 0 % de CH₃CN (0,1 % TFA) puis avec un gradient en 30 min de 0 à 10 % de CH₃CN (0,1% TFA) et une lecture à 220 nm,
- pour Ac-Ahx-(EC'LC')₆-NH₂ avec un gradient en 30 min de 5 à 20 % de CH₃CN (0,1 % TFA) et une lecture à 220 nm.

Les caractéristiques HPLC et spectrométrie de masse des peptides sont reportées ci-dessous, sous forme de tableau, **Tableau 15**.

Code personnel	Gradient analytique	R _T (min)	m/z calculé [M-H] ⁻	m/z mesuré [M-H] ⁻
RM4-180 : Ac-Ahx-(EC'PC') ₆ -NH ₂	Gradient de 2 à 15% B en 10 min	4,7	3339	3341
RM4-181 : Ac-Ahx-(EC'LC') ₆ -NH ₂	Gradient de 5 à 20% B en 10 min	7,9	3435	3440

Tableau 15. Caractéristiques HPLC et spectrométrie de masse de Ac-Ahx-(EC'PC')₆-NH₂ et Ac-Ahx-(EC'LC')₆-NH₂ (Eluants HPLC : A = H₂O (0,1% TFA) et B = CH₃CN (0,1% TFA))

b) Biacore

Les expériences de Biacore ont été réalisées sur l'appareil Biacore 3000 (GE Healthcare) de la plateforme « Interactions moléculaires » de l'IBPS. Les consommables et solutions utilisées sont :

- Puce de type CM5 de GE Healthcare, Réf. 29-1496-04
- Tampons d'immobilisation, 10 mM sodium acétate pH 4.5 (GE Healthcare)
- Tampon de course, HBS-EP pH 7,4 (GE Healthcare)
- Tampon d'interaction (binding buffer), HBS-EP pH 7,4 (GE Healthcare)

Immobilisation de la streptavidine sur la puce : Les acides carboxyliques de la matrice de dextran carboxyméthylé de la puce sont activés en faisant circuler, pendant 7 minutes, 70 µL d'un mélange de 0,4 M de 1-ethyl-3-(3-diméthylaminopropyl)-carbodiimide (EDC) et 0,1 M de *N*-hydroxysuccinimide (NHS) (débit de 10 µL/min). La Streptavidine (Streptavidin from *Streptomyces avidinii*, Sigma Aldrich ref. 85878-1MG) est alors injectée, pour immobilisation, à la concentration de 50 µg/mL dans le tampon d'immobilisation, à 10 µL/min

durant 7 minutes. Enfin une solution d'éthanolamine à 1 M est injectée afin de désactiver les fonctions réactives restantes (acides carboxyliques activés).

Immobilisation des mimes biotinylés : Après préconditionnements par injection d'un mélange de NaCl 1 M et NaOH 50mM, 10 µL de mime biotinylé (500 nM dans le tampon d'interaction) sont injectés en 1 minute puis la puce est rincée par un flux continu de tampon.

Circulation de la protéine NLS-HD Otx2 : En analyse directe, la protéine est injectée à différentes concentrations, dans le tampon d'interaction, à 5 µL/min durant 2 minutes puis du tampon de course est utilisé, durant 15 minutes, pour dissocier NLS-HD Otx2 de la puce portant le mime biotinylé. En analyse par compétition, la protéine (à 10 nM dans le tampon d'interaction) est injectée durant 5 minutes, seule ou après avoir été incubée au préalable 5 minutes à température ambiante avec un inhibiteur à différentes concentrations. Le complexe est dissocié par le passage d'un flux continue de tampon durant 15 minutes.

CHAPITRE V : Travaux et résultats préliminaires

A. Introduction

Les travaux en cours et les résultats préliminaires que je vais présenter dans ce chapitre final auraient certainement eu droit à un meilleur sort, des chapitres spécifiques, si un virus n'était pas venu perturber nos vies !

Sans attendre les résultats des études d'interaction *in vitro* décrites dans le chapitre IV, nous avons commencé à explorer d'autres aspects. Je savais par expérience que les études par RMN, qui plus est impliquant la surexpression et la production d'une protéine, seraient longues, mais c'était la stratégie qui s'avérait la plus rationnelle pour avancer vers de nouvelles générations de sulfopeptides. Cependant, entre mon apprentissage d'un domaine loin de mon champ de compétences et l'encadrement du Dr Ludovic Carlier, Maître de conférences qui a vu sa vie d'enseignant, et donc de chercheur, grandement modifiée, ce fut encore plus long que prévu initialement.

Les pistes que j'ai explorées en parallèle correspondent à :

- La synthèse d'un analogue avec agrafe, sur la base du térapeptide EC'AC',
- la synthèse de polymères non linéaires de la séquence térapeptide EC'AC',
- la synthèse de sondes fluorescentes pour des études biophysiques d'interaction par polarisation de fluorescence.

Je présenterai ici les résultats préliminaires de ces pistes explorées qui constituent sans nul doute le début de nouveaux travaux.

B. Travaux en cours et résultats préliminaires

1. Synthèse d'un analogue linéaire avec agrafe.

Comme déjà évoqué dans le chapitre IV-C.1, avec la synthèse des sulfopeptides Ac-Ahx-(EC'PC')₆-NH₂ et Ac-Ahx-(EC'LC')₆-NH₂, il nous a semblé intéressant d'explorer de nouvelles stratégies de synthèse visant à obtenir des sulfopeptides plus structurés en solution. Une stratégie en particulier, la chimie des peptides agrafes (« stapled peptides »), a retenu notre attention. Elle vise à stabiliser des structures hélicoïdales par réduction de l'entropie conformationnelle du peptide, en le contraignant par une agrafe formée par la liaison covalente entre deux chaînes latérales d'acides aminés, formant ainsi un macrocycle

peptidique. Nous ne citerons que deux articles de référence dans le domaine des peptides agrafes[215], [216].

Pour une étude préliminaire de faisabilité, j'ai travaillé à partir de la séquence du mime court Ac-Ahx-EC'**AC'**-EC'**AC'**-EC'**AC'**-NH₂. Afin de pouvoir contraindre sa structure, j'ai remplacé les Alanines (en gras dans la séquence, **A** en position i et i+8) par des Lysines pour pouvoir ensuite « relier » leurs chaînes latérales par couplage des fonctions ε-amines avec du chlorure de biphényl-4,4'-disulfonyl. Le chlorure de biphényl-4,4'-disulfonyl présente l'avantage de résister aux conditions d'oxydation et sa taille doit permettre de rapprocher, de manière adéquate, les deux lysines et forcer le peptide à adopter une structure hélicoïdale. La formule du peptide attendu est reportée dans la **Figure 114**.

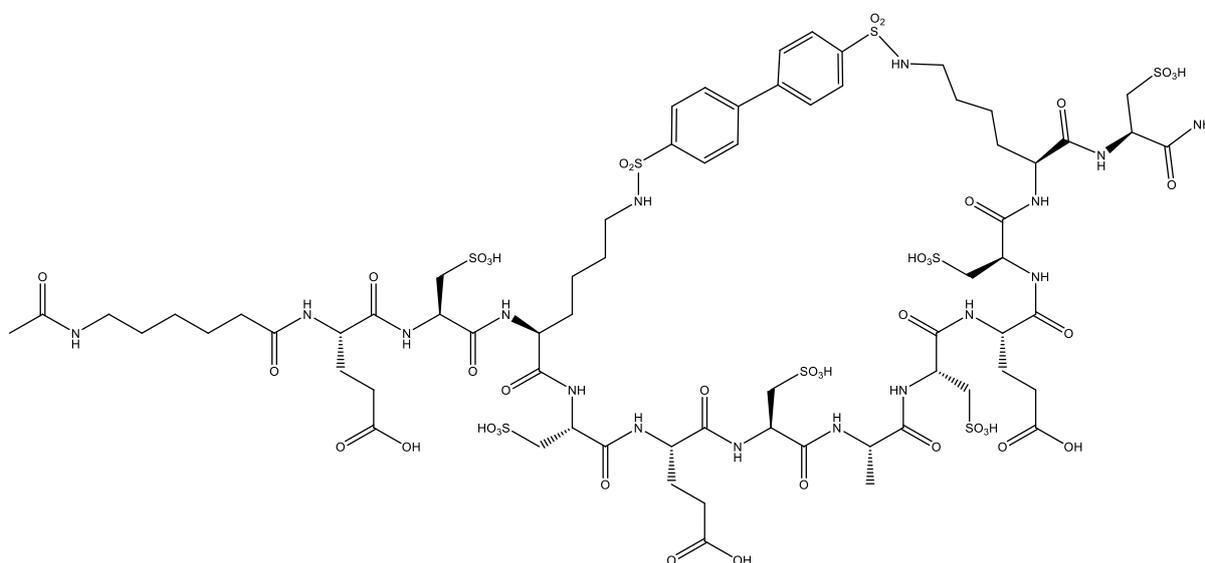


Figure 114. Formule du peptide « agrafe » Ac-Ahx-EC'KC'**-EC'**AC'**-EC'**KC'**-NH₂**

La stratégie de synthèse employée repose sur l'utilisation de Fmoc-Lys(Mmt)-OH, les chaînes latérales des Lysines portant le groupement protecteur fortement acido-labile Mmt (méthoxytrityl). A la fin de la synthèse du peptide protégé, réalisée sur support solide en stratégie Fmoc (pour plus de détails sur la synthèse, se référer au chapitre II-C.4), les groupements Mmt ont été retirés par action d'un mélange AcOH/TFE/DCM (1/2/7 en volume). Ces conditions n'altèrent ni le lien du peptide avec le support solide ni les autres groupes protecteurs utilisés. Ainsi il devient possible de coupler, sur les deux fonction ε-amines libérées, le chlorure de biphényl-4,4'-disulfonyl.

La coupure du support et l'oxydation du mélange réactionnel ont été réalisées selon le mode opératoire décrit pour la synthèse des analogues linéaires sans agrafe (*cf.* chapitre V-B.1.a). Le mélange final a ensuite été analysé par HPLC en phase inverse, **Figure 115**, et chaque pic observé a été caractérisé par spectrométrie de masse MALDI-TOF, avec une matrice *p*-nitroaniline. Le spectre MALDI du produit attendu (indiqué par une flèche sur la **Figure 115**) est reporté **Figure 116**.

Rg : Au fil des analyses par MALDI-TOF MS de ces peptides nous avons constaté que l'utilisation de pmg n'était pas nécessaire et que la matrice *p*-nitroaniline permettait à elle seule d'analyser les sulfopeptides par MALDI-TOF MS.

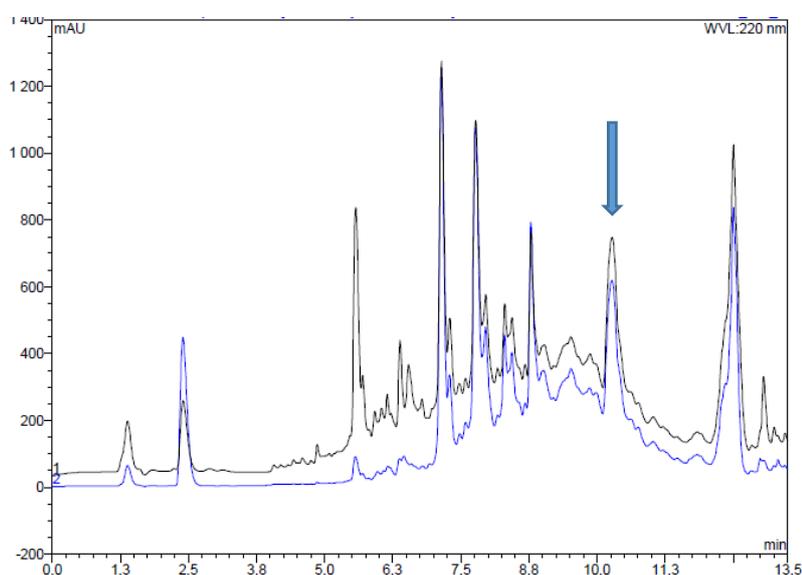


Figure 115. Profils HPLC en phase inverse du peptide « agrafe » Ac-Ahx-EC'KC'-EC'AC'-EC'KC'-NH₂ brut. La flèche bleue indique le pic correspondant au produit attendu.
Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 5 μm, 200 Å, 4,6x100 mm - Elution avec un gradient en 10 min de 1 à 20% de CH₃CN (0,1% TFA) - Débit : 1 mL/min - Lecture à 220 (en noir) et 245 nm (en bleu).

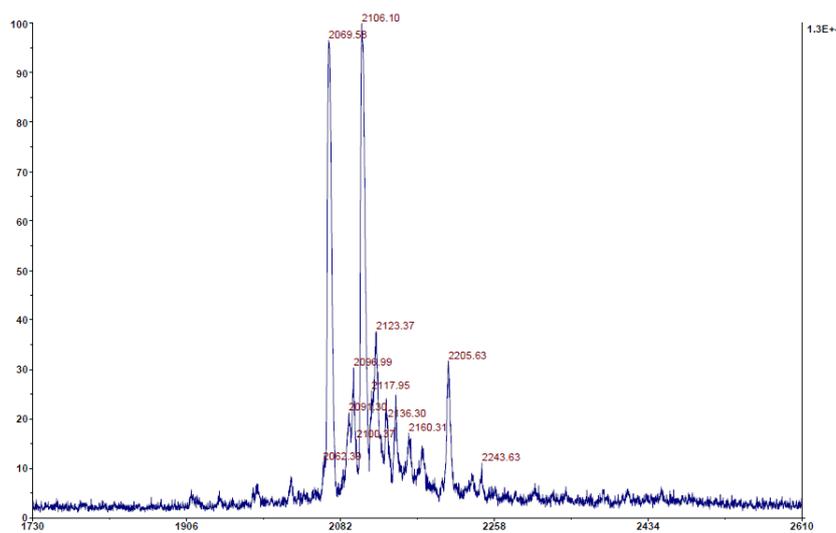


Figure 116. Profil MALDI-TOF du peptide « agrafe » Ac-Ahx-EC'KC'-EC'AC'-EC'KC'-NH₂ (pic ciblé par une flèche bleue sur le spectre HPLC précédent, collecté en sortie d'HPLC analytique), avec matrice *p*-nitroaniline, en mode linéaire négatif.

Le produit attendu, peptide « agrafe », a bien été identifié dans ce brut réactionnel, cependant le rendement de synthèse est très faible avec beaucoup de sous-produits de synthèse, difficiles à identifier, malgré plusieurs tentatives d'optimisation. J'ai cependant purifié ce pic par HPLC préparative mais n'ai pu en récupérer qu'une quantité « infinitésimale ».

2. Synthèse de polymères non linéaires

Parallèlement à ces travaux de synthèse d'analogue linéaire, nous nous sommes intéressés à d'autres moyens d'améliorer l'affinité de nos mimes de GAGs sans devoir allonger à l'excès leur séquence : synthétiser des polymères de type dendrimérique à partir d'un noyau poly-Lysine et/ou de type branché en liant chimiquement des sulfopeptides sur un support peptidique linéaire.

a) Polymères de type dendrimère

Les dendrimères peptidiques sont des polymères composés d'un noyau central et de groupements fonctionnels de surface. Parmi les différentes formes de dendrimères peptidiques, le MAP (Multiple Antigenic Peptide) est le plus courant [217], [218]. Son noyau central ainsi que ses groupements fonctionnels sont de type peptidique. Si le noyau d'un MAP

peut être composé de différents acides aminés, l'acide aminé le plus couramment utilisé est la Lysine, en raison de la présence de deux fonctions amine en α et ϵ , **Figure 117**.

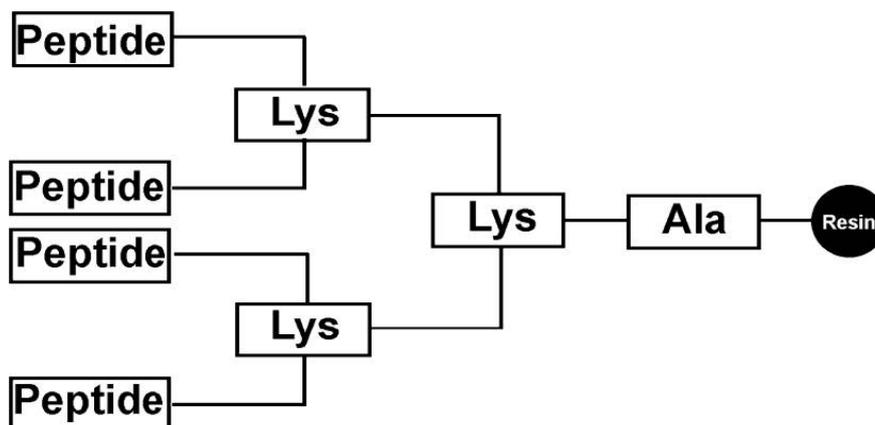


Figure 117. Conception classique d'un MAP (image extraite de[219])

Afin d'améliorer la présentation de nos meilleurs mimes de GAGs, j'ai donc étudié la possibilité de synthétiser, sur support solide, une version dendrimérique, de type MAP, du motif « EC'AC' », répété 4 fois, voir **Figure 118**.

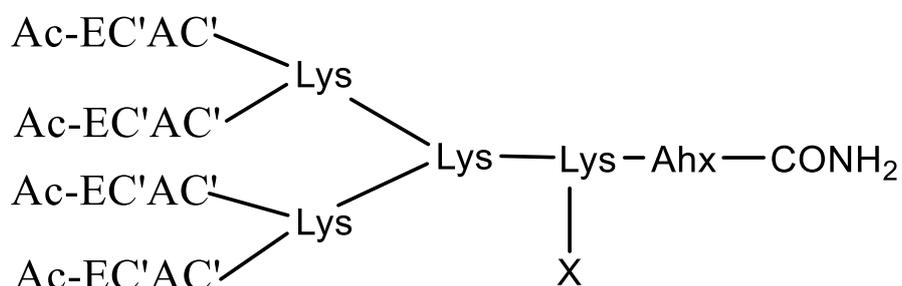


Figure 118. Schéma d'une version dendrimérique à 4 répétitions, de type MAP, d'un mime de GAGs. Le groupement X peut être un acétyle, une biotine, un fluorophore...

Pour se faire, la synthèse a été réalisée, manuellement, par la stratégie Fmoc déjà décrite, avec une résine Fmoc-Rink-Amid MBHA (Fmoc-Rink-Amid-4-méthylbenzhydramine, substituée à 0,62 mmol/g) en utilisant, dans l'ordre de la synthèse, Fmoc-Ahx-OH, Fmoc-Lys(Mmt)-OH, Fmoc-Lys(Fmoc)-OH, Fmoc-Lys(Fmoc)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Ala-OH, Fmoc-Cys(Trt)-OH et Fmoc-Glu(OtBu)-OH. Les couplages ont été réalisés avec le couple d'agents activateurs HBTU/DIEA et la déprotection du Fmoc par action de pipéridine. Pour les derniers couplages, le nombre d'équivalents d'acides aminés et d'agents de couplage employés a été augmenté pour compenser l'augmentation du nombre d'amines libres à

coupler. Les extrémités *N*-terminales ont été acétylées puis le groupement protecteur Mmt a été retiré sélectivement par action d'un mélange AcOH/TFE/DCM (1/2/7 en volume) pour permettre le couplage, sur support solide, d'un groupement X sur la chaîne latérale de la lysine. Pour la preuve de concept, nous avons choisi ici X = Ac mais, à terme, l'idée est de pouvoir coupler des groupements fonctionnels comme une biotine(O₂) ou un fluorophore.

La coupure du support a été réalisée selon le mode opératoire déjà décrit (*cf.* chapitre II-C.3.a). Après lyophilisation, il est apparu que le peptide dendrimérique brut non oxydé montrait une meilleure solubilité dans les solvants utilisés en HPLC en phase inverse que ses analogues linéaires non oxydés puisqu'une partie de ce brut a pu être solubilisée dans un mélange à 50 % H₂O (0,1 % TFA)/50% CH₃CN (0,1 % TFA) avant d'être analysée par HPLC analytique, **Figure 119**.

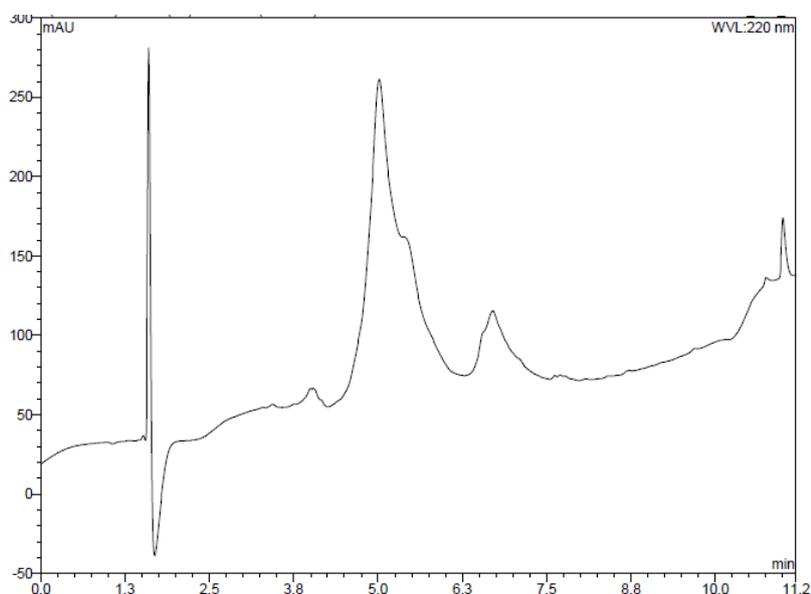


Figure 119. Profil HPLC en phase inverse du dendrimère brut non oxydé solubilisé dans un mélange à 50 % H₂O (0,1 % TFA)/50% CH₃CN (0,1 % TFA). Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 5 μm, 200 Å, 4,6x100 mm – Gradient en 10 min de 5 à 100 % de CH₃CN (0,1 % TFA) - Débit : 1 mL/min - lecture à 220 nm.

L'analyse par MALDI-TOF MS, avec une matrice HCCA, en mode linéaire positif, du pic principal observé sur le spectre HPLC a permis de confirmer qu'il s'agissait bien du produit attendu (MW attendue : 2478 g/mol⁻¹). Sur le spectre MS, nous observons principalement deux signaux, à [M+Na]⁺ = 2500 g/mol⁻¹ et [M+K]⁺ = 2518 g/mol⁻¹, correspondant au peptide

attendu porteur respectivement d'un contre-ion sodium (+ 23 g/mol⁻¹) et potassium (+ 39 g/mol⁻¹), voir **Figure 120**.

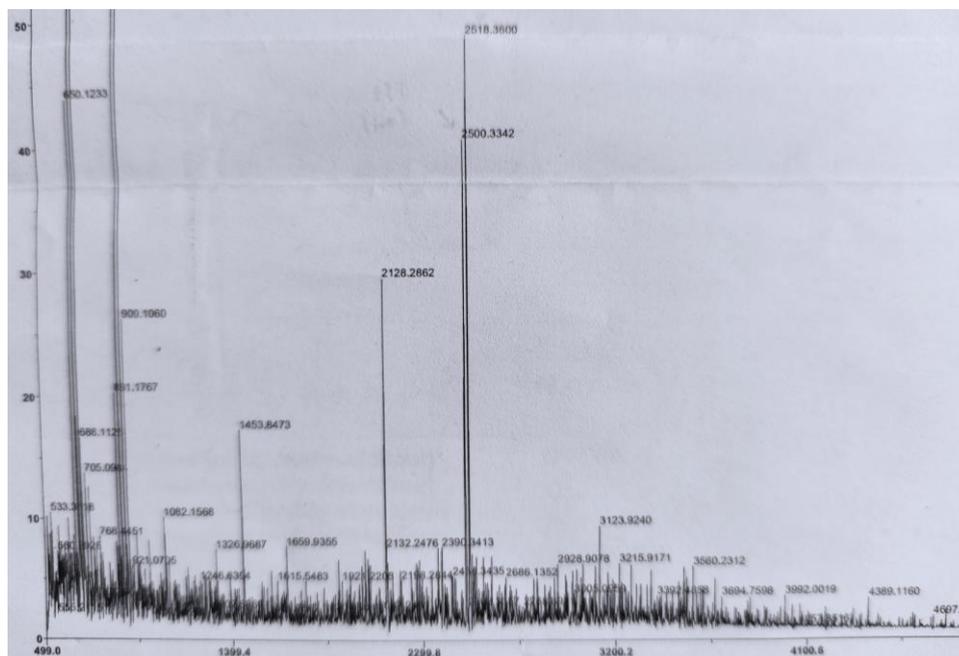


Figure 120. Profil MALDI-TOF du pic principal du brut de synthèse du dendrimère non oxydé, avec matrice HCCA, en mode linéaire positif.

Une partie du peptide dendrimérique brut non oxydé (environ 30 mg) a été purifiée par HPLC phase inverse, voir le profil HPLC après purification **Figure 121**, et les fractions pures ont été lyophilisées.

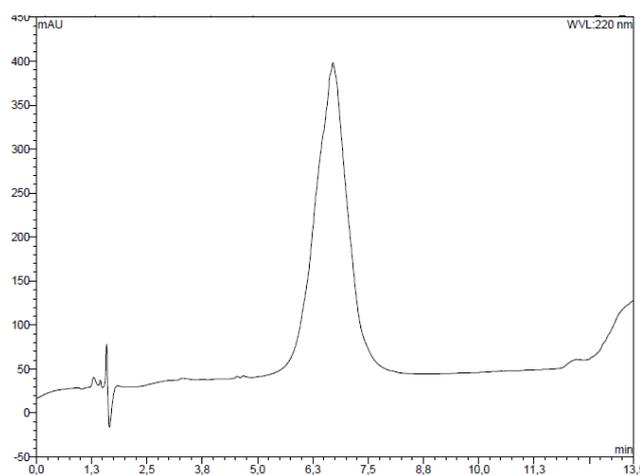


Figure 121. Profil HPLC en phase inverse du dendrimère non oxydé purifié. Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 5 μ m, 200 \AA , 4,6x100 mm – Gradient en 10 min de 10 à 50 % de CH₃CN (0,1 % TFA) - Débit : 1 mL/min - lecture à 220 nm.

Ce peptide dendrimérique purifié (10 mg environ) ainsi que 30 mg de peptide dendrimérique brut ont été oxydés selon les conditions vues précédemment (Cf. chapitre II-C.3.b). Les profils HPLC obtenus, après oxydation et dessalage, sont reportés **Figure 122**.

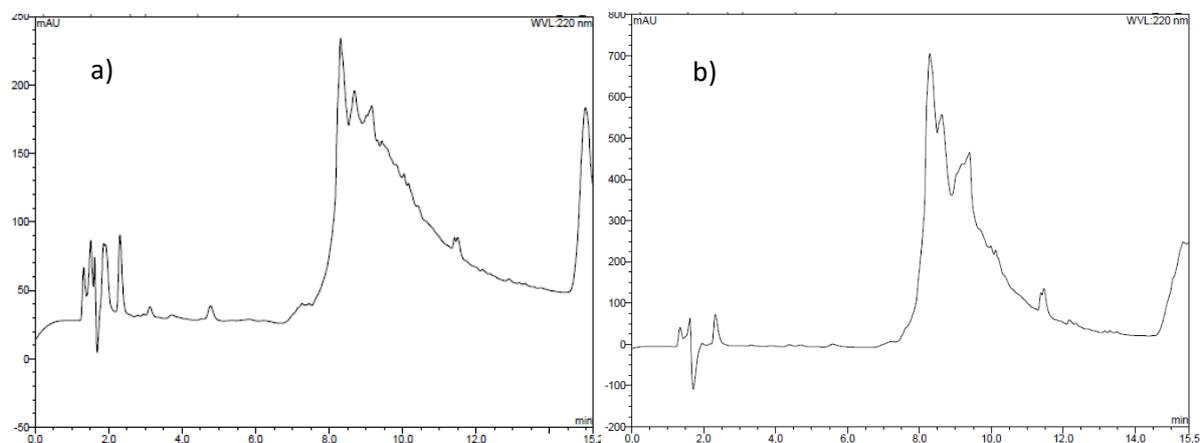


Figure 122. Profil HPLC en phase inverse a) du dendrimère brut oxydé/dessalé et b) du dendrimère purifié puis oxydé/dessalé. Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 5 μm , 200 \AA , 4,6x100 mm – Elution en mode isocratique pendant 2,5 min à 1 % de CH_3CN (0,1 % TFA) puis avec un gradient en 10 min de 1 à 20 % de CH_3CN (0,1 % TFA) - Débit : 1 mL/min - lecture à 220 nm.

Les profils très similaires obtenus suggèrent que l'étape de purification du brut avant oxydation n'est pas nécessaire, il semble que ce soit l'étape d'oxydation qui génère de nombreux sous-produits difficiles à séparer du peptide attendu.

Plutôt que de « perdre » du temps à purifier un brut de synthèse aussi complexe, nous avons estimé qu'il serait préférable de relancer une nouvelle synthèse, en utilisant cette fois une résine Fmoc-Rink-Amid MBHA moins substituée (de l'ordre de 0,2 mmol/g au lieu de 0,6) pour améliorer les couplages d'acides aminés, rendus plus difficiles avec l'encombrement. Cette nouvelle synthèse reste à faire.

b) Polymères de type branchés

Une autre conception, qui n'est pour le moment qu'à l'état d'idée mais qui permettrait également d'améliorer la présentation de nos meilleurs mimes de GAGs, a été envisagée. De même que les protéoglycanes de la matrice extracellulaire sont composés de protéines sur lesquelles sont accrochés des GAGs, voir **Figure 14** du chapitre I-A.2.b, nous synthétiserons un support peptidique linéaire sur lequel viendront s'ancrer, en solution, nos mimes préalablement synthétisés et purifiés. Pour réaliser cet ancrage, nous utiliserons des mimes

porteurs d'un atome de brome à leur extrémité *N*-terminale et les fonctions thiols de Cystéines sur le support peptidique linéaire. Un mécanisme de substitution de type SN_2 , en conditions basiques, permettra ainsi d'obtenir des liaisons thioéther entre les mimes de GAGs et le support peptidique, pour exemple *cf.* **Figure 123**.

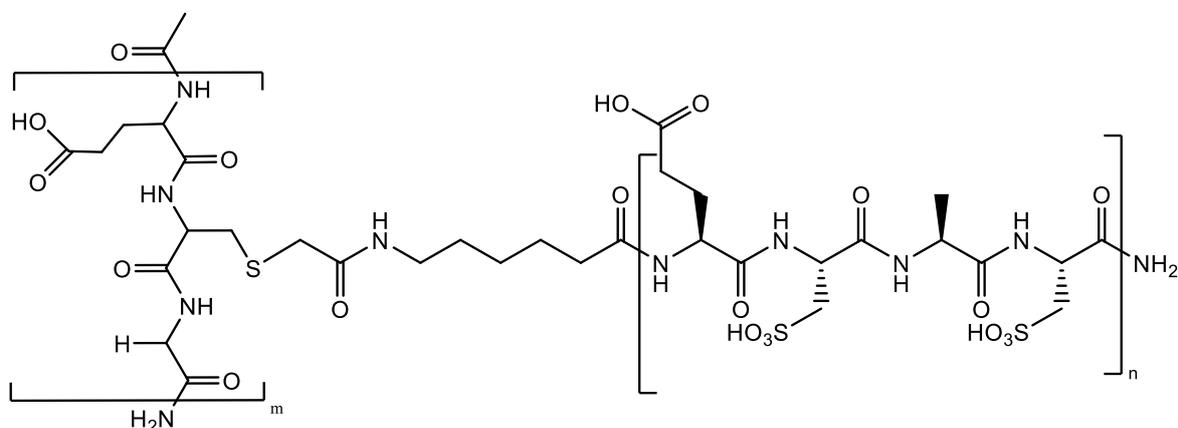


Figure 123. Exemple d'un polymère branché de mimes de GAGs. Les liaisons thioéther sont obtenues, en solution, par réaction de type SN_2 entre les thiols du support peptidique pur (motif ECG répété *m* fois) et les carbones porteurs de brome du mime de GAGs pur (motif EC'AC' répété *n* fois + bras espaceur Ahx)

Ce polymère n'a pas encore été synthétisé, cependant le mécanisme de substitution de type SN_2 , entre un thiol et un carbone portant un brome, a été employé pour la préparation de mimes fluorescents, *cf.* paragraphe 3.b) ci-après.

3. Exploration de nouveaux protocoles de mesures d'affinité.

Nous avons vu, dans le chapitre IV, l'intérêt que pouvaient avoir l'utilisation de la RMN et de la SPR de type Biacore pour l'étude des interactions entre nos mimes de GAGs et le domaine NLS-HD d'Otx2. Cependant la première méthode offre des résultats assez lourds à traiter tandis que la deuxième nécessite l'utilisation d'un appareil et d'un savoir-faire qui n'appartiennent pas au laboratoire. C'est pourquoi il nous a semblé souhaitable de trouver des approches plus rapides et pouvant être réalisées au sein du laboratoire. Pour cela nous nous sommes tournés vers des techniques basées sur la fluorescence.

a) Fluorescence du Tryptophane

De par la présence d'un hétérocycle indole, chromophore dont le spectre de fluorescence est sensible à la polarité de son environnement[220], le Tryptophane des protéines est largement utilisé pour mettre en évidence et mesurer des interactions protéine-ligand. Si le Tryptophane d'une protéine est impliqué dans l'interaction de cette protéine avec un ligand, ce changement d'environnement pourra altérer le spectre de fluorescence du Tryptophane et entraîner une variation de l'intensité de fluorescence et/ou un décalage de la longueur d'onde d'émission maximale[221].

Le domaine NLS-HD d'Otx2 possédant un Tryptophane en position 85 (Hélice 3 de l'homéodomaine), nous avons émis l'hypothèse qu'une interaction de la protéine avec un ligand mime de GAGs pourrait altérer la fluorescence de ce Tryptophane, nous fournissant ainsi un moyen rapide pour quantifier les interactions. Malheureusement, la fluorescence du Tryptophane, enfoui dans le noyau hydrophobe de l'homéodomaine, est « quenchée » par la Phénylalanine en position 45, diminuant ainsi son rendement quantique[222]. De ce fait, je n'ai observé qu'une très faible émission autour de 312 nm du domaine NLS-HD d'Otx2 (pourtant à 110 μ M), incompatible avec des expériences de liaison.

b) Polarisation de fluorescence

Puisque nous ne pouvions utiliser la fluorescence intrinsèque de la protéine NLS-HD Ox2, nous avons envisagé d'ajouter une sonde fluorescente aux mimes peptidiques de GAGs, selon deux approches distinctes.

Nous pouvions, selon le même principe que pour la fluorescence du Tryptophane, ajouter aux sulfopeptides une sonde sensible à son environnement et ainsi espérer mesurer des variations de fluorescence entre les mimes libres et les mimes liés à la protéine cible afin de déterminer l'affinité (K_d). Des essais préliminaires ont été réalisés mais ces travaux sont trop récents pour permettre de les discuter dans ce manuscrit.

L'autre approche était de coupler aux sulfopeptides une sonde fluorescente dont la fluorescence ne serait ici pas modifiée par une interaction avec la protéine (la fluorescéine) mais permettrait de réaliser des expériences de mesure d'affinité par la technique de polarisation de fluorescence (PF).

Polarisation de fluorescence - Principe

Lorsque qu'une lumière polarisée dans un plan excite dans une solution une petite molécule fluorescente on observe l'émission de lumière dans différents plans de polarisation (lumière non polarisée) car cette petite molécule tourne rapidement durant le laps de temps séparant l'excitation de l'émission (mouvements browniens). Par contre, si cette petite molécule fluorescente est liée à une plus grosse, le complexe formé a une rotation moindre durant le temps séparant l'excitation et l'émission et on observe alors l'émission d'une lumière davantage polarisée dans le même plan que la source d'excitation, **Figure 124**. Une courbe représentant la valeur de polarisation de fluorescence (PF) de la petite molécule en fonction de la concentration en grosse molécule peut ainsi être tracée, en mesurant la fluorescence émise par la petite molécule dans le même plan ($F_{||}$) et dans le plan perpendiculaire à la source d'excitation (F_{\perp}) et en appliquant la formule $PF = ((F_{||} - F_{\perp}) / (F_{||} + F_{\perp}))$. La constante d'affinité, K_d , des molécules est alors obtenue en déterminant la concentration en grosse molécule nécessaire pour obtenir 50 % de petites molécules fluorescentes sous forme liée et 50 % sous forme libre.

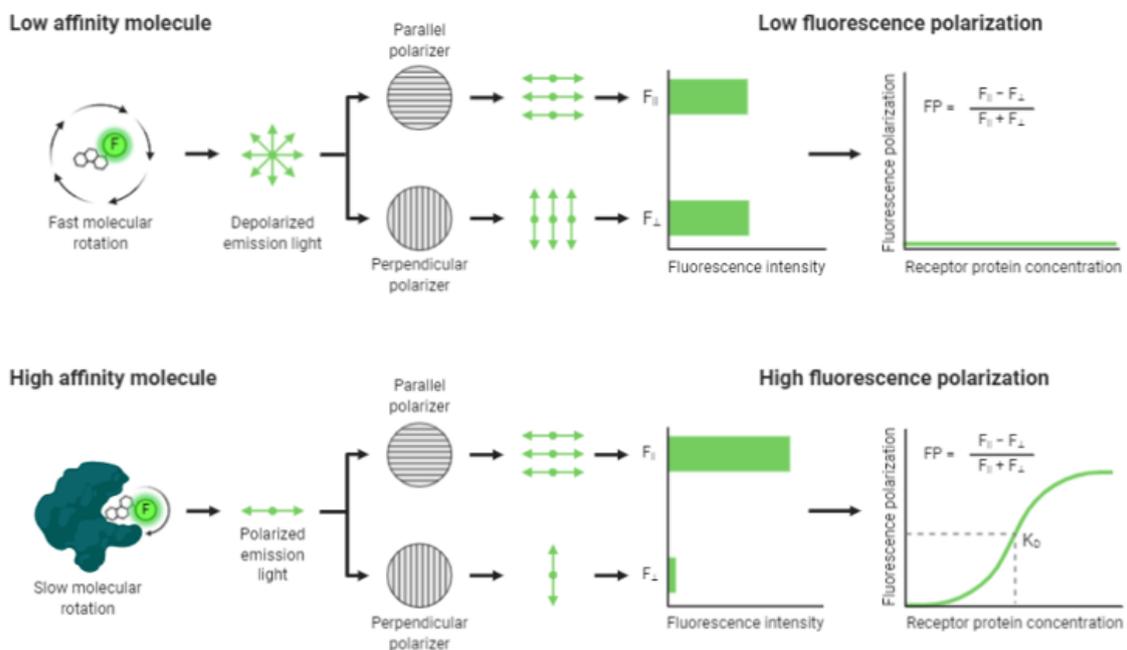


Figure 124. Principe de la polarisation de fluorescence. Figure extraite de[223]

Nous pourrions donc par PF quantifier les interactions entre un de nos mimes, modifié par une sonde fluorescente, et la protéine NLS-HD Otx2, en déterminant l'affinité, K_d , puis,

dans un deuxième temps, connaissant ce K_d , réaliser des mesures d'affinité relative (CI_{50}) d'analogues non fluorescents par des expériences de compétition.

Synthèse de peptides fluorescents

Afin de rendre nos mimes de GAGs fluorescents, nous avons choisi d'utiliser la fluorescéine, molécule soluble dans l'eau et dont les caractéristiques physiques (coefficient d'extinction molaire élevé, rendement quantique proche de 1 à pH neutre...) en font un bon fluorophore, classiquement utilisé en PF. Toutefois, il était impossible de coupler la fluorescéine lors de la synthèse du peptide sur support solide car l'étape d'oxydation qui s'ensuit aurait dégradé le fluorophore. Il nous a donc fallu réaliser la synthèse du mime fluorescent en trois étapes :

- 1) la synthèse, l'oxydation et la purification du mime peptidique de GAG (Bloc 1)
- 2) la synthèse et la purification de la partie fluorescente (Bloc 2)
- 3) le couplage des deux blocs (Bloc 1 + 2)

Pour « assembler » les deux blocs, nous avons choisi d'utiliser une réaction de substitution de type SN_2 , en conditions basiques, entre un carbone portant un atome de brome présenté par le mime peptidique (Bloc 1) et la fonction thiol de Cystéine portée par la partie fluorescente (Bloc 2), **Figure 125**.

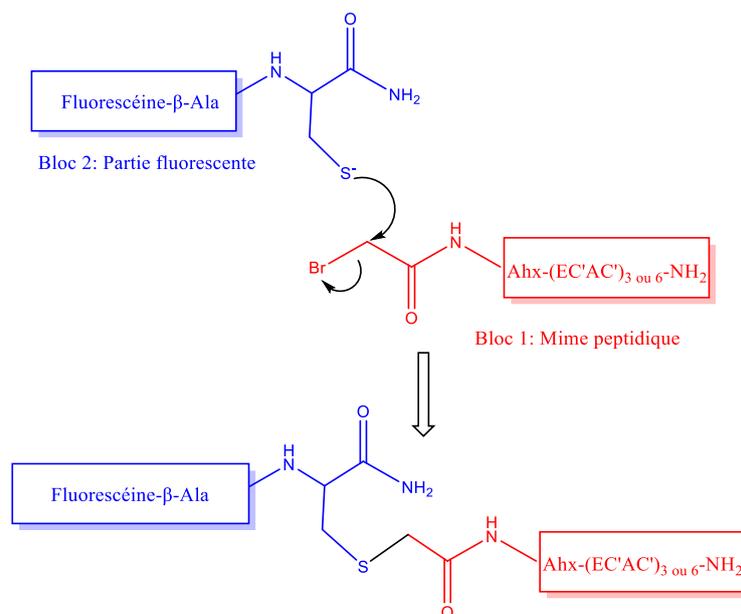


Figure 125. Représentation schématique de l'« assemblage » d'un mime de GAGs fluorescent, par une réaction de type SN_2 .

Le Bloc 1 a été synthétisé, pour la partie « ECAC » ((ECAC)₃ et (ECAC)₆), par commande au niveau de la plateforme de synthèse peptidique de l'Institut de Biologie Paris-Seine (IBPS) (cf. chapitre II-C.4.) puis, en N-terminal, ont été ajoutés manuellement un bras espaceur 6-aminohexanoïque, Ahx, après activation de Fmoc-Ahx-OH et un acide bromoacétique activé par l'agent de couplage DIC (*N,N'*-Diisopropylcarbodiimide). Après coupure puis oxydation (cf. chapitre V-B.1.a), le mélange final a été purifié par HPLC préparative puis analysé par HPLC analytique et par spectrométrie de masse MALDI-TOF avec une matrice *p*-nitroaniline.

Le Bloc 2 a été synthétisé manuellement sur résine Fmoc-Rink-Amid MBHA (Fmoc-Rink-Amid-4-méthylbenzhydrylamine) en utilisant, dans l'ordre des couplages, Fmoc-Cys(Trt)-OH, Fmoc-β-Ala-OH et 5(6)-carboxyfluorescéine. Les couplages ont été réalisés avec le couple d'agents activateurs HBTU/DIEA et la déprotection du Fmoc par action de pipéridine. La coupure du support a été réalisée selon le mode opératoire déjà décrit (cf. chapitre II-C.3.a). Le peptide brut a ensuite été solubilisé dans un mélange H₂O/CH₃CN à 50 % avant d'être lyophilisé. Après lyophilisation, le peptide a été purifié par HPLC préparative puis analysé par HPLC analytique et par spectrométrie de masse MALDI-TOF avec une matrice HCCA.

Le Bloc 1 (version avec (EC'AC')₃ ou (EC'AC')₆) et le Bloc 2, purs et lyophilisés, ont été solubilisés, ensemble et en quantité équimolaire (1 mM chacun après solubilisation), dans un tampon carbonate de sodium/hydrogénocarbonate de sodium à 100 mM, pH 8,9, dégazé sous vide. Le couplage entre les deux blocs s'est déroulé durant la nuit, sous atmosphère inerte et un suivi de la réaction a été effectué par HPLC analytique, pour exemple, cf. **Figure 126**.

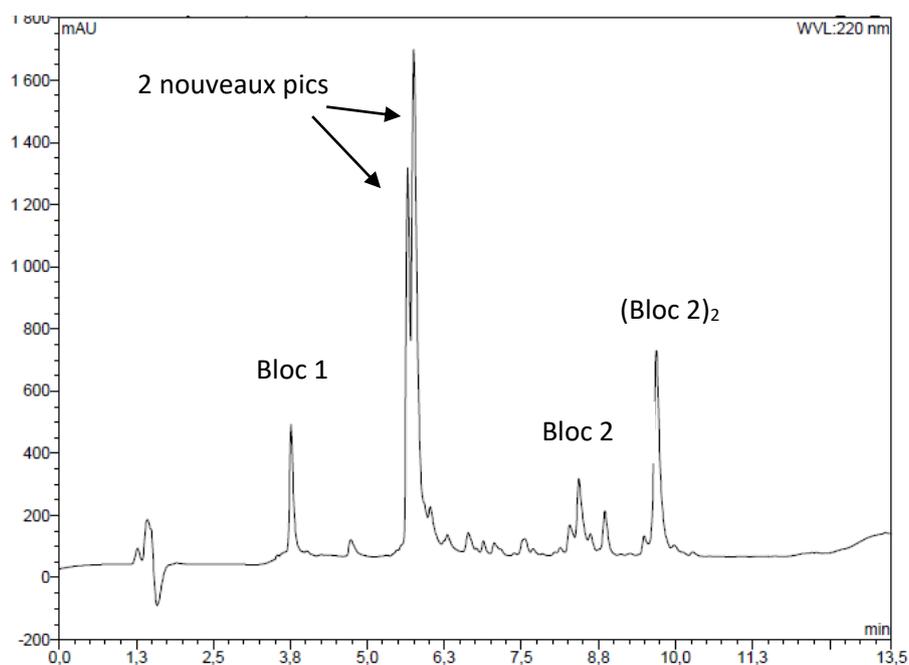


Figure 126. Exemple de suivi, par HPLC analytique, du couplage entre Fluorescéine-β-Ala-Cys-NH₂ (Bloc 2) et Br-Ahx-(EC'AC')₃-NH₂ (Bloc 1). Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 3 μm, 200 Å, 4,6x100 mm – Elution avec un gradient en 10 min de 1 à 50 % de CH₃CN (0,1 % TFA) - Débit : 1 mL/min - Lecture à 220 nm.

Rq : L'apparition de dimères (Fluorescéine-β-Ala-Cys-)₂ ou (Bloc 2)₂, résultant de la formation en milieu basique d'un pont disulfure entre les Cystéines, nous a contraint à ajouter, dans le milieu réactionnel, une quantité plus importante de Bloc 2, jusqu'à observer en HPLC la disparition totale du Bloc 1.

Nous avons observé l'apparition de deux nouveaux pics sur les spectres HPLC, qui, après analyses par MALDI-TOF MS, **Figures 127a et b**, se sont avérés avoir tous deux la masse moléculaire attendue. Ces deux produits, sans doute des isomères obtenus en raison du mélange 5(6)-carboxyfluorescéine utilisé, ont été purifiés par HPLC préparative puis analysés par HPLC analytique avant d'être lyophilisés, voir **Figures 128a et b**.

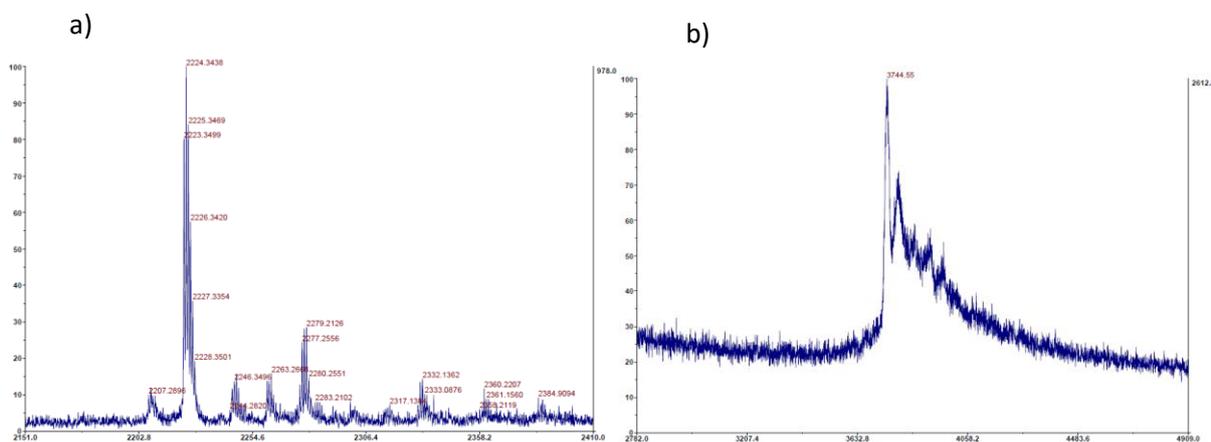


Figure 127. Profil MALDI-TOF des deux nouveaux pics observés sur les spectres HPLC de suivi du couplage entre Fluorescéine- β -Ala-Cys-NH₂ (Bloc 2) et a) Br-Ahx-(EC'AC')₃-NH₂ (Bloc 1) avec matrice *p*-nitroaniline, en mode réflecteur négatif ou b) Br-Ahx-(EC'AC')₆-NH₂ (Bloc 1), avec matrice *p*-nitroaniline, en mode linéaire négatif.

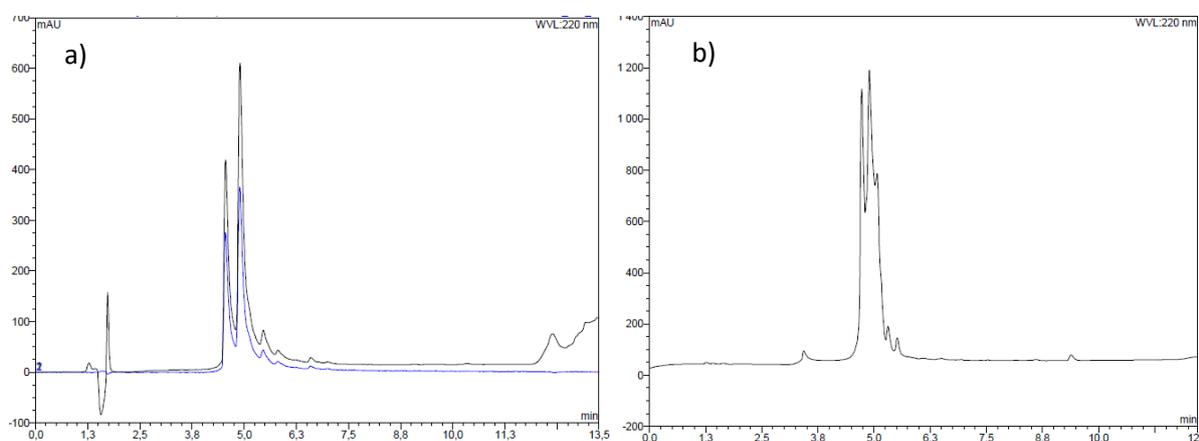


Figure 128. Profil HPLC de : a) Fluo- β -Ala-Cys-Ahx-(EC'AC')₃-NH₂ purifié et b) Fluo- β -Ala-Cys-Ahx-(EC'AC')₆-NH₂ purifié. Conditions : Colonne C18 Proto200 (Higgins Analytical, Inc.), 3 μ m, 200 Å, 4,6x100 mm - Débit : 1 mL/min – a) Elution avec un gradient en 10 min de 10 à 30 % de CH₃CN (0,1 % TFA), lecture à 220 (en noir) et 434 nm (en bleu) et b) Elution avec un gradient en 10 min de 1 à 50 % de CH₃CN (0,1 % TFA), lecture à 220 nm.

Les mimes fluorescents obtenus, Fluo- β -Ala-Cys-Ahx-(EC'AC')₃-NH₂ et Fluo- β -Ala-Cys-Ahx-(EC'AC')₆-NH₂, ont ensuite été mis en solution dans de l'eau distillée et leur pH a été amené autour de 7 par ajout d'une solution de NaOH à 1M. La mesure de leur absorbance à 490 nm, longueur d'onde d'absorption maximale de la fluorescéine à ce pH, a alors permis de déterminer la concentration des solutions, en utilisant la loi de Lambert-Beer.

Résultats préliminaires de la polarisation de fluorescence

Des expériences préliminaires de polarisation de fluorescence ont été menées en utilisant une solution de Fluo- β -Ala-Cys-Ahx-(EC'AC')₆-NH₂ à 10 nM final en présence d'une concentration croissante en NLS-HD Otx2. Nous avons opéré de façon à ce que le volume de l'échantillon analysé ne varie pas au cours des expériences. Les valeurs moyennes de polarisation de la sonde fluorescente sont reportées **Figure 129**, en fonction de la concentration en protéine.

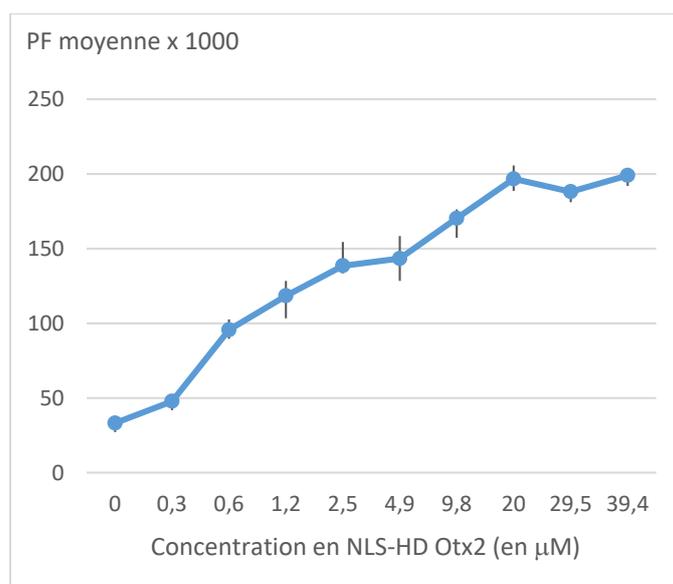


Figure 129. Courbe représentant la valeur moyenne (obtenue à partir de 3 réplicats) de la polarisation de fluorescence (PF) de Fluo- β -Ala-Cys-Ahx-(EC'AC')₆-NH₂ à 10 nM final en fonction de la concentration en protéine NLS-HD Otx2 dans le milieu.

Nous observons bien une augmentation de la polarisation de fluorescence de Fluo- β -Ala-Cys-Ahx-(EC'AC')₆-NH₂ en fonction de la concentration en NLS-HD Otx2. Cependant la courbe obtenue n'est pas de type sigmoïde, comme représenté théoriquement **Figure 124**. Cette expérience préliminaire nécessite certainement des expériences complémentaires, en ajoutant par exemple des concentrations supplémentaires de protéine pour affiner la courbe et tenter de visualiser les plateaux. Toutefois, au vu de ces résultats, nous sommes en droit de nous poser certaines questions quant à cette approche :

- La différence de masse moléculaire entre le mime fluorescent (3730 Da) et la protéine (8480 Da), ici un facteur 2,3, est-elle suffisante pour permettre une immobilisation efficace de la sonde ? Il est en effet d'usage de travailler avec des

protéines de plusieurs dizaines de kDa et avec une masse moléculaire environ 10 fois supérieure à celle de la petite molécule fluorescente[224].

- La concentration en Fluo- β -Ala-Cys-Ahx-(EC'AC')₆-NH₂ utilisée (10 nM) est-elle suffisante pour obtenir des mesures fiables et reproductibles de fluorescence ou sommes-nous à la limite de la sensibilité de l'appareil utilisé ?
- La fluorescence totale de Fluo- β -Ala-Cys-Ahx-(EC'AC')₆-NH₂ est-elle bien la même quelle que soit la concentration de protéine NLS-HD Otx2 employée, point essentiel pour une bonne utilisation de la technique ? Pour tenter de répondre à cette question, la fluorescence des solutions employées lors des expériences de polarisation de fluorescence (mime fluorescent + protéine) a été mesurée (excitation à 485 nm et émission à 520 nm), les résultats des mesures sont reportés **Figure 130**.

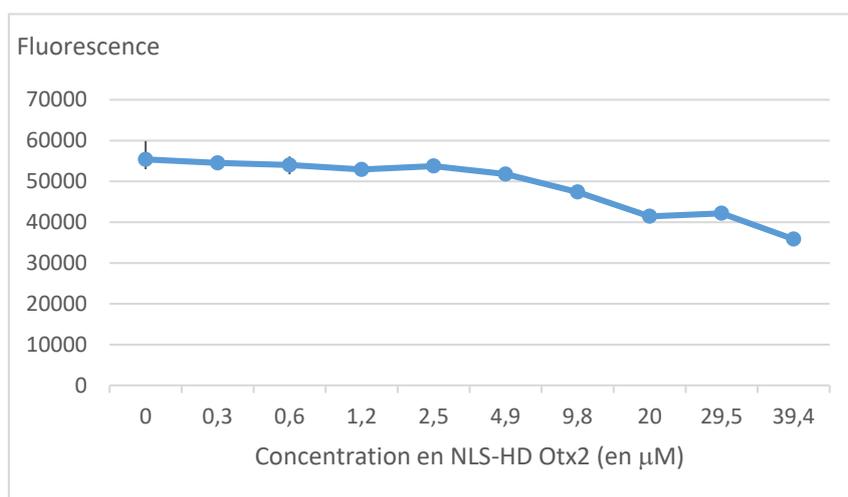


Figure 130. Evolution de la valeur moyenne (obtenue à partir de 3 réplicats) de la fluorescence totale de Fluo- β -Ala-Cys-Ahx-(EC'AC')₆-NH₂ à 10 nM final en fonction de la concentration en protéine NLS-HD Otx2 dans le milieu.

Au-delà d'une concentration de 5 μ M environ de NLS-HD Otx2, la fluorescence de la sonde diminue. La fluorescence de la fluorescéine serait-elle sensible à cet environnement protéique concentré ? A moins qu'il ne s'agisse simplement d'une erreur de manipulation ? Tout cela nécessitera des vérifications afin de pouvoir apporter les réponses à ces questions.

C. Partie expérimentale

Le protocole général de synthèse, l'étape d'oxydation, les étapes de purification et les conditions de stockage des peptides ayant été décrits dans le chapitre II-C3 et C4 et dans le « Supplementary Information » du manuscrit pour publication inséré chapitre III, je ne détaillerai, dans cette partie expérimentale, que les éléments présentant un caractère nouveau.

1. Synthèse des analogues linéaires avec agrafe

Déprotection sélective de la chaîne latérale de la Lysine (Mmt) sur support solide : Les groupements protecteurs de la chaîne latérale des Lysines ont été retirés en effectuant plusieurs passages de la résine dans un mélange AcOH/TFE/DCM en proportion volumique 1/2/7, jusqu'à disparition de la coloration jaune provenant de la libération du Mmt dans le milieu : 5 x 1 minute, 2 x 5 minutes, 3 x 10 minutes puis 1 x 1 heure. Entre chaque passage, la résine a été filtrée puis lavée au dichlorométhane (DCM).

Couplage du chlorure de biphényl-4,4'-disulfonyl : 1 équivalent, par rapport au peptide, de chlorure de biphényl-4,4'-disulfonyl a été couplé, en présence de 4 équivalents de DIEA, dans du DMF, durant 15 heures à température ambiante et sous agitation.

2. Synthèse des polymères de type dendrimère

Le peptide de type dendrimère a été synthétisé manuellement sur support solide, par stratégie Fmoc, sur 0,1 mmol de résine Fmoc-Rink-Amid MBHA (Iris Biotech, Ref BR-1300, Fmoc-Rink-Amid-4-méthylbenzhydrylamine, 100-200 Mesh, substituée à 0,62 mmol/g). Les acides aminés ont été activés durant 3 min par le couple d'agent HBTU/DIEA dans le DMF, couplés à température ambiante, sous agitation, durant 3 heures et les groupements protecteurs Fmoc ont été retirés par action de pipéridine à 20 % dans le DMF (3 min, 5 min puis 10 min). Entre chaque étape la résine a été lavée au DMF. Les 3 premiers couplages, de Fmoc-Ahx-OH, Fmoc-Lys(Mmt)-OH puis Fmoc-Lys(Fmoc)-OH, ont été réalisés en utilisant 3 équivalents d'acide aminé (0,3 mmol), 2,8 équivalents de HBTU (0,28 mmol) et 6 équivalents de DIEA (0,6 mmol). Le couplage de l'acide aminé suivant, Fmoc-Lys(Fmoc)-OH a nécessité l'utilisation de 6 équivalents d'acide aminé (0,6 mmol), 5,6 équivalents de HBTU (0,56 mmol)

et 12 équivalents de DIEA (1,2 mmol). Les couplages des 4 acides aminés suivants (Fmoc-Cys(Trt)-OH, Fmoc-Ala-OH, Fmoc-Cys(Trt)-OH puis Fmoc-Glu(OtBu)-OH) ont été réalisés en utilisant 16 équivalents d'acide aminé (1,6 mmol), 15,2 équivalents de HBTU (1,56 mmol) et 32 équivalents de DIEA (3,2 mmol). Le groupement protecteur Mmt a été retiré sélectivement par action d'un mélange AcOH/TFE/DCM (1/2/7 en volume) durant 8 x 15 min, avec un lavage de la résine au DCM entre chaque ajout de mélange frais. Après leur déprotection, les extrémités *N*-terminales ainsi que l'amine de la chaîne latérale de la Lysine ont été acétylées par action d'anhydride acétique à 20 % dans le DMF et de 10 équivalents de DIEA (1 mmol) durant 2 heures. La résine a ensuite été lavée avec du DMF, du DCM puis du MeOH avant d'être séchée.

3. Purifications par HPLC préparative

Les peptides bruts discutés dans ce chapitre V ont été purifiés par HPLC préparative sur une colonne C18 Ace (AIT France), 5 µm, 300 Å, 10x250 mm, à un débit de 6 mL/min :

- pour le peptide « agrafe » avec un gradient en 30 min de 1 à 20 % de CH₃CN (0,1 % TFA) et une lecture à 220 et 245 nm,
- pour le peptide dendrimérique avec un gradient en 30 min de 5 à 100 % de CH₃CN (0,1 % TFA) et une lecture à 220 nm,
- pour le « Bloc 1 » de façon isocratique pendant 5 min à 0 % de CH₃CN (0,1 % TFA) puis avec un gradient en 30 min de 0 à 10 % de CH₃CN (0,1% TFA) et une lecture à 220 nm,
- pour le « Bloc 2 » avec un gradient en 30 min de 10 à 40 % de CH₃CN (0,1 % TFA) et une lecture à 220 nm,
- pour Fluo-β-Ala-Cys-Ahx-(EC'AC')₃-NH₂ et Fluo-β-Ala-Cys-Ahx-(EC'AC')₆-NH₂ avec un gradient en 30 min de 1 à 30 % de CH₃CN (0,1 % TFA) et une lecture à 220 nm.

4. Caractéristiques HPLC et spectrométrie de masse

Les caractéristiques HPLC et spectrométrie de masse des peptides du chapitre V sont reportées ci-dessous, sous forme de tableau, **Tableau 16**.

Code personnel	Gradient analytique	R _T (min)	m/z calculé [M-H] ⁻	m/z mesuré [M-H] ⁻
Agrafe	Gradient de 1 à 20% B en 10 min	10,2	2069	2069
RM4-47A : Dendrimère	Isocratique 1 % B pendant 2,5 min et gradient de 1 à 20 % B en 10 min	8,2	2860	2867
RM4-71 : Fluo-β-Ala-Cys-Ahx-(EC'AC') ₃ -NH ₂	Gradient de 10 à 30% B en 10 min	4,7 et 5	2224	2223
RM4-99 : Fluo-β-Ala-Cys-Ahx-(EC'AC') ₆ -NH ₂	Gradient de 1 à 50% B en 10 min	4,7 et 5	3730	3747

Tableau 16. Caractéristiques HPLC et spectrométrie de masse des différents peptides synthétisés (Eluants HPLC : A = H₂O (0,1% TFA) et B = CH₃CN (0,1% TFA))

5. Fluorescence du Tryptophane

Les mesures de fluorescence de la protéine NLS-HD Otx2 ont été réalisées avec un fluorimètre MOS 200M (Biologic SA) dans une cuve en quartz de 100 μL, en utilisant une longueur d'onde d'excitation de 280 nm et une gamme d'émission entre 300 et 400 nm. La mesure du blanc a été réalisée sur un tampon PBS (Tampon Phosphate Salin, comprimé dissout dans 200 mL d'eau distillée, pH 7,4 à 25 °C – Sigma-Aldrich – Ref P4417). Les mesures de fluorescence ont été réalisées sur la protéine NLS-HD Otx2, diluée dans du tampon PBS ou utilisée pure (110 μM dans le tampon de FG – cf. chapitre IV-E.1.a).

6. Polarisation de fluorescence

Les expériences de polarisation de fluorescence ont été menées sur le lecteur de plaque POLARstar OPTIMA (BMG LABTECH), avec des microplaques noires « non binding » de 384 puits de 130 μL (Greiner Bio-One, ref 781900). Chaque condition expérimentale a été réalisée en triplicate. Dans les puits, 10 μL d'une solution de Fluo-β-Ala-Cys-Ahx-(EC'AC')₆-NH₂ (diluée à 40 nM dans un tampon PBS à pH 7,4) ont été déposés puis 30 μL de solution de protéine NLS-HD Otx2 à différentes concentrations (de 0,4 à 52 μM, dans le tampon PBS) ou 30 μL de tampon PBS (témoin sans protéine). Après 5 minutes d'incubation à 37 °C puis 1 minute d'agitation, dans le noir, la lecture a été réalisée, avec une longueur d'onde d'excitation à 485 nm et une longueur d'onde d'émission à 520 nm.

CHAPITRE VI : Conclusion générale et perspectives

Dans le cadre de mon travail de thèse, en collaboration avec l'équipe du Pr Alain Prochiantz au Collège de France, ***j'ai conçu et synthétisé des mimes de GAGs à partir d'un squelette peptidique.*** Les synthèses de peptides, plus rapides que celles de saccharides analogues de GAGs, devaient permettre de moduler plus facilement des modifications sur les chaînes latérales des acides aminés.

Les sulfopeptides que j'ai préparés présentent des ***motifs répétés de type (EC'AC')_n***, avec les fonctions acide de l'acide Glutamique, et sulfonique obtenue par oxydation de la Cystéine. ***J'ai mis au point les différentes étapes de synthèse, et ainsi pu obtenir quelques analogues comme preuve de concept.***

En collaboration avec Damien Testa, qui a préparé et soutenu une thèse sous la direction du Dr Ariel Di Nardo au Collège de France, ***nous avons étudié les activités de cette première génération de sulfopeptides comme inhibiteurs de l'interaction d'une protéine à homéodomaine (Otx2) avec le réseau de GAGs*** présents autour d'un type cellulaire cible de cette protéine. ***Le peptide comportant six motifs « EC'AC' » a montré une activité in vivo très prometteuse.*** Nous avons en conséquence déposé un brevet (WO2019243272A1) et une demande ANR qui a été acceptée (<https://anr.fr/Projet-ANR-18-CE16-0013>).

J'ai réalisé une étude structurale par RMN avec l'aide du Dr Ludovic Carlier du LBM pour analyser les interactions des sulfopeptides avec la protéine Otx2, que j'ai exprimée et produite avec un double marquage ¹³C, ¹⁵N. J'ai ainsi pu déterminer les zones de contact entre ces ligands et Otx2, ainsi que leurs affinités pour les acides aminés d'Otx2 affectés par la liaison de ces sulfopeptides. ***En accord avec notre hypothèse de départ, ces sulfopeptides interagissent avec le domaine de liaison aux GAGs d'Otx2, ils affectent aussi l'hélice 3 de l'homéodomaine*** qui est à l'origine de l'internalisation d'Otx2 dans les cellules.

Nous avons complété cette étude structurale par une approche de type Biacore, pour développer une stratégie d'analyse plus rapide d'activité de nos analogues et nous permettre de développer une étude structure-activité plus efficace. Je n'ai décrit que quelques-uns des analogues que nous avons développés ces derniers mois et d'ores et déjà j'ai montré que modifier l'Alanine au cœur du motif initial par une Leucine ou une Proline conduisait à des sulfopeptides 2,5 à 5 fois plus affins pour Otx2 que notre motif de départ.

J'ai ouvert la porte à d'autres perspectives de synthèse et d'analyses physicochimiques, comme exposé dans le dernier chapitre.

Il demeure deux points qui resteront à analyser en parallèle « des portes que j'ai ouvertes », à savoir la stabilité métabolique de nos meilleurs analogues et l'immunogénicité de ces molécules, deux points cruciaux pour leur éventuelle utilisation en thérapeutique. Concernant le premier point, une expérience préliminaire encourageante a d'ores et déjà été réalisée, au Collège de France, par la nouvelle doctorante de l'équipe, Rachel Gibel-Russo. Elle a ainsi observé, par Dot-blot, que la pré-incubation du mime Biot(O₂)-Ahx-(EC'AC')₆-NH₂ avec une endopeptidase à large spectre de digestion, la protéinase K, n'altérerait pas son interaction avec la protéine Otx2, ce qui nous fournit un bon indice quant à la stabilité de nos composés.

Bibliographie

- [1] Université d'Angers. La synthèse des protéines. *Cours en ligne : figure extraite d'une publication de E. Jaspard (2013)* biochimej.univ-angers.fr/Page2/COURS/7RelStructFonction/2Biochimie/1SyntheseProteines/1SyntheseProt.htm.
- [2] Anfinsen, C. B., Haber, E., Sela, M. & White, F. H. The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain. *Proc Natl Acad Sci USA* **47**, 1309 (1961).
- [3] Anfinsen, C. B. Studies on the principles that govern the folding of proteins chains. *Nobel Lect.* (1972).
- [4] Uversky, V. N. Intrinsically Disordered Proteins and Their "Mysterious" (Meta)Physics. *Front. Phys.* **7**, (2019).
- [5] Pauling, L., Corey, R. B. & Branson, H. R. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proc Natl Acad Sci USA* **37**, 205–211 (1951).
- [6] https://simple.wikipedia.org/wiki/Protein_structure. *Simple English Wikipedia, the free encyclopedia* https://simple.wikipedia.org/w/index.php?title=Protein_structure&oldid=6899034 (2020).
- [7] Feuillet bêta. *Wikipédia* https://fr.wikipedia.org/wiki/Feuillet_b%C3%AAta (2020).
- [8] Hélice alpha. *Wikipédia* https://fr.wikipedia.org/wiki/H%C3%A9lice_alpha (2021).
- [9] Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.* **7**, 95–99 (1963).
- [10] Diez, M., Petuya, V., Martínez-Cruz, L. A. & Hernández, A. Biokinematic protein simulation by an adaptive dihedral angle approach. *Mech. Mach. Theory* **69**, 105–114 (2013).
- [11] Université d'Angers. Les structures des protéines - Introduction. *Cours en ligne : figure extraite d'une publication de Kuriyan et al. (2013)* <http://biochimej.univ-angers.fr/Page2/COURS/7RelStructFonction/3Structure/1StructPrimQuat/1Introduction.htm>.
- [12] Whitby, L. R., Ando, Y., Setola, V., Vogt, P. K., Roth, B. L. & Boger, D. L. Design, Synthesis, and Validation of a β -Turn Mimetic Library Targeting Protein–Protein and Peptide–Receptor Interactions. *J. Am. Chem. Soc.* **133**, 10184–10194 (2011).
- [13] de Brevern, A. G. Extension of the classical classification of Beta-turns. *Sci. Rep.* **6**, (2017).
- [14] Camerman, A., Mastropaolo, D., Karle, I., Karle, J. & Camerman, N. Crystal structure of leucine-enkephalin. *Nature* **306**, 447–450 (1983).
- [15] Doi, M., Tanaka, M., Ishida, T., Inoue, M., Fujiwara, T., Tomita, T., Kimura, T., Sakakibara, S. & Sheldrick, G. M. Crystal Structures of [Met5] and [(4-Bromo)Phe4,Met5]Enkephalins : Formation of a Dimeric Antiparallel Beta-Structure. *J Biochem* **101**, 485–490 (1987).
- [16] Griffin, J. F., Langs, D. A., Smith, G. D., Blundell, T. L., Tickle, I. J. & Bedarkar, S. The crystal structures of [Met5]enkephalin and a third form of [Leu5]enkephalin: Observations of a novel pleated Beta-sheet. *Proc Natl Acad Sci USA* **83**, 3272–3276 (1986).
- [17] Pawlak, T. & Potrzebowski, M. J. Fine Refinement of Solid-State Molecular Structures of Leu- and Met-Enkephalins by NMR Crystallography. *J. Phys. Chem. B* **118**, 3298–3309 (2014).
- [18] du Vigneaud, V., Ressler, C., Swan, J. M., Roberts, C. W. & Katsoyannis, P. G. The Synthesis of Oxytocin. *J. Am. Chem. Soc.* **76**, 3115–3121 (1954).
- [19] du Vigneaud, V., Gish, D. T. & Katsoyannis, P. G. A synthetic preparation possessing biological properties associated with arginine-vasopressin. *J Am Chem Soc* **76**, 4751–4752 (1954).
- [20] Zhang, Y. The first protein ever synthesized in vitro—a personal reminiscence of the total synthesis of crystalline insulin. *Sci. China Life Sci.* **53**, 16–18 (2010).
- [21] Brunel, F. M., Mayer, J. P. & DiMarchi, R. D. Recent advances in the chemical synthesis of insulin and related peptides. *Future Med. Chem.* **12**, 649–654 (2020).
- [22] Fischer, E. & Fourneau, E. Ueber einige Derivate des Glykocolls. *Berichte Dtsch. Chem. Ges.* **34**, 2868–2877 (1901).
- [23] Bergmann, M. & Zervas, L. Über ein allgemeines Verfahren der Peptid-Synthese. *Berichte Dtsch. Chem. Ges. B Ser.* **65**, 1192–1201 (1932).
- [24] du Vigneaud, V., Ressler, C., Swan, C. J. M., Roberts, C. W., Katsoyannis, P. G. & Gordon, S. The synthesis of an octapeptide amide with the hormonal activity of oxytocin. *J. Am. Chem. Soc.* **75**, 4879–4880 (1953).
- [25] Merrifield, R. B. Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
- [26] Merrifield, R. B. Solid phase synthesis. *Science* **232**, 341–347 (1986).
- [27] Merrifield, R. B., Stewart, J. M. & Jernberg, N. Instrument for automated synthesis of peptides. *Anal Chem* **38**, 1905–1914 (1966).

- [28] Merrifield, R. B. New Approaches to the Chemical Synthesis of Peptides. in *Recent Progress in Hormone Research* vol. 23 451–482 (Elsevier, 1967).
- [29] Merrifield, R. B. Solid phase synthesis. *Nobel Lecture*
<https://www.nobelprize.org/uploads/2018/06/merrifield-lecture.pdf> (1984).
- [30] Dawson, P., Muir, T., Clark-Lewis, I. & Kent, S. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).
- [31] Agouridas, V., El Mahdi, O., Diemer, V., Cargoët, M., Monbaliu, J.-C. M. & Melnyk, O. Native Chemical Ligation and Extended Methods: Mechanisms, Catalysis, Scope, and Limitations. *Chem. Rev.* **119**, 7328–7443 (2019).
- [32] Muir, T. W., Sondhi, D. & Cole, P. A. Expressed protein ligation: A general method for protein engineering. *Proc. Natl. Acad. Sci.* **95**, 6705–6710 (1998).
- [33] Rostovtsev, V. V., Green, L. G., Fokin, V. V. & Sharpless, K. B. A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective “Ligation” of Azides and Terminal Alkynes. *Angew. Chem. Int. Ed.* **41**, 2596–2599 (2002).
- [34] Tornøe, C. W., Christensen, C. & Meldal, M. Peptidotriazoles on Solid Phase: [1,2,3]-Triazoles by Regiospecific Copper(I)-Catalyzed 1,3-Dipolar Cycloadditions of Terminal Alkynes to Azides. *J. Org. Chem.* **67**, 3057–3064 (2002).
- [35] Saxon, E. Cell Surface Engineering by a Modified Staudinger Reaction. *Science* **287**, 2007–2010 (2000).
- [36] Sharpless, K. B. Searching for new reactivity. *Nobel Lecture*
<https://www.nobelprize.org/uploads/2018/06/sharpless-lecture.pdf> (2001).
- [37] Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H. *Essentials of Glycobiology, 3rd edition*. (Cold Spring Harbor Laboratory Press, 2015).
- [38] Esko, J. D., Kimata, K. & Lindahl, U. Chapter 16 : Proteoglycans and Sulfated Glycosaminoglycans. in *Essentials of Glycobiology- 2nd edition* (eds. Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E.) (Cold Spring Harbor Laboratory Press, 2009).
- [39] Fischer, E. Syntheses in the purine and sugar group. *Nobel Lect.* (1902).
- [40] Haworth, W. N. The structure of carbohydrates and of vitamin C. *Nobel Lect.* (1937).
- [41] Agirre, J. Strategies for carbohydrate model building, refinement and validation. *Acta Crystallogr. Sect. Struct. Biol.* **73**, 171–186 (2017).
- [42] Reeves, R. E. Cuprammonium--Glycoside Complexes. III. The Conformation of the D-Glucopyranoside Ring in Solution. *J Am Chem Soc* **71**, 215–217 (1949).
- [43] Jeanloz, R. W. The nomenclature of mucopolysaccharides. *Arthritis Rheum.* **3**, 233–237 (1960).
- [44] Gandhi, N. S. & Mancera, R. L. The Structure of Glycosaminoglycans and their Interactions with Proteins. *Chem. Biol. Drug Des.* **72**, 455–482 (2008).
- [45] Fischer, G. & Boedeker, C. Künstliche Bildung von Zucker aus Knorpel (Chondrogen), und über die Umsetzung des genossenen Knorpels im menschlichen Körper. *Justus Liebigs Ann. Chem.* **117**, 111–118 (1861).
- [46] Levene, P. A. & Laforge, F. B. On chondroitin sulphuric acid. *J Biol Chem* **20**, 433–444 (1915).
- [47] Djerbal, L., Lortat-Jacob, H. & Kwok, J. Chondroitin sulfates and their binding molecules in the central nervous system. *Glycoconj. J.* **34**, 363–376 (2017).
- [48] Solís, D., Bovin, N. V., Davis, A. P., Jiménez-Barbero, J., Romero, A., Roy, R., Smetana, K. & Gabius, H.-J. A guide into glycosciences: How chemistry, biochemistry and biology cooperate to crack the sugar code. *Biochim. Biophys. Acta BBA - Gen. Subj.* **1850**, 186–235 (2015).
- [49] Gabius, H.-J. The sugar code: Why glycans are so important. *Biosystems* **164**, 102–111 (2018).
- [50] Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W. & Etzler, M. E. *Essentials of Glycobiology, 2nd Edition*. (Cold Spring Harbor Laboratory Press, 2009).
- [51] Rees, D. A. & Welsh, E. J. Secondary and Tertiary Structure of Polysaccharides in Solutions and Gels. *Angew. Chem. Int. Ed. Engl.* **16**, 214–224 (1977).
- [52] UNISCIEL. Diversité de formes des oligosaccharides. *Université en ligne*
http://uel.unisciel.fr/biologie/biochimie1/biochimie1_ch02/co/apprendre_ch2_03.html.
- [53] Zaccheus, M. *Structural and conformational studies of Oligo- and Polysaccharides*. (Department of Organic Chemistry, Stockholm University, 2012).
- [54] Pérez, S., Kouwijzer, M., Mazeau, K. & Engelsen, S. B. Modeling polysaccharides: Present status and challenges. *J. Mol. Graph.* **14**, 307–321 (1996).
- [55] Fittolani, G., Seeberger, P. H. & Delbianco, M. Helical polysaccharides. *Pept. Sci.* **112**, (2020).
- [56] Capila, I. & Linhardt, R. J. Heparin-Protein Interactions. *Angew. Chem. Int. Ed.* **41**, 390–412 (2002).

- [57] Lim, T. C., Cai, S., Huber, R. G., Bond, P. J., Siew Chia, P. X., Khou, S. L., Gao, S., Lee, S. S. & Lee, S.-G. Facile saccharide-free mimetics that recapitulate key features of glycosaminoglycan sulfation patterns. *Chem. Sci.* **9**, 7940–7947 (2018).
- [58] Plante, O. J. Automated Solid-Phase Synthesis of Oligosaccharides. *Science* **291**, 1523–1527 (2001).
- [59] Zhang, Z., Ollmann, I. R., Ye, X.-S., Wischnat, R., Baasov, T. & Wong, C.-H. Programmable One-Pot Oligosaccharide Synthesis. *J. Am. Chem. Soc.* **121**, 734–753 (1999).
- [60] Kulkarni, S. S., Wang, C.-C., Sabbavarapu, N. M., Podilapu, A. R., Liao, P.-H. & Hung, S.-C. “One-Pot” Protection, Glycosylation, and Protection–Glycosylation Strategies of Carbohydrates. *Chem. Rev.* **118**, 8025–8104 (2018).
- [61] Panza, M., Pistorio, S. G., Stine, K. J. & Demchenko, A. V. Automated Chemical Oligosaccharide Synthesis: Novel Approach to Traditional Challenges. *Chem. Rev.* **118**, 8105–8150 (2018).
- [62] Seeberger, P. H. The Logic of Automated Glycan Assembly. *Acc. Chem. Res.* **48**, 1450–1463 (2015).
- [63] Mende, M., Bednarek, C., Wawryszyn, M., Sauter, P., Biskup, M. B., Schepers, U. & Bräse, S. Chemical Synthesis of Glycosaminoglycans. *Chem. Rev.* **116**, 8193–8255 (2016).
- [64] Li, J., Sparkenbaugh, E. M., Su, G., Zhang, F., Xu, Y., Xia, K., He, P., Baytas, S., Pechauer, S., Padmanabhan, A., Linhardt, R. J., Pawlinski, R. & Liu, J. Enzymatic Synthesis of Chondroitin Sulfate E to Attenuate Bacteria Lipopolysaccharide-Induced Organ Damage. *ACS Cent. Sci.* **6**, 1199–1207 (2020).
- [65] Cheng, C.-W., Zhou, Y., Pan, W.-H., Dey, S., Wu, C.-Y., Hsu, W.-L. & Wong, C.-H. Hierarchical and programmable one-pot synthesis of oligosaccharides. *Nat. Commun.* **9**, (2018).
- [66] Li, J., Su, G. & Liu, J. Enzymatic Synthesis of Homogeneous Chondroitin Sulfate Oligosaccharides. *Angew. Chem. Int. Ed.* **56**, 11784–11787 (2017).
- [67] Xiao, X., Zeng, J., Fang, J., Sun, J., Li, T., Song, Z., Cai, L. & Wan, Q. One-Pot Relay Glycosylation. *J. Am. Chem. Soc.* **142**, 5498–5503 (2020).
- [68] Despras, G., Bernard, C., Perrot, A., Cattiaux, L., Prochiantz, A., Lortat-Jacob, H. & Mallet, J.-M. Toward Libraries of Biotinylated Chondroitin Sulfate Analogues: From Synthesis to In Vivo Studies. *Chem. - Eur. J.* **19**, 531–540 (2013).
- [69] Sugahara, K. Recent advances in the structural biology of chondroitin sulfate and dermatan sulfate. *Curr. Opin. Struct. Biol.* **13**, 612–620 (2003).
- [70] Wong, E. V. Cells - Molecules and Mechanisms/Extracellular Matrix and Cell Adhesion /Introduction to Extracellular Matrix and Cell Adhesion. *LibreTexts* [https://bio.libretexts.org/Bookshelves/Cell_and_Molecular_Biology/Book%3A_Cells_-_Molecules_and_Mechanisms_\(Wong\)/13%3A_Extracellular_Matrix_and_Cell_Adhesion/13.01%3A_Introduction_to_Extracellular_Matrix_and_Cell_Adhesion](https://bio.libretexts.org/Bookshelves/Cell_and_Molecular_Biology/Book%3A_Cells_-_Molecules_and_Mechanisms_(Wong)/13%3A_Extracellular_Matrix_and_Cell_Adhesion/13.01%3A_Introduction_to_Extracellular_Matrix_and_Cell_Adhesion).
- [71] Wong, E. V. Cells - Molecules and Mechanisms/Extracellular Matrix and Cell Adhesion /Proteoglycans. *LibreTexts* [https://bio.libretexts.org/Bookshelves/Cell_and_Molecular_Biology/Book%3A_Cells_-_Molecules_and_Mechanisms_\(Wong\)/13%3A_Extracellular_Matrix_and_Cell_Adhesion/13.03%3A_Proteoglycans](https://bio.libretexts.org/Bookshelves/Cell_and_Molecular_Biology/Book%3A_Cells_-_Molecules_and_Mechanisms_(Wong)/13%3A_Extracellular_Matrix_and_Cell_Adhesion/13.03%3A_Proteoglycans).
- [72] Theocharis, A. D., Skandalis, S. S., Gialeli, C. & Karamanos, N. K. Extracellular matrix structure. *Adv. Drug Deliv. Rev.* **97**, 4–27 (2016).
- [73] Sherman, L. S. & Back, S. A. A ‘GAG’ reflex prevents repair of the damaged CNS. *Trends Neurosci.* **31**, 44–52 (2008).
- [74] Pomin, V. & Mulloy, B. Glycosaminoglycans and Proteoglycans. *Pharmaceuticals* **11**, 27 (2018).
- [75] Rowlands, D., Sugahara, K. & Kwok, J. Glycosaminoglycans and Glycomimetics in the Central Nervous System. *Molecules* **20**, 3527–3548 (2015).
- [76] Griffith, A. R., Rogers, C. J., Miller, G. M., Abrol, R., Hsieh-Wilson, L. C. & Goddard, W. A. Predicting glycosaminoglycan surface protein interactions and implications for studying axonal growth. *Proc. Natl. Acad. Sci.* **114**, 13697–13702 (2017).
- [77] Kim, S. Y., Jin, W., Sood, A., Montgomery, D. W., Grant, O. C., Fuster, M. M., Fu, L., Dordick, J. S., Woods, R. J., Zhang, F. & Linhardt, R. J. Characterization of heparin and severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) spike glycoprotein binding interactions. *Antiviral Res.* **181**, 104873 (2020).
- [78] Palaniappan, K. K. & Bertozzi, C. R. Chemical Glycoproteomics. *Chem. Rev.* **116**, 14277–14306 (2016).
- [79] Vallet, S. D., Clerc, O. & Ricard-Blum, S. Glycosaminoglycan–Protein Interactions: The First Draft of the Glycosaminoglycan Interactome. *J. Histochem. Cytochem.* (2020) doi:10.1369/0022155420946403.
- [80] Lichtenhaler, F. W. 100 Years “Schlüssel-Schloss-Prinzip”: What Made Emil Fischer Use this Analogy? *Angew. Chem. Int. Ed. Engl.* **33**, 2364–2374 (1994).

- [81] Koshland, D. E. Application of a theory of enzyme specificity to protein synthesis. *Proc. Natl. Acad. Sci.* **44**, 98–104 (1958).
- [82] Csermely, P., Palotai, R. & Nussinov, R. Induced fit, conformational selection and independent dynamic segments: an extended view of binding events. *Trends Biochem. Sci.* **35**, 539–546 (2010).
- [83] Cardin, A. D. & Weintraub, H. J. Molecular modeling of protein-glycosaminoglycan interactions. *Arterioscler. Off. J. Am. Heart Assoc. Inc* **9**, 21–32 (1989).
- [84] Torrent, M., Nogués, M. V., Andreu, D. & Boix, E. The “CPC Clip Motif”: A Conserved Structural Signature for Heparin-Binding Proteins. *PLoS ONE* **7**, e42692 (2012).
- [85] Fawcett, J. W., Oohashi, T. & Pizzorusso, T. The roles of perineuronal nets and the perinodal extracellular matrix in neuronal function. *Nat. Rev. Neurosci.* **20**, 451–465 (2019).
- [86] Beurdeley, M., Spatazza, J., Lee, H. H. C., Sugiyama, S., Bernard, C., Di Nardo, A. A., Hensch, T. K. & Prochiantz, A. Otx2 Binding to Perineuronal Nets Persistently Regulates Plasticity in the Mature Visual Cortex. *J. Neurosci.* **32**, 9429–9437 (2012).
- [87] Di Nardo, A. A., Fuchs, J., Joshi, R. L., Moya, K. L. & Prochiantz, A. The Physiology of Homeoprotein Transduction. *Physiol. Rev.* **98**, 1943–1982 (2018).
- [88] Sugiyama, S., Di Nardo, A. A., Aizawa, S., Matsuo, I., Volovitch, M., Prochiantz, A. & Hensch, T. K. Experience-Dependent Transfer of Otx2 Homeoprotein into the Visual Cortex Activates Postnatal Plasticity. *Cell* **134**, 508–520 (2008).
- [89] Spatazza, J., Lee, H. H. C., Di Nardo, A. A., Tibaldi, L., Joliot, A., Hensch, T. K. & Prochiantz, A. Choroid-Plexus-Derived Otx2 Homeoprotein Constrains Adult Cortical Plasticity. *Cell Rep.* **3**, 1815–1823 (2013).
- [90] Prochiantz, A. & Di Nardo, A. A. Homeoprotein Signaling in the Developing and Adult Nervous System. *Neuron* **85**, 911–925 (2015).
- [91] Billeter, M., Qian, Y. Q., Otting, G., Muller, M., Gehring, W. & Wuthrich, K. Determination of the nuclear magnetic resonance solution structure of an Antennapedia homeodomain-DNA complex. *J. Mol. Biol.* **234**, 1084–1097 (1993).
- [92] Joliot, A., Pernelle, C., Deagostini-Bazin, H. & Prochiantz, A. Antennapedia homeobox peptide regulates neural morphogenesis. *Proc. Natl. Acad. Sci.* **88**, 1864–1868 (1991).
- [93] Joliot, A. & Prochiantz, A. Transduction peptides: from technology to physiology. *Nat. Cell Biol.* **6**, 189–196 (2004).
- [94] Spatazza, J., Di Lullo, E., Joliot, A., Dupont, E., Moya, K. L. & Prochiantz, A. Homeoprotein Signaling in Development, Health, and Disease: A Shaking of Dogmas Offers Challenges and Promises from Bench to Bed. *Pharmacol. Rev.* **65**, 90–104 (2013).
- [95] Derossi, D., Joliot, A. H., Chassaing, G. & Prochiantz, A. The Third Helix of the Antennapedia Homeodomain Translocates through Biological membranes. *J Biol Chem* **269**, 10444–10450 (1994).
- [96] Miyata, S., Komatsu, Y., Yoshimura, Y., Taya, C. & Kitagawa, H. Persistent cortical plasticity by upregulation of chondroitin 6-sulfation. *Nat. Neurosci.* **15**, 414–422 (2012).
- [97] Hou, X., Yoshioka, N., Tsukano, H., Sakai, A., Miyata, S., Watanabe, Y., Yanagawa, Y., Sakimura, K., Takeuchi, K., Kitagawa, H., Hensch, T. K., Shibuki, K., Igarashi, M. & Sugiyama, S. Chondroitin Sulfate Is Required for Onset and Offset of Critical Period Plasticity in Visual Cortex. *Sci. Rep.* **7**, (2017).
- [98] UniProt. Homeobox protein OTX2. *UniProtKB - P32243 (OTX2_HUMAN)*
<https://www.uniprot.org/uniprot/P32243>.
- [99] Bernard, C. F. Otx2-glycosaminoglycan interaction to regulate visual cortex plasticity. *HAL*
<https://tel.archives-ouvertes.fr/tel-01084780> (2014).
- [100] Testa, D. Contrôler la plasticité du cortex cérébral adulte à travers l’action non-autonome de l’homéoprotéine Otx2. *HAL* <https://tel.archives-ouvertes.fr/tel-02479650> (2020).
- [101] Karamanos, N. K., Piperigkou, Z., Theocharis, A. D., Watanabe, H., Franchi, M., Baud, S., Brézillon, S., Götte, M., Passi, A., Vigetti, D., Ricard-Blum, S., Sanderson, R. D., Neill, T. & Iozzo, R. V. Proteoglycan Chemical Diversity Drives Multifunctional Cell Regulation and Therapeutics. *Chem. Rev.* **118**, 9152–9232 (2018).
- [102] Griffin, M. E. & Hsieh-Wilson, L. C. Synthetic probes of glycosaminoglycan function. *Curr. Opin. Chem. Biol.* **17**, 1014–1022 (2013).
- [103] Esko, J. D., H. Prestegard, J. & Linhardt, R. J. Chapter 38 : Proteins That Bind Sulfated Glycosaminoglycans. in *Essentials of Glycobiology - 3rd edition* (eds. Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Darvill, A. G., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L. & Seeberger, P. H.) (Cold Spring Harbor Laboratory Press, 2015).
- [104] Scott, R. A. & Panitch, A. Glycosaminoglycans in biomedicine: Glycosaminoglycans in biomedicine. *Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol.* **5**, 388–398 (2013).

- [105] Desai, U. R. The promise of sulfated synthetic small molecules as modulators of glycosaminoglycan function. *Future Med. Chem.* **5**, 1363–1366 (2013).
- [106] Paderi, J., Prestwich, G. D., Panitch, A., Boone, T. & Stuart, K. Glycan Therapeutics: Resurrecting an Almost Pharma-Forgotten Drug Class. *Adv. Ther.* **1**, 1800082 (2018).
- [107] Mohamed, S. & Coombe, D. Heparin Mimetics: Their Therapeutic Potential. *Pharmaceuticals* **10**, 78 (2017).
- [108] Ayerst, B. I., Merry, C. L. R. & Day, A. J. The Good the Bad and the Ugly of Glycosaminoglycans in Tissue Engineering Applications. *Pharmaceuticals* **10**, 54 (2017).
- [109] Zhang, X., Lin, L., Huang, H. & Linhardt, R. J. Chemoenzymatic Synthesis of Glycosaminoglycans. *Acc. Chem. Res.* **53**, 335–346 (2020).
- [110] Wei, M., Huang, L., Liu, Y., Jin, W., Yao, X., Rong, J., Bai, F., Song, X. & Wang, Z. Strategy for Isolation, Preparation, and Structural Analysis of Chondroitin Sulfate Oligosaccharides from Natural Sources. *Anal. Chem.* **92**, 11644–11653 (2020).
- [111] Tully, S. E., Mabon, R., Gama, C. I., Tsai, S. M., Liu, X. & Hsieh-Wilson, L. C. A Chondroitin Sulfate Small Molecule that Stimulates Neuronal Growth. *J. Am. Chem. Soc.* **126**, 7736–7737 (2004).
- [112] Solera, C., Macchione, G., Maza, S., Kayser, M. M., Corzana, F., de Paz, J. L. & Nieto, P. M. Chondroitin Sulfate Tetrasaccharides: Synthesis, Three-Dimensional Structure and Interaction with Midkine. *Chem. - Eur. J.* **22**, 2356–2369 (2016).
- [113] Friand, V., Haddad, O., Papy-Garcia, D., Hlawaty, H., Vassy, R., Hama-Kourbali, Y., Perret, G.-Y., Courty, J., Baleux, F., Oudar, O., Gattegno, L., Sutton, A. & Charneau, N. Glycosaminoglycan mimetics inhibit SDF-1/CXCL12-mediated migration and invasion of human hepatoma cells. *Glycobiology* **19**, 1511–1524 (2009).
- [114] Maza, S., Mar Kayser, M., Macchione, G., López-Prados, J., Angulo, J., de Paz, J. L. & Nieto, P. M. Synthesis of chondroitin/dermatan sulfate-like oligosaccharides and evaluation of their protein affinity by fluorescence polarization. *Org. Biomol. Chem.* **11**, 3510 (2013).
- [115] Place, L. W., Sekyi, M. & Kipper, M. J. Aggrecan-Mimetic, Glycosaminoglycan-Containing Nanoparticles for Growth Factor Stabilization and Delivery. *Biomacromolecules* **15**, 680–689 (2014).
- [116] Daley, E. L. H., Coleman, R. M. & Stegemann, J. P. Biomimetic microbeads containing a chondroitin sulfate/chitosan polyelectrolyte complex for cell-based cartilage therapy. *J. Mater. Chem. B* **3**, 7920–7929 (2015).
- [117] Liu, Y., Wang, S., Sun, D., Liu, Y., Liu, Y., Wang, Y., Liu, C., Wu, H., Lv, Y., Ren, Y., Guo, X., Sun, G. & Ma, X. Development of a Biomimetic Chondroitin Sulfate-modified Hydrogel to Enhance the Metastasis of Tumor Cells. *Sci. Rep.* **6**, (2016).
- [118] Paluck, S. J., Nguyen, T. H. & Maynard, H. D. Heparin-Mimicking Polymers: Synthesis and Biological Applications. *Biomacromolecules* **17**, 3417–3440 (2016).
- [119] Mhanna, R., Becher, J., Schnabelrauch, M., Reis, R. L. & Pashkuleva, I. Sulfated Alginate as a Mimic of Sulfated Glycosaminoglycans: Binding of Growth Factors and Effect on Stem Cell Behavior. *Adv. Biosyst.* **1**, 1700043 (2017).
- [120] Zong, C., Venot, A., Li, X., Lu, W., Xiao, W., Wilkes, J.-S. L., Salanga, C. L., Handel, T. M., Wang, L., Wolfert, M. A. & Boons, G.-J. Heparan Sulfate Microarray Reveals That Heparan Sulfate–Protein Binding Exhibits Different Ligand Requirements. *J. Am. Chem. Soc.* **139**, 9534–9543 (2017).
- [121] Lee, E. C., Nguyen, C. T. H., Strounina, E., Davis-Poynter, N. & Ross, B. P. Structure–Activity Relationships of GAG Mimetic-Functionalized Mesoporous Silica Nanoparticles and Evaluation of Acyclovir-Loaded Antiviral Nanoparticles with Dual Mechanisms of Action. *ACS Omega* **3**, 1689–1699 (2018).
- [122] Al Nahain, A., Ignjatovic, V., Monagle, P., Tsanaktsidis, J. & Ferro, V. Heparin mimetics with anticoagulant activity. *Med. Res. Rev.* **38**, 1582–1613 (2018).
- [123] Zhang, G.-L. & Ye, X.-S. Synthetic Glycans and Glycomimetics: A Promising Alternative to Natural Polysaccharides. *Chem. - Eur. J.* **24**, 6696–6704 (2018).
- [124] Liu, Q., Chen, G. & Chen, H. Chemical synthesis of glycosaminoglycan-mimetic polymers. *Polym. Chem.* **10**, 164–171 (2019).
- [125] Morla, S. Glycosaminoglycans and Glycosaminoglycan Mimetics in Cancer and Inflammation. *Int. J. Mol. Sci.* **20**, 1963 (2019).
- [126] Afosah, D. K. & Al-Horani, R. A. Sulfated Non-Saccharide Glycosaminoglycan Mimetics as Novel Drug Discovery Platform for Various Pathologies. *Curr. Med. Chem.* **27**, 3412–3447 (2020).
- [127] Yang, C., Gao, L., Shao, M., Cai, C., Wang, L., Chen, Y., Li, J., Fan, F., Han, Y., Liu, M., Linhardt, R. J. & Yu, G. End-functionalised glycopolymers as glycosaminoglycan mimetics inhibit HeLa cell proliferation. *Polym. Chem.* **11**, 4714–4722 (2020).

- [128] Jouault, T., Fradin, C., Dzierszinski, F., Borg-Von-Zepelin, M., Tomavo, S., Corman, R., Trinel, P.-A., Kerckaert, J.-P. & Poulain, D. Peptides that mimic *Candida albicans*-derived β -1,2-linked mannosides. 9.
- [129] Ziebell, M. R., Zhao, Z.-G., Luo, B., Luo, Y., Turley, E. A. & Prestwich, G. D. Peptides that mimic glycosaminoglycans: high-affinity ligands for a hyaluronan binding domain. *Chem. Biol.* **8**, 1081–1094 (2001).
- [130] Tansik, G., Kilic, E., Beter, M., Demiralp, B., Kiziltas Sendur, G., Can, N., Ozkan, H., Ergul, E., Guler, M. O. & Tekinay, A. B. A glycosaminoglycan mimetic peptide nanofiber gel as an osteoinductive scaffold. *Biomater. Sci.* **4**, 1328–1339 (2016).
- [131] Yasa, O., Uysal, O., Ekiz, M. S., Guler, M. O. & Tekinay, A. B. Presentation of functional groups on self-assembled supramolecular peptide nanofibers mimicking glycosaminoglycans for directed mesenchymal stem cell differentiation. *J. Mater. Chem. B* **5**, 4890–4900 (2017).
- [132] Vázquez-Campos, S., St. Hilaire, P. M., Damgaard, D. & Meldal, M. GAG Mimetic Libraries: Sulphated Peptide as Heparin-like Glycosaminoglycan Mimics in Their Interaction with FGF-1. *QSAR Comb. Sci.* **24**, 923–942 (2005).
- [133] Maynard, H. D. & Hubbell, J. A. Discovery of a sulfated tetrapeptide that binds to vascular endothelial growth factor. *Acta Biomater.* **1**, 451–459 (2005).
- [134] Kim, S. H. & Kiick, K. L. Heparin-mimetic sulfated peptides with modulated affinities for heparin-binding peptides and growth factors. *Peptides* **28**, 2125–2136 (2007).
- [135] Mammadov, R., Mammadov, B., Guler, M. O. & Tekinay, A. B. Growth Factor Binding on Heparin Mimetic Peptide Nanofibers. *Biomacromolecules* **13**, 3311–3319 (2012).
- [136] Vázquez-Campos, S., Miranda, L. P. & Meldal, M. Preparation of novel O-sulfated amino acid building blocks with improved acid stability for Fmoc-based solid-phase peptide synthesis. *J. Chem. Soc. Perkin 1* **24**, 682–686 (2002).
- [137] Sakakibara, S., Shimonishi, Y., Kishida, Y., Okada, M. & Sugihara, H. Use of Anhydrous Hydrogen Fluoride in Peptide Synthesis. I. Behavior of Various Protective Groups in Anhydrous Hydrogen Fluoride. *Bull. Chem. Soc. Jpn.* **40**, 2164–2167 (1967).
- [138] Carpino, L. A. & Han, G. Y. 9-Fluorenylmethoxycarbonyl amino-protecting group. *J. Org. Chem.* **37**, 3404–3409 (1972).
- [139] Carpino, L. A., Sadat-Aalae, D., Chao, H. G. & DeSelms, R. H. (9-Fluorenylmethyl)oxy]carbonyl (Fmoc) amino acid fluorides. Convenient new peptide coupling reagents applicable to the Fmoc/*tert*-butyl strategy for solution and solid-phase syntheses. *J Am Chem Soc* **112**, 9651–9652 (1990).
- [140] Behrendt, R., White, P. & Offer, J. Advances in Fmoc solid-phase peptide synthesis. *J. Pept. Sci.* **22**, 4–27 (2016).
- [141] Mitchell, A. R. Invited revue - Bruce Merrifield and solid-phase peptide synthesis: A historical assessment. *Pept. Sci.* **90**, 175–184 (2008).
- [142] Vaino, A. R. & Janda, K. D. Solid-Phase Organic Synthesis: A Critical Understanding of the Resin [†]. *J. Comb. Chem.* **2**, 579–596 (2000).
- [143] Sheehan, J. C. & Hess, G. P. A New Method of Forming Peptide Bonds. *J. Am. Chem. Soc.* **77**, 1067–1068 (1955).
- [144] König, W. & Geiger, R. A new method for synthesis of peptides: activation of the carboxyl group with dicyclohexylcarbodiimide using 1-hydroxybenzotriazoles as additives. *Chem. Ber.* **103**, 788–798 (1970).
- [145] Castro, B., Dormoy, J.-R., Dourtoglou, B., Evin, G., Selve, C. & Ziegler, J.-C. Peptide coupling reagents. A novel, cheaper preparation of benzotriazolylxytris[*dimethylamine*]phosphonium hexafluorophosphate (BOP reagent). *Synthesis* **1976**, 751–752 (1976).
- [146] Dourtoglou, V., Ziegler, J.-C. & Gross, B. L'hexafluorophosphate de O-benzotriazole-N,N-tetraméthyluronium : Un réactif de couplage peptidique nouveau et efficace. *Tetrahedron Lett.* **15**, 1269–1272 (1978).
- [147] Abdelmoty, I., Albericio, F., Carpino, L. A., Foxman, B. M. & Kates, S. A. Structural studies of reagents for peptide bond formation: Crystal and molecular structures of HBTU and HATU. *Let. Pept. Sci.* **1**, 57–67 (1994).
- [148] Montalbetti, C. A. G. N. & Falque, V. Amide bond formation and peptide coupling. *Tetrahedron* **61**, 10827–10852 (2005).
- [149] El-Faham, A. & Albericio, F. Peptide Coupling Reagents, More than a Letter Soup. *Chem. Rev.* **111**, 6557–6602 (2011).
- [150] Albericio, F. & El-Faham, A. Choosing the Right Coupling Reagent for Peptides: A Twenty-Five-Year Journey. *Org. Process Res. Dev.* **22**, 760–772 (2018).

- [151] Hollanders, K., Maes, B. & Ballet, S. A New Wave of Amide Bond Formations for Peptide Synthesis. *Synthesis* **51**, 2261–2277 (2019).
- [152] Anuradha, M. V. & Ravindranath, B. Ultrasound in peptide synthesis. 3 Zinc-salt assisted anchoring of carboxylic acids to Merrifield resin. *Tetrahedron* **51**, 5671–5674 (1995).
- [153] Nissen, F., Kraft, T. E., Ruppert, T., Eisenhut, M., Haberkorn, U. & Mier, W. Hot or not—the influence of elevated temperature and microwave irradiation on the solid phase synthesis of an affibody. *Tetrahedron Lett.* **51**, 6216–6219 (2010).
- [154] Pedersen, S. L., Tofteng, A. P., Malik, L. & Jensen, K. J. Microwave heating in solid-phase peptide synthesis. *Chem Soc Rev* **41**, 1826–1844 (2012).
- [155] Merlino, F., Tomassi, S., Yousif, A. M., Messere, A., Marinelli, L., Grieco, P., Novellino, E., Cosconati, S. & Di Maro, S. Boosting Fmoc Solid-Phase Peptide Synthesis by Ultrasonication. *Org. Lett.* **21**, 6378–6382 (2019).
- [156] Mant, C. T., Chen, Y., Yan, Z., Popa, T. V., Kovacs, J. M., Mills, J. B., Tripet, B. P. & Hodges, R. S. HPLC Analysis and Purification of Peptides. *Methods Mol. Biol.* **386**, (2007).
- [157] Henninot, A., Collins, J. C. & Nuss, J. M. The Current State of Peptide Drug Discovery: Back to the Future? *J. Med. Chem.* **61**, 1382–1414 (2018).
- [158] Isidro-Llobet, A., Kenworthy, M. N., Mukherjee, S., Kopach, M. E., Wegner, K., Gallou, F., Smith, A. G. & Roschangar, F. Sustainability Challenges in Peptide Synthesis and Purification: From R&D to Production. *J. Org. Chem.* **84**, 4615–4628 (2019).
- [159] Al Musaimi, O., de la Torre, B. G. & Albericio, F. Greening Fmoc/tBu solid-phase peptide synthesis. *Green Chem.* **22**, 996–1018 (2020).
- [160] Agouridas, V., El Mahdi, O., Cargoët, M. & Melnyk, O. A statistical view of protein chemical synthesis using NCL and extended methodologies. *Bioorg. Med. Chem.* **25**, 4938–4945 (2017).
- [161] Nguyen, G. K. T., Wang, S., Qiu, Y., Hemu, X., Lian, Y. & Tam, J. P. Butelase 1 is an Asx-specific ligase enabling peptide macrocyclization and synthesis. *Nat. Chem. Biol.* **10**, 732–738 (2014).
- [162] Thompson, R. E. & Muir, T. W. Chemoenzymatic Semisynthesis of Proteins. *Chem. Rev.* **120**, 3051–3126 (2020).
- [163] Vlachy, N., Jagoda-Cwiklik, B., Vácha, R., Touraud, D., Jungwirth, P. & Kunz, W. Hofmeister series and specific interactions of charged headgroups with aqueous ions. *Adv. Colloid Interface Sci.* **146**, 42–47 (2009).
- [164] Skvarnavičius, G., Dvareckas, D., Matulis, D. & Petrauskas, V. Thermodynamics of Interactions Between Charged Surfactants and Ionic Poly(amino acids) by Isothermal Titration Calorimetry. *ACS Omega* **4**, 17527–17535 (2019).
- [165] Bodanszky, M. Side Reactions in Peptide Synthesis. in *Principles of Peptide Synthesis* vol. 16 158–201 (Springer, 1984).
- [166] Bachem. Tips and tricks for solid phase peptide synthesis from the experts at Bachem. *Solid Phase Peptide Synthesis*
https://www.bachem.com/fileadmin/user_upload/pdf/Catalogs_Brochures/Solid_Phase_Peptide_Synthesis.pdf.
- [167] Tam, J. P., Wong, T.-W., Riemen, M. W., Tjoeng, F.-S. & Merrifield, R. B. Cyclohexyl ester as a new protecting group for aspartyl peptides to minimize aspartimide formation in acidic and basic treatments. *Tetrahedron Lett.* **20**, 4033–4036 (1979).
- [168] Michels, T., Dölling, R., Haberkorn, U. & Mier, W. Acid-Mediated Prevention of Aspartimide Formation in Solid Phase Peptide Synthesis. *Org. Lett.* **14**, 5218–5221 (2012).
- [169] Neumann, K., Farnung, J., Baldauf, S. & Bode, J. W. Prevention of aspartimide formation during peptide synthesis using cyanosulfurylides as carboxylic acid-protecting groups. *Nat. Commun.* **11**, (2020).
- [170] Wade, N. *La course au Nobel*. (S. Messinger, 1981).
- [171] Dimarchi, R. D., Tam, J. P., Kent, S. B. H. & Merrifield, R. B. Weak acid-catalyzed pyrrolidone carboxylic acid formation from glutamine during solid phase peptide synthesis: Minimization by rapid coupling. *Int. J. Pept. Protein Res.* **19**, 88–93 (2009).
- [172] Dadová, J., Galan, S. R. & Davis, B. G. Synthesis of modified proteins via functionalization of dehydroalanine. *Curr. Opin. Chem. Biol.* **46**, 71–81 (2018).
- [173] Hirs, C. H. W. The oxidation of ribonuclease with performic acid. *J Biol Chem* **219**, 611–621 (1956).
- [174] Du, P., Liu, W., Cao, H., Zhao, H. & Huang, C.-H. Oxidation of amino acids by peracetic acid: Reaction kinetics, pathways and theoretical calculations. *Water Res. X* **1**, 100002 (2018).
- [175] Medzihradzsky, K. F., Darula, Z., Perlson, E., Fainzilber, M., Chalkley, R. J., Ball, H., Greenbaum, D., Bogoy, M., Tyson, D. R., Bradshaw, R. A. & Burlingame, A. L. O-Sulfonation of Serine and Threonine: Mass

- Spectrometric Detection and Characterization of a New Posttranslational Modification in Diverse Proteins Throughout the Eukaryotes. *Mol. Cell. Proteomics* **3**, 429–440 (2004).
- [176] Nick Pace, C. & Martin Scholtz, J. A Helix Propensity Scale Based on Experimental Studies of Peptides and Proteins. *Biophys. J.* **75**, 422–427 (1998).
- [177] Ohara, K., Jacquinet, J.-C., Jouanneau, D., Helbert, W., Smietana, M. & Vasseur, J.-J. Matrix-assisted laser desorption/ionization mass spectrometric analysis of polysulfated-derived oligosaccharides using pyrenemethylguanidine. *J. Am. Soc. Mass Spectrom.* **20**, 131–137 (2009).
- [178] Anthis, N. J. & Clore, G. M. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm: Sequence-Specific Protein Concentration at 205 nm. *Protein Sci.* **22**, 851–858 (2013).
- [179] Saidel, L. J., Goldfarb, A. R. & Waldman, Sheldon. The absorption spectra of amino acids in the region two hundred to two hundred and thirty millimicrons. *J. Biol. Chem.* **197**, 285–291 (1952).
- [180] Green, N. M. A Spectrophotometric Assay for Avidin and Biotin Based on Binding of Dyes by Avidin. *Biochem. J.* **94**, 23C (1965).
- [181] Lavielle, S., Chassaing, G. & Marquet, A. Avidin binding of biotinylated analogues of substance P. *Biochim. Biophys. Acta BBA - Gen. Subj.* **759**, 270–277 (1983).
- [182] Carlini, A. S., Adamiak, L. & Gianneschi, N. C. Biosynthetic Polymers as Functional Materials. *Macromolecules* **49**, 4379–4394 (2016).
- [183] Wang, K. *et al.* Advanced functional polymer materials. *Mater. Chem. Front.* **4**, 1803–1915 (2020).
- [184] Mazzoleni, A. Peptides glucosylés dans les maladies auto-immunes : stratégies de synthèse et leur application à la détection et à la capture d'anticorps. (Thèse PSL, 2020).
- [185] Carlier, L., Balayssac, S., Cantrelle, F.-X., Khemtémourian, L., Chassaing, G., Joliot, A. & Lequin, O. Investigation of Homeodomain Membrane Translocation Properties: Insights from the Structure Determination of Engrailed-2 Homeodomain in Aqueous and Membrane-Mimetic Environments. *Biophys. J.* **105**, 667–678 (2013).
- [186] Cardon, S. Étude quantitative du rôle spécifique de glycosaminoglycanes dans le mécanisme d'internalisation de l'homéoprotéine engrailed 2. (Thèse PSL, 2017).
- [187] Bolanos-Garcia, V. M. & Davies, O. R. Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *Biochim. Biophys. Acta* **1760**, 1304–1313 (2006).
- [188] Havel, T. F. & Wüthrich, K. An evaluation of the combined use of nuclear magnetic resonance and distance geometry for the determination of protein conformations in solution. *J. Mol. Biol.* **182**, 281–294 (1985).
- [189] Williamson, M. P., Havel, T. F. & Wüthrich, K. Solution conformation of proteinase inhibitor IIA from bull seminal plasma by ¹H nuclear magnetic resonance and distance geometry. *J. Mol. Biol.* **182**, 295–315 (1985).
- [190] RCSB PDB. Growth of Structures from NMR Experiments Released per Year. *PDB Statistics* <https://www.rcsb.org/stats/growth/growth-nmr>.
- [191] Cavanagh, J., Fairbrother, W. J., Palmer III, A. G. & Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice.* (1995).
- [192] Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158 (1999).
- [193] Kanelis, V., Forman-Kay, J. D. & Kay, L. E. Multidimensional NMR Methods for Protein Structure Determination. *IUBMB Life* **52**, 291–302 (2001).
- [194] Poulsen, F. M. A brief introduction to NMR spectroscopy of proteins. <https://users.cs.duke.edu/~brd/Teaching/Bio/asmb/Papers/Intro-reviews/flemming.pdf> (2002).
- [195] Charlier, C., Bouvignies, G., Pelulessy, P., Walrant, A., Marquant, R., Kozlov, M., De Ioannes, P., Bolik-Coulon, N., Sagan, S., Cortes, P., Aggarwal, A. K., Carlier, L. & Ferrage, F. Structure and Dynamics of an Intrinsically Disordered Protein Region That Partially Folds upon Binding by Chemical-Exchange NMR. *J. Am. Chem. Soc.* **139**, 12219–12227 (2017).
- [196] Williamson, M. P. Using chemical shift perturbation to characterise ligand binding. *Prog. Nucl. Magn. Reson. Spectrosc.* **73**, 1–16 (2013).
- [197] Fielding, L. NMR Methods for the Determination of Protein-Ligand Dissociation Constants. *Curr. Top. Med. Chem.* **3**, 39–53 (2003).
- [198] Wishart, D. S. & Sykes, B. D. The ¹³C Chemical-Shift Index: A simple method for the identification of protein secondary structure using ¹³C chemical-shift data. *J. Biomol. NMR* **4**, 171–180 (1993).

- [199] Wishart, D. S., Sykes, B. D. & Richards, F. M. The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **31**, 1647–1651 (1992).
- [200] Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S. & Sykes, B. D. ¹H, ¹³C and random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* **5**, 67–81 (1995).
- [201] Ohnishi, S., Saito, K., Koshiba, S., Inoue, M., Kigawa, T. & Yokoyama, S. Solution structure of the homeobox domain of Homeobox protein OTX2. *RCSB PDB* <https://www.rcsb.org/structure/2dms> (2006).
- [202] Erdozain, A. M., De Gois, S., Bernard, V., Gorgievski, V., Pietrancosta, N., Dumas, S., Macedo, C. E., Vanhoutte, P., Ortega, J. E., Meana, J. J., Tzavara, E. T., Vialou, V. & Giros, B. Structural and Functional Characterization of the Interaction of Snapin with the Dopamine Transporter: Differential Modulation of Psychostimulant Actions. *Neuropsychopharmacology* **43**, 1041–1051 (2018).
- [203] Chakrabartty, A., Kortemme, T. & Baldwin, R. L. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Sci.* **3**, 843–852 (1994).
- [204] Lyu, P. C., Sherman, J. C., Chen, A. & Kallenbach, N. R. α -Helix stabilization by natural and unnatural amino acids with alkyl side chains. *Proc Natl Acad Sci USA* **88**, 4 (1991).
- [205] Lyu, P., Liff, M., Marky, L. & Kallenbach, N. Side chain contributions to the stability of α -helical structure in peptides. *Science* **250**, 669–673 (1990).
- [206] Dai, J., Zhang, Y., Wang, J., Li, X., Lu, Z., Cai, Y. & Qian, X. Identification of degradation products formed during performic oxidation of peptides and proteins by high-performance liquid chromatography with matrix-assisted laser desorption/ionization and tandem mass spectrometry. *Rapid Commun. Mass Spectrom.* **19**, 1130–1138 (2005).
- [207] How Does Surface Plasmon Resonance Work? *Biosensing Instrument* <https://biosensingusa.com/technologies/surface-plasmon-resonance/surface-plasmon-resonance-work/>.
- [208] GE Healthcare. Biacore Sensor Surface Handbook. <https://timothyspringer.org/files/tas/files/biacore3000-sensorsurface.pdf>.
- [209] Hill equation (biochemistry). *Wikipedia* [https://en.wikipedia.org/w/index.php?title=Hill_equation_\(biochemistry\)&oldid=1041741984](https://en.wikipedia.org/w/index.php?title=Hill_equation_(biochemistry)&oldid=1041741984) (2021).
- [210] Karoyan, P. & Chassaing, G. New strategy for the synthesis of 3-substituted prolines. *Tetrahedron Lett.* **38**, 85–88 (1997).
- [211] Karoyan, P., Quancard, J., Vaissermann, J. & Chassaing, G. Amino–Zinc–Enolate Carbometalation Reactions: Application to Ring Closure of Terminally Substituted Olefin for the Asymmetric Synthesis of cis- and trans-3-Prolinoleucine. *J. Org. Chem.* **68**, 2256–2265 (2003).
- [212] Quancard, J., Karoyan, P., Sagan, S., Convert, O., Lavielle, S., Chassaing, G. & Lequin, O. Characterization of the bioactive conformation of the C-terminal tripeptide Gly-Leu-Met-NH₂ of substance P using [3-prolinoleucine¹⁰]SP analogues. *Eur. J. Biochem.* **270**, 2869–2878 (2003).
- [213] Sagan, S., Quancard, J., Lequin, O., Karoyan, P., Chassaing, G. & Lavielle, S. Conformational analysis of the C-terminal Gly-Leu-Met-NH₂ tripeptide of substance P bound to the NK-1 receptor. *Chem. Biol.* **12**, 555–565 (2005).
- [214] Marley, J., Lu, M. & Bracken, C. A method for efficient isotopic labeling of recombinant proteins. *J. Biomol. NMR* **20**, 71–75 (2001).
- [215] Schafmeister, C. E., Po, J. & Verdine, G. L. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *J. Am. Chem. Soc.* **122**, 5891–5892 (2000).
- [216] Verdine, G. L. & Hilinski, G. J. Stapled Peptides for Intracellular Drug Targets. in *Methods in Enzymology* vol. 503 3–33 (Elsevier, 2012).
- [217] Sadler, K. & Tam, J. P. Peptide dendrimers: applications and synthesis. *Rev. Mol. Biotechnol.* **90**, 195–229 (2002).
- [218] Posnett, D. N., McGrath, H. & Tam, J. P. A novel method for producing anti-peptide antibodies. Production of site-specific antibodies to the T cell antigen receptor beta-chain. *J. Biol. Chem.* **263**, 1719–1725 (1988).
- [219] Joshi, V. G., Dighe, V. D., Thakuria, D., Malik, Y. S. & Kumar, S. Multiple antigenic peptide (MAP): a synthetic peptide dendrimer for diagnostic, antiviral and vaccine strategies for emerging and re-emerging viral diseases. *Indian J. Virol.* **24**, 312–320 (2013).
- [220] van DUUREN, B. L. Solvent Effects in the Fluorescence of Indole and Substituted Indoles ¹. *J. Org. Chem.* **26**, 2954–2960 (1961).
- [221] Vivian, J. T. & Callis, P. R. Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophys. J.* **80**, 2093–2109 (2001).

- [222] Nanda, V. & Brand, L. Aromatic interactions in homeodomains contribute to the low quantum yield of a conserved, buried tryptophan. *Proteins Struct. Funct. Bioinforma.* **40**, 112–125 (2000).
- [223] BioRender Templates. <https://app.biorender.com/biorender-templates/figures/>.
- [224] Molecular Devices. Establishing and optimizing a fluorescence polarization assay (Technical note). (2020).

Mimes peptidiques de glycosaminoglycanes : synthèses, structures et activités biologiques

Les Glycosaminoglycanes (GAGs) sont des polysaccharides linéaires, plus ou moins sulfatés, présents à la surface des cellules et dans la matrice extracellulaire. Ils interagissent avec un grand nombre de protéines (cytokines, facteurs de croissance, enzymes ...) et sont impliqués dans divers processus physiologiques et pathologiques.

Malgré ces rôles clés, leurs fonctions au niveau moléculaire demeurent encore mal comprises, pour diverses raisons :

- Leurs très grandes diversités de familles (héparine, chondroïtine, ...), de degré de sulfatation et de longueur de chaînes
- Leurs modifications rapides *in vivo* (épimérisation, *N* et *O*-sulfatation)
- Leurs difficultés de purification et d'analyses...

Travailler sur des GAGs ou analogues purs et chimiquement définis n'est donc pas chose aisée, même en employant des techniques classiques de synthèse (chimique ou enzymatique). Par exemple, cinquante étapes sont nécessaires pour synthétiser un hexasaccharide - taille qui est le plus souvent minimale dans une interaction. Dans le cadre de mon travail de thèse, nous avons conçu et synthétisé des molécules mimes de GAGs plus simples : des peptides anioniques de 12 à 24 acides aminés, portant des motifs répétés de type (EC'AC')_n et présentant les fonctions caractéristiques des GAGs naturels, à savoir les fonctions acides, apportées par l'acide Glutamique, et sulfoniques, obtenues par oxydation des Cystéines. Une première génération de ces sulfopeptides a été étudiée, au Collège de France, comme inhibiteurs de l'interaction d'une protéine à homéodomaine (Otx2) avec le réseau de GAGs présents autour d'un type cellulaire cible de cette protéine. Le peptide comportant six motifs « EC'AC' » a montré une activité *in vivo* très prometteuse. Une étude structurale par RMN a ensuite été réalisée pour analyser les interactions de ces sulfopeptides avec la protéine Otx2 produite avec un double marquage ¹³C, ¹⁵N. Cette étude a confirmé que l'interaction se faisait au niveau du domaine de liaison aux GAGs de la protéine. Cette interaction affecte aussi l'hélice 3 de l'homéodomaine, à l'origine de l'internalisation d'Otx2 dans les cellules. Enfin nous avons complété cette étude structurale par une approche de type Biacore, pour développer une stratégie d'analyse plus rapide de l'activité de nos analogues et nous permettre de développer une étude structure-activité efficace.

Mots clés : glycosaminoglycanes, sulfopeptides, Otx2, homéodomaine, mimes, RMN, Biacore

Peptide mimics of glycosaminoglycans: synthesis, structures and biological activities

Glycosaminoglycans (GAGs) are linear polysaccharides, more or less sulfated, present on the surface of cells and in the extracellular matrix. They interact with a large number of proteins (cytokines, growth factors, enzymes, etc.) and are involved in various physiological and pathological processes.

Despite these key roles, their functions at the molecular level are poorly understood, for various reasons:

- Their very large diversity of families (heparin, chondroitin, ...), degree of sulfation and length of chains
- Their rapid modifications *in vivo* (epimerization, *N* and *O*-sulfation)
- Their difficulties of purification and analyzes ...

Working on pure and chemically defined GAGs or analogues is therefore not easy, even when using conventional synthesis techniques (chemical or enzymatic). For example, fifty steps are necessary to synthesise an hexasaccharide - a size which is most often minimal for interaction.

As part of my thesis work, we designed and synthesized simpler GAGs mimetic molecules: anionic peptides of 12 to 24 amino acids, carrying repeated motifs (EC'AC')_n and exhibiting the characteristic functions of natural GAGs, namely the acid functions, provided by glutamic acid, and sulfonic functions, obtained by oxidation of cysteines. A first generation of these sulfopeptides was studied, at the Collège de France, as inhibitors of the interaction of a homeodomain protein (Otx2) with the network of GAGs present around a target cell type of this protein. The peptide comprising six "EC'AC" motifs has shown very promising *in vivo* activity. A structural study by NMR was then carried out to analyze the interactions of these sulfopeptides with the Otx2 protein produced with a double ¹³C, ¹⁵N labeling. This study confirmed that the interaction takes place at the GAG binding domain of the protein. This interaction also affects the homeodomain helix 3, which is responsible for the internalization of Otx2 into cells. Finally, we completed this structural study with a Biacore-type approach, to develop a strategy for faster analysis of the activity of our analogues and allow us to develop an efficient structure-activity study.

Keywords: glycosaminoglycans, sulfopeptides, Otx2, homeodomain, mimics, NMR, Biacore