HAL
open science

Inhibition d'interactions protéine-protéine par des foldamères mixtes oligoamide/oligourée

Léonie Cussol

To cite this version:

Léonie Cussol. Inhibition d'interactions protéine-protéine par des foldamères mixtes oligoamide/oligourée. Chimie organique. Université de Bordeaux, 2018. Français. NNT : 2018BORD0394 . tel-03196598

HAL Id: tel-03196598

https://theses.hal.science/tel-03196598v1

Submitted on 13 Apr 2021

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.

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

SPÉCIALITÉ CHIMIE ORGANIQUE

Par Léonie CUSSOL

INHIBITION D'INTERACTIONS PROTEINE-PROTEINE PAR DES FOLDAMERES MIXTES OLIGOAMIDE/OLIGOUREE

Sous la direction de : Dr. Gilles GUICHARD

Soutenue le 18 Décembre 2018

Membres du jury :

Mme VOISIN-CHIRET, Anne-Sophie	Professeur des Universités, Université de Caen Normandie	Président
M. BOTURYN, Didier	Directeur de recherche CNRS, Université de Grenoble-Alpes	Rapporteur
Mme FAURE, Sophie	Chargé de recherche CNRS, Université de Clermont-Ferrand	Rapporteur
Mme ROCHEL-GUIBERTEAU, Natacha	Directeur de recherche CNRS, Université de Strasbourg	Examinateur
Mme FISCHER, Lucile	Chargé de recherche CNRS, Université de Bordeaux	Examinateur
Mme DOLAIN, Christel	Maître de conférences, Université de Bordeaux	Examinateur
M. GOUDREAU, Sébastien	Directeur scientifique d'UREkA	Membre Invité
M. GUICHARD, Gilles	Directeur de recherche CNRS, Université de Bordeaux	Directeur

Remerciements

Rédiger une thèse est loin d'être un travail solitaire...
Ainsi, je tiens à remercier toutes les personnes qui m'ont accompagnée.

Je remercie Madame le Docteur Sophie Lecomte, directrice du CBMN, et Monsieur le Docteur Jean-Louis Mergny, directeur de l'IECB, de m'avoir permis de réaliser mes travaux de recherches doctorales dans leurs Instituts.

J'adresse mes vifs remerciements à Monsieur le Docteur Didier Boturyn, directeur de recherche CNRS, et Madame le Docteur Sophie Faure, chargée de recherche CNRS, pour avoir accepté d'évaluer ce travail en qualité de rapporteurs, et à Madame le Docteur Lucile Fischer, chargée de recherche CNRS, à Monsieur le Docteur Sébastien Goudreau, directeur scientifique d'UREKA, à Madame le Docteur Natacha Rochel-Guiberteau, directrice de recherche CNRS, et à Madame le Docteur Anne-Sophie Voisin-Chiret, professeur d'université, pour leur contribution en tant que membres du jury.

Je n'aurais jamais pu réaliser ce travail de recherche sans mon directeur de thèse, Monsieur le Docteur Gilles Guichard, principal guide de mon parcours et de mes recherches. Je lui exprime toute ma gratitude pour m'avoir proposé cette thèse suite à mon stage de M2 et pour m'avoir fait confiance pendant ces trois années de doctorat. Je le remercie pour tous les conseils avisés qu'il m'a prodigués. Je lui sais gré d'avoir fait preuve d'exigence, même quand il m'a semblé difficile d'y répondre : cette expérience a sans nul doute été pour moi formatrice, enrichissante.

J'exprime ma profonde reconnaissance à Madame le Docteur Christel Dolain, maître de conférences, pour son encadrement durant ces trois années, pour ces bons conseils et sa bienveillance.

Je tiens aussi à remercier l'ensemble du personnel des plateformes de l'IECB avec qui j'ai eu la chance de travailler : Axelle et Estelle pour la RMN, Frédéric et Loïc pour la spectrométrie de masse, et Brice pour la résolution de certaines structures RX.

J'ai aussi une pensée pour toutes les personnes rencontrées à l'IECB, et notamment pour Mariline, personnalité souriante de l'accueil, toujours disponible, toujours à l'écoute.

Comme la vie d'une thésarde est aussi riche en rencontres, j'ai une pensée particulière pour tous ces anciens et nouveaux compagnons de thèse qui, de près ou de loin, m'ont témoigné leur attention et leur soutien, et avec qui j'ai partagé de très bons moments de travail ou de détente. Merci à Adrien, Alba, Albano, Alice, Ana, Antoine, Antoine A., Antoine M., Antoine S., Arthur, Arup, Brigitte, Camille F., Camille P., Camille Z., Clémence, Clément, Christophe, Daniela, Guillaume C., Guillaume N, Ivan, Joan, Joséphine, Julen, Juliette, Johanne, Laura A., Laura B., Ly, Nassima, Mégane, Morgane, Noureddine, Rich, Stefanie, Stefanio, Stéphanie, Sonia, SunHyung, Thierry, Thomas, Valerio, Victor, et Vincent. Merci à tous ces compagnons de route.

Je remercie vivement Claire, Jérémie et Laura pour leur disponibilité et leur soutien indéfectible lors de la rédaction, et pour leur précieuse contribution à la relecture.

Cette thèse n'aurait pas été la même sans la présence de Diane. Ses conseils, son énergie et son humour m'ont été précieux, secourables, et ont sans aucun doute contribué à cette belle amitié qui nous lie.

Je tiens également à remercier mes amis Cindy, Gary, Julie, Léa, Mathieu et Sébastien pour leurs diverses attentions.

Enfin, j'exprime ma profonde reconnaissance à ma famille pour son soutien inconditionnel et ses multiples encouragements. Mes derniers remerciements s'adressent plus particulièrement à ma Maman et à Thibault.

Abréviations

2D	deux dimensions
Δδ	écart de déplacement chimique (ppm)
AF	domaine à fonction de transactivation
Abs	absorbance
ACN	acétonitrile
Aib	acide α-aminoisobutyrique
ASU	unité cristalline asymétrique
Bn	benzyle
Boc	*tert*-butoxycarbonyle
Bu	butyle
CCM	chromatographie sur couche mince
CD	dichroïsme circulaire
Cha	cyclohexylalanine
Cin	cinnamoyle
Coa	co-activateur
COSY	correlated spectrocopy
DCM	dichlorométhane
DBD	domaine de liaison à l'ADN
ddH_2O	didistilled water
DBP	vitamin D binding protein
DIAD	diisopropyl azodicarboxylate
DIEA	N,N'-diisopropyléthylamine
DippDB	dimer at protein-protein interfaces database
DMF	diméthylformamide
DMSO	diméthylsulfoxide
DR	répétition directe
DRIP	vitamin D interaction protein
DSC	carbonate de disuccinimide
DSF	fluorimétrie à balayage différentiel
dsHBS	disulfide hydrogen bond surrogate
DRX	diffraction des rayons X
equiv.	Equivalent
ER	récepteur à œstrogène
ESI-MS	spectrométrie de masse par ionisation electrospray

ESRF	installation européenne de rayonnement synchrotron
Fmoc	fluorénylméthoxycarbonyle
GR	récepteur des glucocorticoïdes
HBS	hydrogen bond surrogate
HBH	hélice-bouble-hélice
HCH	hélice-coude-hélice
HDX	échange hydrogène-deutérium
HMDI	hexaméthylène diisocyante
HOBT	1-hydroxybenzotriazole
HPLC	chromatographie liquide haute performance
HTS	high throughput screening
IBCF	chloroformiate d'isobutyle
$\mathbf{IC_{50}}$	concentration d'inhibition à 50%
ITC	titration calorimétrique isotherme
$\mathbf{K_d}$	constante de dissociation à l'équilibre
$\mathbf{K_i}$	constante d'inhibition
LBD	domaine de liaison au ligand
MDI	méthylène diphényl diisocyanate
MALDI-MS	spectrométrie de masse par désorption-ionisation laser assistée par matrice
MBHA	méthoxybenzhydrilamine
Me	méthyle
MST	thermophorèse à micro-échelle
$\mathbf{M_W}$	masse molaire
NMM	N-méthylmorpholine
Nle	Norleucine
NR	récepteur nucléaire
PDB	protein data bank
PEG	polyéthylène glycol
PPAR	récepteur activé par les proliférateurs de peroxysomes
PS	polystyrène
quant.	quantitatif
RCM	réaction de métathèse cyclisante
RID	domaine d'interaction au récepteur
RMN	résonance magnétique nucléaire
RMSD	root mean square deviation
ROESY	rotating-frame nuclear overhauser effect spectroscopy
RX	rayons X

SAR	relation structure-affinité
SANS	small-angle neutron scattering
SAXS	small-angle X-ray scattering
SCCL	side-chain crosslinking
SPS	synthèse sur support solide
SRC	co-activateur des récepteurs stéroïdiens
RAR	récepteur de l'acide rétinoïque
RXR	récepteur X de retinoïdes
VEGFR	facteur de croissance de l'endothélium vasculaire
TA	température ambiante
term	terminal
TFA	acide trifluoroacétique
TFE	2,2,2-trifluoroéthanol
THF	tétrahydrofurane
TOCSY	total correlation spectroscopy
TIS	triisopropylsilane
TR	récepteur des hormones thyroïdiennes
TR-FRET	fluorescence resonance energy transfer
TtMDI	tétra méthylène diisocyanate
VDR	récepteur de la vitamine D
VitD$_3$	vitamine D$_3$

Nomenclature utilisée

Les acides α aminés seront indifféremment désignés sous leur nom complet ou sous leur code à une ou trois lettres. Par exemple, l'alanine sera nommée A ou Ala.

Les résidus de type urée seront désignés par le code à une ou trois lettres de l'acide aminé dont ils sont issus suivi en exposant de la lettre « U ». Ainsi le dérivé de l'alanine sera nommé A^U ou Ala^U (**Fig. 1**). La position de la chaîne latérale sur le carbone α d'un résidu sera désignée par la lettre « α » en indice.

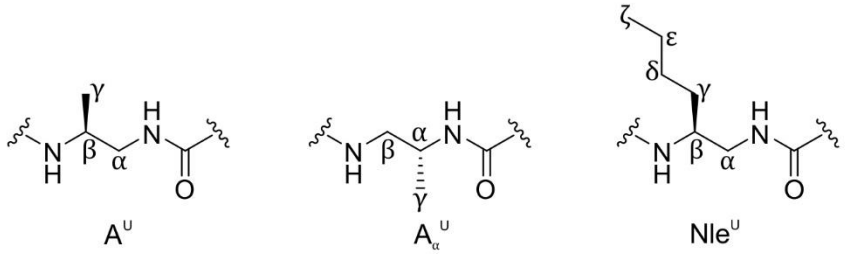


Figure 1 : Nomenclature des résidus de type urée.

Les séquences des oligourées synthétisées dans ce manuscrit seront désignées par une notation spécifique. Un trimère protégé par un groupement tert-butoxycarbonyle et comportant trois résidus urée dérivés de la valine, de l'alanine et de la leucine sera noté Boc-$V^UA^UL^U$-NH_2 (**Fig. 2a**).

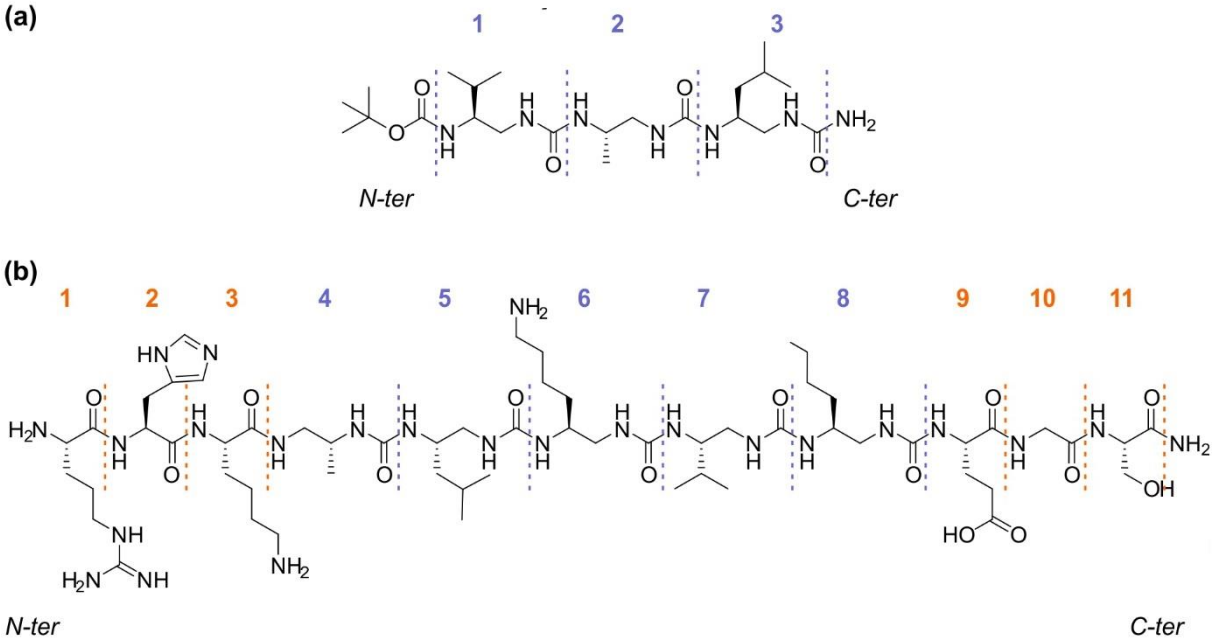


Figure 2 : Notation et numérotation des résidus **(a)** dans une oligourée et **(b)** dans une chimère.

De la même manière un oligomère chimère comportant onze résidus : cinq résidus urées dérivés de l'alanine D, de la leucine, de la lysine, de la valine et de la norleucine insérés

dans une séquence peptidique de six résidus acides aminés arginine, histidine, lysine, acide glutamique, glycine et sérine sera notée ainsi : RHKA$_{\alpha}$^{U}L^{U}K^{U}V^{U}Nle^{U}EGS. Les résidus seront numérotés en partant de l'extrémité N-terminale vers l'extrémité C-terminale. Ainsi la position 1 correspondra au résidu dont l'amine terminale est libre (**Fig. 2b**).

Les angles dièdres (Φ, Θ1, Θ_2 et Ψ) seront définis par Φ = C_{-1}O-NH-C_{β}-C_{α}, Θ_1 = NH-C_{β}-C_{α}-N'H, Θ_2 = C_{β}-C_{α}-N'H-CO et Ψ = C_{α}-N'H-CO-N_{+1}H (**Fig. 3**).

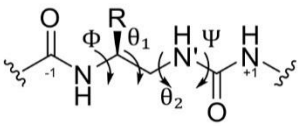


Figure 3 : Angles dièdres des résidus.

La plupart des composés décrits font l'objet d'une numérotation systématique, suivant leur ordre d'apparition dans ce manuscrit, réinitialisée au début de chaque nouveau chapitre. Les chimères, monomères et oligourées synthétisés seront désignés par les lettres **C**, **M** et **OL** suivies de leur ordre d'apparition dans le manuscrit.

Table des matières

Avant-propos.....21

Chapitre 1

Introduction bibliographique.....23

I. Stabilisation de structures secondaires hélicoïdales par le développement de peptidomimes artificiels.....25

1. Développement de peptidomimes artificiels via des modifications ponctuelles du squelette carboné pour stabiliser la structure secondaire.....26

a. *Structure peptidique contrainte via des acides aminés α,α-dialkylés*.....26

b. *Réticulation de chaînes latérales sur une face de l'hélice – Méthode Side-chain crosslinking (SCCL)*.....26

c. *Liaisons hydrogène de substitution : méthode Hydrogen Bond Surrogate (HBS)*.....28

2. Développement de structures secondaires artificielles très stables : les foldamères.....30

a. *Généralités sur cette famille de peptidomimes*.....30

b. *Les foldamères d'oligourée : une structure hélicoïdale proche de l'hélice α*.....31

c. *Concept de chimère : foldamère mixte oligoamide/oligourée*.....33

II. Les récepteurs nucléaires : une cible de choix pour le développement de mimes de petite portion peptidique.....35

1. Interactions protéine-protéine avec les récepteurs nucléaires.....36

a. *Une structure commune à plusieurs récepteurs nucléaires*.....36

b. *Fonctionnement des récepteurs nucléaires sous forme de complexe*.....37

c. *Un mécanisme de régulation complexe, concerté et contrôlé*.....39

2. Le récepteur de la vitamine D (VDR), une cible thérapeutique de choix.....42

a. *Le récepteur de la vitamine D (VDR)*.....42

b. *Structure de VDR*.....43

c. *Mécanisme d'action et voie de signalisation de l'activation de VDR*.....44

3. Une famille de protéines coactivatrices communes : les protéines SRC.....45

a. *Généralités sur cette famille de coactivateurs*.....45

b. *SRC et impact sur la santé*.....46

c. *Rôle du motif LXXLL responsable de l'interaction et étude de la zone d'interaction avec le récepteur nucléaire*.....47
4. Inhibiteurs de VDR décrits dans la littérature.....49
a. *Antagonistes de VDR*.....49
b. *Inhibiteurs de l'interaction VDR/corégulateurs*.....50
i. Des petites molécules pour inhiber l'interaction VDR/corégulateurs.....51
ii. Développement d'inhibiteurs peptidiques synthétiques.....52
iii. Développement d'inhibiteurs peptidomimes avec des stapled peptides.....53
iv. Développement de peptidomimes à base de foldamères d'oligourée hybrides.....54

III. Mime de zones étendues et vastes : de la structure secondaire aux structures tertiaires et quaternaires.....54
1. Une structure quaternaire particulière : le *coiled coil*.....55
a. *Origine de l'assemblage quaternaire sous forme de coiled coil*.....55
b. *Principes gouvernant l'assemblage des coiled coils*.....55
2. Conception rationnelle de structures quaternaires.....57
a. *Approche descendante top-down pour la conception de foldamères hybrides capables de s'assembler en bundles*.....57
b. *Approche ascendante bottom-up pour la conception de novo de foldamères hybrides amphipatiques*.....58
3. Conception rationnelle de mimes de HCH pour l'inhibition d'interactions protéine-protéine.....60
a. *Stabilisation de dimères d'hélices impliqués dans des IPP par formation d'un lien covalent entre deux hélices*.....61
b. *Application de l'approche top-down de foldamères hybrides α/β pour la conception d'inhibiteurs de l'interaction du facteur de croissance de l'endothélium vasculaire et son domaine Z*.....62
c. *Application de l'approche bottom-up pour la conception de novo de foldamères repliés en structure HCH*.....63

IV. Objectifs de la thèse.....64

Références bibliographiques.....65

Chapitre 2

Développement de foldamères mixtes pour l'inhibition d'interactions protéine-protéine : application à l'interaction entre le récepteur nucléaire de la vitamine D et un peptide coactivateur.....75

I. Génération de foldamères mixtes oligoamide/oligourée comme inhibiteurs peptidomimétiques de l'interaction VDR/SRC1.....77

1. Conception des chimères.....77
   - *a. Stratégie de conception des séquences chimères oligoamide/oligourée*.....79
   - *b. Remplacement d'une portion du motif hélicoïdal d'interaction LXXLL*.....80
   - *c. Remplacement de l'intégralité du motif hélicoïdal d'interaction LXXLL*.....83
2. Préparation des monomères azoture/carbamate de succinimide.....85
3. Synthèse sur support solide des chimères **C2**-**C11**.....90
   - *a. Principe de la synthèse sur support solide*.....90
   - *b. Méthodologie de synthèse sur support solide de foldamères mixtes oligoamide/oligourée*.....91
   - *c. Résultats des synthèses sur support solide des chimères **C2**-**C11***.....92
4. Conclusion sur la génération de foldamères mixtes oligoamide/oligourée.....94

II. Caractérisation bio-structurale des foldamères mixtes **C2**-**C11**.....94

1. Caractéristiques et statistiques d'affinements des données cristallographiques.....95
2. Effet du remplacement de la portion LXXL de SRC1-2 par trois urées : chimère **C2**.....99
3. Effet du remplacement de la portion XXLL de SRC1-2 par trois urées : chimères **C3** à **C8**.102
   - *a. Impact de la modification de la position de la chaîne latérale en C_α*.....103
   - *b. Impact de la modification sur deux résidus urée aux positions 2 et 4*.....107
4. Effet du remplacement de l'intégralité du motif d'interaction LXXLL de SRC1-2 par cinq urées : chimères **C9** à **C11**.....109
5. Conclusion de l'étude de relation structure-affinité sur les chimères **C2** - **C11**.....112

III. Génération d'une nouvelle série de chimères mixtes.....112

1. Optimisation de la chimère tête de série **C11**.....112
   - *a. Suppression des portions peptidiques*.....113
   - *b. Modification du résidu urée en position 2*.....114
2. Synthèse des séquences optimisées **C12-C21**.....114

a. Préparation de monomères activés.....114
*b. Synthèse sur support solide des chimères **C12**-**C21***.....115
i. Synthèse des chimères **C12** et **C13**.....115
ii. Synthèse de la chimère **C14**.....116
iii. Synthèse des chimères **C15**-**C21**.....118
c. Problèmes de synthèses identifiés et solutions envisagées.....120
i. Pureté des monomères urées et impact sur la synthèse supportée.....120
ii. Suivi de réaction par micro-clivage.....123
iii. Simulation de réaction sur support solide.....125
3. Caractérisation bio-structurale des foldamères mixtes de nouvelle génération.....126
4. Conclusion sur la génération d'une seconde série de foldamères mixtes.....131

IV. Application à d'autres récepteurs nucléaires.....131

V. Conclusions du chapitre et perspectives.....132

Références bibliographiques.....135

Chapitre 3

Développement de mimes de structures tertiaires hélicoïdales à base de foldamères d'oligourée.....137

I. Conception et caractérisation des oligourées capable de se replier en structure hélice-coude-hélice en milieu organique.....140
1. Modification de la nature de l'espaceur d'une séquence d'oligourée modèle soluble en milieu organique.....141
2. Synthèse en solution des oligourées.....142
3. Caractérisation structurale des oligourées **OL4** et **OL5**.....143
a. Cristallisation racémique comme outil de cristallisation et résolution de structure..143
*b. Structure cristalline de l'oligourée **OL4***.....144
*c. Structure cristalline de l'oligomère **OL5***.....146
d. Détermination d'une règle de conception des foldamères d'oligourées pour former des structures de type Hélice coude hélice.....147

II. Conception de structure repliée en hairpin à base de foldamères hydrosolubles.....148
1. Conception d'oligourées amphiphiles.....149
2. Préparation du monomère activés dérivés de l'acide glutamique **M12**.....151
3. Synthèse sur support solide.....151
a. Dimérisation sur support solide.....151
*b. Protocole de dimérisation sur support solide des oligourées **OL8** et **OL9***.....153
c. Résultat des synthèses sur support solide.....154
i. Synthèse des séquences « contrôle » **OL10**, **OL11** et **OL12**.....154
ii. Synthèse des dimères d'hélices **OL8** et **OL9**.....156
4. Caractérisation structurale des oligourées hydrosolubles.....157
a. Caractérisation à l'état solide par diffraction des rayons X.....157
b. Caractérisation en solution par dichroïsme circulaire.....159

III. Utilisation de l'iodation des résidus tyrosyles pour résoudre les problèmes de cristallisation et de phase.....161
1. Iodation des résidus tyrosyles comme outil pour résoudre les structures RX.....161
a. Iodation chimique in situ de foldamères oligourées amphiphiles.....161
*b. Résultats de l'iodation in situ et post cristallogenèse des oligourées **OL8** et **OL9***.....162
2. Iodation d'oligourée en amont de la cristallogenèse.....163
a. Iodation des résidus tyrosines des protéines et peptides décrits dans la littérature.....163
*b. Premières applications de cette méthode d'iodation sur les foldamères d'oligourées **OL9** et **OL10***.....164

IV. Conclusion du chapitre et perspectives.....168

Références bibliographiques.....169

Conclusion et perspectives.....173

Partie expérimentale.....179

I. General methods.....181

II. General procedure A for solid phase oligourea synthesis under microwave irradiation using azide strategy.....184

III. Experimental data of Chapter 2.....185
1. Preparation of the azide protetected O-succinimidyl carbamates **M1**-**M11**.....185
*a. Synthesis of O-succinimidyl carbamate **M6***.....185
*b. Synthesis of O-succinimidyl carbamate **M7***.....188
*c. Synthesis of O-succinimidyl carbamate **M9***.....190
*d. Synthesis of O-succinimidyl carbamate **M10***.....192
*e. Synthesis of O-succinimidyl carbamate **M11***.....194
2. Synthesis of oligomers **C2**-**C21**.....195
3. Simulation of solid phase synthesis of chimeras.....210
4. Biostructural studies.....211
a. General procedure for nanoDSF experiments.....211
b. Results of nanoDSF experiments.....212
c. General procedure for MST measurements.....213
d. Results of MST experiments.....214
e. X-ray crystallography.....218
i. General methods for crystallization.....218
ii. X-ray crystallography results for **C2**-**C11**.....218
iii. X-ray crystallography results for **C14** and **C19**.....220

IV. Experimental data of Chapter 3.....221
1. Synthesis of oligoureas **OL1**-**OL5**.....221
a. General procedure B for oligourea synthesis in solution.....221
*b. Synthesis and characterization of oligoureas **OL1**-**OL5** and their enantiomers **OLent-1**–**OL-ent-4***.....221
2. X-ray crystallography results of oligourea **OL1**-**OL5**.....226
3. Preparation of the azide protected O-sucinimidyl carbamate monomer **M12**.....227

4. Synthesis of oligourea **OL8**-**OL12**.....229
a. General procedure C for on-resin dimerization.....229
b. Synthesis of oligoureas.....229
5. X-ray crystallography results of watersoluble oligourea **OL12**.....235
6. Iodination of oligourea **OL9** and **OL10**.....235
7. Circular dichroism of oligourea **OL8**-**OL12**.....238

Références bibliographiques.....244

Avant-propos

Ce manuscrit de thèse a pour objet de rendre compte d'études menées sur les interactions protéine-protéine en s'intéressant à la conception de foldamères d'oligourée et de chimère oligoamide/oligourée pour cibler des surfaces de protéine.

Il comprend trois chapitres, le premier présentant une introduction bibliographique et les deux autres décrivant les résultats de deux axes de recherche développés indépendamment.

Le Chapitre 1 est divisé en trois parties. La première partie est une étude bibliographique centrée sur le développement et la synthèse de peptidomimes visant à former ou stabiliser des éléments de structures secondaires et plus particulièrement des motifs hélicoïdaux. Cette partie générale est complétée par un état des lieux plus spécifique en lien avec les deux axes de recherche poursuivi au cours de la thèse, à savoir le développement de mimes d'un motif en hélice-α pour cibler l'interaction entre un récepteur nucléaire et une protéine coactivatrice, avec une emphase sur le récepteur de la vitamine D et les protéines coactivatrices SRC, et d'autre part le développement de mimes de structures tertiaires.

Le Chapitre 2 décrit les travaux réalisés pour développer des foldamères mixtes oligoamide/oligourée pour l'inhibition d'interactions protéine-protéine avec une application à l'interaction entre le récepteur de la vitamine D et un peptide coactivateur.

Le Chapitre 3 présente un travail préliminaire de développement de mimes de structures tertiaires hélicoïdales à base de foldamères d'oligourée en se focalisant sur le motif hélice-coude-hélice.

Ces trois chapitres sont suivis d'une conclusion.

Puis une partie expérimentale est développée.

Chapitre 1

Introduction bibliographique

Les interactions protéine–protéine sont au cœur de la plupart des processus biologiques – de la communication intercellulaire à la mort programmée des cellules[1,2]– et représentent donc des cibles thérapeutiques prometteuses pour le développement de nouveaux traitements de certaines maladies. La conception rationnelle de molécules antagonistes visant à moduler ou inhiber ces interactions[3] par reconnaissance spécifique se heurte cependant à des problèmes majeurs. De grandes surfaces de contact –typiquement 800 $Å^2$ par protéine– sont impliquées dans ces phénomènes de reconnaissance intermoléculaire qui mettent en jeu différents modes d'interaction (liaisons H, électrostatique, hydrophobe), et la plupart de ces interfaces protéine–protéine ne sont pas adjacentes. Les petites molécules organiques (M < 500 Da) peuvent difficilement interagir avec des surfaces aussi importantes et se lier aussi fortement que le partenaire naturel.

Une analyse statistique récente sur les complexes protéine–protéine référencés dans la Protein Data Bank[5] montre que dans 62% des cas, l'interaction implique un motif en hélice α, d'où l'idée d'isoler le fragment peptidique en hélice α pour venir interagir avec la protéine à la place du partenaire naturel. Or les propriétés intrinsèques des peptides naturels en font de mauvais candidats médicaments. Pris isolément, les courts peptides ne sont généralement pas structurés en hélice α, problème auquel s'ajoutent une faible biodisponibilité, un haut poids moléculaire, une difficulté au passage membranaire et une sensibilité à la dégradation enzymatique. Des approches visant à stabiliser ces structures en hélice ou bien développer de nouvelles molécules capables de mimer ces hélices (peptidomimes) ont ainsi été développées pour ensuite pouvoir cibler des biomacromolécules d'intérêt.

I. Stabilisation de structures secondaires hélicoïdales par le développement de peptidomimes artificiels

La stabilisation des hélices α représente un défi fondamental dans la découverte de molécules innovantes à visée thérapeutique. Depuis quelques décennies, plusieurs équipes de recherche se sont intéressées à cette thématique[6].

Les premières méthodologies de stabilisation élaborées reposent toutes sur le choix rationnel des acides aminés constituant la séquence peptidique afin d'obtenir un meilleur repliement. L'insertion dans la séquence peptidique d'acides aminés possédant une forte propension à l'hélicité comme l'alanine et la leucine[7] conduit à la formation d'une hélice α plus stable. L'effet hydrophobe, le capping aux extrémités des séquences d'acides aminés, les interactions de π-stacking entre noyaux aromatiques, les interactions polaires, la formation de ponts salins[8], et la création de paires d'ions entre charges opposées entre les chaînes latérales

de résidus séparés par un tour de l'hélice (en position i, i+3 ou i+4) sont d'autres voies conduisant à la stabilisation des hélices α.

1. Développement de peptidomimes artificiels via des modifications ponctuelles du squelette carboné pour stabiliser la structure secondaire

a. Structure peptidique contrainte via des acides aminés α,α-dialkylés

La structure secondaire en hélice α est peu stable à l'état naturel lorsqu'une courte portion d'acides aminés (moins de dix résidus) est isolée de la protéine. Contraindre l'espace conformationnel via des analogues d'acides aminés α,α-dialkylés est une méthode permettant de stabiliser la structure hélicoïdale[9] (**Fig. 4**). L'acide α-aminoisobutyrique (Aib) (**Fig. 4a)** est l'un des analogues les plus utilisés dans cette approche de restriction conformationnelle.

(a) **(b)**

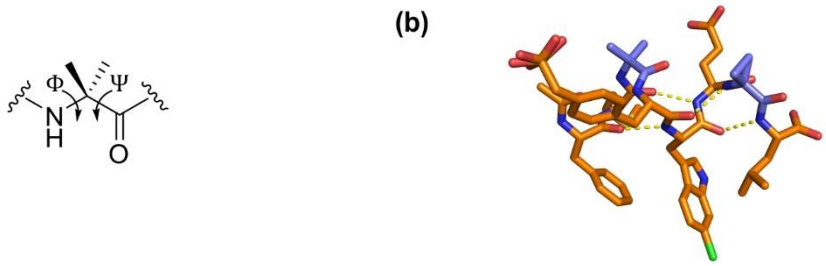


Figure 4 : (a) Nomenclature du résidu d'acide 2-aminoisobutyrique (Aib), **(b)** structure cristalline d'un peptide de huit résidus d'acides aminés possédant deux résidus Aib (en bleu) replié en hélice (PDB : **2GV2**).

L'introduction de ce résidu achiral permet d'utiliser l'effet Thorpe-Ingold – tétrasubstitution du Cα – en imposant les valeurs des angles dièdres ϕ et ψ de (+60 ; +30) et (-60 ; -30) respectivement, ce qui en fait un fort inducteur d'hélicité. De cette façon, de courtes séquences peptidiques (typiquement 7-8 acides aminés) possédant quelques résidus Aib sont capables de se structurer en hélice α stable (**Fig. 4b**). Bien que cette méthode conduise à une stabilisation de la structure hélicoïdale, la présence de gêne stérique rend le couplage d'Aib et de ses analogues difficiles en solution. Le développement de la synthèse sur support solide a permis de surmonter la difficulté de la synthèse en solution.

b. Réticulation de chaînes latérales sur une face de l'hélice – Méthode Side-chain crosslinking (SCCL)

La méthode de « Side-chain crosslinking » repose sur un principe de stabilisation simple : la formation de ponts covalents par cyclisation des chaînes latérales sur une face de

l'hélice α. Cette méthode nécessite donc l'insertion d'acides aminés naturels et non naturels à des positions bien définies dans la séquence, généralement en position i, i+3 ou i+4 à un tour d'hélice ou en position i, i+7 à deux tours d'hélice (**Fig. 5a**).

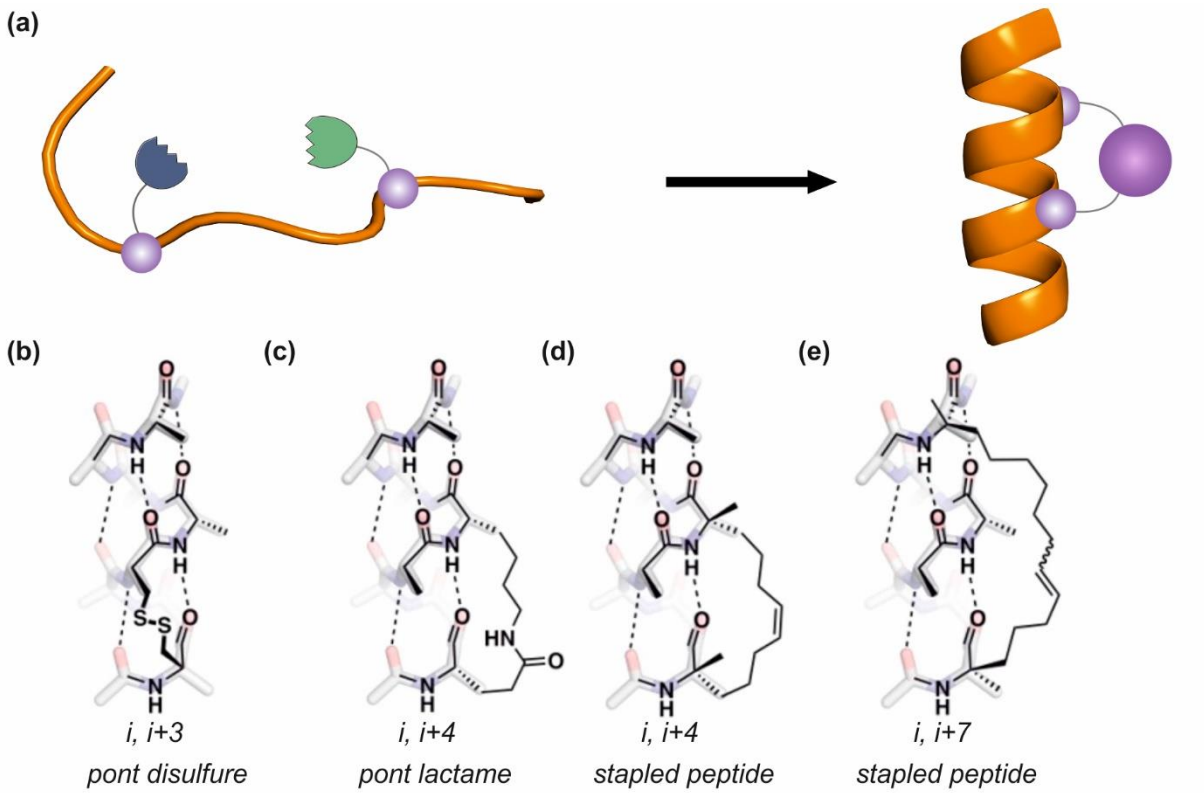


Figure 5 : Schéma Side-chain cross linking (SCCL) : **(a)** principe général ; **(b)** pont disulfure entre les chaînes latérales i, i+3 ; **(c)** pont lactame entre deux chaînes latérales en i, i+4 ; **(d)** ponts hydrocarbonés entre deux résidus non naturels aux positions **(d)** i, i+4 et **(e)** i, i+7

Plusieurs types de ponts covalents sont décrits dans la littérature. Les plus courants sont les ponts disulfures[11] (**Fig. 5b**), les ponts lactames[12,13] (**Fig. 5c**) et les ponts hydrocarbonés (**Fig. 5d-e**).

Les premiers exemples de ponts hydrocarbonés ont été décrits par Robert Grubbs en 1998 qui s'est intéressé à la cyclisation par métathèse d'oléfine catalysée au ruthénium entre les chaînes latérales d'acides aminés non naturels de type sérine O-allyl en (i, i+4) d'une séquence peptidique en hélice[14]. Ces travaux ont inspiré l'équipe de Gregory Verdine pour le développement de liaisons stables pour remplacer les ponts disulfures et lactames sensibles à la protéolyse[15]. Verdine introduit le concept de « stapled peptide » en 2000. Ces peptides sont reliés par des chaînes hydrocarbonées à des positions déterminées (i, i+3, i+4 ou i+7) pour conduire à des macrocycles via une métathèse cyclisante catalysée au ruthénium. La synthèse de ces peptides est compatible avec la stratégie Fmoc et se fait directement sur support solide que ce soit pour l'insertion des acides aminés modifiés ou pour l'étape de

cyclisation par métathèse (**Fig. 6a**). Il a utilisé cette technique pour coupler deux acides aminés chiraux α,α-disubstitués portant une oléfine sur deux chaînes alkyles pour former un peptide cyclique possédant une chaîne carbonée allant de 8 à 11 carbones reliés respectivement en position (i, i+3), (i, i+4) et (i, i+7)[16] (**Fig. 6b**).

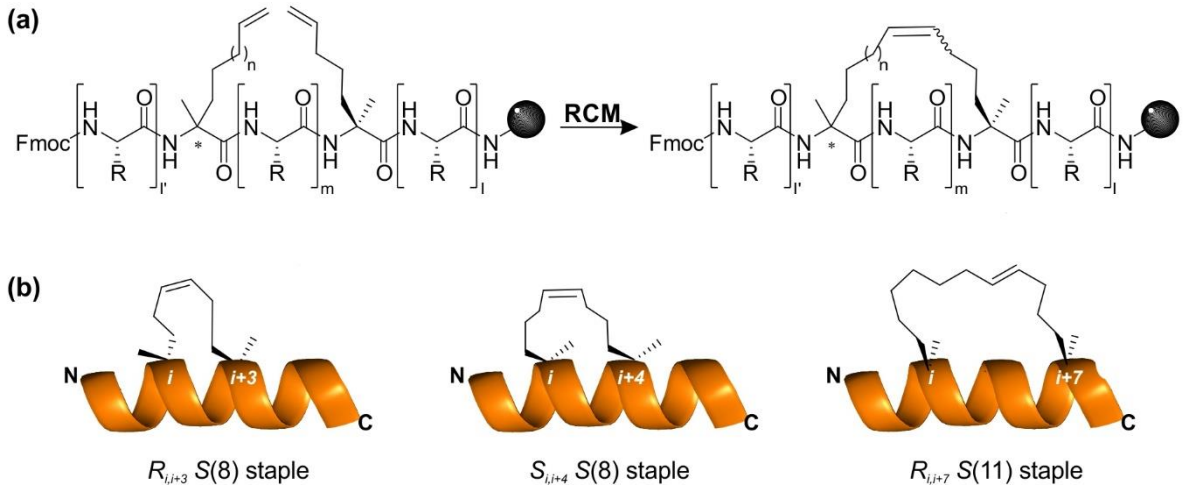


Figure 6 : (a) Schéma général de préparation de stapled peptide sur support solide utilisant la stratégie Fmoc via une réaction de métathèse cyclisante [16] avec un catalyseur de Grubbs de première génération. **(b)** Exemple de trois stapled peptides hydrocarbonés[16]

L'ensemble de ces travaux a inspiré Gregory Bird, qui en 2010 décrit des peptides en hélice contenant deux stapled peptides aux positions (i, i+4) aux extrémités N et C-terminales de la séquence peptidique[17] via une réaction de métathèse cyclisante. Il a montré que la présence d'une cyclisation supplémentaire dans la séquence peptidique permet d'améliorer la stabilité de la structure en hélice.

L'approche de réticulation des chaînes latérales sur une face de l'hélice a ainsi permis d'augmenter significativement l'hélicité et la stabilité métabolique par rapport à des peptides naturels non-modifiés. Cependant, le pont hydrocarboné doit être correctement positionné pour ne pas encombrer la face de l'hélice impliquée dans l'interaction.

c. Liaisons hydrogène de substitution : méthode Hydrogen Bond Surrogate (HBS)

L'approche HBS est l'une des approches clés pour la stabilisation des hélices α. Cette méthode repose sur la pré-organisation des quatre premières liaisons amide de la séquence peptidique via un lien covalent de façon à initier le repliement en hélice. La première liaison hydrogène intramoléculaire, C=O--H–N, présente entre le groupe C=O du premier résidu et la fonction NH du cinquième résidu à l'extrémité N-terminale de l'hélice est remplacée par une

liaison covalente de type C=X–Y–N. Ce lien covalent est de type alcane, alcène ou thioéther[18]. Cette approche est analogue à la méthode SCCL présentée précédemment mais elle se fait à l'extrémité de l'hélice ce qui permet de ne pas modifier les chaînes latérales du peptide impliquées dans l'interaction protéine-protéine et de ne pas encombrer une face de l'hélice

Le premier exemple d'HBS a été décrit par Cabezas et Satterhwait[19] : ils font état de l'utilisation d'une hydrazone pour stabiliser l'hélice α d'un petit peptide. Puis Arora et son équipe a stabilisé des peptides via HBS en utilisant une réaction de métathèse cyclisante[18a] (**Fig. 7**).

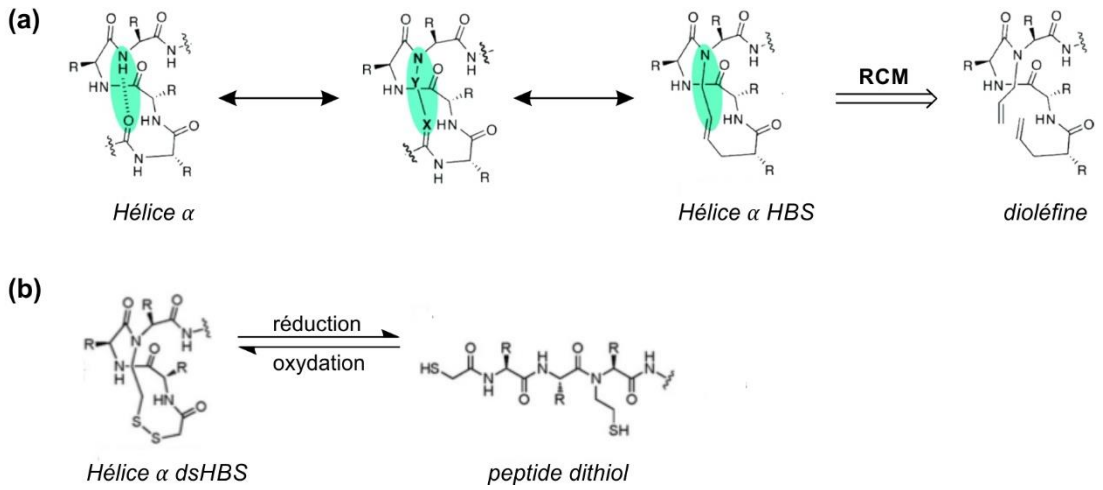


Figure 7 : (a) Nucléation de l'hélice α en remplaçant la liaison hydrogène i, i+4 du N-terminal (C=O---H–N) par une liaison covalente (C=X–Y–N). Liaison hydrogène de substitution (HBS) de l'hélice α via une liaison carbone–carbone obtenue par une réaction de métathèse cyclisante (RCM) entre deux oléfines[20]. **(b)** Hélice α dsHBS en i, i+4 et réversibilité de l'hélice α dsHBS

Récemment, une technologie dérivée de HBS a aussi été développée. La dsHBS (disulfide hydrogen bond surrogate) a la particularité d'introduire un pont disulfure réversible selon les conditions oxydante ou réductrice du milieu (**Fig. 7b**).

Des analyses structurales par RMN, dichroïsme circulaire (*circular dichroism*, CD) et diffraction des rayons X (DRX) ont permis de confirmer le repliement de ces structures en hélices stables. Les avantages de cette technologie de stabilisation sont nombreux : prédictibilité des conformations en hélice, capacité à apporter une rigidité conformationnelle dans de petits peptides et conservation des fonctions organiques des chaînes latérales proches de l'hélice.

2. Développement de structures secondaires artificielles très stables : les foldamères

a. Généralités sur cette famille de peptidomimes

Les foldamères ont été définis pour la première fois en 1998 par Samuel Gellman[21] comme tout polymère ou oligomère ayant une forte propension à adopter une conformation spécifique, compacte et bien définie en solution.

Depuis, de nombreuses familles de foldamères aliphatiques ont été développées, les principales étant celles des β-peptides[22], des peptides mixtes α/β[23] ou β/γ[24] ainsi que les peptoïdes[25] (**Fig. 8**). D'autres stratégies consistent à remplacer le lien amide par des fonctions hydrazide ou urée pour conduire aux foldamères d'hydrazino-peptide et d'oligourée[26] respectivement (**Fig. 8**). Une analogie aux biopolymères a conduit au développement de ces foldamères aliphatiques dit « biotiques » avec lesquels ils partagent des principes similaires de repliement. En parallèle, une autre classe de foldamères dits « abiotiques » et dont le squelette possède des noyaux aromatiques[27] a été largement étudiée en raison de leurs nombreuses qualités intrinsèques (forte stabilité structurale, prédictibilité du mode de repliement, propension à cristalliser et accessiblilité par synthèse).

L'objectif des foldamères est de concevoir des molécules capables d'adopter des structures repliées mimant les structures rencontrées dans les protéines et possédant des fonctions rivalisant avec celles des protéines. Leurs structures stables et bien définies, leur résistance à la dégradation enzymatique, leur diversité structurale et la prédictibilité de leurs structures font des foldamères des composés de choix pour le développement de molécules innovantes à visée thérapeutique.

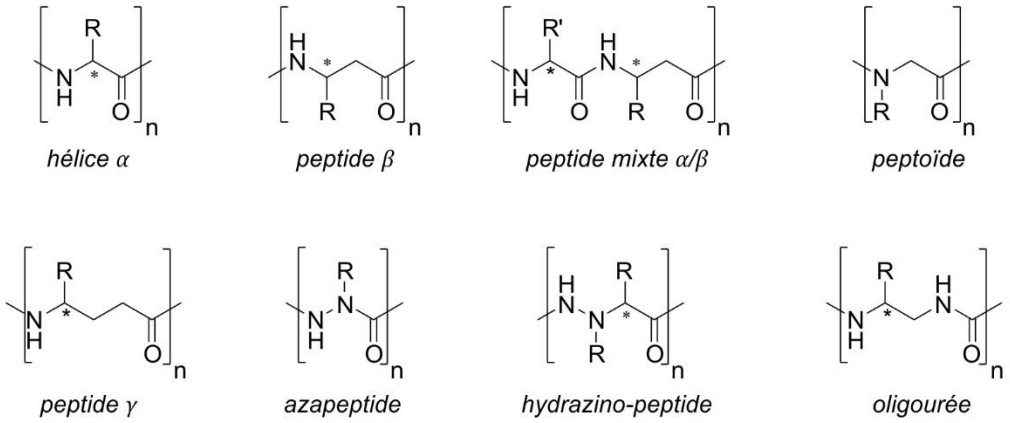
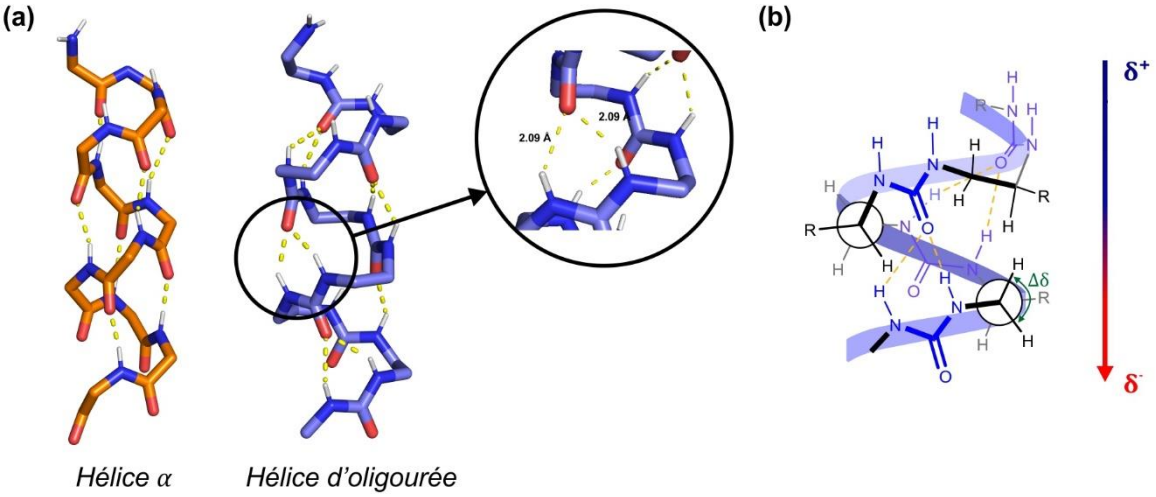
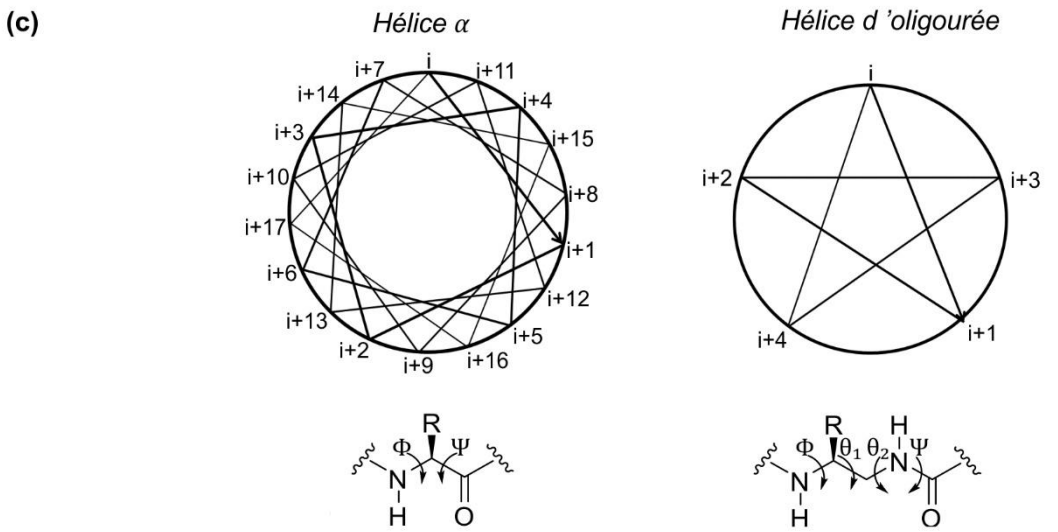


Figure 8: Nomenclature de différentes familles de foldamères possédant des structures aliphatiques.

b. Les foldamères d'oligourée : une structure hélicoïdale proche de l'hélice α





	hélice α[32]	hélice d 'oligourée
Φ (°)	-58	-100
θ_1 (°)	-	+55
θ_2 (°)	-	+85
Ψ (°)	-47	-170
Diamètre (Å)	4,6	5,4
Pas de l 'hélice (Å)	5,5	5,1
Résidu/tour	3,6	2,5
Pas (Å)/résidu	1,5	2,0

Figure 9 : (a) Comparaison de structures cristallines d'une hélice α et d'une oligourée. **(b)** Représentation du pseudo-cycle à 12 et 14 atomes et macrodipôle de l'hélice d'oligourée. **(c)** Comparaison structurale de l'hélice d'oligourée et de l'hélice α.

Depuis de nombreuses années, l'équipe de Gilles Guichard s'est intéressée au développement de foldamères d'oligourée aliphatiques hélicoïdaux. Le domaine d'application des foldamères développés au sein du groupe est très vaste : de l'organocatalyse, aux inhibiteurs d'interactions protéine–protéine et mimes de peptides antimicrobiens. Les foldamères d'oligourée sont des analogues de γ-peptides (**Fig. 8**) dont les monomères activés sous forme de succinimide sont synthétisés à partir d'acides aminés α naturels. L'hélice d'oligourée et l'hélice α naturelle diffèrent dans le nombre de résidus par tour d'hélice, le pas par résidu, le diamètre et pas de l'hélice (**Fig. 9c**). Des études structurales en solution par RMN et dans le solide par diffraction des rayons X ont permis de montrer que ces oligourées adoptent des structures en hélice droite 2,5 bien définies[28], stabilisées par un réseau de liaisons hydrogène intramoléculaires à trois centres[29] et à longue distance entre fonctions urées (**Fig. 9a**). Ce réseau se fait entre le carbonyle C=O du résidu i et les NHs de l'urée en position i+2 et i+3, formant ainsi un pseudo-cycle à 12 et 14 atomes (**Fig. 9b**). L'hélicité est indépendante de la nature des chaînes latérales, mais elle est dépendante de la chiralité des résidus. L'hélice d'oligourée est généralement droite, elle est caractérisée en RMN par une différence de déplacement chimique (Δδ) significative des protons $^{\alpha}CH_2$ diastéréotopiques[29, 30] (**Fig. 10a**), et en CD par la présence d'un maximum du signal à 203 nm, un signal nul à 193 nm et un minimum à 188 nm dans le trifluoroéthanol[31] (**Fig. 10b**). La résolution des structures cristallines obtenues par diffraction des rayons X a permis de compléter les données structurales de l'hélice identifiées par RMN. L'hélice canonique d'oligourée possède des angles de torsion Φ, ϑ_1, ϑ_2 et ψ de respectivement -100°, +55°, +85° et -170° (**Fig. 9c**).

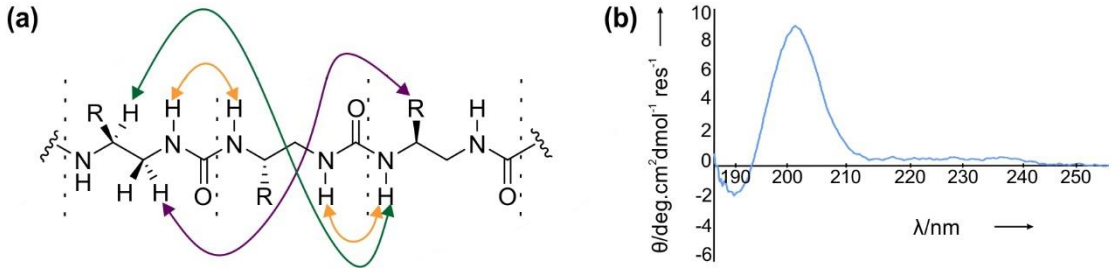


Figure 10 : (a) Connectivités NOE dans un foldamère d'oligourée. **(b)** Spectre de dichroïsme circulaire d'un foldamère d'oligourée dans le trifluoroéthanol.

Les valeurs proches des diamètres des hélices d'oligourée et α naturelles[32], des pas d'hélice ainsi qu'une polarité commune des hélices, montrent une forte similarité structurale entre l'hélice d'oligourée et l'hélice α naturelle, permettant d'envisager l'utilisation des foldamères d'oligourée comme peptidomimes de l'hélice α.

c. Concept de chimère : foldamère mixte oligoamide/oligourée

Des séquences hybrides foldamère/peptide (également appelés chimères) mimes de l'hélice α ont été décrites dans la littérature dans lesquelles un résidu d'acide aminé est remplacé ponctuellement par un monomère non naturel[23,24,33]. Cependant, peu d'études sur l'influence de l'insertion d'un segment artificiel sur la stabilisation d'un segment peptidique ont été rapportées[34]. L'équipe de Gilles Guichard s'est récemment intéressée à étudier l'impact d'un insert de séquence de foldamère oligourée dans une séquence peptidique sur la stabilité de la structure secondaire de l'ensemble[35] (**Fig. 11**).

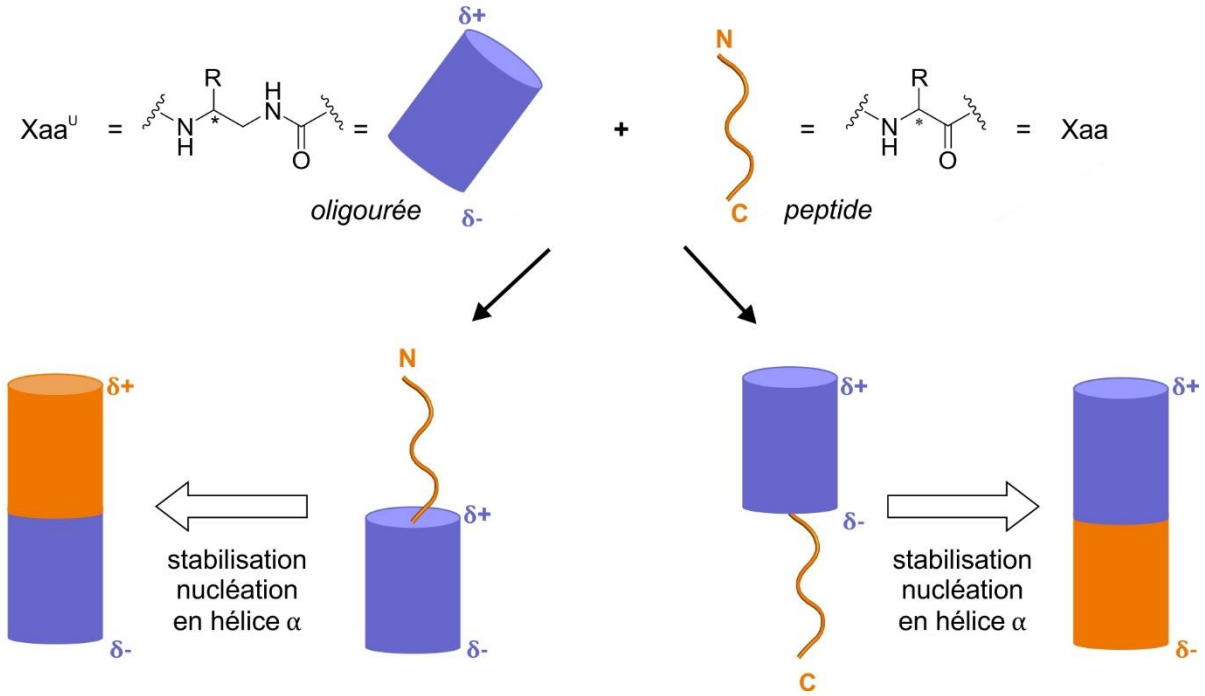


Figure 11 : Représentation des foldamères mixtes oligoamide/oligourée, chimères. Principe de stabilisation et de la nucléation en hélice α du fragment peptidique par un fragment urée.

Des séquences oligourée de différentes longueurs ont été introduites soit en position N-terminale soit en C-terminale d'une séquence peptidique modèle alternant des résidus alanine et leucine. Les résultats obtenus par RMN, CD et DRX à haute résolution ont montré qu'un segment oligourée de 3 résidus était suffisant pour stabiliser le repliement en hélice d'un heptapeptide et qu'un réseau continu de liaisons hydrogène était maintenu à la jonction peptide/urée faisant le lien entre les deux types d'hélice[35] (**Fig. 12**).

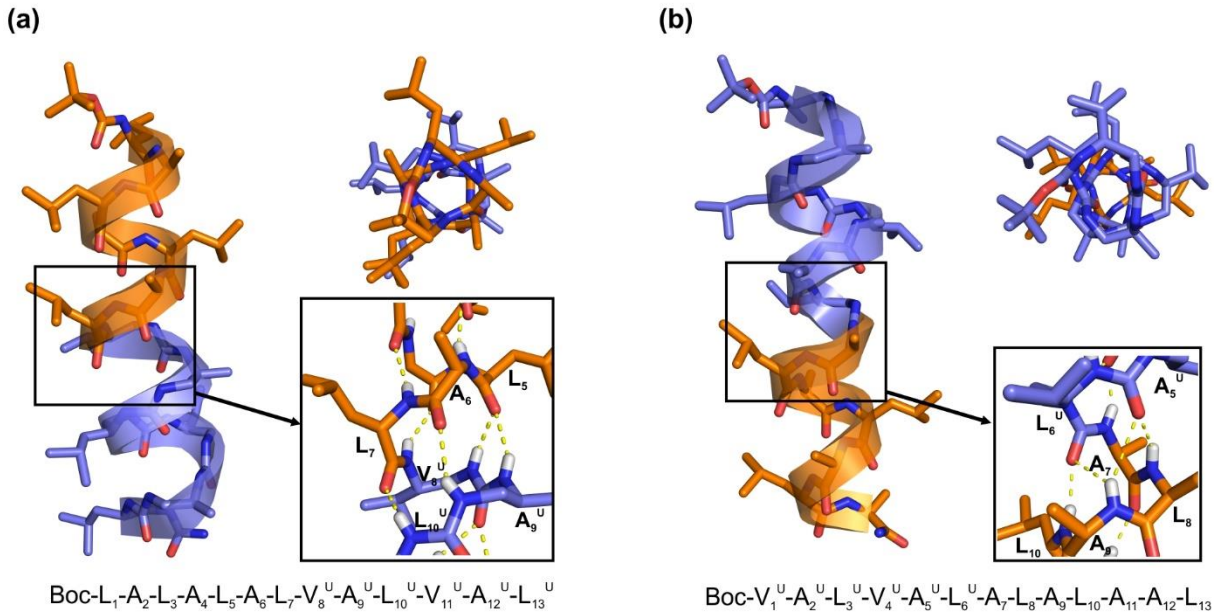


Figure 12 : Structures cristallines des foldamères mixtes avec la portion oligourée à l'extrémité **(a)** C-terminale et **(b)** N-terminale de la séquence peptidique. Observation de la jonction peptide/urée et de la conservation d'un réseau continu de liaisons hydrogène (CCDC : **1058821** et **1040765**)

L'équipe s'est ensuite intéressée à insérer une portion de trois urées aux deux extrémités de la séquence peptidique (**Fig. 13a**). Les données obtenues à haute résolution ont permis d'observer la conservation du repliement en hélice de la structure (**Fig. 13b**). Celle-ci est également observée lorsqu'une séquence de six résidus urées est introduite dans une séquence peptidique[36]. L'ensemble de ces résultats a permis de conclure que la fusion de séquences oligourée/oligoamide est bien supportée et que l'hélice globale est bien conservée quelle que soit la nature des chaînes latérales[36]. Il est donc possible de remplacer une séquence de plusieurs acides aminés par une séquence d'oligourée non naturelle sans modifier la structure secondaire et en améliorant sa stabilité.

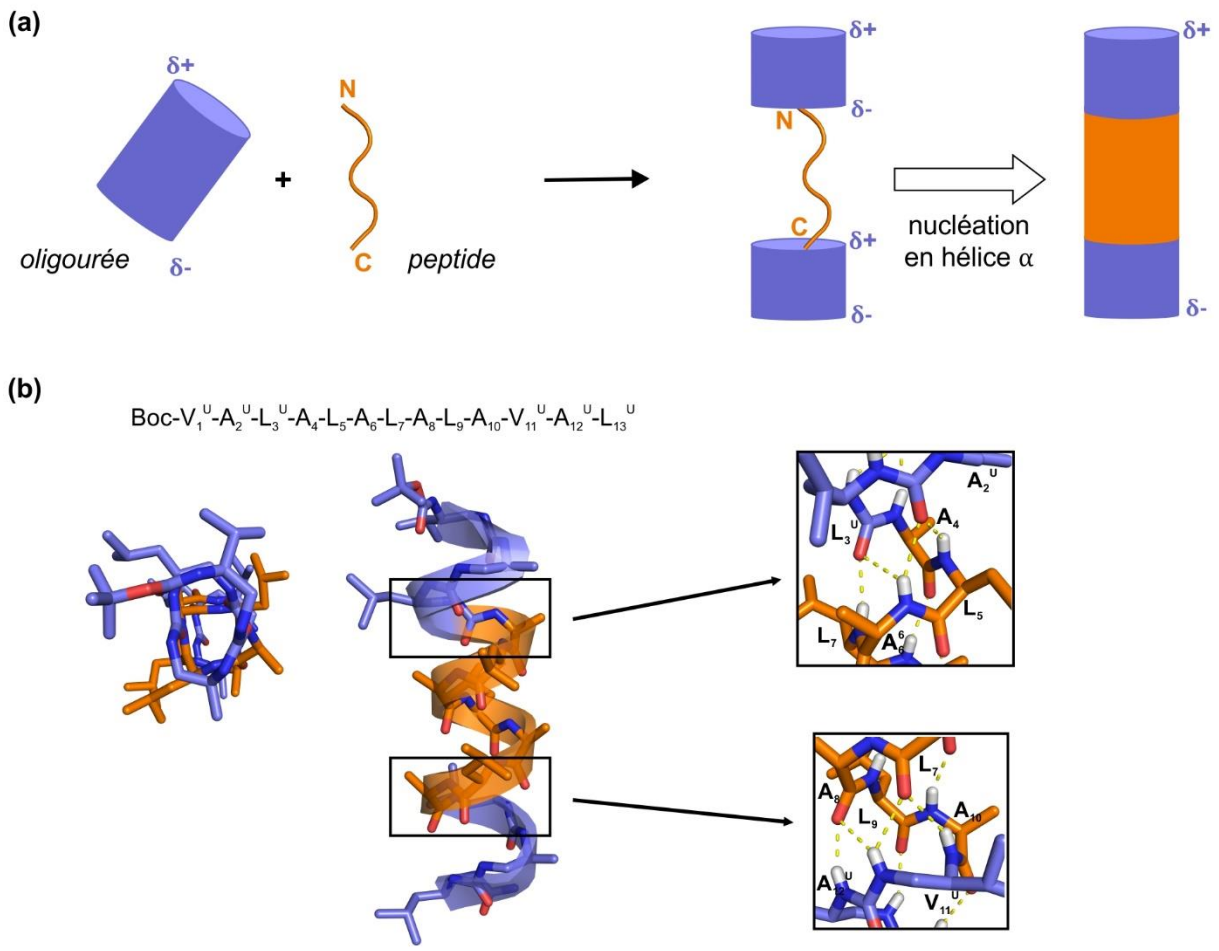


Figure 13 : **(a)** Représentation d'une chimère avec l'insertion d'un fragment urée aux extrémités d'une séquence peptidique pour stabiliser l'hélice α du fragment peptidique. **(b)** Structure cristalline de la chimère : vue de haut, de face et du réseau de liaisons hydrogène à trois centres aux jonctions urée/peptide (CCDC : **1041244**)

II. Les récepteurs nucléaires : une cible de choix pour le développement de mimes de petite portion peptidique

Les récepteurs nucléaires (RN) appartiennent à la superfamille des facteurs de transcription. Il existe plus de cent récepteurs nucléaires décrits dans la littérature, formant ainsi la plus large famille de facteurs de transcription des organismes métazoaires[37]. Les récepteurs nucléaires sont ligand-dépendants, impliqués dans la régulation de l'expression des gènes dans des cellules cibles. Ils se différencient des autres facteurs de transcription en se liant avec des petites molécules hydrophobes telles que les stéroïdes sexuels (androgènes, œstrogène, progestérone), les stéroïdes surrénaliens (gluco- et minéralcorticoïde), la vitamine D, les hormones thyroïdiennes et rétinoïdes[37, 38]. Les RN contribuent à une multitude de phénomènes physiologiques : homéostase, reproduction, développement cellulaire,

apoptose... Ils constituent ainsi une cible de choix pour le développement de nouveaux traitements pour de nombreuses maladies[39] comme le cancer, l'ostéoporose, l'obésité et le diabète (type I et II). Il est donc important de connaitre leurs fonctionnements, structurations et mécanismes d'actions afin de pouvoir développer des molécules pertinentes à visée thérapeutique.

Dans cette partie, nous nous intéresserons dans un premier temps aux récepteurs nucléaires stéroïdiens et non-stéroïdiens (récepteurs à œstrogène *ER*, récepteur de l'acide rétinoïque *RAR*, récepteur X de rétinoïdes *RXR*, récepteur de la vitamine D *VDR* et récepteur des hormones thyroïdiennes *TR*) qui ont été décrits dans la littérature (II-1). Une emphase sera réalisée sur un récepteur nucléaire, le récepteur de la vitamine D (VDR) qui a été la cible thérapeutique choisie dans ce projet de thèse de développement d'inhibiteurs d'interactions protéine-protéine (II-2). Puis nous terminerons cette partie par une étude des protéines coactivatrices des récepteurs stéroïdiens (SRC) (II-3), celle-ci étant commune à de nombreux récepteurs nucléaires.

1. Interactions protéine-protéine avec les récepteurs nucléaires

a. Une structure commune à plusieurs récepteurs nucléaires

Une étude phylogénétique a permis de classer les récepteurs nucléaires en trois sous familles[37, 40] :

- les récepteurs stéroïdiens (ER, RXR, récepteurs à la progestérone PR, aux androgènes AR, aux glucocorticoïdes GR et aux minéralocorticoïdes MR)
- les récepteurs non-stéroïdiens (RAR, TR, VDR, récepteur activé par les proliférateurs de peroxysomes PPAR), strictement nucléaires, ils hétérodimérisent avec RXR, ou homodimérisent et se lient à leurs ligands correspondants
- les récepteurs orphelins, pour lesquels les ligands ne sont pas encore connus et identifiés.

Bien que ces récepteurs nucléaires aient des fonctions biologiques caractéristiques et différentes[37] les unes des autres, ils partagent tous une architecture commune qui se divise en cinq domaines (**Fig. 14**) :

Figure 14 : Structure générale d'un récepteur nucléaire

Le domaine N-terminal, appelé également domaine A/B est le domaine le plus variable en terme de longueur et de séquence dans la structure des récepteurs nucléaires. Il contient le domaine à fonction de transactivation indépendant du ligand, appelé domaine AF-1 (*Activation Function* 1). Ce domaine N-terminal est suivi du domaine C, qui est le domaine de liaison à l'ADN (*DNA Binding Domain* DBD), celui-ci contient deux doigts zinc[41], il est central et est le domaine dont la séquence est la plus conservée parmi les récepteurs nucléaires. Ce domaine conservé est suivi du domaine D, aussi appelé *Hinge* (charnière en français), qui est une partie peptidique assez flexible contenant le signal de localisation nucléaire. Il permet au dimère de se lier à divers éléments de réponse de l'ADN[42]. Celui-ci est suivi du domaine E qui correspond au domaine de liaison au ligand (*Ligand Binding Domain* LBD). Ce domaine est l'un des plus importants car il contient la surface de dimérisation (homo ou hétérodimérisation) et des sites d'interactions en C-terminal avec des protéines corégulatrices via le domaine d'activation de transcription ligand-dépendant (AF-2).

b. Fonctionnement des récepteurs nucléaires sous forme de complexe

Bien qu'il ait été rapporté dans la littérature que les RN peuvent fonctionner en tant que monomères, la plupart d'entre eux réalisent leur activité de transcription des gènes sous forme de dimères (homo- ou hétérodimères). Ils lient leur DBD respectifs à une séquence d'hexanucléotide de l'ADN 5'-AGGTCA-3' (dans le cas des RN non-stéroïdiens et ER) répétée deux fois et qui est espacée par un nombre de nucléotides variant de 1 à 5. Cette séquence répétée constitue un élément de réponse de l'ADN, aussi appelée répétition directe (*direct repeat*, DR). La séquence d'hexanucléotide est appelée hémi-site. Les RN sont également similaires par leur mode de liaison à l'ADN[43]. Ils diffèrent simplement par le nombre de nucléotides séparant les hémi-sites via la règle décrite par Evans en 1995[39a]. Il a montré dès 1991 que cette différence dans le nombre de nucléotides espaçant les hémi-sites des DR est reliée à la sélectivité des RN sur l'ADN, et elle est ainsi sélective de la réponse hormonale réalisée. Evans et al ont permis d'identifier les RN sur l'ADN en fonction du nombre de nucléotides espaceurs des hémi-sites sur les DR[39a,43] (**Tableau 1**). Des études structurales en solution des récepteurs entiers ou tronqués de leur domaine AB ont été réalisées par Rochel *et al.*[44] en 2011 (par SAXS, *Small-Angle Xray Scattering*, SANS *Small-Angle Neutron Scattering*, spectroscopie de fluorescence et ajustement aux structures cristallographiques), à l'exception du complexe PPAR/RXR/DR1 dont la structure cristalline a été résolue en 2008 par Chandra *et al.*[45] (PDB : **3DZU**).

5'	*3'*	*Complexe*
RAR	RXR	RAR/RXR/DR1
PPAR		PPAR/RXR/DR1
RXR	RAR	RXR/RAR/DR2
	VDR	RXR/VDR/DR3
	TR	RXR/TR/DR4
	RAR	RXR/RAR/DR5

Tableau 1 : Complexes de RXR/RN avec l'élément de réponse à l'ADN.

Les DRs étant asymétriques, les hétérodimères se lient à l'ADN également de manière asymétrique[39a]. RXR se lie à l'hémi-site en 5' et son RN partenaire (RAR/VDR/TR) en position 3' de l'autre hémi-site[46] (**Fig. 15**) excepté dans le cas des complexes RAR/RXR/DR1 et PPAR/RXR/DR1 où RXR se lie en 3' de l'ADN.

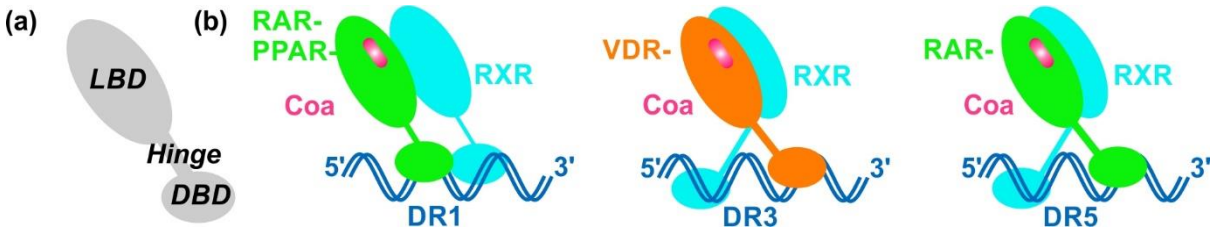


Figure 15 : (a) Structure schématique d'un récepteur nucléaire avec les domaines de liaison à l'ADN (DBD) et de liaison au ligand (LBD) reliés par la partie *Hinge* **(b)** Orientations des LBDs et DBDs de différents complexes de récepteurs nucléaires qui montrent la flexibilité et l'importance du domaine peptidique *Hinge*.

Une première structure en solution du complexe des LBD RXR/VDR/DR3 a été résolue par SAXS[44] (*Small-Angle Xray Scattering*). Celle-ci montre une structure étendue et asymétrique avec deux DBD et LBD différents reliés par le fragment peptidique Hinge. Ces observations sont également présentes pour les complexes des LBD de RAR/RXR/DR5, PPAR/RXR/DR1. L'orientation des LBD des hétérodimères est identique pour tous les RN[44]. Evans a également montré que RXR, lié à son ligand, perd son activité de transcription lorsqu'il se lie à un autre récepteur (RAR, TR, VDR), c'est son partenaire qui va pouvoir réaliser son activité de transcription[39a]. Ceci suggère la présence d'une communication moléculaire entre les deux récepteurs une fois liés à l'ADN.

c. Un mécanisme de régulation complexe, concerté et contrôlé

En 2013, des études menées sur l'activation ou la répression des récepteurs des glucocorticoïdes (GR) ont montré que l'interaction de ces RN sous forme de dimère, ainsi qu'une coopération positive et/ou négative entre les deux RN, sont déterminantes pour l'activation ou la répression des GR[47].

En 2015, Osz *et al.* se sont intéressés à la reconnaissance des hémi-sites par RXR et plus particulièrement à son élément de réponse DR1[48]. L'équipe a proposé un modèle moléculaire où la séquence 3' de l'hémi-site de DR1 exercerait une régulation allostérique sur le DBD dans le cas des homo- et hétérodimères. Cette régulation se ferait via un changement du positionnement du LBD de RXR, qui provoquerait une propagation de la modification structurelle à d'autre domaines du récepteur, et qui contribuerait ainsi au bon déroulement de la transcription. Un contrôle allostérique par l'ADN sur le recrutement de protéines coactivatrices a également été observé dans le cas des dimères RAR/RXR et RXR/RXR, ce qui confirme la présence d'une communication au niveau moléculaire entre RN une fois liés à l'ADN[49]. Cependant, le mécanisme d'action des RN ne se concentre pas essentiellement sur un contrôle allostérique de l'ADN. En effet, la régulation de la transcription des gènes requiert une action coordonnée des RN et des corégulateurs en interaction.

Il est admis depuis 1995 et les travaux de Chen *et al.* qu'en absence de ligand, le récepteur nucléaire est en interaction avec un corépresseur SMRT, ou NCoR, qui va maintenir la chromatine dans un état dit « silencieux » et ainsi empêcher une quelconque transcription des gènes[50]. Ils ont également montré qu'en présence de ligand le récepteur nucléaire va libérer le corépresseur et recruter un coactivateur pour réaliser la transcription[50a-b]. D'un point de vue structural, cette échange corépresseur – coactivateur se fait via un changement conformationnel de l'hélice H11. Cette hélice H11 se situe dans la LBD du récepteur et a été précédemment identifiée comme étant le capteur de ligand. Celle-ci va alors passer d'une structure désordonnée en une hélice α[46a]. Cette modification de structure de H11 va également impacter le domaine qui la suit, AF-2 située sur l'hélice H12, qui est initialement projetée hors de la LBD et va alors pivoter de 180° pour entrer en contact avec la LBD de son récepteur nucléaire. Ainsi cette hélice H12 devient disponible et accessible pour pouvoir se lier à une protéine coactivatrice, telle SRC1 qui est un coactivateur de récepteur stéroïdien (*Steroid Receptor Coactivator* SRC) appartenant à la famille des coactivateur p160[51] (**Fig. 16**).

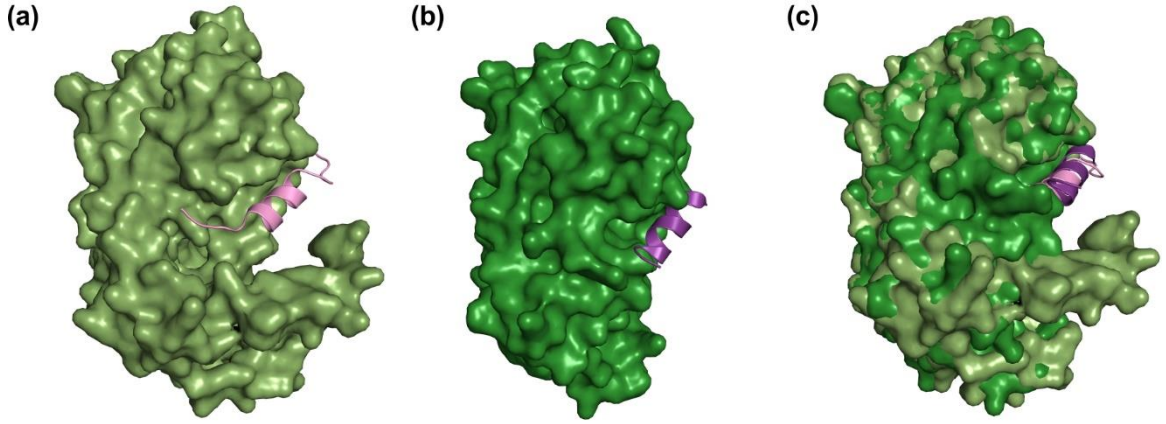


Figure 16 : Exemple de l'impact de la liaison du ligand sur la structure du récepteur nucléaire PPARα avec une protéine corégulatrice : **(a)** PPARα et SMRT (en rose), un corépresseur (PDB : **1KKQ**), **(b)** PPARα et SRC1 (en violet), un coactivateur (PDB : **2P54**). **(c)** Superposition des deux structures cristallines pour montrer la modification de la projection de la partie C-terminale du récepteur nucléaire.

Osz *et al.* ont réalisé des études structurales et thermodynamiques (par ITC et SAXS) sur un homodimère de RARβ et son coactivateur SRC1 pour tenter de prouver l'existence du mécanisme moléculaire de l'allostérie entre les deux récepteurs nucléaires d'un dimère[52], et ainsi apporter un élément de réponse à la perte d'activité de transcription de RXR lorsqu'il est sous forme d'hétérodimère. SRC1 possède un domaine d'interaction au récepteur (*Receptor Interacting Domain* RID) responsable de l'interaction avec le récepteur nucléaire, via la présence de trois copies d'un motif conservé LXXLL[53] (L pour leucine et X pour tout autre acide aminé). Ils ont montré qu'une protéine coactivatrice se lie à un dimère (stœchiométrie 1:1 RN :Coa). En utilisant un fragment de 25 acides aminés dérivé du motif 2 de SRC-1 (**SRC-1 NR2**), Osz *et al.* ont constaté que la stœchiométrie du complexe peptide/dimère avait une relation 2:1, soit deux peptides coactivateurs pour un dimère. Des mesures d'affinité par ITC ont été réalisées avec le LBD de RARβ et le peptide ou la protéine coactivatrice. Le peptide a une affinité plus forte pour RARβ (K_d = 1,5 µM) que la protéine (K_d = 2,5 µM), ce qui suggère qu'il y a un changement conformationnel au second site de liaison d'un des monomères. En comparant les structures de RARβ sous forme d'homodimère et d'hétérodimère, ils ont démontré que la liaison du coactivateur sur le site de liaison d'un des monomères avait modifié la conformation du second monomère et de son site de liaison au coactivateur au niveau de l'interface de dimérisation (**Fig. 17a**). Elle implique les résidus des hélices H7, H9, H10 et des boucles L8-9, L9-10 de chaque monomère (**Fig. 17b**). Un mécanisme moléculaire du contrôle allostérique a ainsi été proposé (**Fig. 17c**) : le coactivateur (SRC1) se lie à la LBD de $RAR\beta_1$ qui va promouvoir sa dimérisation via les résidus en N-terminal d'un des motifs de LXXLL de SRC1 en contact avec les boucles L8-9 du récepteur. Cette dimérisation va provoquer une asymétrie dans la structure qui va se propager jusqu'au second monomère via l'hélice H10 et

qui induit une modification de la conformation à l'extrémité C-terminale (H11 et H12) du second monomère RARβ$_2$ déstabilisant ainsi sa liaison au second site de liaison de la protéine coactivatrice.

En utilisant RAR comme modèle pour déterminer le mécanisme moléculaire de l'allostérie, ils ont pu déterminer sans ambiguïté l'origine du changement conformationnel. Ils ont également montré que les résidus entourant le motif LXXLL du coactivateur impliqué dans l'interface de dimérisation sont les plus efficaces et importants pour contrôler ce mécanisme d'allostérie. L'existence de cette anti-coopérativité entre les RN a une explication structurale : les changements de conformation dans le second monomère affectent la surface de liaison et donc l'affinité pour le même domaine sensible au coactivateur. Osz *et al.* suggèrent également que les différences structurales entre monomères provoquées par ce phénomène d'allostérie seraient favorables pour le recrutement de protéines coactivatrices. Ils ont également observé cette allostérie dans le cas de l'hétérodimère RXR/RARβ, ce mécanisme peut alors être appliqué aux hétérodimères. Ils suggèrent que la liaison entre le coactivateur et RARβ régulerait la liaison du coactivateur à RXR, ce qui donnerait une explication structurale sur la perte d'activité de transcription de RXR lorsqu'il est sous forme d'hétérodimère.

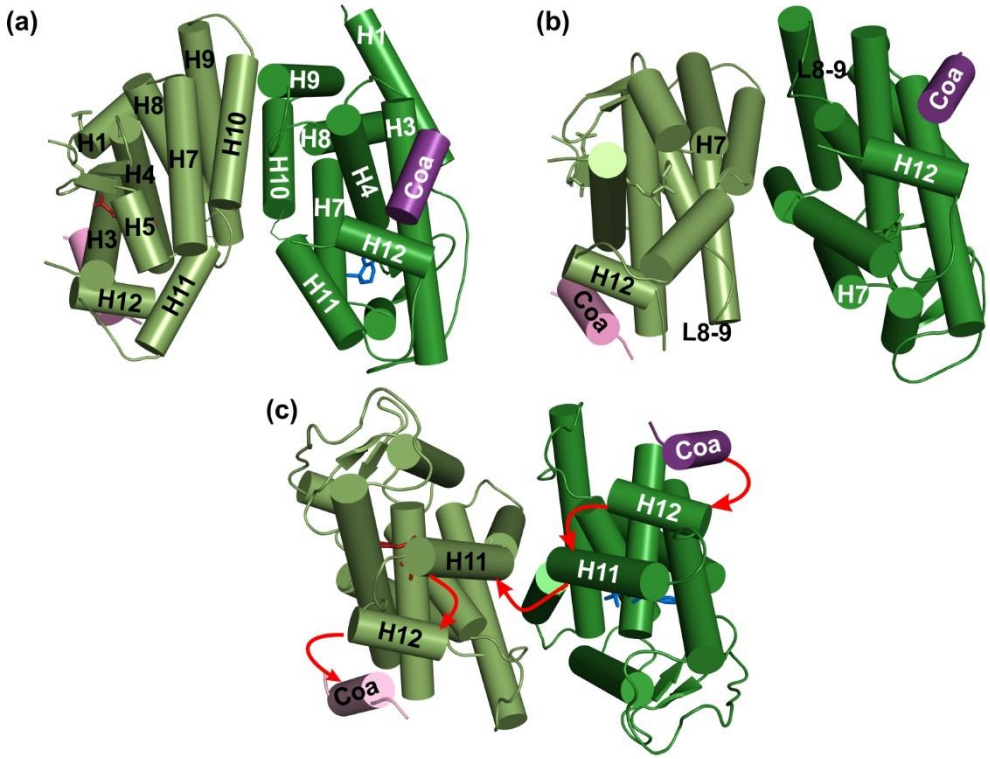


Figure 17 : Structures cristallines des LBD des récepteurs nucléaires en interaction **(a) (b)** entre RXR (vert clair) et RARβ (vert foncé) (PDB 1XDK) avec les coactivateurs (Coa) **(c)** Mécanisme proposé du contrôle allostérique dans l'homodimère de RARβ (PDB : **4DM6**).

De précédents travaux sur le récepteur à la thyroïde (TR) suggéraient déjà un rôle de l'ADN sur le contrôle allostérique de la liaison du coactivateur et de la dimérisation puisque la liaison du domaine d'interaction de SRC1 était directement influencée par la nature de la séquence des éléments de réponse de l'ADN[54]. Ce phénomène de communication allostérique entre SRC1 et l'ADN a également été décrit dans la littérature pour RXR/TR[55] et RXR/VDR[56], et il peut être étendu à d'autres récepteurs nucléaires. La combinaison des mécanismes allostériques induits par l'ADN, les protéines coactivatrices et le ligand permettent de réaliser la complexité qu'implique la réalisation de la transcription des gènes.

2. Le récepteur de la vitamine D (VDR), une cible thérapeutique de choix

a. Le récepteur de la vitamine D (VDR)

Le récepteur de la vitamine D (VDR) est un récepteur nucléaire découvert en 1974[57]. C'est une protéine de 48 kDa qui ne contient pas de domaine AF-1[58]. VDR interagit avec 2700 sites de l'ADN chez l'homme et affecte l'expression de 229 gènes[59]. Il permet de réaliser l'activité biologique de la forme active de la vitamine D_3, la dihydroxyvitamine D_3[60] (calcitriol ou $1\alpha,25\text{-}(OH)_2D_3$). Le calcitriol est le métabolite le plus actif[61] de la vitamine D, il est issu de l'hydroxylation de la 25-hydroxyvitamine D_3 par l'enzyme 1α-hydroxylase[62]. La réaction d'hydroxylation a lieu principalement dans les reins, mais elle est également réalisée dans le cerveau, le pancréas et les ganglions lymphatiques[63].

Figure 18 : Structure du récepteur de la vitamine D

VDR a d'abord été découvert dans les intestins par Haussler[57], puis dans les glandes parathyroïdiennes, les reins et les os[64]. Il a également été identifié dans beaucoup d'organes et tissus non impliqués dans l'homéostase du calcium comme les muscles, le pancréas, la prostate, le cerveau, le cœur et le système immunitaire[57, 65]. Cette diversité d'organes cibles de la vitamine D et de VDR donne une diversité d'actions biologiques à $1\alpha,25\text{-}(OH)_2D_3$ et VDR telles que la modulation de l'immunité, le contrôle de la sécrétion hormonale, la prolifération cellulaire, l'apoptose et l'angiogenèse[66]. Un dérèglement du fonctionnement de VDR et de $1\alpha,25\text{-}(OH)_2D_3$ est responsable de nombreuses maladies[67] : cancer, maladies auto-immunes, diabète… Par conséquent VDR est une cible thérapeutique très intéressante pour l'industrie pharmaceutique et le développement de nouvelles molécules à visée thérapeutique.

b. Structure de VDR

La structure du récepteur de la vitamine D diffère de la structure commune des récepteurs nucléaires par la petite taille de son domaine N-terminal (17 acides aminés) et l'absence de fonction de transactivation AF-1. Le domaine de liaison au ligand du récepteur de la vitamine D contient une interface de dimérisation[42] avec celle du récepteur X des rétinoïdes (RXR). Et bien que les structures cristallographiques de LBD et DBD de VDR (PDB : **1YNW** pour DBD, PDB : **1DB1** pour LBD) soient connues, celles du récepteur VDR et de son complexe avec RXR ne sont, à ce jour, pas déterminées. Des études SAXS[42] ont permis de caractériser la structure du complexe RXR/VDR/DR3. Les structures cristallographiques des LBD et DBD de VDR et RXR liés à l'ADN ont été ajustées dans une carte obtenue par cryo-microscopie électronique, ceci a permis de construire une enveloppe de la structure en trois dimensions du complexe RXR/VDR/DR3.

La confrontation de ces données à celles obtenues par SAXS a permis de résoudre et de caractériser la structure du complexe RXR/VDR/DR3 avec une résolution de 10-12 Å (**Fig. 19**). La structure obtenue révèle une architecture en forme de L (**Fig. 19a**) dont l'une des extrémités a une partie en forme de V (**Fig. 19b**). Cette partie en forme de V (40x50x70 Å) a été assimilée au dimère LBD, alors que le coté le plus long de la partie en forme de L (25x35x32 Å) a été caractérisé comme étant le DBD lié à un segment d'ADN[42]. Cette structure montre que le complexe a une architecture ouverte, dont l'une des unités de l'hétérodimère de LBD est orientée à la perpendiculaire de la partie DBD-ADN (**Fig. 19a**). Les LBD de VDR et RXR peuvent être facilement identifiés dans la carte obtenue par cryo-microscopie par les caractéristiques asymétriques du LBD de VDR. Le segment VDR comprend une hélice H4, une boucle L8/L9, RXR, elle, comprend l'hélice H7 et la région avec les hélices H2, H3 et un feuillet β (marqué par une étoile sur les **Fig.19b** et **Fig.19c**).

Outre l'hétérodimérisation, la LBD de VDR peut aussi homodimériser[68] et se lier à DR3 (comme l'hétérodimère RXR/VDR). L'arrangement du récepteur nucléaire sur DR3 est similaire qu'il soit homodimère VDR/VDR ou hétérodimère[68] RXR/VDR.

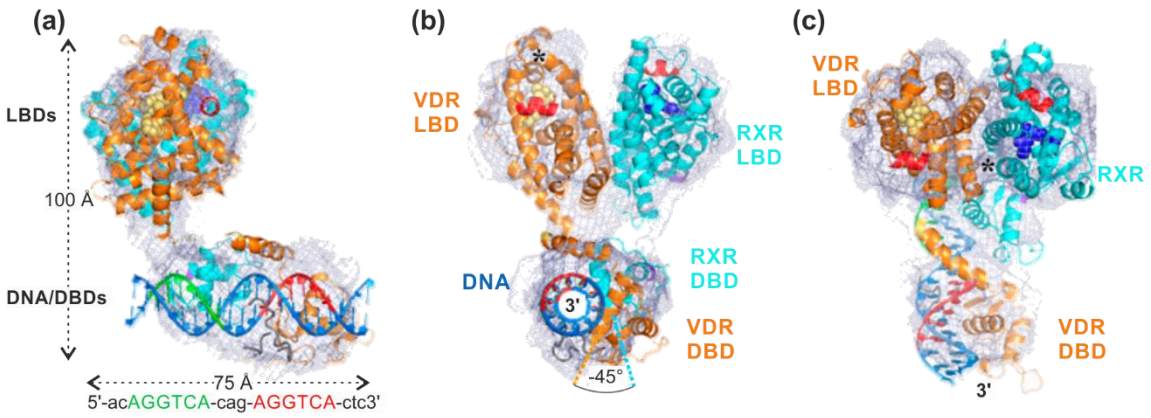


Figure 19 : Structure du complexe RXR/VDR/DR3 : **(a)** coupe de profil de la structure révélant la forme en L du complexe, les LBD de RXR et VDR sont orientés perpendiculairement à l'ADN. **(b) (c)** coupe de face et de haut du complexe et observation de la forme en V des extrémités.

c. Mécanisme d'action et voie de signalisation de l'activation de VDR

Moins d'un pourcent du calcitriol se lie à la DBP (*vitamin D Binding Protein*, protéine spécifique liant la vitamine D dans le sérum) et l'albumine, et est ainsi diffusé dans les cellules et se lient à VDR**.** Lorsque VDR est libre (non lié au calcitriol), il est en équilibre entre le noyau et le cytoplasme **(Fig. 20)**. Le calcitriol entre dans la cellule, se lie à VDR via des contacts de Van der Waals avec H11 de VDR[69]. Cette liaison va provoquer une modification interne du récepteur qui va ensuite entrer rapidement dans le noyau via une liaison entre la partie charnière *Hinge*, contenant le signal de localisation nucléaire du récepteur nucléaire et des protéines de transport nucléo-cytoplasmiques[70]. VDR et RXR vont simultanément se lier à l'ADN par des éléments de réponse et dimériser **(Fig. 20).** Des études d'échange hydrogène-deutérium (HDX) réalisées par Zhang *et al.* montrent la présence du mécanisme de communication allostérique entre RXR et VDR qui rend l'hélice H12 (AF-2) disponible pour recruter des protéines coactivatrices[69] (SRC, DRIP) et des protéines de modification de la chromatine (CBP/P300) simultanément pour décondenser la structure de la chromatine en acétylant les histones. Une fois l'acétylation réalisée, SRC se délie de VDR[71] et d'autres protéines coactivatrices, les protéines DRIP, se lient à la partie AF-2 de VDR également via un motif LXXLL et interagissent avec l'ARN polymérase II pour initier la transcription[71,72]. VDR code de nombreux gènes[72c] dont CYP24A1 (codant 25-hydroxylase), BGLAP (ostéocalcine exprimée dans les ostéoblastes), CDKN1A (inhibiteur p21 de kinases dépendantes des cyclines CDK) et SPP1 (ostéopontine). Il a également été suggéré que VDR serait un récepteur membranaire, mais à ce jour aucune donnée n'a été rapportée dans ce sens.

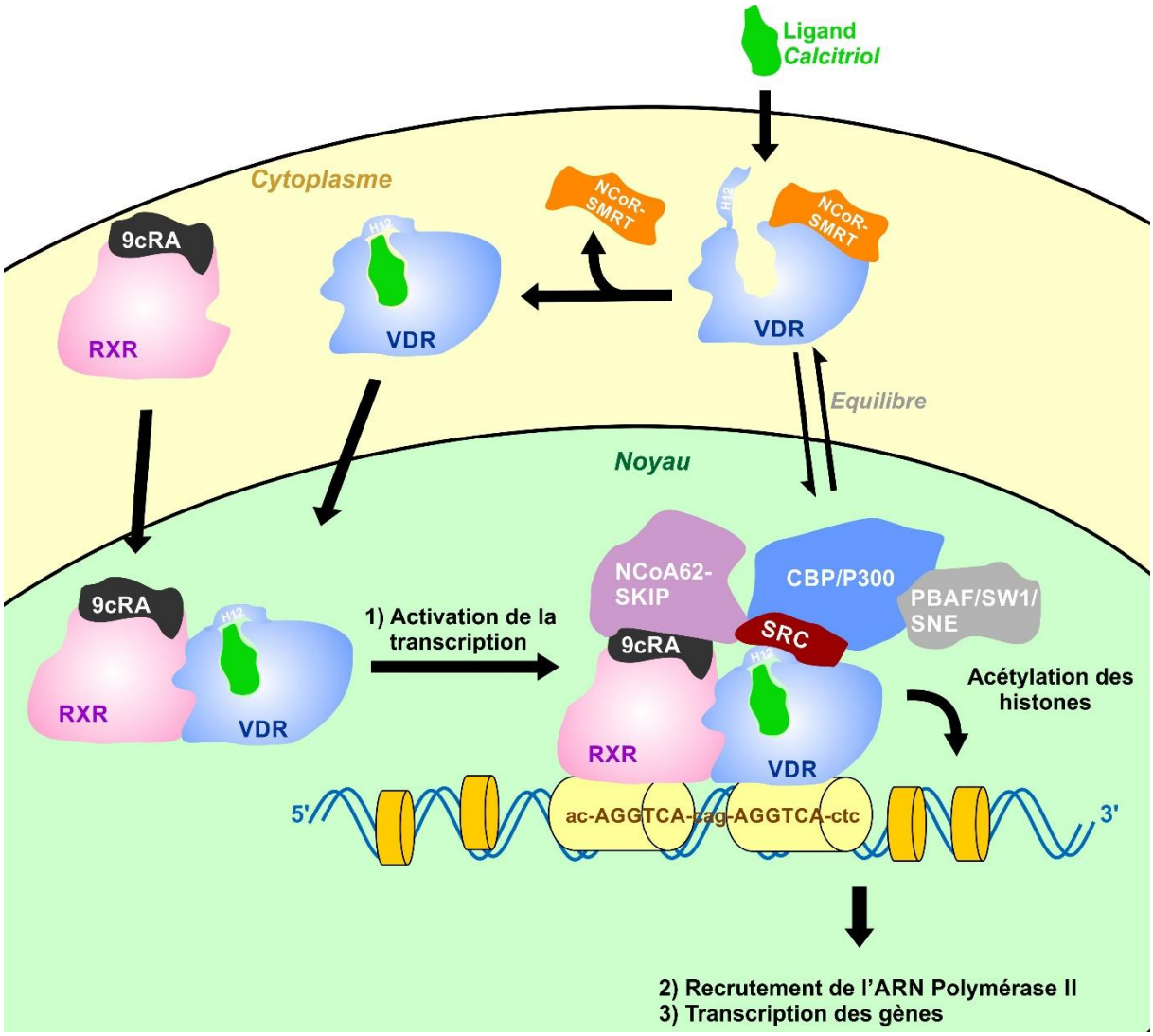


Figure 20 : Voie de signalisation simplifiée de l'activation du récepteur de la vitamine D VDR

3. Une famille de protéines coactivatrices communes : les protéines SRC

a. Généralités sur cette famille de coactivateurs

A ce jour, plus de 250 protéines corégulatrices ont été décrites dans la littérature pour les récepteurs nucléaires[73]. Comme décrit en II-1-*c*, la famille des coactivateurs de récepteurs stéroïdiens (SRC) est la plus grande famille de protéines coactivatrices. Elle comporte trois isoformes[37] : SRC1, SRC2 et SRC3. Ceci suggère que les SRC possèderaient des fonctions et des spécificités différentes les unes des autres pour des récepteurs nucléaires donnés. Parmi ces trois protéines, SRC1 a été identifié comme étant un des coactivateurs communs à beaucoup de récepteurs nucléaires[51], et de nombreuses études ont été réalisées sur ce coactivateur. Des études de double-hybride ont montré que SRC1 interagit à la fois avec des

récepteurs nucléaires stéroïdiens (GR, ER, AR, PR) et non-stéroïdiens[53] (RAR, RXR, TR, VDR).

b. SRC et impact sur la santé

Des études génétiques sur des souris ont permis de mettre en évidence que les SRC ont un rôle distinct dans la reproduction, les cancers et le métabolisme de l'énergie[74]. Ces études révèlent la présence d'un lien étroit entre les SRC et l'existence d'anomalies génétiques chez l'homme. Il a été rapporté dans la littérature que dans le cas de la maladie de Von Gierke, les phénotypes des SRC2 humains et murins sont similaires[75]. Bien que SRC1, SRC2 et SRC3 servent à activer la transcription et qu'ils permettent de stabiliser le complexe préformé via un ciblage du facteur de transcription basale, il a également été montré que les SRC n'étaient pas tous présents aux mêmes endroits dans le corps et qu'ils n'ont ainsi pas les mêmes fonctions biologiques. SRC1 et SRC3 ont une activité d'acétylation de l'histone en modifiant l'histone et les interactions histone-histone, ils permettent de décondenser le brin d'ADN pour réaliser la transcription. Ils sont également présents dans le cerveau et sont impliqués dans divers aspects du fonctionnement cérébral[73]. SRC3 est également exprimé à haute dose dans le cœur et le placenta[40] et SRC2 est exprimé dans le foie. Il fonctionne comme un régulateur de l'expression de la glucose-6-phosphatase, le coactivateur permet de réguler l'absorption des graisses et l'accumulation d'énergie dans le corps[74c]. Des études sur les souris ont montré le lien entre la présence de ce coactivateur et l'obésité chez les souris[73]. En absence de la protéine SRC2 dans leur organisme, les souris sont protégées de l'obésité causée par un régime riche en graisse[76].

Dû à leurs capacités de régulation de l'expression des gènes via les hormones et les facteurs de croissance, les corégulateurs peuvent être des promoteurs de pathologies lorsque le système d'expression se dérègle durant la croissance cellulaire ou la différenciation[77]. Différents types de cancer peuvent être des exemples de ces pathologies via une prolifération cellulaire excessive et la formation de métastases[74b].

De nombreux exemples de développement d'analogues agonistes ou antagonistes aux petites molécules se liant naturellement aux récepteurs nucléaires (tel que le calcitriol à VDR) ont été rapportés dans la littérature. Les corégulateurs, dont les coactivateurs, ne bénéficient pas du même intérêt de recherche car ils sont considérés comme trop complexes dans leur sélectivité et dans leur fonctionnement pour être des cibles thérapeutiques prometteuses[78].

c. Rôle du motif LXXLL responsable de l'interaction et étude de la zone d'interaction avec le récepteur nucléaire

Osz *et al.* se sont intéressés au mode de liaison de la protéine coactivatrice SRC1 aux récepteurs nucléaires en utilisant une portion peptidique en interaction **SRC1 NR2**. L'étude de la structure cristalline de la LBD de l'homodimère RARβ lié au peptide leur a permis de montrer que le motif LXXLL contenu dans ce peptide forme une hélice α amphipatique qui interagit avec une poche hydrophobe formée en position C-terminale de H3, L3-4, H4 et H12 de RARβ (**Fig. 21**). Ce mode d'interaction est similaire pour les autres récepteurs nucléaires dont VDR[73].

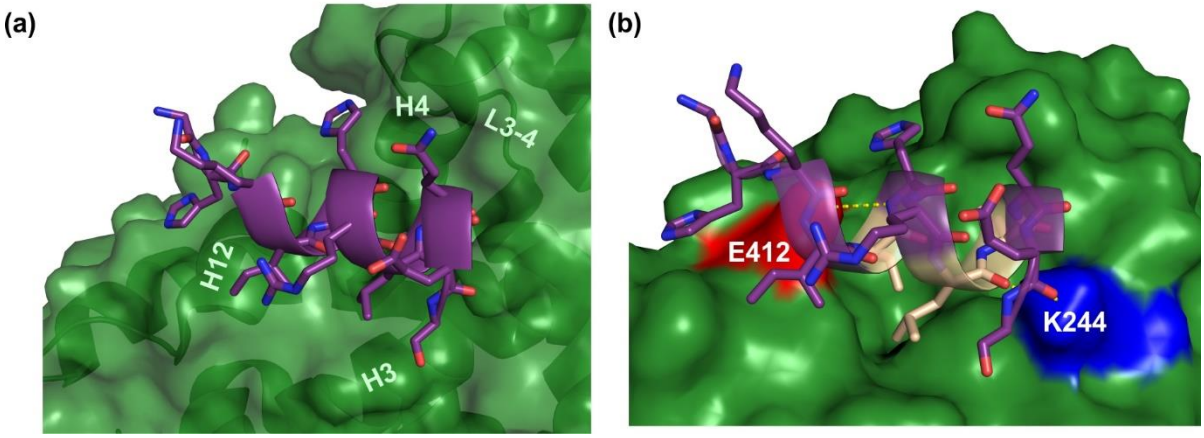


Figure 21 : (a) Hélice α amphipatique du motif LXXLL de **SRC1 NR2** (en violet) en interaction avec la poche hydrophobe de RARβ (en vert), les résidus d'acides aminés leucines sont représentés en beige ; **(b)** identification de la pince chargée entre **SRC1 NR2** et RARβ (PDB :**4DM6**).

Leur étude permet de révéler la présence d'une pince chargée via la formation de liaisons hydrogène entre les fonctions amide de la chaîne principale du peptide **SRC1 NR2** et les résidus d'acides aminés K244 de H3 et E412 de H12 (**Fig. 21b**). La présence de cette pince chargée est également observable pour l'homodimère de l'hormone stéroïdienne ERα et VDR[73]. Il y a ainsi une analogie et une forte similarité des récepteurs nucléaires entre eux via leurs modes de liaison au ligand, à l'ADN et aux protéines coactivatrices.

Les SRC possèdent plusieurs motifs LXXLL responsables de l'interaction avec le récepteur[79]. Ce motif, aussi appelé NR box (*nuclear receptor box*) est nécessaire et suffisant pour former la liaison du coactivateur à un récepteur donné[37], il correspond au concept de hot-spot. Chaque SRC contient trois motifs de cette NR box, excepté SRC1 qui en contient une supplémentaire en C-terminal de la protéine.

Les zones responsables de l'interaction avec le récepteur nucléaire étant identifiées et localisées, il est alors supposé que les surfaces de liaison des motifs LXXLL aux récepteurs

sont différentes. Des modifications dans la nature des résidus entourant ces motifs conservés déterminent l'affinité de ces coactivateurs à la poche hydrophobe du domaine AF-2 des récepteurs nucléaires[80]. Cela a été démontré par des analyses de mutation des NR box de SRC1 et SRC2, qui ont révélé une spécificité d'interaction où différents récepteurs nucléaires nécessitent différents hot-spots pour interagir avec le co-activateur[81]. Par conséquent, les résidus entourant les hot-spots sont importants et contribuent à la sélectivité et à la spécificité de l'interaction récepteur nucléaire/SRC.

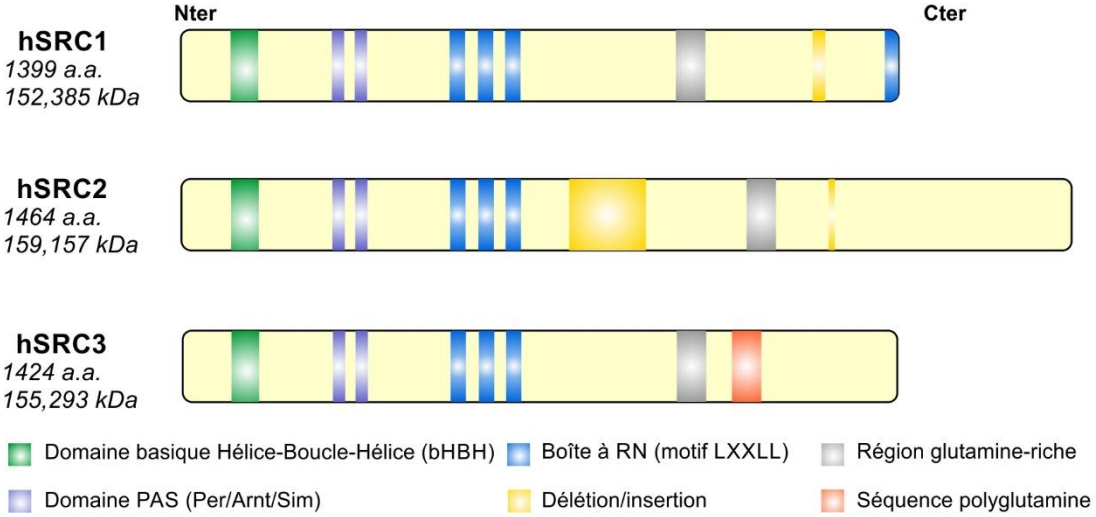


Figure 22 : Structures des protéines SRC humaines.

Les structures complètes des protéines SRC n'ont pas été résolues à haute résolution. La partie N-terminale contient un domaine Hélice-boucle-Hélice (HBH) /PAS (Per/Arnt/Sim : Per, *period*, protéine intervenant dans le rythme circadien ; Arnt protéine nucléaire de translocation du récepteur d'aryle hydrocarboné, *Aryl hydrocarbon receptor nuclear translocator* ; et protéine Sim, *Single-minded*, intervenant dans le développement neuronal). Ce domaine PAS est impliqué dans les interactions homo- et hétérodimère entre les protéines qui contiennent ce motif[82]. Les protéines SRC contiennent également une partie C-terminale qui active la transcription, un domaine de liaison au récepteur et un, voire deux, domaine d'activation autonome[38a]. Leur différence en termes de nombre et de nature des résidus indique une diversité fonctionnelle qui serait décisive pour déterminer leurs interactions avec les récepteurs nucléaires, mais également avec d'autres corégulateurs de transcription[37]. En effet les récepteurs nucléaires et les coactivateurs interagissent avec d'autres coactivateurs de transcription comme CBP/P300 et PCAF[83], un large complexe protéique semble nécessaire pour activer la transcription d'un gène cible.

4. Inhibiteurs de VDR décrits dans la littérature

Le développement de molécules ciblant l'action de VDR est en plein essor. Diverses approches ont été décrites dans la littérature : développement de petites molécules antagonistes au calcitriol, peptidomimes et petites molécules ciblant l'interaction VDR/coactivateurs.

a. Antagonistes de VDR

La stratégie des antagonistes au calcitriol repose sur le design de petites molécules avec un squelette sécostéroïde (**Fig. 23**). Parmi eux, *TEI-9647*[84], *ZK168281*[85], *ZK159222*[84, 85b, 86], *AD47*[86,87] et des dérivés branchés[88] ont été rapportés dans la littérature. Ces antagonistes ont une bonne activité d'inhibition de l'activité de transcription de VDR avec des IC_{50} de l'ordre du nano molaire. Malgré une bonne activité d'inhibition de la transcription de VDR, certains de ces antagonistes ont été identifiés comme promoteurs de dysfonctionnement du récepteur nucléaire. Certains d'entre eux ont également tendance à montrer un certain degré d'agonisme. L'ensemble de ces phénomènes n'a pas permis leur développement en thérapeutique.

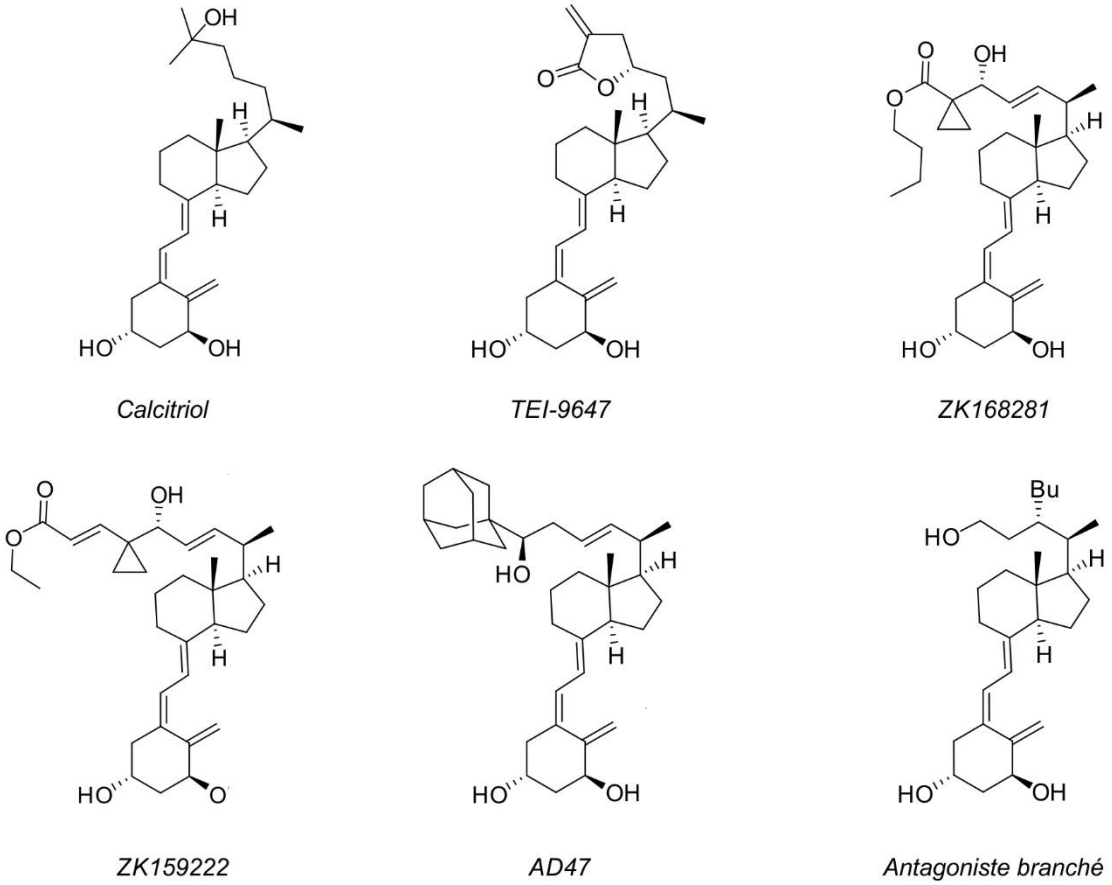


Figure 23 : Structures du calcitriol et d'inhibiteurs antagonistes sécostéroïdes.

Une autre famille de molécules antagonistes à VDR non-sécostéroïdes ont été développées[89] : les analogues de *GW0742* (**Fig. 24**).

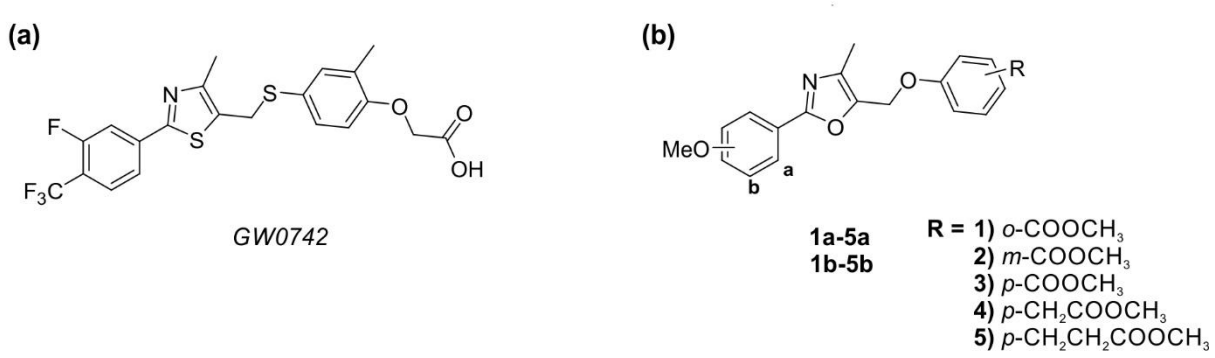


Figure 24 : Structures de *GW0742* **(a)** et du squelette du dérivé oxazole **(b)**

L'équipe de Nandhikonda *et al.* a identifié par *High Throughput Screening* (HTS) *GW0742* comme un nouvel antagoniste non-sécostéroïde de VDR en 2013 (**Fig. 24a**). *GW0742* est à l'origine un agoniste de PPARδ [89a, 90] et il inhibe la transcription de VDR avec une IC_{50} de 20,7 µM. Puis Arnold *et al.* ont développé une série d'analogues de *GW0742* en introduisant un noyau oxazole[89b] (**Fig. 24b**). Les oxazoles contenant une fonction ester en ortho ou para sont spécifiques à l'inhibition de l'activité issus de l'interaction VDR/SRC1 avec une IC_{50} de 6,7 µM pour le plus actif **5b**. L'inhibition de l'activité de VDR/SRC1 est plus forte lorsque l'ester est en position méta (IC_{50} = 2,5 µM pour **2a**) et en para lorsque la chaîne est plus longue (IC_{50} = 0,66 µM pour **4b**). L'équipe s'est ensuite intéressée à la spécificité de leurs antagonistes pour VDR en étudiant l'inhibition de la transcription de leurs composés avec d'autres récepteurs nucléaires ER, PPAR, RXR et TR. L'inhibition étant plus forte pour VDR, ils ont pu conclure que leurs analogues oxazoles de *GW0742* sont sélectifs[89a] de VDR mais avec des affinités modérées.

b. Inhibiteurs de l'interaction VDR/corégulateurs

Le développement de molécules antagonistes à VDR est une première stratégie pour inhiber la transcription de VDR. Cependant, une stratégie innovante visant à inhiber l'interaction entre le récepteur nucléaire et le corégulateur est en plein essor. Diverses approches ont été considérées pour inhiber l'interaction : de la petite molécule aux peptidomimes du corégulateur.

i. Des petites molécules pour inhiber l'interaction VDR/corégulateurs

Le développement de petites molécules ciblant l'interaction VDR/coactivateur est une première approche pour l'inhibition de l'interaction avec le récepteur nucléaire. En 2010, Mita *et al.* ont développé un inhibiteur réversible de l'interaction VDR/coactivateur (SRC, DRIP) avec une molécule de type benzodiazépine[91] (**Fig. 25a**). Celle-ci a la particularité de mimer les pharmacophores du motif d'interaction LXXLL de la protéine coactivatrice DRIP. Les positions i, i+3 et i+4 des leucines de l'hélice α amphipatique sont mimées par des chaînes alkyles branchés. La pince chargée entre l'inhibiteur et les résidus E412 et K244 de VDR est conservée avec la présence de l'amine en position 8 et de l'amide en position 1. L'inhibition de la transcription de VDR par **6** a été déterminée par TR-FRET (transfert d'énergie entre molécules fluorescentes, *Fluorescene Resonance Energy Transfer*) avec une IC_{50} de 17 µM (**Fig. 25b**). Ils ont tenté d'optimiser **6** en réalisant une étude de relation structure-activité (SAR) et ont développé le composé **7** qui a une IC_{50} de 14 µM[92] (**Fig. 25**).

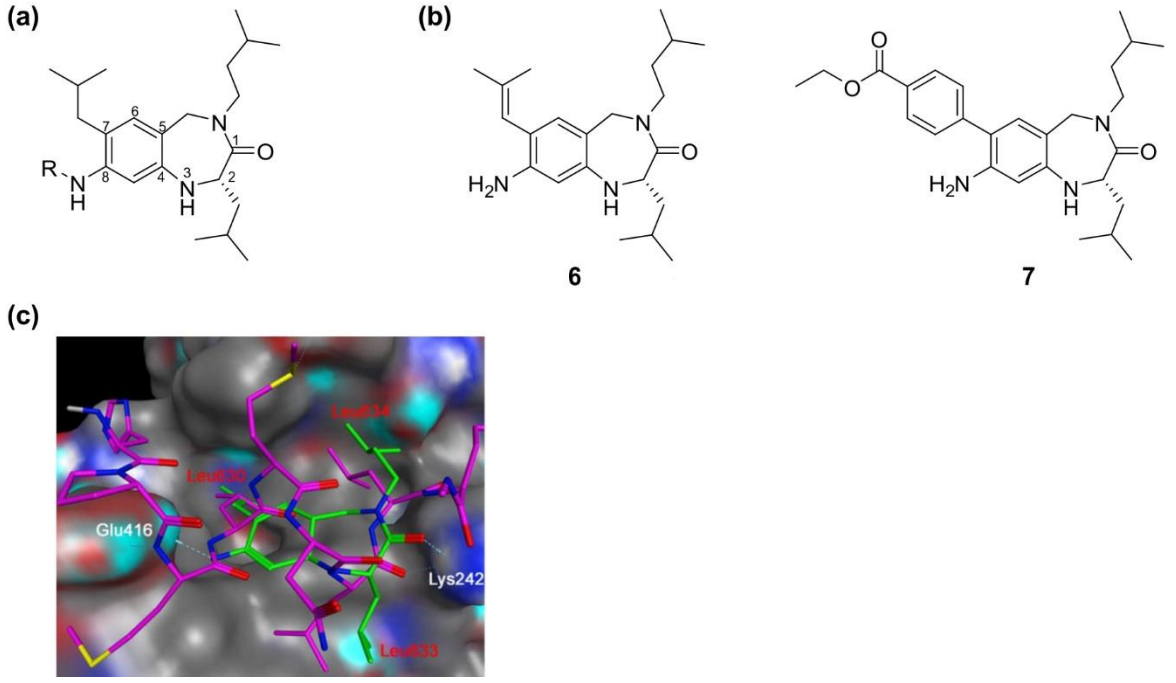


Figure 25 : Structures des inhibiteurs à VDR dérivé de benzodiazépine : **(a)** squelette général, **(b)** structures des inhibiteurs **6** et **7**. **(c)** Docking de l'inhibiteur **6** (en vert) superposé à la structure cristalline de VDR (en gris)/DRIP205 (en violet).

Une autre famille d'inhibiteurs de l'interaction VDR/SRC a été développée par Arnold *et al.*[93] : les 3-indolylméthanamines (**Fig. 26**). En 2012, ils ont identifié par HTS le premier inhibiteur irréversible de l'interaction VDR/coactivateur. Sur 275 000 composés, 140 inhibiteurs ont été identifiés avec une activité cellulaire, et, parmi eux la famille des

3-indolylméthanamines[93] (**Fig. 26a**). Une étude SAR sur ce squelette a permis d'identifier 31b (**Fig. 26b**) comme l'antagoniste le plus actif avec une IC_{50} de 4,2 µM et le plus sélectif de l'interaction VDR/SRC2-3. L'irréversibilité a été confirmée par une relation linéaire entre l'énergie libre et le taux d'inhibition des 3-indolylméthanamines portant des substituants électroniques différents[93]. Un second inhibiteur sélectif à l'interaction VDR/SRC2-3 a été identifié par une étude SAR, *PS121912* (**Fig. 26b**) qui inhibe l'interaction avec une IC_{50} de 590 nM [94].

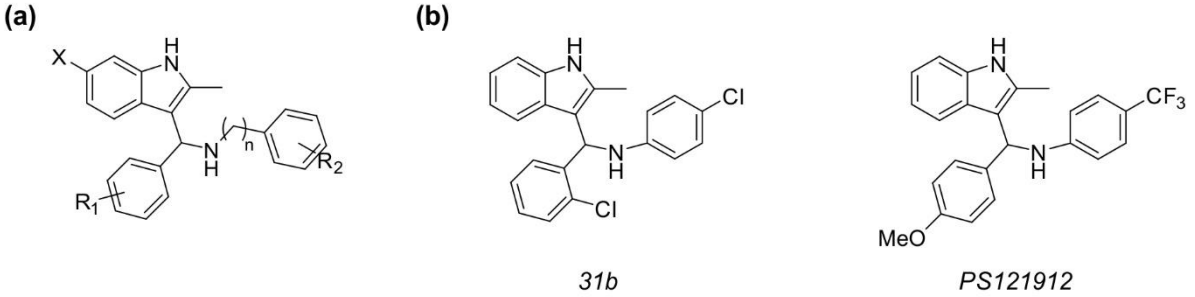


Figure 26 : Structures des inhibiteurs à VDR 3-indolylméthanamine : **(a)** squelette général, **(b)** structures des inhibiteurs *31b* et *PS121912*.

ii. Développement d'inhibiteurs peptidiques synthétiques

Une seconde approche d'inhibition de l'interaction VDR/coactivateur repose sur le développent d'inhibiteurs peptidiques basés sur la séquence des coactivateurs de VDR. Les premiers exemples de peptides synthétiques inhibiteurs de l'interaction VDR/SRC ont été décrit par McDonnell *et al.*[95] (**Tableau 2**). Les peptides **C33**, **D47**, **EBIP41** et **EBIP44** ont été obtenus par phage-display et ils contiennent tous le motif d'interaction LXXLL.

	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12	**Cible**
C33	H	V	E	M	H	P	L	L	M	G	L	L	M	E	S	Q	W	G	A	VDR
D47	H	V	Y	Q	H	P	L	L	L	S	L	L	S	S	E	H	E	S	G	VDR
EBIP41	R	R	D	D	F	P	L	L	I	S	L	L	K	D	G	A	L	S	Q	VDR
EBIP44	Y	G	L	K	M	S	L	L	E	S	L	L	R	E	D	I	S	T	V	VDR
F6	G	H	E	P	L	T	L	L	E	R	L	L	S	G	T	S	V	A	E	RXR
PI1	L	S	E	T	H	P	L	L	W	T	L	L	S	S	E	G	D	S	M	VDR
PI2	M	Q	E	R	F	P	M	L	W	D	L	L	D	L	P	S	P	T	S	VDR
PI3	L	G	E	S	H	P	L	L	M	Q	L	L	T	E	N	V	G	T	H	VDR

Tableau 2 : Séquences des peptides inhibiteurs de l'interaction VDR/coactivateurs. Les résidus en jaune correspondent aux résidus d'acides aminés leucine du motif d'interaction LXXLL.

Le peptide **F6** qui est sélectif à RXR a révélé une activité d'inhibition pour la transcription de VDR, ce qui permet de confirmer la présence d'un phénomène de

transactivation entre VDR et RXR. En 2007, un second phage-display a permis d'identifier trois nouvelles séquences peptidiques contenant le motif d'interaction LXXLL (**PI1**, **PI2** et **PI3**) comme inhibiteurs de VDR[96]. Ces études par phage-display ont contribué à la validation de l'interaction entre VDR et ses coactivateurs comme cible pour la modulation de la transcription et ont permis d'identifier les premiers peptides inhibiteurs sélectifs de l'interaction VDR/corégulateurs.

iii. Développement d'inhibiteurs peptidomimes avec des stapled peptides

La troisième approche d'inhibition de l'interaction VDR/coactivateur décrite dans la littérature repose sur le développement de peptidomimes. Demizu *et al.*[97] ont utilisé la technologie des stapled peptides pour stabiliser la structure en hélice d'une courte séquence peptidique d'un motif LXXLL du coactivateur DRIP en introduisant une chaîne hydrocarbonée. Les acides aminés X sont une O-allyl-homosérine et un monomère Aib. Ils ont développé deux séries de stapled peptides où deux hydroxyles sont introduits sur le lien hydrocarboné (**Fig. 27**). Ils ont synthétisé une dizaine de stapled peptides mimes de DRIP2052-2 et DPI-08 a été identifié comme inhibiteur de VDR/DRIP205-2 avec une IC_{50} de 3,2 µM. Des études par dichroïsme circulaire ont permis de confirmer que leurs stapled peptides conservent leur structure en hélice α droite. A ce jour, aucune structure cristalline des complexes peptidomimes stapled peptides/VDR n'a été rapportée dans la littérature. Les seules structures cristallines disponibles de peptidomimes stapled peptides en interaction avec un récepteur nucléaire sont ceux développés par Spletz et al[98]. Ces peptidomimes cycliques inhibent l'interaction ERα/SRC avec une IC_{50} supérieure à 90 nM.

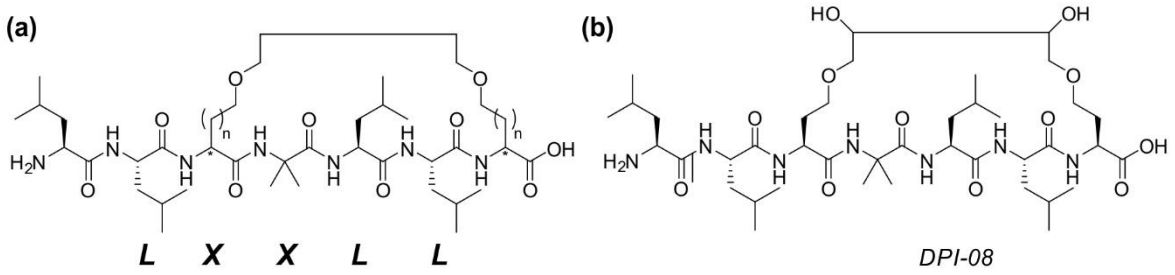


Figure 27 : Structure des inhibiteurs antagonistes à VDR **(a)** squelette des stapled peptides mimes de DRIP205-2, **(b)** structure de l'inhibiteur DPI-08.

iv. Développement de peptidomimes à base de foldamères d'oligourée hybrides

Ces derniers exemples d'inhibition de l'interaction VDR/coactivateur montrent qu'il est possible d'utiliser des technologies de stabilisation d'hélice α pour mimer la partie peptidique d'une protéine coactivatrice responsable de l'interaction avec le récepteur nucléaire. A ce jour, seuls ces exemples contenant une modification ponctuelle de la séquence ont été rapportés dans la littérature. Il est ainsi possible d'imaginer une modification d'une portion de la séquence peptidique par une séquence non naturelle. L'utilisation de la technologie de foldamère mixte oligoamide/oligourée a été la méthode considérée dans mon projet de thèse pour le développement de peptidomimes inhibiteurs de l'interaction VDR/SRC1-2 via une étude de relation structure-affinité.

III. Mime de zones étendues et vastes : de la structure secondaire aux structures tertiaires et quaternaires

Les structures secondaires en hélice sont fréquemment présentes et impliquées dans l'interaction entre protéines[5]. Cependant, de nombreuses interfaces d'interaction protéine-protéines montrent un mode d'interaction plus complexe qu'une simple hélice. Des motifs hélicoïdaux de structures tertiaires et quaternaires de type *coiled coils* (superhélices en français) ont été identifiées[99], ces motifs sont aussi nommés *bundles* d'hélices. Cela suggère la possibilité de concevoir et de développer des mimes de structures tertiaires et quaternaires comme candidats d'une nouvelle classe d'inhibiteurs d'IPP[100]. Il est estimé que 5% des séquences de protéine de transcription codent des domaines sous forme de *coiled coils* hélicoïdaux[101].

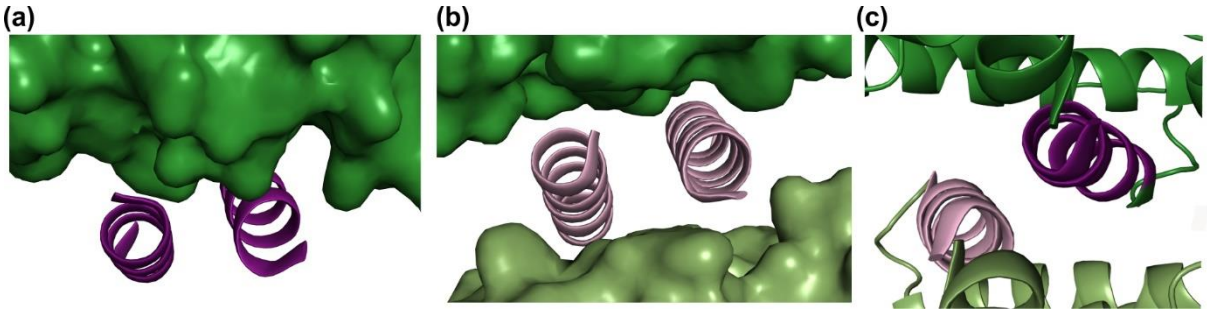


Figure 28 : Classification des interactions protéine-protéine faisant intervenir un dimère d'hélices. **(a)** Cas 1 où le dimère d'hélices (en violet) interagit avec la protéine (en vert). **(b)** Cas 2 où un dimère d'hélices interagit avec deux protéines différentes. **(c)** Cas 3 où le dimère d'hélices (en violet et rose pâle) se forme à l'interface entre les deux protéines.

En 2015, Arora s'est intéressé à cette problématique et a développé une base de données[99] qui répertorie toutes les cibles d'IPP impliquant plus qu'une simple hélice : la base de données de dimères hélicoïdaux aux interfaces protéine-protéine (*dimer at protein-protein interfaces database DippDB*). Trois cas d'interaction protéine-protéine ont été identifiés (**Fig. 28**) : interaction d'un dimère d'hélices avec une protéine (cas 1), interaction d'un dimère d'hélices avec deux protéines (cas 2) et formation d'un dimère d'hélices à l'interface par interaction des deux protéines (cas 3).

1. Une structure quaternaire particulière : le *coiled coil*

a. Origine de l'assemblage quaternaire sous forme de coiled coil

La caractérisation structurale de la kératine α initiée par Pauling et Corey en 1951[102] a permis à Francis Crick de définir la structure de cette protéine fibreuse comme une superhélice, qu'il nomma *coiled coil*[103] (**Fig. 29**).

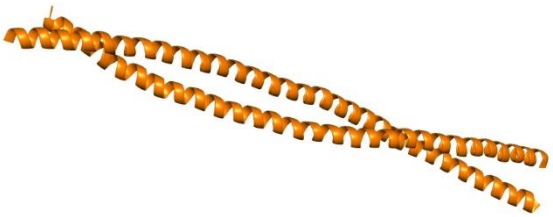


Figure 29 : Structure cristalline de la kératine assemblée en bundle sous forme de dimère d'hélices (PDB : **4ZRY**).

Les *coiled coils* consistent en des assemblages supramoléculaires d'au moins deux hélices α amphipatiques qui s'enroulent les unes autour des autres pour former une structure sous forme de fibre[103,104]. L'analyse des séquences naturelles des *coiled coils*, l'émergence des structures cristallines disponibles ainsi que les outils bio-informatiques ont permis de mieux comprendre la relation séquence-structure de ces systèmes et a mis en évidence les principes gouvernant l'auto-assemblage de ces structures quaternaires.

b. Principes gouvernant l'assemblage des coiled coils

Les hélices peuvent être orientées de façon parallèle ou antiparallèle avec des chaînes latérales identiques (homo-) ou différentes (hétéro). L'effet hydrophobe gouverne le repliement et l'auto-assemblage des protéines. Francis Crick proposa un mécanisme d'interaction de type *knobs-into-holes* entre deux hélices amphipatiques d'un *coiled coil*[103-104]. Les chaînes latérales

d'une hélice sont enfouies dans la cavité formée par quatre chaînes latérales de la seconde hélice. Le décalage entre le nombre de résidus par tour d'hélice (3,6) et la périodicité des résidus hydrophobes au sein des séquences de type *coiled coils* conduisent les deux hélices à s'enrouler l'une autour de l'autre afin de maximiser les contacts entre les surfaces hydrophobes au cœur de l'assemblage. Les séquences *coiled coils* contiennent une répétition régulière de sept résidus (heptade), notée *abcdefg* (**Fig. 30**). Les résidus hydrophobes sont situés aux positions *a* et *d* et forment le cœur hydrophobe à l'interface entre les deux hélices. La nature des résidus à ces positions clés influence la stabilité, l'état d'oligomérisation, la sélection de la protéine partenaire et l'orientation des hélices entre elles. Des résidus aliphatiques (Ile, Leu, Met, Val) sont le plus souvent présents à ces positions. Des études in silico effectuées par Arora et al[99] confirment que le cœur hydrophobe des *coiled coils* est naturellement riche en leucine et isoleucine. Le volume et les contraintes stériques des résidus aromatiques hydrophobes (Phe, Tyr, Trp) expliquent qu'ils soient peu ou non présents à ces positions dans les séquences naturelles[105]. Les résidus aux positions *b c e f g* sont polaires et/ou promoteurs d'hélicité (Ala, Glu, Lys, Gln). La présence de ponts salins, d'interactions électrostatiques et une complémentarité de charges des résidus situés aux positions périphériques *e* et *g* jouent un rôle significatif pour la sélection de l'hélice partenaire[106] et l'auto-assemblage des superhélices. Les interactions inter-hélices à ces positions impliquent souvent des résidus de type lysine et acide glutamique. Les acides aminés de type glycine, proline et cystéine sont peu présents dans les séquences naturelles[105-107]. Leur utilisation lors de la conception de superhélices est limitée sauf si des éléments spécifiques sont nécessaires à la séquence. Par exemple, la cystéine permet de former des ponts disulfures pour relier de façon covalente des oligomères. Le cœur hydrophobe et les interactions ioniques permettent de stabiliser la structure quaternaire. Les séquences des hélices donnant lieu à un assemblage de type coiled coil sont constituées d'au moins trois répétitions du motif en heptade, soit 21 acides aminés, pour garantir le bon repliement et la stabilité de l'assemblage[108].

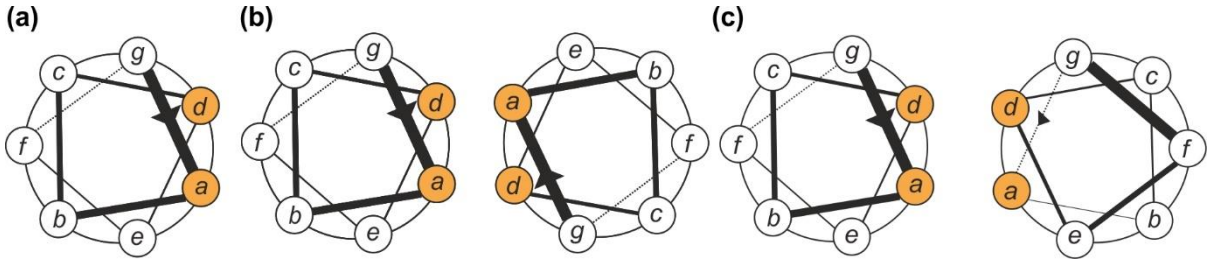


Figure 30 : Diagramme de roue hélicoïdale : **(a)** représentation de l'heptade *abcdefg* avec les résidus hydrophobes aux positions *a* et *d* (en orange). Représentation du cœur hydrophobe formé par les résidus hydrophobes a et *d* pour des hélices **(b)** parallèles et **(c)** antiparallèles d'un *coiled coil* sous forme de dimère d'hélices.

2. Conception rationnelle de structures quaternaires

Deux approches ont été décrites dans la littérature pour la conception rationnelle de structures quaternaires artificielles à base de *coiled coils* : l'approche descendante « top-down » dans laquelle la séquence d'un peptide naturel connu pour former une structure en *coiled coil* est utilisée comme point de départ pour construire une séquence non naturelle hybride foldamère/peptide qui s'auto-assemble en *bundles*, ou une approche ascendante « bottom-up », dans laquelle l'auto-assemblage est basé sur une conception *de novo* de foldamères hélicoïdaux amphipatiques avec une face hydrophobe et hydrophile.

a. Approche descendante top-down pour la conception de foldamères hybrides capables de s'assembler en bundles

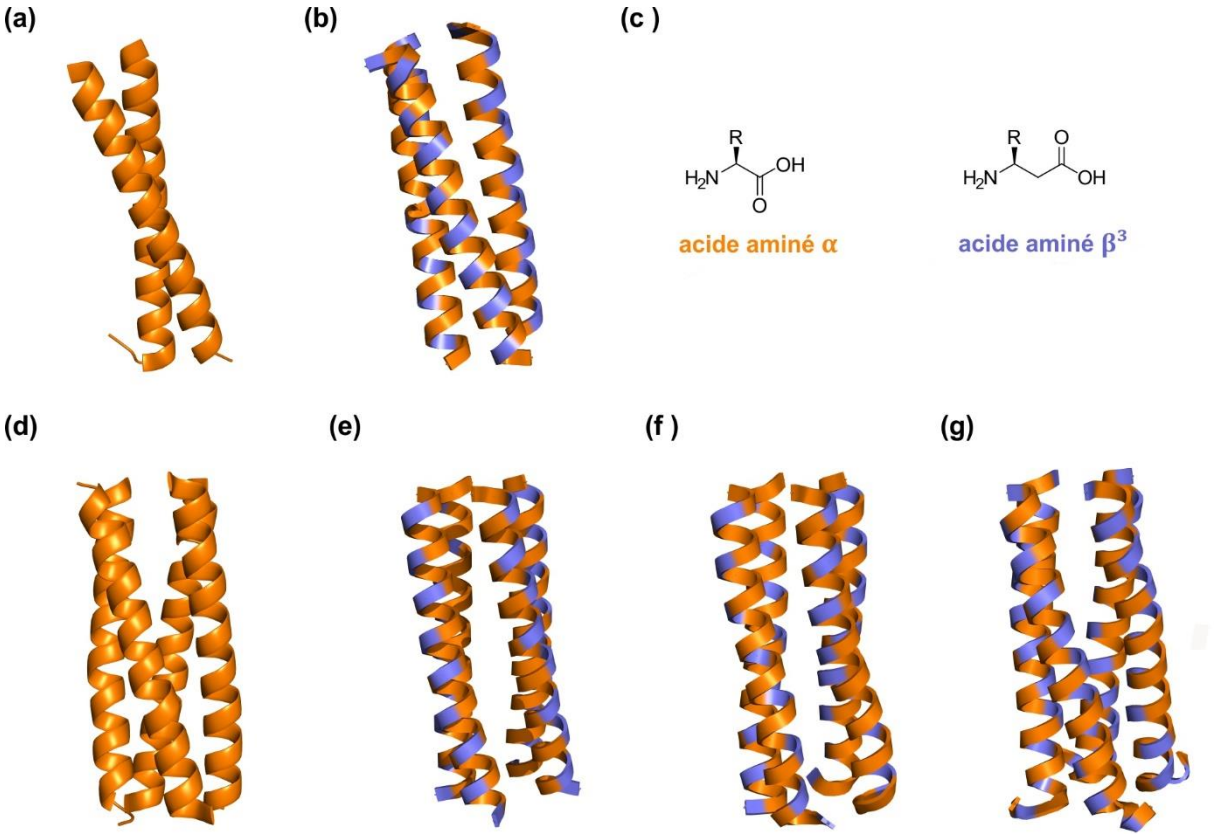


Figure 31 : Structures cristallines de *coiled coils* **(a)** d'une séquence naturelle de GCN4-p1 (PDB : **4DM4**), **(b)** du peptidomime α/β de GCN4-p1 (PDB : **2XOJ**). **(c)** Acide aminé α naturel et acide aminé β^3. Structures cristallines de *coiled coil* **(d)** d'une séquence naturelle du peptide mutant GCN4-pLI (PDB : **1GCL**) et de peptidomimes contenant divers motifs de substitution : **(e)** ααβαααβ (PDB : **2OXK**), **(f)** αααβ (PDB **:3C3F**) et **(g)** αβ (PDB **: 3C3H**). Les acides aminés β^3 sont en bleu et les acides aminés naturels α sont en orange.

Samuel Gellman s'est intéressé au développement de structures quaternaires artificielles[109]. Les premières études[109a,b] réalisées par son équipe se sont intéressées à l'auto-assemblage d'hélices mixtes α/β dont la séquence dérive de celle d'un segment peptidique du domaine de dimérisation du régulateur de transcription GCN4[98g,101]: GNC4-p1. Le remplacement d'acides aminés α par des acides aminés β^3 dans la séquence de GCN4-p1 a conduit à la formation d'un assemblage quaternaire caractérisé à l'état solide comme étant un trimère d'hélices, tandis que la séquence peptidique naturelle privilégie la forme dimérique (**Fig. 31a, b**). La modification de la stœchiométrie du *coiled coil* est un phénomène connu pour des peptides dérivés de GCN4-p1. Ainsi, l'équipe de Gellman s'est intéressée à étudier l'impact du remplacement d'acide aminé α par des acides aminés β^3 dans la séquence d'un peptide mutant plus stable[109c], GCN4-pLI, formant un tétramère d'hélices très stable où les positions *a* et *d* de l'heptade sont occupées par des leucines ou des isoleucines (**Fig. 31d**). Les structures cristallines de différentes séquences hybrides révèlent que leur capacité à s'auto-assembler est bien conservée et tolère différents motifs de substitution α/β : ααβαααβ, αααβ et αβ (**Fig. 31e-g**).

b. Approche ascendante bottom-up pour la conception de novo de foldamères hybrides amphipatiques

La première structure quaternaire artificielle de foldamères hélicoïdaux amphipatiques obtenue à haute résolution par l'approche ascendante a été décrite par Alana Schepartz[110] en 2007 (**Fig. 32**). L'équipe a utilisé les propriétés des peptides β^3 qui adoptent une structure hélicoïdale stable dans l'eau[111] pour concevoir un décapeptide Zwit-1F qui s'auto-assemble spontanément en solution aqueuse. La structure cristallographique de Zwit-1F a été résolue et met en évidence la formation d'une structure quaternaire constituée de huit hélices (**Fig. 32**) de type capsule. Chaque hémicapsule est constituée de quatre hélices. Ces portions sont constituées de deux dimères d'hélices parallèles orientés entre eux de manière antiparallèle. L'obtention d'un assemblage similaire de type capsule avec des dimères d'hélices parallèles liées par liaisons covalentes a été considérée[112] et représente une autre stratégie pour générer de tels assemblages. La distribution des chaînes latérales de Zwit-1F a été conservée pour maintenir les interactions inter-hélices, ce qui a nécessité la modification de l'orientation des hélices initialement parallèles en structures hélice-boucle-hélice antiparallèles pour simplifier la synthèse. Aucune donnée cristallographique n'est disponible, mais des études par dichroïsme circulaire et centrifugation analytique ont confirmé la présence des tétramères et leurs propriétés d'assemblages[112].

Figure 32 : Assemblage quaternaire de deux faisceaux d'hélices de peptides β^3 Zwit-1F. Les dimères d'hélices parallèles et antiparallèles sont de couleurs différentes (bleu ou blanc).

Plus récemment, notre équipe a développé une série de foldamères oligourée hydrosolubles qui s'auto-assemblent spontanément dans l'eau[113]. Il a été montré que la topologie de l'assemblage quaternaire des hélices peut être dirigée selon la nature de la séquence primaire et principalement par les chaînes latérales hydrophobes. Deux topologies d'assemblage ont été observées à l'état solide et caractérisées en solution (CD, résonance magnétique et spectrométrie de masse native) : une capsule et un canal (**Fig. 33**). En changeant la taille de la surface hydrophobe et la distribution des résidus chargés, il a en effet été possible de passer d'un assemblage discret sous forme de bundle hexamérique à une nanostructure tubulaire.

D'autres exemples de structures tertiaires artificielles à haute résolution ont été décrites dont des squelettes à base de peptoïdes[114] et d'oligoamides aromatiques[115]. A ce jour, la littérature contient quelques descriptions précises du phénomène d'auto-assemblage tertiaire et quaternaire de polymères non naturels sous forme de *bundle* et de *coiled coil*, mais les principes sont peu connus et maîtrisés pour les foldamères par rapport à ceux décrits pour les polypeptides. La recherche menée dans ce domaine est en plein essor, les premiers résultats sont prometteurs et permettent d'imaginer un large champ des possibles.

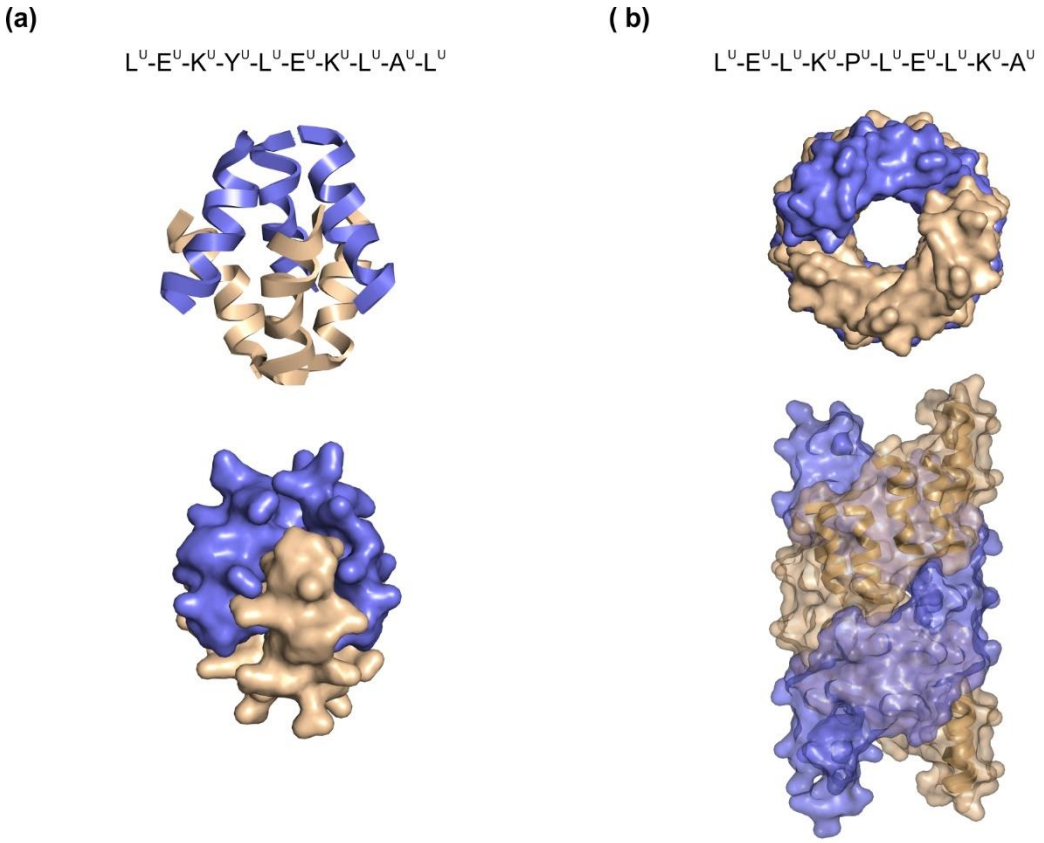


Figure 33 : Impact de la séquence primaire sur l'auto-assemblage des foldamères d'oligourée. Deux topologies sont obtenues **(a)** une capsule (CCDC : **1030455**) et **(b)** un canal (CCDC : **1030456**). Chaque *bundle* d'hélices est coloré d'une couleur différente (bleu et blanc).

3. <u>Conception rationnelle de mimes de HCH pour l'inhibition d'interactions protéine-protéine</u>

Lors de la création de la base de données DippDB, Arora *et al.* ont supposé que les motifs en hélice-coude-hélice (HCH), aussi nommés hairpin d'hélices (*helical hairpins* en anglais), seraient les éléments les plus présents aux interfaces entre protéines[99]. Le motif HCH est un motif très simple qui consiste en deux hélices α connectées par un coude contenant un petit nombre de résidus, entre 2 et 9 résidus (**Fig. 34a**). Arora a montré dans la DippDB que ce motif est très présent dans les interactions protéine-protéine des cas 1 et 2, ce qui suggère que ces structures tertiaires pourraient être mimées par des inhibiteurs spécialement élaborés pour elles. Ce motif HCH semble être le meilleur pour le développement de mimes inhibiteurs ciblant des surfaces d'IPP larges[99, 116].

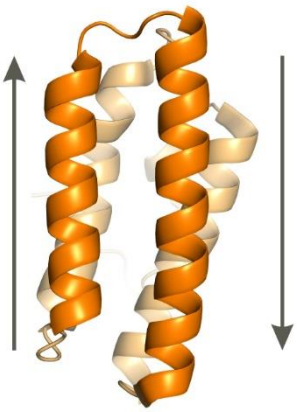


Figure 34 : Structure hélice-coude-hélice antiparallèle d'une métalloprotéine oligomérique assurant le transport de dioxygène chez le vers marin *Phascolopsis gouldii* (PDB :**1I4Y**)

La conception d'inhibiteurs de type HCH repose sur les propriétés des superhélices, ainsi que les règles de conception qui relieraient leurs séquences à leurs assemblages. Les dimères de *coiled coils* jouent un rôle biologique essentiel en dirigeant la multimérisation ainsi que la stabilisation des protéines impliquées dans les complexes de facteurs de transcription[117]. Le développement d'inhibiteurs d'IPP capables de mimer des structures tertiaires voire quaternaires semble être une approche pertinente comme nouvelle stratégie d'inhibition.

a. Stabilisation de dimères d'hélices impliqués dans des IPP par formation d'un lien covalent entre deux hélices.

Fort de ses découvertes sur l'implication de dimères d'hélices dans la reconnaissance moléculaire entre protéines, Arora s'est intéressé au développement de dimères d'hélices peptidiques courtes et stables[116a, 118] ciblant l'interaction entre le domaine NHR2 de la protéine AML1-ETO et le domaine de liaison à NHR2 (N2B) des protéines E (**Fig. 35a**). NHR2 se lie à N2B via une superhélice antiparallèle et joue un rôle essentiel dans la leucémogénèse[119].

Arora *et al.* ont ainsi développé un motif HCH en connectant deux hélices courtes antiparallèles de 10 résidus par leurs extrémités via un lien bistriazole[116a] (**Fig. 35b**). Ce lien covalent remplace un pont salin situé aux positions *e* de chaque hélice. La formation du bistriazole est réalisée sur support solide par une réaction de cycloaddition 1,3-dipolaire catalysée au cuivre I entre un alcyne et un azoture[120]. L'équipe a synthétisé de nombreuses séquences contenant ce lien bistriazole inter hélices en faisant varier la longueur de la chaîne des azido-lysines. Des mesures d'affinité par polarisation de fluorescence ont permis d'identifier *CHD-NHR2-2* comme le plus affin, avec une constante de dissociation de 236 µM (contre 356 µM pour la protéine naturelle). Cette séquence a été optimisée grâce à l'insertion

d'un point disulfure dans le cœur hydrophobe (*CHD-NHR2-3*). Cette rigidification de la structure renforce sa stabilité et améliore drastiquement l'affinité avec N2B qui est alors de 53 µM.

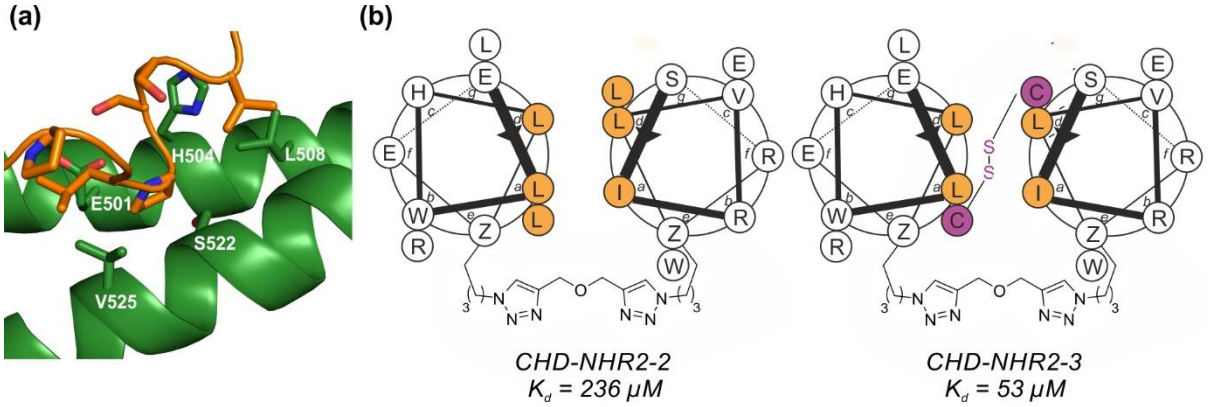


Figure 35 : Développement d'inhibiteurs de l'interaction NHR2/N2B. **(a)** Zone d'interaction entre la superhélice NHR2 (en vert) et N2B (en orange) (PDB : **4JOL**). **(b)** Diagrammes de roues hélicoïdales des séquences des inhibiteurs *CHD-NHR2-2* et *CHD-NHR2-3.*

b. Application de l'approche top-down de foldamères hybrides α/β pour la conception d'inhibiteurs de l'interaction du facteur de croissance de l'endothélium vasculaire et son domaine Z

La stabilisation de structures secondaires par des foldamères mixtes, décrit au paragraphe III-2-*a,* a permis de concevoir des assemblages supramoléculaires artificiels. Ainsi, Samuel Gellman s'est intéressé à appliquer les résultats obtenus sur la stabilisation de superhélices avec des foldamères mixtes α/β à la reconnaissance de surfaces de protéines[116b]. La cible choisie a été l'interaction du facteur de croissance de l'endothélium vasculaire (*vascular endothelial growth factor* VGEF) et son domaine Z. Ce domaine est composé de 59 résidus qui forment un *bundle* à trois hélices (**Fig. 36b**). La troisième hélice n'étant pas en interaction avec la surface de $VGEF_{8-109}$, Samuel Gellman s'est intéressé à développer une séquence contenant les deux premières hélices dont la structure ressemble à un motif HCH (**Fig. 36a**). Les deux hélices antiparallèles sont séparées par un coude composé de quatre résidus d'acides aminés α. La structure est stabilisée par l'ajout d'un pont disulfure inter hélices entre les résidus des positions N et C-terminales, et par l'insertion contrôlée de résidus Aib et d'acides β aminés cycliques et aliphatiques dans la séquence (**Fig. 36a** et **c**). Les mesures d'affinité réalisées par polarisation de fluorescence ont montré que la structure non naturelle est plus affine (K_i = 0,11 µM) à la protéine que la séquence naturelle (K_i = 0,4 µM).

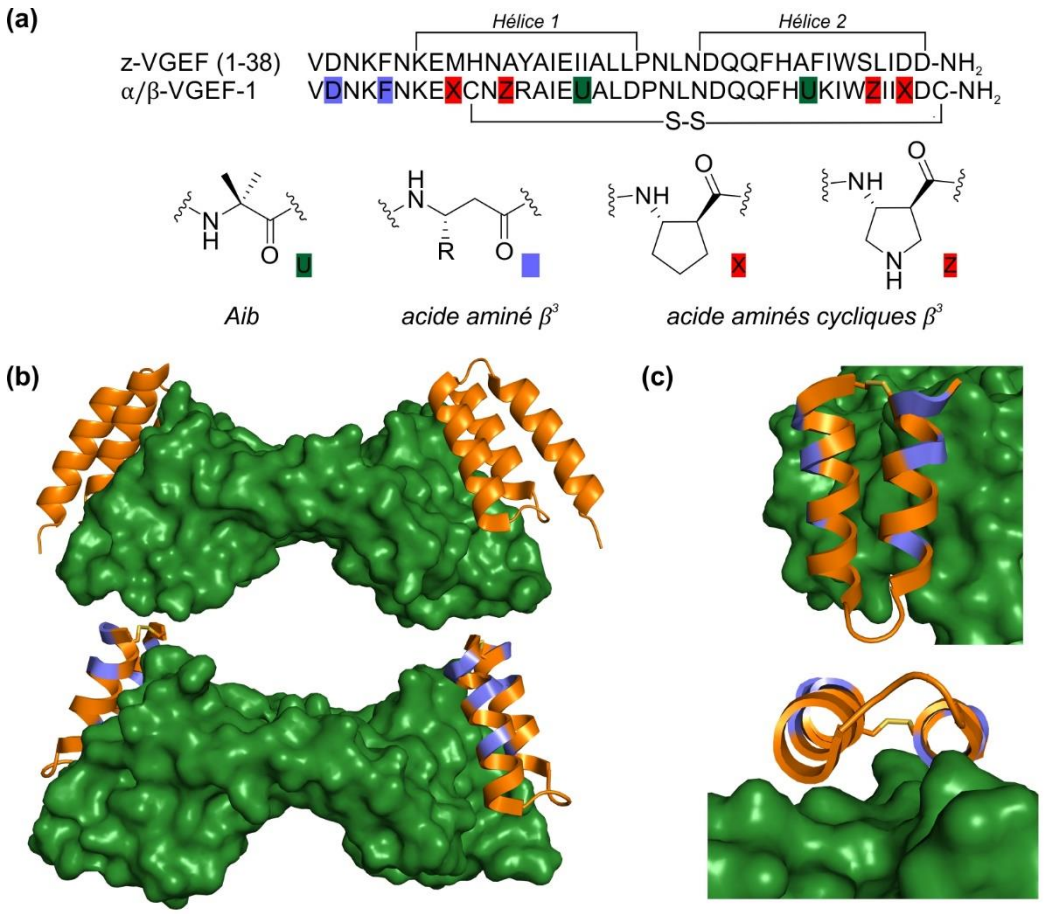


Figure 36 : Développement d'un inhibiteur de type HCH artificiel ciblant l'interaction entre la protéine VGEF et son domaine Z. **(a)** Séquence peptidique tronquée du domaine Z et séquence d'un inhibiteur α/β-VGEF-1 dont la séquence contient un pont disulfure, des résidus Aib et des acides aminés β. **(b, haut)** Structure cristalline du peptide Z-VGEF (en orange) en liaison avec $VGEF_{8-109}$ (en vert) (**PDB :3S1K**) (**b, bas**) structure cristalline de α/β-VGEF-1 en complexe avec $VGEF_{8-109}$, les résidus non naturels sont en bleu et les naturels en orange (**PDB :4WPB**). **(c)** Structure cristalline en vue de profil et de haut de α/β-VGEF-1 en complexe avec $VGEF_{8-109}$

c. *Application de l'approche bottom-up pour la conception de novo de foldamères repliés en structure HCH*

Le développement de dimère d'hélices non naturel pour effectuer la reconnaissance de surfaces de protéines est assez récent et les premiers résultats sont très encourageants. Cependant, il est nécessaire de comprendre les forces motrices qui guident l'auto-assemblage des structures artificielles. Mon projet de thèse s'intéresse à la compréhension des principes gouvernant le repliement de foldamères d'oligourée en *hairpin* en solvant organique pour pouvoir concevoir des séquences amphipatiques en solution aqueuse. La caractérisation à haute résolution de ces séquences d'oligourée amphipatiques est nécessaire pour mieux

comprendre la relation séquence-structure de ces systèmes et envisager le développement de foldamères mixtes repliés en HCH comme inhibiteurs d'interactions protéine-protéine.

IV. Objectifs de la thèse

Les stratégies d'inhibition d'interactions protéine-protéine par le développement de peptidomimes sont nombreuses : 1) développement de mimes d'une simple hélice et 2) développement de mimes de structure tertiaire via un motif hélice-coude-hélice. Dans ce projet de thèse, nous nous sommes intéressés à explorer ces deux approches d'inhibition.

Dans un premier temps, le développement d'inhibiteurs de type foldamères mixtes oligoamide/oligourée sera discuté pour inhiber l'interaction entre le récepteur de la vitamine D et les protéine SRC. Les données structurales disponibles ont été utilisées pour concevoir une série de foldamères chimères. Une étude de relation structure-affinité *(structure affinity relationship* SAR), réalisée en collaboration avec le Dr. Natacha Rochel de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) à Strasbourg, a permis d'identifier une chimère « tête de série » dont la structure cristalline a été résolue. Les données structurales de ce foldamère ont conduit à la conception d'une nouvelle génération de séquences chimères dont l'étude SAR est en cours.

Le développement et l'étude d'une structure tertiaire artificielle de type hélice-coude-hélice seront ensuite présentés. Comme décrit précédement, peu de structures tertiaires artificielles à haute résolution ont été rapportées dans la littérature et peu d'études sur ces structures permettent de généraliser les principes d'assemblage et de conception de celles-ci sous forme de HCH artificiels. Ce projet de thèse s'est intéressé à l'étude des règles de conception de dimère d'hélices de type HCH en solvant organique. La stratégie envisagée repose sur une connexion par liaison covalente de deux hélices d'oligourée par leurs extrémités. Des hélices de différente nature et longueur ont été utilisées, ainsi que divers connecteurs pour étudier l'effet de ces trois paramètres sur la structure en HCH. Les données structurales obtenues par analyses RMN et par DRX ont permis d'identifier un facteur primordial pour la conception de motifs HCH stables. Ces informations ont été appliquées pour élaborer des HCH hydrosolubles. La caractérisation de ces strutures à haute résolution est nécéssaire pour envisager le développement de foldamères oligourée en HCH comme inhibiteurs d'interactions protéine-protéine.

Références bibliographiques

1. J. D. Jordan, E. M Landaue, R. Iyengar *Cell* **2000**, *5*, 193-200
2. D. R. Hipfner, S. M. Cohen *Nat. Rev. Mol. Cell. Bio.* **2004**, *5*, 805-815
3. (a) M. W. Peczuh, A. D. Hamilton *Chem. Rev.* **2000**, *42*, 2479-2494 ; (b) P. L. Toogood *J. Med. Chem.* **2002**, *45*, 1545-1558 ; (c) T. Berg *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 2462-2481 ; (d) M. R. Arkin, J. A. Wells *Nat. Rev. Drug. Discov.* **2004**, *3*, 301-317 ; (e) L. Zhao, J. Chmielewski *Curr. Opin. Struct. Biol.* **2005**, *15*, 31-33 ; (f) H. Yin, A. D. Hamilton *Angew. Chem. Int. Ed.* **2005**, *44*, 4130-4163
4. D. P. Fairlie, M. L. West, A. K. Wong *Curr. Med. Chem.* **1998**, *5*, 29-62
5. A. L. Jochim, P. S. Arora *ACS Chem. Biol.* **2010**, *5*, 919-923
6. (a) M. K. I. Andrews, A. B. Tabor *Angew. Tetrahedron* **1999**, *55*, 11711-11743 ; (b) L. K. Henchey, A. L. Jochim, P. S. Arora *Curr. Opin. Chem. Biol.* **2008**, *12*, 692-697 ; (c) T. Edwards, A. Wilson *Amino acids* **2011**, *41*, 743-754
7. A. Charkrabartty, R. L. Baldwin *Adv. Protein Chem.* **1995**, *46*, 141-176
8. J. E. Donald, D. W. Kulp, W. F. DeGrado *Proteins* **2011**, *79*, 898-915
9. K. Sakurai, C. Schubert, D. Khane *J. Am. Chem. Soc.* **2006**, *128*, 11000-11001
10. C. Toniolo, M. Crisma, F. Fromaggio, C. Peggion *Biopolymers.* **2001**, *60*, 396-419
11. (a) A. Ravi, B. V. V. Prasad, P. Balaram *J. Am. Chem. Soc.* **1983**, *105*, 105-109 ; (b) A. M. Leduc, J. O. Trent, J. L. Wittliff *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 11273-11278
12. (a) G. Osapay, J. W. Taylor *J. Am. Chem. Soc.* **1992**, *114*, 6966-6973 ; (b) C. Braken, J. Gulyas, J. W. Taylor, J. Braum *J. Am. Chem. Soc.* **1994**, *116*, 6431-6432 ; (c) R. S. Harrison, N E. Shepherd, H. N. Hoang *Proc. Natl. Acad. Sci. USA.* **2010**, *107*, 11686-11691
13. (a) D. Y. Jackson, D. S. King, J. Chmielewski *J. Am. Chem. Soc.* **1991**, *113*, 9391-9392 ; (b) J. C. Phelan, N. J. Skelton, A. C. Braisted, R. C. McDowell *J. Am. Chem. Soc.* **1997**, *119*, 455-460 ; (c) C. Yu, J. W. Taylor *Bioorg. Med. Chem.* **1999**, *7*, 161-175
14. H. E. Blackwell, R. H Grubbs *Angew. Chem. Int. Ed.* **1998**, *37*, 3281-3284 ; (b) L. D. Walensky, A. L. Kung, I. Escher *Science* **2004**, *305*, 1466-1470 (c) S. Bhattacharya, H. T. Zhang, A. K. Debnath, D. Cowburn *J. Biol. Chem.* **2008**, *283*, 16274-16278
15. C. E. Schafmeister, J. Po, G. L. Verdine *J. Am. Chem. Soc.* **2000**, *122*, 5891-5892
16. (a) Y. W. Kim, G. L. Verdine *Bioorg. Med. Chem. Lett.* **2009**, *19*, 2533-2536 ; (b) Y-W. Kim, T. N. Grossmann, G. L. Verdine *Nat. Protoc.* **2011**, *6*, 761-771
17. G. H. Bird, N. Madani, A. F. Perry, A. M. Princiotto, J. G. Supko, X. He, E. Gavathiotis, J. G. Sodroski, L. D. Walensky *Proc. Natl. Acad. Sci. USA.* **2010**, *107*, 14093-14098

18. (a) R. N. Chapman, G. Dimartino, P. S. Arora *J. Am. Chem. Soc.* **2004**, *126*, 12252-12253 ; (b) D. Wang, K. Chen, G. Dimartino, P. S. Arora *Org. Biomol. Chem.* **2006**, *4*, 4074-4081
19. E. Cabezas, A. C. Satterhwait *J. Am. Chem. Soc.* **1999**, *121*, 3862
20. S. Wang, K. Chen, J. L. Kulp, P. S. Arora *J. Am. Chem. Soc.* **2006**, *126*, 9248-9256
21. S. H. Gellman *Acc. Chem. Res.* **1998**, *31*, 173-180
22. (a) G. Guichard, I. Huc *Chem. Commun.* **2011**, *47*, 5933-5941 ; (b) C. M. Goodman, S. Choi, S. Shandler, W. F. DeGrado *Chem. Biol.* **2007**, *3*, 252-262
23. W. S Horne, S. H. Gellman *Acc. Chem. Res.* **2008**, 1399-1408
24. (a) L. Guo, Y. Chi, A. M. Almeida, I. A. Guzei, B. K. Parker, S. H. Gellman *J. Am. Chem. Soc.* **2009**, *131*, 16018-16020 ; (b) L. Guo, A. M. Almeida, W. Zhang, A. G. Reidenbach, S. H. Choi, I. A. Guzei, S. H. Gellman *J. Am. Chem. Soc.* **2010**, *132*, 7868-7869S
25. A. Fowler, H. E. Blackwell *Org. Biomol. Chem.* **2009**, *7*, 1508-1524
26. (a) V. Semetey, D. Rognan, C. Hemmerlin, R. Graff, J.-P. Briand, M. Marraud, G. Guichard *Angew. Chem. Int. Ed.* **2002**, *41*, 1893-1895 ; (b) A. Salaun, M. Potel, T. Roisnel, P. Gall, P. Le Grel *J. Org. Chem.* **2005**, *70*, 6499-6502 ; (c) X. Li, D. Yang *Chem. Commun.* **2006**, *0*, 3367-3379
27. (a) J. C. Nelson, J. G. Saver, J. S. Moore, P. G. Wolynes *Science* **1997**, *277*, 1793-1796 ; (b) R. S. Lokey, B. L. Iverson *Nature* **1995**, *375*, 303-305 ; (c) D. M. Bassani, J.-M. Lehn, G. Baum, D. Fenske *Angew. Chem. Int. Ed.* **1997**, *36*, 1845-1847 ; (d) Y. Hamuro, S. J. Greib, A. D. Hamilton *J. Am. Chem. Soc.* **1996**, *118*, 7529-7541 ; (e) V. Berl, I. Huc, R. Khoury, M. Krische, J.-M. Lehn *Nature* **2000**, *407*, 720-723 ; (f) J. Zhu, R. D. Parra, H. Zeng, E. Skrzypczak-Jankun, X. Cheng Zeng, B. Gong *J. Am. Chem. Soc.* **2000**, *122*, 4219-4220 ; (g) H. Jiang, J.-M. Léger, I. Huc *J. Am. Chem. Soc.* **2003**, *125*, 3448-3449
28. A. Violette, N. Lancelot, A. Poschalko, M. Piotto, J-P. Briand, J. Raya, K. Elbayed, A. Bianco, G. Guichard *Chemistry* **2008**, *14*, 3874-3882
29. N. Pendem, C. Douat, P. Claudon, M. Laguerre, S. Castano, B. Desbat, D. Cavagnat, E. Ennifar, B. Kauffmann, G. Guichard *J. Am. Chem. Soc.* **2013**, *135*, 4884-4892
30. V. Semetey, C. Didierjean, J-P. Briand, A. Aubry, G. Guichard *Angew. Chem. Int. Ed.* **2002**, *41*, 1895-1897
31. (a) A. Violette, M-C. Averlant-Petit, V. Semetey, C. Hemmerlin, R. Casimir, R. Graff, M. Marraud, J-P. Briand, D. Rognan, G. Guichard *J. Am. Chem. Soc.* **2005**, *127*, 2156-2164 ; (b) J. Fremaux, L. Fischer, T. Arbogast, B. Kauffmann, G. Guichard *Angew. Chem. Int. Ed.* **2011**, *50*, 11382-11385
32. M. Crisma, F. Formaggio, A. Moretto, C. Toniolo *Biopolymers* **2006**, *84*, 3-12

33. (a) W. S Horne, S. H. Gellman *Acc. Chem. Res.* **2008**, 1399-1408 ; (b) A. Roy, P. Prabhakaran, P. K. Baruah, G. J. Sanjayann *Chem. Commun.* **2011**, *47*, 11593-11611 ; (c) M. Kudo, V. Maurizot, B. Kauffmann, A. Tanatani, I. Huc *J. Am. Chem. Soc.* **2013**, *135*, 9628-9631

34. (a) J. D. Sadowsky, W. D. Fairlie, E. B. Hadley, H-S. Lee, N. Umezawa, Z. Nikolovska-Coleska, S. Wang, D. C. S. Huang, Y. Tomita, S. H. Gellman *J. Am. Chem. Soc.* **2006**, *129*, 139-154 ; (b) J. D. Sadowsky, M. A. Schmitt, H-S. Lee, N. Umezawa, S. Wang, Y. Tomita, S. H. Gellman *J. Am. Chem. Soc.* **2005**, *127*, 11966-11968

35. J. Fremaux, L. Mauran, K. Pulka-Ziach, B. Kauffmann, B. Odaert, G. Guichard *Angew. Chem. Int. Ed.* **2015**, *54*, 9816-9820

36. L. Mauran, B.Kauffmann, B. Odaert, G. Guichard *C. R. Chimie* **2016**, *19*, 123-131

37. N. McKenna, R. B. Lanz, B. W. O'Malley *Endocr. Rev.* **1999**, *20*, 321-344

38. (a) J. J. Voegel, M. J. Heine, M. Tini, V. Vivat, P. Chambon, H. Gronemeyer *Embo J.* **1998**, *17*, 507-519 ; (b) Y. Brelivet, S. Krammerer, N. Rochel, O. Poch, D. Moras *Embo Rep.* **2004**, *5*, 423-429

39. (a) D. J. Mangelsdorf, R. M. Evans *Cell* **1995**, *83*, 841-850 ; (b) L Altucci, H. Gronemeyer *Trends Endicronol. Metab.* **2001**, *12*, 460-468

40. C. Leo, H. Li, J. D. Chen *J. Mol. Chem.* **2000**, *275*, 5976-5982

41. N. D'Abreo, A. H. Hindenburg *Vitam. Horm.* **2013**, *93*, 99-133

42. I. Orlov, N. Rochel, D. Moras, B. P. Klahoz *Embo J.* **2012**, *31*, 291-300

43. K. Umesono, K. K Murakami, C. C. Thompson, R. M. Evans *Cell* **1991**, *65*, 1255-1265

44. N. Rochel, F. Ciesielski, J. Godet, E. Moman, M. Roessle, C. Peluso-Iltis, M. Moulin, M. Haertlein, P. Callow, Y. Mély, D. Svergun, D. Moras *Nat. Struct. Mol. Biol.* **2011**, *18*, 564-596

45. V. Chandra, P. Huang, Y. Hamuro, S. Raghuram, Y. Wang, T. P. Burris, F. Rastinejad *Nature* **2008**, *456*, 350-356

46. (a) T. Perlmann, P. N. Rangarajan, K. Umesono, R. M. Evans *Genes Dev.* **1993**, *7*, 1411-1422 ; (b) R. Kurokawa, V. C. Yu, A. Näär, S. Kyakumoto, Z. Hans, S. Silverman, M. G. Rosenfeld, C. K. Glass *Genes Dev.* **1993**, *7*, 1423-1435 ; (c) C. Zechel, X-Q. Shen, J-Y. Chen, Z-P. Chen, P. Chambon, H. Gronemeyer *Embo J.* **1994**, *13*, 1425-1433

47. (a) W. H. Hudson, C. Youn, E. A. Ortlund *Nat. Struct. Mol. Biol.* **2013**, *20*, 53-58 ; (b) L. C. Watson, *Nat. Struct. Mol. Biol.* **2013**, *20*, 876-883

48. J. Osz, A. G. McEwen, P. Poussin-Courmontagne, E. Moutier, C. Birck, I. Davidson, D. Moras, N. Rochel *Sci. Rep.* **2015**, *5*, 8216

49. A. Mouchon, M. H. Delmotte, P. Formstecher, P. Lefebvre *Mol. Cell. Biol.* **1999**, *19*, 3073-3085
50. (a) D. J. Mangelsdorf, C. Thummel, M. Beato, P. Herrlich, G. Schütz, K. Umesono, B. Blumberg, P. Kastner, M. Mark, P. Chambon, R. M. Evans *Cell* **1995**, *83*, 835-839 ; (b) J. D. Chen, R. M. Evans *Nature* **1995**, *377*, 454-457 ; (c) L. Nagy *Genes Dev.* **1999**, *13*, 3209-3216 ; (d) C. K. Glass, M. G. Rosenfeld *Genes Dev.* **2000**, *14*, 121-141
51. S. A. Oñate, S. Y. Tsai, M. J. Tsai, B. J. O'Malley *Science* **1995**, *270*, 1354-1357
52. J. Osz, Y. Brelivet, C. Peluso-Iltis, V. Cura, S. Eiler, M. Ruff, W. Bourguet, N. Rochel, D. Moras *Proc. Natl. Acad. Sci. USA.* **2012**, *109*, E588-E594
53. X. F. Ding, C. M. Anderson, H. Ma, H. Hong, R. M. Uht, P. J. Kushner, M. R. Stallcup *Mol. Endocrinol.* **1998**, *12*, 302-313
54. (a) L. F. Velasco *et al. J. Biol. Chem.* **2007**, *282*, 12458-12466 ; (b) A. Takeshita *et al. J. Biol. Chem.* **1998**, *273*, 21554-21562
55. B. D. Putcha, E. J. Fernandez *J. Biol. Chem.* **2009**, *284*, 22514-22524
56. J. Zhang *et al. Nat. Struct. Mol. Biol.* **2011**, *18*, 556-563
57. M. R. Haussler, J. F. Myrtle, A. W. Norman *J. Biol. Chem.* **1968**, *243*, 4055-4064
58. T. Sone, S. Kerner, J. W. Pike *J. Biol. Chem.* **1991**, *266*, 23296-23305
59. S. V.Ramagopolan, A. Heger, A. J. Berlanga, N. J. Maugeri, M. R. Lincoln, A. Burrell, L. Handunnetthi, A. E. Handel, G. Disanto, S. M. Orton, C. T. Watson, J. M. Morahan, G. Giovannoni, C. P. Ponting, G. C. Ebers, J. C. Knight *Genome Res.* **2010**, *20*, 1352-1360
60. . W. Blunt, H. F. DeLuca, H. K. Schnoes **1968**, *7*, 3317-3322
61. (a) D. R. Fraser, E. Kodicek *Nature* **1970**, *228*, 764-766 ; (b) M. F. Holick, H. K. Schnoes, H. F. DeLuca, T. Suda, R. J. Cousins *Biochemistry* **1971**, *10*, 2799-2804 ; (c) D. E. Lawson, D. R. Fraser, Z. Kodicek, H. R. Morris, D. H. Morris *Nature* **1971**, *230*, 228-230
62. A. Murayama *Biochem. Biophys. Res. Commun.* **1998**, *249*, 11-16
63. D. Zehnder *J. Clin. Endocrinol. Metab.* **2001**, *86*, 888-894
64. J Pike *Annual Review of Nutrition* **1991**, *11*, 189-216
65. (a) U. Berger, P. Wilson, R. A. McClelland, K. Colston, M. R. Haussler, J. W. Pike *J. Clin. Endocrinol. Metab.* **1988**, *67*, 607-613 ; (b) T. L. Clemens, K. P. Garrett, X. Y. Zhou, J. W. Pike, M. R. Haussler, D. W. Dempster *Endocrinology* **1988**, 122, 1224-1230 ; (c) P. Banerjee, M. Chatterjee *Mol. Cell. Biochem.* **2003**, *253*, 247-254
66. (a) R. Losel, M. Wehling *Nature Rev. Mol. Cell. Biol.* **2003**, *4*, 46-56 ; (b) X. Wang, G. P. Studzinski *J. Cell Biochem.* **2001**, *80*, 471-482 ; (c) M. Simboli-Campbell, C. J. Narvaez, M. Tenniswood, J. Welsh *J. Steroid Biochem. Mol. Biol.* **1996**, *58*, 367-376

67. (a) R. Bouillon, G. Eelen, L. Verlinden, C. Mathieu, G. Carmeliet, A. Verstuyf *J. Steroid Biochem. Mol. Biol.* **2006**, 102, 156-162 ; (b) K. V. Pinette, Y. K. Yee, B. Y. Amegadzie, S. Nagpal *Mini. Rev. Med. Chem.* **2003**, 3, 193-204
68. P. L. Shaffer, D. T. Gewirth *J. Steroid Biochem. Mol. Biol.* **2004**, *89*, 215-219
69. J. Zhang, M. J. Chalmers, K. R. Stayrook, L. L. Burris, Y. Wang, S. A. Busby, B. D. Pascal, R. D. Garcia-Ordonez, J. B. Bruning, M. A. Istrate, D. Kojetin, J. A. Dodge, T. P. Burris, P. R. Griffin *Nat. Struct. Mol. Biol.* **2011**, *18*, 556-563
70. (a) A. S. Dusso, A. J. Brown *Am. J. Kidney Dis.* **1998**, *32*, 23-24 ; (b) J. C. Hsieh *J. Cell Biochem.* **1998**, *70*, 94-109 (b) T. Michigami *J. Biol. Chem.* **1999**, *274*, 33531-33538 ; (c) R. Yasmin, R. M. Williams, M. Xu, N. Noy *J. Biol. Chem.* **2005**, *280*, 40152-40160
71. K. A. Teske, O. Yu, L. A. Arnold *Vitam. Horm.* **2016**, 100, 45-82
72. (a) B. Li, M. Carey, J. L. Workman *Cell* **2007**, *128*, 707-719 ; (b) M. R. Haussler *J. Bone Miner. Res.* **1998**, *13*, 325-349
73. D. M. Lonard, B. W. O'Malley *Nat. Rev. Endocrinol.* **2012**, *8*, 598-604
74. (a) D. M. Lonard, R. B. Lanz, B. W. O'Malley *Endocr. Rev.* **2007**, *28*, 575-587 ; (b) D. M. Lonard, B. W. O'Malley *Mol. Cell.* **2007**, *27*, 691-700 ; (c) J. Xu, Q. Li *Mol. Endocrinol.* **2003**, *17*, 1681-1692
75. A. R. Chopra *et al. Science* **2008**, *322*, 1395-1399
76. (a) F. Picard *et al. Cell* **2002**, *111*, 931-941 ; (b) A. R. Chopra *et al. Cell. Metab.* **2011**, *13*, 35-43
77. (a) S. J. Han, D. M. Lonard, B. W. O'Malley *Trends Endocrinol. Metab.* **2009**, *20*, 8-15 ; (b) D. M. Lonard, R. Kumar, B. W. O'Malley *Mol. Endocrinol.* **2010**, *24*, 279-285
78. T. Chen *Curr. Opin. Chem. Biol.* **2008**, *12*, 418-426
79. (a) P. R. Housley, E. R. Sanchez, M. Danielsen, G. M. Ringod, W. Pratt *J. Biol. Chem.* **1990**, *265*, 12778-12781 ; (b) D. M. Heery, E. Kalkhoven, S. Hoare, M. G. Parker *Nature* **1997**, *387*, 733-736
80. B. D. Darimont, R. L. Wagner, J. W. Apriletti, M. R. Stallcup, P. J. Kushner, J. D. Baxter, R. J. Fletterick, K. R. Yamamoto *Genes Dev.* **1998**, *12,* 3343-3356
81. (a) E. M. McInerney, D. W. Rose, S. E. Flynn, S. Westin, T. M. Mullen, A. Krones, J. Inostroza, J. Torchia, R. T. Nolte, N. Assa-Munt, M. V. Milburn, C. K. Glass, M. G. Rosenfeld *Genes Dev.* **1998**, *12,* 3357-3368 ; (b) J. Leers, E. Treuter, J. A. Gustafsson *Mol. Cell. Biol.* **1998**, *18,* 6001-6013
82. O. Hankinson *Annu. Rev. Pharmacol. Toxicol.* **1995**, *35,* 307-340
83. (a) J. C. Blanco, S. Minucci, J. Lu, X. J. Yang, K. K. Walker, H. Chen, R. M. Evans, Y. Nakatani, K. Ozato *Genes Dev.* **1998**, *12,* 1638-1651 ; (b) H. Li, J. D. Chen, *J. Biol. Chem.* **1998**, *273*, 5948-5954

84. D. Miura, K. Manabe, K. Ozono *J. Biol. Chem.* **1999**, *274*, 16392-16399
85. (a) Y. Bury, A. Steinmeyer, C. Carlberg *Mol. Pharmacol.* **2000**, *58*, 1067-1074 ; (b) L. A. Zella, C. Y. Chang, D. P. McDonnell, J. W. Pike *Arch. Biochem. Biophys.* **2007**, *460*, 206-212
86. K. Yamamoto, D. Abe, N. Yoshimoto, M. Choi, K. Yamagishi, H. Tokiwa *J. Med. Chem.* **2006**, *49*, 1313-1324
87. (a) Y. Inaba, N. Yoshimoto, N. Yoshimoto, M. Matsunawa, S. Uno, S. Yamada *Mol. Pharmacol.* **2007**, *71*, 1298-1311 ; (b) M. Igarashi, N. Yoshimoto, K. Yamamoto *Arch. Biochem. Biophys.* **2007**, *460*, 240-253
88. Y. Inaba, N. Yoshimoto, Y. Sakamaki *J. Med. Chem.* **2009**, *52,* 1438-1449
89. (a) P. Nandhikonda, A. Yasgar, A. M. Baranowski, P. S. Sidhu, M. M. McCallum, A. J. Pawlak *Biochemistry* **2013**, *52*, 4193-4203 ; (b) K. A. Teske, J. W. Bogart, L. A. Arnold *Bioorg. Med. Chem.* **2018**, 28, 351-354
90. M. L. Sznaidman, C. D. Haffner, P. Malooney *Bioorg. Med. Chem. Lett.* **2003**, 13, 1517-1521
91. Y. Mita, K. Dodo, T. Noguchi-Yachide, H. Miyachi, M. Makishima, Y. Hashimoto, M. Ishikawa *Bioorg. Med. Chem. Lett.* **2010**, *20*, 1712-1717
92. Y. Mita, K. Dodo, T. Noguchi-Yachide, H. Miyachi, Y. Hashimoto, M. Ishikawa *Bioorg. Med. Chem. Lett.* **2013**, *21*, 993-1005
93. P. Nandhikonda, W. Z. Lynt, M. M. McCallum, T. Ara, A. M. Baranowski, N. Y. Yuan, D. Pearson, D. D. Bikle, R. K. Guy, L. A. Arnold *J. Med. Chem.* **2012**, *55*, 4640-4651
94. (a) P. S. Sidhu, N. Nassif, M. M. McCallum, K. Teske, B. Feleke, N. Y. Yuan, P. Nandhikonda, J. M. Cook, R. K. Singh, D. D. Bikle, L. A. Arnold *ACS Med. Chem. Lett.* **2014**, *5*, 199-204 ; (b) P. S. Sidhu, K. Teske, B. Feleke, N. Y. Yuan, M. L. Guthrie, G. B. Fernstrum *Cancer Chemother. Pharmacol.* **2014**, *74*, 787-798
95. (a) C. Chang, J. D. Norris, H. Gron, L. A. Paige, P. T. Hamilton, D. J. Kenan *Mol. Cell. Bio.* **2004**, *19*, 8226-8239 (b) D. P. McDonnell, C. Y. Chang, J. D. Norris *J. Steroid Biochem. Mol. Biol.* **2000**, *74*, 327-335
96. L. A. Zella, C. Y. Chang, D. P. McDonnell, J. W. Pike *Arch. Biochem. Biophys.* **2007**, *460*, 206-212
97. (a) Y Demizu, S. Nagoya, M. Shirakawa, M. Kawamura, N. Yamagata, Y. Sato, M. Doi, M. Kurihara *Bioorg. Med. Chem. Lett.* **2013**, 23, 4292-4296 ; (b) T. Misawa, Y. Demizu, M. Kawamura, N. Yamagata, M. Kurihara *Bioorg. Med. Chem.* **2015**, 23, 1055-1061
98. T. E. Speltz, S. W. Fanning, C. G. Mayne, C. Fowler, E. Tajkhorshid, G. L. Greene, T. W. Moore *Angew. Chem. Int. Ed.* **2016**, *55*, 4252-4255
99. A. M. Watkins, M. G Wuo, P. S. Arora *J. Am. Chem. Soc.* **2015**, *137*, 11622-11630

100. M. Raj, B. N. Bullock, P. S. Arora *Bioorg. Med. Chem*. 2013, *32*, 4051-4057
101. E. Wolf, P. S. Kim, B. Berger *Prot. Sci*. **1997**, *6*, 1179-1189
102. L. Pauling, R. B. Corey, H. R. Branson *Proc. Natl. Acad. Sci. U. S. A.* **1951**, *37*, 205-211
103. F. H.Crick *Nature* **1952**, *170***,** 882-883
104. (a) F. H. Crick *Acta Cryst.* 1953, *6*, 689-697 ; (b) D. N. Woolfson *Adv. Protein Chem.* **2005**, *70*, 79-112
105. (a) D. A. D. Parry *Biosci. Rep.* **1982**, *2*, 1017-1024 ; (b) D. N. Woolfson, T. Albert *Prot. Sci.* **1995**, *4*, 1596-1607
106. (a) E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber *Science* **1991**, *254*, 539-544 ; (b) E. K. O'Shea, R. Rutkowski, P. S. Kim *Cell* **1992**, *68*, 699-708 ; (c) E. K. O'Shea, R. Rutkowski, W. Stafford, P. S. Kim *Science* **1989**, *244*, 646-648
107. (a) A. Lupas, M. Vandyke, J. Stock *Science* **1991**, *252*, 1162-1164 ; (b) J. F. Conway, D. A. D. Parry *Intl. J. Biol. Macromol.* **1990**, *12*, 328-334
108. (a) K. J. Lumb, C. M. Carr, P. S. Kim *Biochemistry* **1994**, *33*, 7361-7367 ; (b) R. S. Hodges, A. K. Saund, P. C. S. Chong, S. A. Stpierre, R. E. Reid *J. Biol. Chem.* **1981**, *256*, 1214-1224 ; (c) J. Y.Su, R. S. Hodges, C. M. Kay *Biochemistry* **1994**, *33*, 15501-15510
109. (a) W. S. Horne, J. L. Prince, J. L. Keck, S. H. Gellman *J. Am. Chem. Soc.* **2007**, *129*, 4178-4180 ; (b) W. S. Horne, M. D. Boersma, M. A. Windsor, S. H. Gellman *Angew. Chem. Int. Ed. Engl*. **2008**, *47*, 2850-2856 ; (c) W. S. Horne, J. L. Price, S. H. Gellman *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 9151-9156
110. D. S. Daniels, E. J. Petersson, J. X. Qiu, A. Schepartz *J. Am. Chem. Soc.* **2007**, *129*, 1532-1533
111. (a) R. P. Chen, W. F. DeGrado *J. Am. Chem. Soc.* **2001**, *123*, 5162-5163 ; (b) P. I. Arvidsson, M. Rueping, D. Seebach *Chem. Commun.* **2001**, 649-650 ; (c) S. A. Hart, A. B. F. Bahadoor, E. E. Matthews, X. Y. J. Qiu, A. Schepartz *J. Am. Chem. Soc.* **2003**, *125*, 9468-9469
112. P. S. Wang, A. Schepartz *Chem. Commun.* **2016**, *52*, 7420-7432
113. G. W. Collie, K. Pulka-Ziach, C. M. Lombardo, J. Fremaux, F. Rosu, M. Decossas, L. Mauran, O. Lambert, V. Gabelica, C. D. Mackereth, G Guichard *Nat. Chem.* **2015**, *7*, 871-878
114. B.-C. Lee, R. N. Zuckermann, K. A. Dill *J. Am. Chem. Soc*. **2005**, *127*, 10999–11009
115. (a) N. Delsuc, J. M. Leger, S. Massip, I. Huc *Angew. Chem. Int. Ed. Engl*. **2007**, *46*, 214–217 ; (b) L. Sebaoun, V. Maurizot, T. Granier, B. Kauffmann, I. Huc *J. Am. Chem. Soc*. **2014**, *136*, 2168–2174

116. (a) M. G. Wuo, A. B. Mahon, P. S. Arora *J. Am. Chem. Soc*. **2015**, *137*, 11618-11621 ; (b) J. W. Checco, D. F. Kreitler, N. C. Thomas, D. G. Belair, N. J. Rettko, W. L. Murphy, K. T. Forest, S. H. Gellman *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 4552-4557

117. (a) P. B. Harbury, T. Zhang, P. S. Kim, T. Alber *Science* **1993**, *262*, 1401-1407 ; (b) C. R. Vinson, J. W. Hai, S. M. Boyd *Genes Dev.* **1993**, *7*, 1047-1058 ; (c) D. L. McClain, H. L. Woods, M. G. Oakley *J. Am. Chem. Soc*. **2001**, *123*, 3151-3152 ; (d) K. J. Lumb, P. S. Kim *Biochemistry* **1995**, *34*, 8642-8648 ; (e) B. Bilgicer; X. Xing, K. Kumar *J. Am. Chem. Soc*. **2001**, *123*, 11815-11816

118. N. Sawyer, A. M. Watkins, P S. Arora *Acc. Chem. Res*. **2017**, *50*, 1313-1322

119. X.-J. Sun, Z. Wang, L. Wang, Y. Jiang, N. Kost, T. D. Soong, W.-Y. Chen, Z. Tang, T. Nakadai, O. Elemento, W. Fischle, A. Melnick, D. J. Patel, S. D. Nimer, R. G. A. Roeder *Nature* **2013**, *500*, 93-97

120. (a) Y. L. Angell, K. Burgess *Chem. Soc. Rev*. **2007**, *36*, 1674-1689 ; (b) O. Torres, D. Yüksel, M. Bernardina, K. Kumar, D. Bong *Chem. Bio. Chem.*. **2008**, *9*, 1701-1705

Chapitre 2

Développement de foldamères mixtes pour l'inhibition d'interactions protéine-protéine : application à l'interaction entre le récepteur de la vitamine D et un peptide coactivateur

L'interaction protéine-protéine entre le récepteur de la vitamine D (VDR) d'une part et la portion peptidique des protéines SRC d'autre part, est impliquée dans de nombreuses pathologies. L'objectif de ce projet de thèse est l'utilisation des propriétés des foldamères mixtes (chimères) afin de mimer cette courte portion peptidique par une hélice stabilisée. Une amélioration de l'affinité du segment peptidique par une séquence hybride oligoamide/oligourée constitue un premier pas dans le développement d'inhibiteurs de l'interaction entre le récepteur nucléaire et la protéine coactivatrice.

Dans un premier temps, la conception et la synthèse de chimères seront détaillées, puis l'étude de relation structure-affinité sera décrite. Cette étude a permis de mettre en évidence une séquence « tête de série » très prometteuse qui a été optimisée dans l'espoir d'améliorer ses propriétés d'interaction avec le récepteur nucléaire VDR.

I. Génération de foldamères mixtes oligoamide/oligourée comme inhibiteurs peptidomimétiques de l'interaction VDR/SRC1

1. Conception des chimères

Les protéines SRC contiennent plusieurs motifs LXXLL responsables de l'interaction avec le récepteur VDR[1]. Ce motif, appelé NR box (*nuclear receptor box*) est nécessaire et suffisant pour former la liaison du coactivateur à un récepteur donné[2]. Afin d'évaluer les interactions de ces NR box au récepteur nucléaire de la vitamine D, leurs constantes d'affinité à l'équilibre (K_d) ont été mesurées par polarisation de fluorescence en présence de calcitriol[3]. Chaque NR box des protéines SRC a été isolée en séquence peptidique contenant deux segments de sept résidus d'acides aminés situés de part et d'autre du motif (**Tableau 3**).

peptides	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	12	K_d (µM)
SRC1-1	C	Y	S	Q	T	S	H	K	L	V	Q	L	L	T	T	T	A	E	Q	Q	> 50
SRC1-2	C	L	T	A	R	H	K	I	L	H	R	L	L	Q	E	G	S	P	S	D	1,80 ± 0,30
SRC1-3	C	E	S	K	D	H	Q	L	L	R	Y	L	L	D	K	D	E	K	D	L	1,40 ± 0,10
SRC2-1	C	D	S	K	G	Q	T	K	L	L	Q	L	L	T	T	K	S	D	Q	M	> 50
SRC2-2	C	L	K	E	K	H	K	I	L	H	R	L	L	Q	D	S	S	S	P	V	1,70 ± 0,20
SRC2-3	C	K	K	K	E	N	A	L	L	R	Y	L	L	D	K	D	D	T	K	D	0,93 ± 0,17
SRC3-1	C	E	S	K	G	H	K	K	L	L	Q	L	L	T	S	S	S	D	D	R	> 20
SRC3-2	C	L	Q	E	K	H	R	I	L	H	K	L	L	Q	N	G	N	S	P	A	7,70 ± 2,10
SRC3-3	C	K	K	E	N	N	A	L	L	R	Y	L	L	D	R	D	D	S	S	D	0,97 ± 0,12

Tableau 3 : Séquences des peptides coactivateurs naturels ainsi que leurs constantes d'affinité à l'équilibre K_d à VDR mesurées par polarisation de fluorescence en présence de calcitriol. Les résidus d'acides aminés leucine conservés du motif LXXLL sont indiqués en jaune. Les acides aminés en orange pâle sont les séquences de peptides coactivateurs utilisées pour cristalliser avec les RN.

Un résidu d'acide aminé cystéine a été ajouté à l'extrémité N-terminale de chaque séquence afin d'être marquée par un fluorophore (Texas Red Maleimide) et ainsi permettre les mesures d'affinité par polarisation de fluorescence[3]. Les résultats obtenus indiquent que les séquences correspondant aux NR box 2 et 3 des protéines SRC ont les meilleures affinités avec VDR. Des structures PDB de VDR en interaction avec des portions de ces peptides coactivateurs ont été obtenues : VDR avec SRC1-2 (**2HC4**) et SRC2-3 (**5H1E**). Ces portions de peptides correspondent aux séquences entre les résidus -4 et 9 des peptides présentés dans le **Tableau 3**, ces treize résidus d'acides aminés sont nécessaires pour avoir l'interaction avec VDR. Dans notre étude, les séquences utilisées comme peptides de références seront composées de ces treize résidus d'acides aminés.

L'étude des structures cristallographiques des complexes VDR/coactivateurs (SRC1-2, SRC2-3) a été le point de départ dans la conception des foldamères mixtes oligoamide/oligourée. L'objectif premier a été d'analyser les interactions à la surface du récepteur nucléaire entre celui-ci et son ligand naturel afin de définir la nature des chaînes latérales indispensables à l'interaction pour ensuite mimer au mieux le ligand naturel. Le segment LXXLL adopte une conformation en hélice α dans laquelle les chaînes latérales des trois leucines en position i, i+3 et i+4 se situent sur une même face de l'hélice et se projettent dans un sillon hydrophobe à la surface du récepteur (**Fig. 37**).

(a) **(b)**

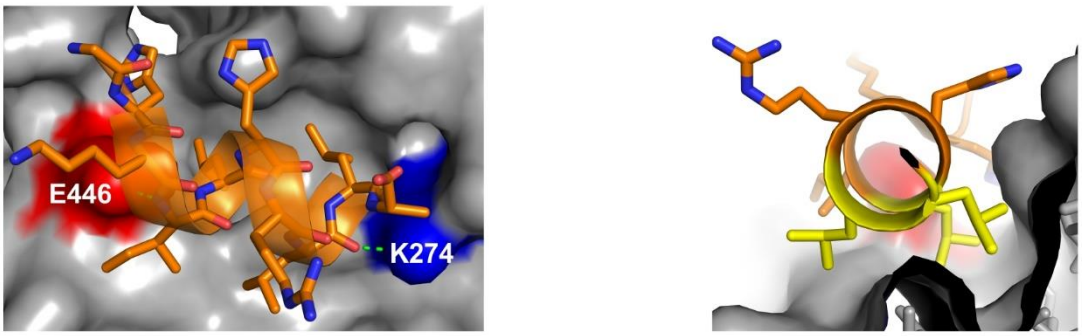


Figure 37 : Interaction protéine-protéine entre VDR (en gris) et le peptide coactivateur SRC1-2 (en orange) (PDB : **2HC4**). **(a)** Représentation de la pince chargée entre l'hélice α de SRC1-2 et les résidus E446 (en rouge) et K274 (en bleu) de VDR, les liaisons hydrogène sont en vert. **(b)** Mise en évidence du motif LXXLL avec la projection des leucines en i, i+3 et i+4 (en jaune) dans la poche hydrophobe.

L'analyse détaillée de la structure suggère par ailleurs que des interactions entre le peptide et des résidus chargés du récepteur se situant à chaque extrémité du sillon stabilisent le complexe. Notamment les carboxylates du résidu E446 de VDR sont impliqués dans des liaisons hydrogène avec les hydrogènes des fonctions amides des résidus aux positions -1 et 1 du peptide coactivateur, tandis que l'amine primaire du résidu K274 de VDR forme une liaison hydrogène avec le groupement carbonyle du résidu L_5 du peptide. L'étude des

structures cristallines montre que seuls les dix (voire onze) premiers résidus du peptide coactivateur sont visibles, les résidus manquants sont situés à l'extrémité C-terminale du peptide. La flexibilité de la chaîne latérale et/ou l'absence d'interaction avec la surface de la protéine peuvent expliquer l'absence de construction de ces résidus lors de la résolution et l'affinement des structures.

a. Stratégie de conception des séquences chimères oligoamide/oligourée

L'approche adoptée pour la conception des foldamères mixtes est basée sur l'étude des structures cristallines disponibles de la protéine VDR en interaction avec un de ses peptides coactivateurs. Les foldamères d'oligourée développés au sein de l'équipe ont la particularité d'adopter une structure en hélice 2,5 très stable grâce à la présence d'un réseau de liaisons hydrogène à trois centres entre fonctions urées. La forte similarité structurale entre l'hélice d'oligourée et l'hélice α naturelle permet d'envisager l'utilisation des foldamères d'oligourée comme peptidomimes d'hélice α (**Fig. 38**).

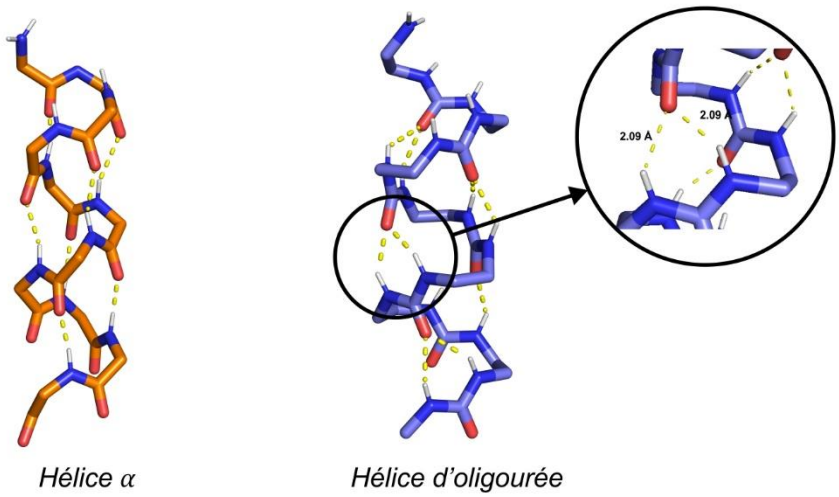


Figure 38 : Structures cristallines d'une hélice α (orange) et d'une oligourée (bleu).

A ce jour, la littérature ne fait pas état de structures cristallines de VDR en complexe avec un foldamère, ou avec des structures de type stapled peptides comme ceux décrits par Demizu *et al*[4]. Au sein de notre laboratoire, des structures cristallines de foldamères mixtes, aussi nommées chimères, d'oligourée/oligoamide sont disponibles. Ces oligomères mixtes ont une structure secondaire très stable en hélice[5] et une approche de type bottom-up a été envisagée pour concevoir les séquences de foldamères hybrides. Une superposition a été réalisée entre la première urée en position N-terminale (V_8^U) d'une séquence chimère **C1** (**Fig. 39**) et la structure du peptide coactivateur SRC1-2 de VDR (PDB : **2HC4**).

C1 Boc-NH-$L_1A_2L_3A_4L_5A_6L_7V_8{}^{U}A_9{}^{U}L_{10}{}^{U}V_{11}{}^{U}A_{12}{}^{U}L_{13}{}^{U}$

Figure 39 : Séquence du foldamère mixte oligoamide/oligourée **C1** développée au laboratoire. La topologie en hélice est conservée entre les portions en acide aminé et oligourée.

La superposition manuelle s'est faite à l'aide du logiciel de visualisation de protéine PyMol en utilisant le programme Pair Fitting. Les superpositions d'une hélice d'oligourée sur une hélice α révèlent que seuls trois et cinq résidus urée sont nécessaires pour mimer des séquences peptidiques de respectivement quatre et sept acides aminés (**Fig. 40**). La superposition entre les deux hélices révèle que parfois la chaîne latérale d'un résidu urée noté X^U, doit être déplacé sur le carbone C_α afin de bien mimer la séquence naturelle (**Fig. 40c**). Cette modification nécessite l'utilisation d'un nouveau résidu urée dit « inversé » noté $X_\alpha{}^U$.

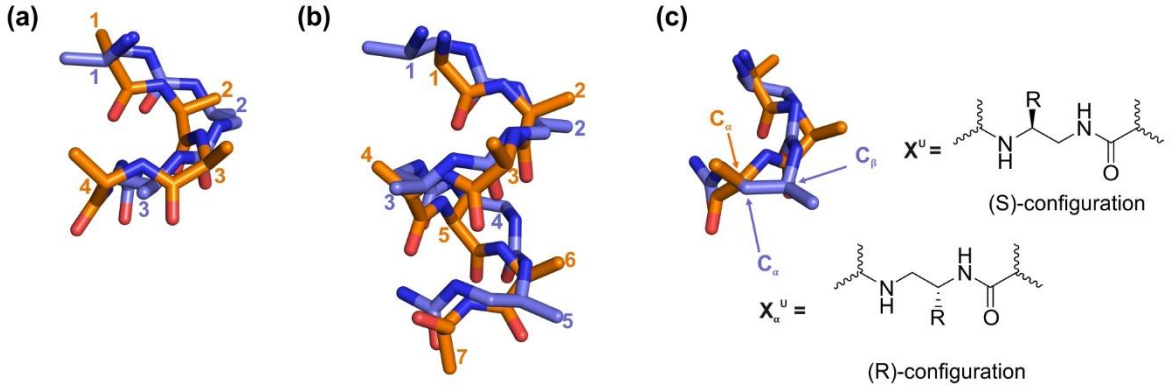


Figure 40 : Superposition des structures cristallines d'une hélice α naturelle (en orange) et d'une oligourée (en bleu). **(a)** Superposition de quatre résidus d'acides aminés avec trois résidus urées. **(b)** Superposition d'une hélice d'un heptapeptide avec celle d'une penta-urée. **(c)** Superposition montrant la modification de la position de la chaîne latérale du résidu urée sur le carbone C_α du résidu pour mimer la chaîne latérale de l'acide aminé.

L'objectif de ces superpositions est de pouvoir imaginer comment la structure du foldamère d'oligourée peut se positionner dans le sillon hydrophobe de la protéine VDR. Deux stratégies ont été envisagées : remplacer une portion du motif d'interaction contenant 4 acides aminés, ou remplacer 7 acides aminés dont les deux résidus adjacents au motif LXXLL.

b. Remplacement d'une portion du motif hélicoïdal d'interaction LXXLL

Cette première stratégie consiste à remplacer quatre résidus d'acides aminés par trois résidus urées. Les portions remplacées sont LXXL ou XXLL. Les acides aminés de ce motif seront numérotés ainsi : $L_1X_2X_3L_4L_5$ et l'ensemble des séquences chimères sont répertoriées dans le **Tableau 4**.

Nous nous sommes d'abord concentrés sur le remplacement des acides aminés de la partie $L_1X_2X_3L_4$ de la séquence de SRC1-2. Le premier résidu urée de **C1** (V_8^U) a été positionné sur le résidu leucine en position 1 du peptide. La superposition des deux structures montre que les chaînes latérales des résidus urées en position 1 et 3 sont bien projetées dans la poche hydrophobe, et miment bien les leucines aux positions 1 et 4 du peptide. Des résidus urées de type leucine ont été choisis pour conserver la nature des chaînes latérales clés de SRC1-2 projetées dans la poche hydrophobe de VDR (**Fig. 41a**). Les acides aminés histidine et arginine correspondant aux résidus X_2 et X_3 pourraient être remplacés par un résidu lysine urée afin de maintenir la nature cationique de ces résidus basiques. Un résidu histidine urée n'a pas été envisagé à cette position à cause des difficultés de synthèse de ce monomère, contrairement à celui de la lysine. Une première séquence chimère a ainsi été conçue : RHKIL^{u}K^{u}L^{u}LQEGS (**C2**).

Nous nous sommes ensuite intéressés au remplacement des acides aminés de la partie $X_2X_3L_4L_5$ (**Fig. 41b**). La superposition du résidu urée V_8^U de **C1** sur l'histidine (X_2) a été réalisée et conduit à une seconde séquence RHKILK^{u}L^{u}L^{u}QEGS (**C3**), l'urée dérivée de la lysine est conservée afin de mimer l'histidine et l'arginine, de même que les résidus leucine urées continuent à mimer les leucines naturelles. La superposition de SRC1-2 et **C1** suggèrent qu'il existe plusieurs séquences possibles où les chaînes latérales des acides aminés leucine L_4 et L_5 pourraient être déplacées sur l'atome de carbone C_α des analogues urées pour conserver l'effet hydrophobe (**Fig. 41c**). Le déplacement de la chaîne latérale sur le carbone C_α du résidu urée à la position 4 semble être la meilleure stratégie afin de mimer l'acide aminé naturel. Ainsi, trois autres séquences chimères **C4**, **C5** et **C6** ont été conçues pour mesurer l'impact de la modification de la position de la chaîne latérale de la leucine urée sur l'affinité et la structure des chimères.

Deux séquences supplémentaires contenant un résidu arginine urée (R^U) à la place des résidus histidine et arginine ont également été envisagées (**C7** et **C8**). Le remplacement du résidu urée lysine en position 2 par une arginine urée a pour objectif de créer une liaison hydrogène avec le résidu E285 de VDR. En effet, la chaîne latérale plus longue du résidu arginine urée pourrait amener les protons de la fonction guanidine à proximité du carboxylate de l'acide aminé E285. La chimère **C7** conserve les leucines urées à la place des acides aminés leucines en position 4 et 5. La séquence **C8** contient une alanine urée à la place du résidu leucine de SRC1-2 car la superposition suggère que la chaîne latérale du résidu urée serait trop longue, et ne pourrait pas être bien positionnée dans la poche hydrophobe, d'où le choix d'un résidu alanine urée (**Fig. 41c**).

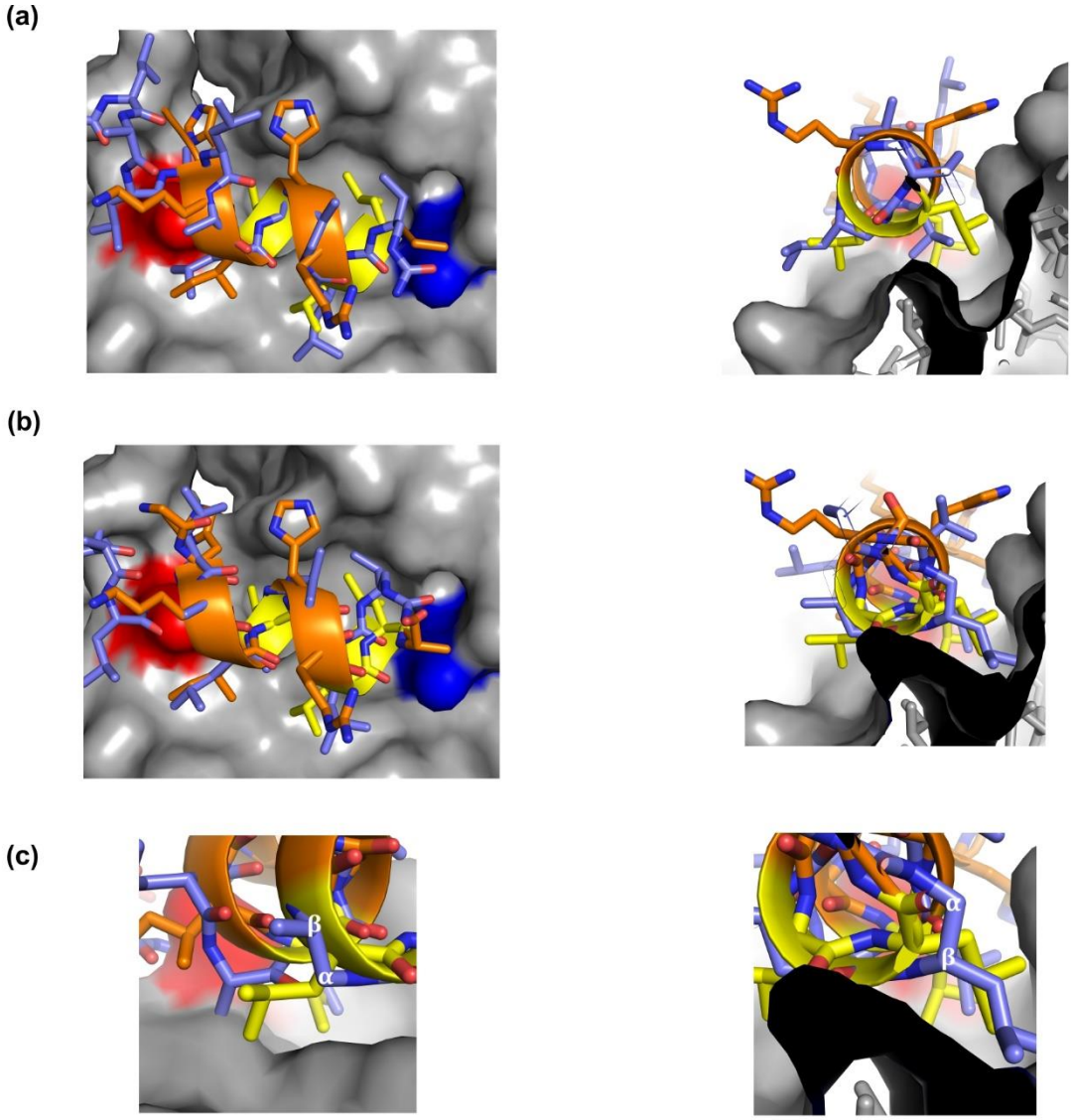


Figure 41 : Superposition de la séquence **C1** (en violet) sur une portion de SRC1-2 (en orange), la pince chargée est conservée avec les résidus E446 en rouge et K274 en bleu, les leucines en i, i+3 et i+4 de SRC-1-2 sont en jaune. **(a)** Superposition de **C1** sur SRC1-2, (droite) vue de face de la superposition du premier résidu urée de **C1** sur l'acide aminé leucine L_1 en position 1 de SRC1-2, (gauche) vue de profil de la superposition qui permet d'imaginer la projection des chaines latérales des résidus urées dans la poche hydrophobe. **(b)** Superposition de la première urée de **C1** sur l'acide aminé histidine en position 2 de SRC1-2, (droite) vue de face de la superposition, (gauche) vue de profil de la superposition permettant d'imaginer la projection des chaines latérales des résidus urées dans la poche hydrophobe. **(c)** Vues plus proche des superpositions des résidus urées sur la séquence naturelle aux positions 4 (droite) et 5 (gauche) des acides aminés leucine L_4 et L_5 permettant d'imaginer un déplacement de la chaine latérale du résidu urée sur le carbone C_α afin de mimer au mieux le résidu naturel.

séquences	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9
SRC1-2	R	H	K	I	L	H	R	L	L	Q	E	G	S
C2	R	H	K	I	L^{U}		K^{U}	L^{U}	L	Q	E	G	S
C3	R	H	K	I	L		K^{U}	L^{U}	L^{U}	Q	E	G	S
C4	R	H	K	I	L		K^{U}	L^{U}	L_{α}^{U}	Q	E	G	S
C5	R	H	K	I	L		K^{U}	L_{α}^{U}	L^{U}	Q	E	G	S
C6	R	H	K	I	L		K^{U}	L_{α}^{U}	L_{α}^{U}	Q	E	G	S
C7	R	H	K	I	L		R^{U}	L_{α}^{U}	L^{U}	Q	E	G	S
C8	R	H	K	I	L		R^{U}	L_{α}^{U}	A^{U}	Q	E	G	S
chimères	-4	-3	-2	-1	1		2	3	4	5	6	7	8

Tableau 4 : Séquences des chimères **C2-C8** conçues par superposition manuelle du motif LXXLL de SRC1-2 avec la chimère modèle **C1**. Les résidus en orange correspondent aux résidus d'acides aminés mimés par les résidus urée en bleu. La nouvelle numérotation des positions des résidus des chimères **C2-C8** est indiquée en bleu en bas du tableau.

c. Remplacement de l'intégralité du motif hélicoïdal d'interaction LXXLL

La seconde approche de remplacement du motif d'interaction consiste à remplacer l'intégralité du motif LXXLL mais en conservant les deux résidus adjacents du motif, soit remplacer la séquence $I_{-1}LHRLLQ_6$ de SRC1-2 par une hélice d'oligourée, soit deux tours d'hélice α par deux tours d'hélice oligourée. La superposition de la première urée N-terminale de **C1** sur l'acide aminé isoleucine I_{-1} suggère que la présence de cinq urées peut être tolérée. Bien que l'hélice d'oligourée ait un diamètre plus important, la projection des chaînes hydrophobes est conservée tout comme la pince chargée formée avec les résidus E446 et K274 de VDR (**Fig. 42a**). Le résidu lysine urée est conservé pour mimer les acides aminés histidine (X_2) et arginine (X_3), la leucine urée est également conservée pour mimer la leucine naturelle de l'acide aminé en position 1. Les acides aminés leucine aux positions 4 et 5 sont remplacés respectivement par des résidus urées de type valine et norleucine afin de conserver les interactions clés dans la poche hydrophobe. L'isoleucine est remplacée par une alanine urée inversée car la superposition suggère qu'il serait plus judicieux de déplacer la chaîne latérale de l'alanine urée en position C_{α} et d'avoir une chaîne latérale plus courte (**Fig. 42c**). Le nombre de résidus par tour étant plus faible que celui de l'hélice α, le résidu glutamine Q_6 ne peut pas être mimé dans la séquence d'oligourée. L'ensemble de ces observations a permis de concevoir la séquence **C9** (**Tableau 5**).

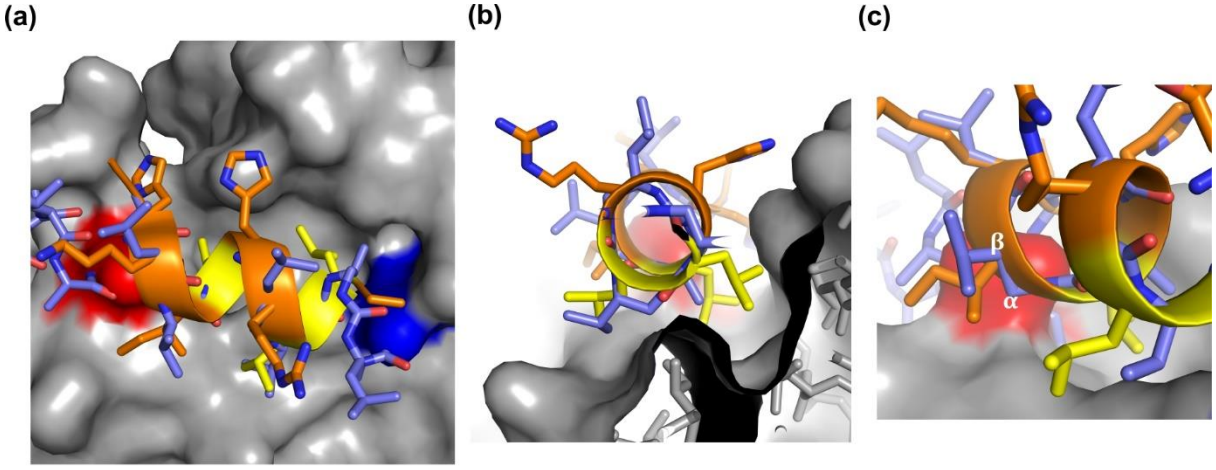


Figure 42 : Superposition de la séquence **C1** (en violet) sur l'intégralité du motif LXXLL de SRC1-2 (en orange), la pince chargée est conservée avec les résidus E446 en rouge et K274 en bleu, les leucines en i, i+3 et i+4 de SRC1-2 sont en jaune. **(a)** Vue de face de la superposition de **C1** sur SRC1-2 où la première urée N-terminale de **C1** est superposée sur l'acide aminé isoleucine de SRC1-2. **(b)** Vue de profil de la superposition permettant de voir la projection des chaînes latérales des résidus urées dans la poche hydrophobe avec des résidus urées de type valine et norleucine pour mimer les acides aminés leucine aux positions 4 et 5. **(c)** Vue rapprochée de la superposition du résidus urées sur la séquence naturelle sur le résidu isoleucine à la position -1 suggérant un déplacement de la chaine latérale du résidu urée sur le carbone C_α pour mimer au mieux ce résidu naturel.

Nous nous sommes enfin intéressés à insérer le motif oligourée de la chimère **C9** dans d'autres séquences de protéines SRC contenant ce motif LXXLL. La séquence de la deuxième NR box de SRC2, SRC2-2, a été retenue. Bien qu'elle contienne des acides aminés adjacents au motif LXXLL différents de SRC1-2, les résidus X_2X_3 sont identiques et la séquence **C10** (**Tableau 5**) a été imaginée. La séquence de la NR box 3 de SRC2 a été également choisie comme structure à mimer. Les résidus adjacents au motif d'interaction, de même que les acides aminés X_2 et X_3 sont différents. La séquence chimère **C11** a été conçue avec un résidu tyrosine urée à la place de l'urée dérivée de la lysine pour remplacer les résidus X_2 et X_3.

Les différentes superpositions de **C1** sur les séquences naturelles de peptides coactivateurs ont permis de concevoir une série de dix séquences chimères. Leur synthèse a récemment été décrite dans la littérature[5b], celle-ci repose sur un couplage séquentiel de monomères dérivés d'éthylène diamine activés sous forme de carbamate de succinimide dans lequel l'amine est masquée sous forme d'azoture. La préparation de ces monomères urée azoture/carbamate de succinimide est donc nécessaire pour réaliser la synthèse des chimères **C2-C11**. L'ensemble de ces travaux a été réalisé par mes soins.

séquences	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9
SRC1-2	R	H	K	I	L	H	R	L	L	Q	E	G	S
C9	R	H	K	A_{α}^{U}	L^{U}	K^{U}		V^{U}	Nle^{U}		E	G	S
SRC2-2	K	H	K	I	L	H	R	L	L	Q	D	S	S
C10	K	H	K	A_{α}^{U}	L^{U}	K^{U}		V^{U}	Nle^{U}		D	S	S
SRC2-3	E	N	A	L	L	R	Y	L	L	D	K	D	D
C11	E	N	A	A_{α}^{U}	L^{U}	Y^{U}		V^{U}	Nle^{U}		K	D	D
chimères	-4	-3	-2	-1	1	2		3	4		5	6	7

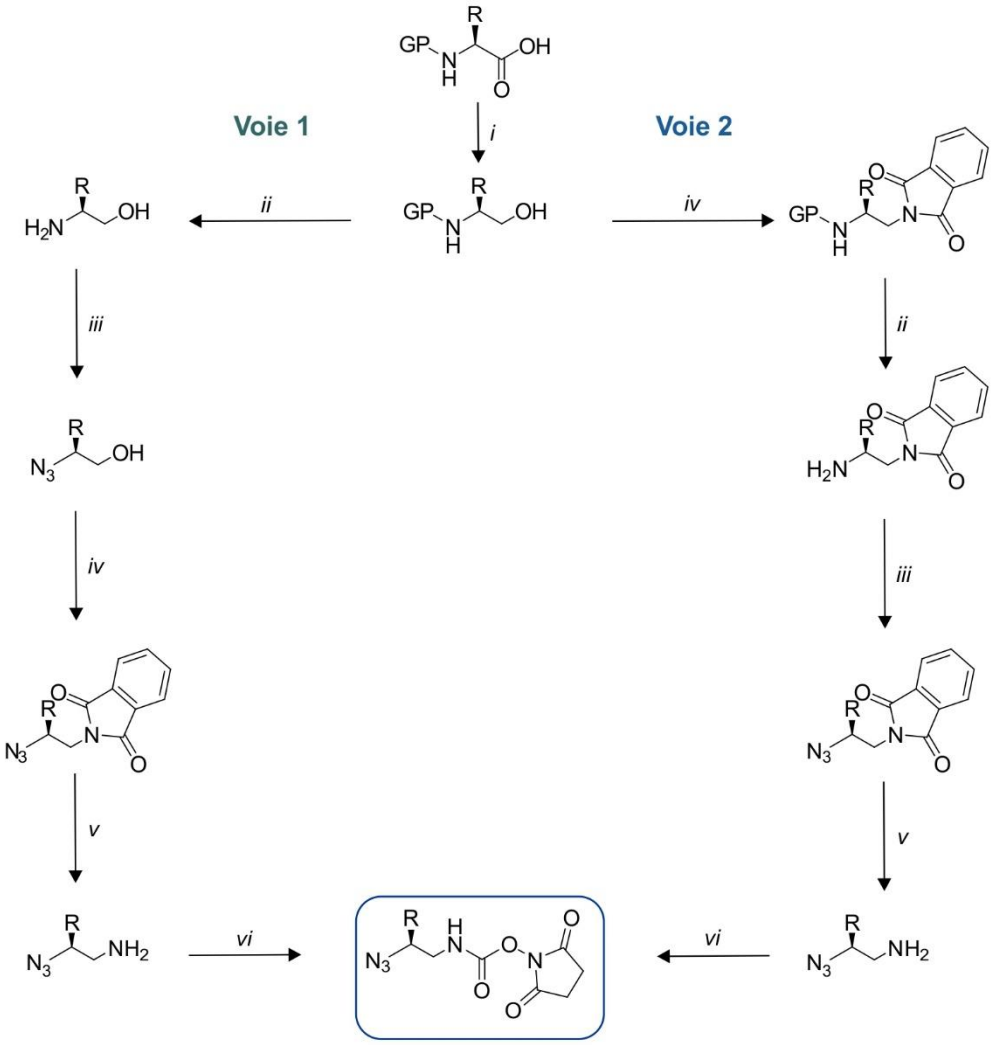
Tableau 5 : Séquences des chimères **C9-C11**. Les résidus en orange correspondent aux résidus d'acides aminés mimés par les résidus urée en bleu. La nouvelle numérotation des positions des résidus des chimères **C9-C11** sont indiqués en bleu en bas du tableau.

2. <u>Préparation des monomères azoture/carbamate de succinimide</u>

La préparation des monomères activés se fait par une synthèse multi-étapes en solution[6] à partir d'acides aminés commerciaux dont l'amine est protégée par un groupement protecteur de type Boc ou Fmoc. Afin de pouvoir accéder aux monomères activés, deux voies de synthèse ont été développées selon la nature de la chaîne latérale du monomère utilisé (**Schéma 1**). La première étape est commune aux deux voies de synthèse et consiste à la réduction de l'acide carboxylique en alcool. Dans la voie 1, la fonction amine du dérivé β-amino-alcool est déprotégée puis convertie en azoture via la méthode de Wong[7] en utilisant le chlorhydrate d'imidazole-1-sulfonyl azoture ($N_3SO_2Im.HCl$) comme agent de transfert[8]. Une réaction de Mitsunobu convertit les alcools primaires des dérivés de type 2-azidoéthanol obtenus en phtalimide[9]. Des chaînes latérales à faible poids moléculaires (isobutyle, isopropyle et méthyle) rendent les intermédiaires amino-alcool volatils et la première voie de synthèse se révèle alors peu efficace. Ainsi la voie 2 a été développée, où l'alcool est d'abord converti en un dérivé phtalimide par la réaction de Mitsunobu, puis l'amine primaire est générée par déprotection du groupement protecteur tert-butoxycarbonyle en milieu acide, et convertie en azoture en présence de l'agent de transfert. Ces deux voies de synthèse conduisent à l'obtention d'un intermédiaire de type azido-phtalimide qui est hydrolysé dans l'hydrazine monohydratée pour conduire au dérivé azido-amine correspondant. L'étape finale consiste à activer ce dérivé azido-amine sous forme de carbamate de succinimide par réaction avec du carbonate *N,N'*-disuccinimide (DSC).

La synthèse des monomères activés dérivés de la lysine (**M1**), de la tyrosine (**M2**), de l'alanine (**M3**), de la leucine (**M4**) et de la valine (**M5**) a été réalisée selon ces deux voies et les monomères activés ont été obtenus avec de bons rendements (**Tableau 6**), en accord avec ceux rapportés dans la littérature[6]. Une RMN du proton des monomères finaux permet

de confirmer la présence du succinimide qui est caractérisé par un singulet intégrant pour quatre hydrogènes à 2,82 ppm dans le chloroforme deutéré.



GP : Groupement protecteur Fmoc (Voie 1) ou Boc (Voie 2) ; i : 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii* : DBU, THF, TA (Voie 1) ou TFA, TA (Voie 2) ; *iii* : $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4.5H_2O$, CH_3CN, H_2O, TA ; *iv* : DIAD, PPh_3, phtalimide, THF, TA ; *v* : N_2H_4, MeOH, TA ; *vi* : DSC, DCM, TA.

Schéma 1 : Synthèse multi-étape des monomères azoture/carbamate de succinimide

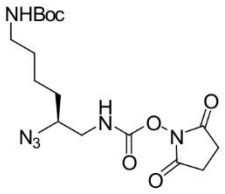
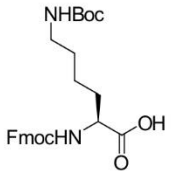
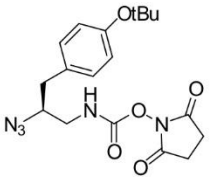
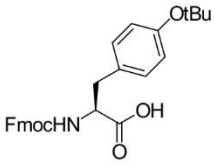
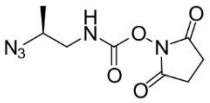
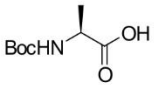
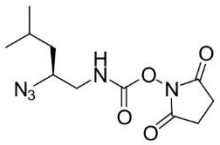
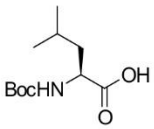
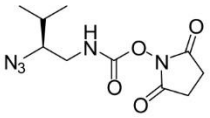
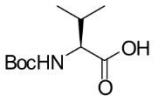
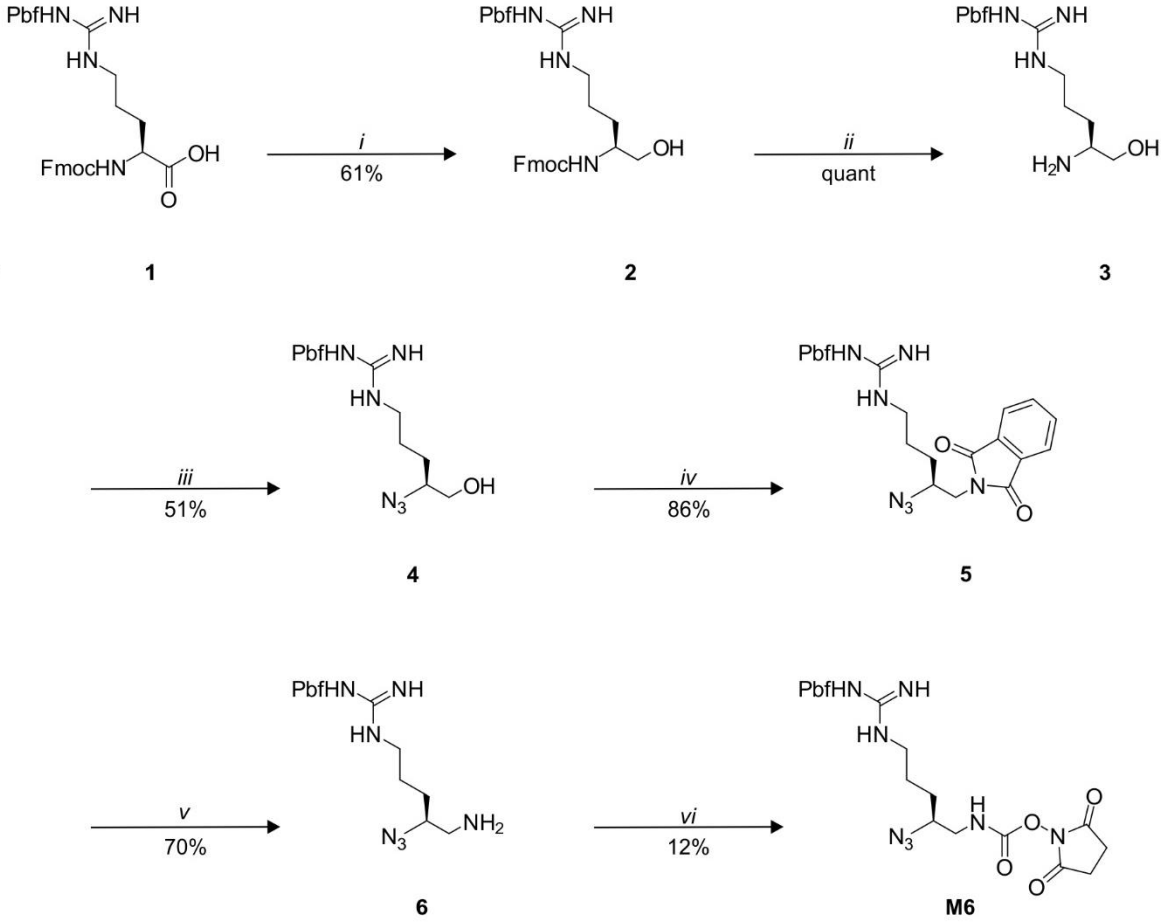
monomère MX		acide aminé de départ	voie de synthèse	Rendement global
M1			**1**	37%
M2				20%
M3			**2**	21%
M4				48%
M5				38%

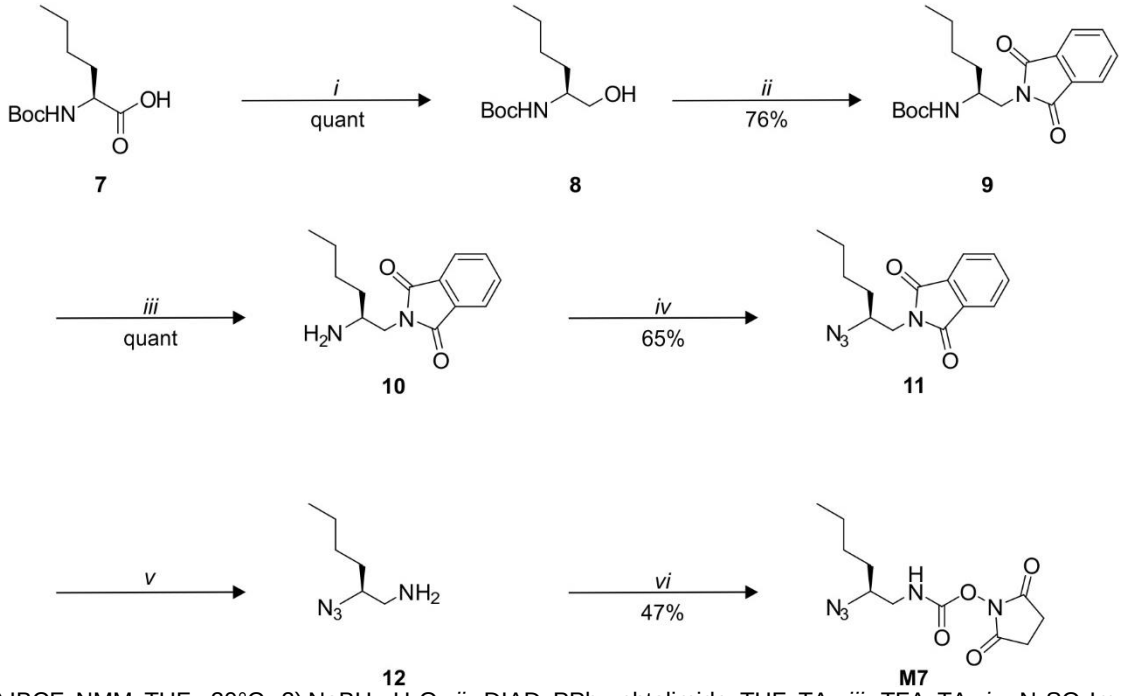
Tableau 6 : Synthèse des monomères **MX** azoture/carbamate de succinimide préparés selon les voies de synthèse décrites dans la littérature[6].

La synthèse des monomères activés dérivés de l'arginine et de la norleucine a été réalisée en suivant la même stratégie de synthèse des monomères décrits précédemment. L'azoture/carbamate de succinimide dérivé de l'arginine (**M6**) a été obtenu selon la voie 1 décrite dans le **Schéma 2** à partir de l'acide aminé commercial Fmoc-L-Arg(Pbf)-OH **1** dont la fonction guanidine est protégée par un groupement 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyle (Pbf). Le monomère **M6** a été obtenu avec un rendement global de 2%. Une purification par chromatographie sur colonne de gel de silice (acétate d'éthyle/dichlorométhane, 2:3) a été réalisée après une première purification par précipitation à froid dans un mélange éther/cyclohexane (8:2). L'analyse RMN du proton dans du chloroforme deutéré après cette précipitation révèle la présence de résidus d'oxyde de triméthylphosphine issus de l'étape de Mitsunobu. L'analyse RMN du proton dans du chloroforme deutéré après la chromatographie sur colonne montre que le monomère est pur. Cette étape de purification supplémentaire peut expliquer ce faible rendement.



i : 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii* : DBU, THF, TA ; *iii* : $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4.5H_2O$, CH_3CN, H_2O, TA ; *iv* : DIAD, PPh_3, phtalimide, THF, TA ; *v* : N_2H_4, MeOH, TA ; *vi* : DSC, DCM, TA.

Schéma 2 : Schéma de synthèse par la voie 1 du monomère azoture/carbamate de succinimide **M6**.

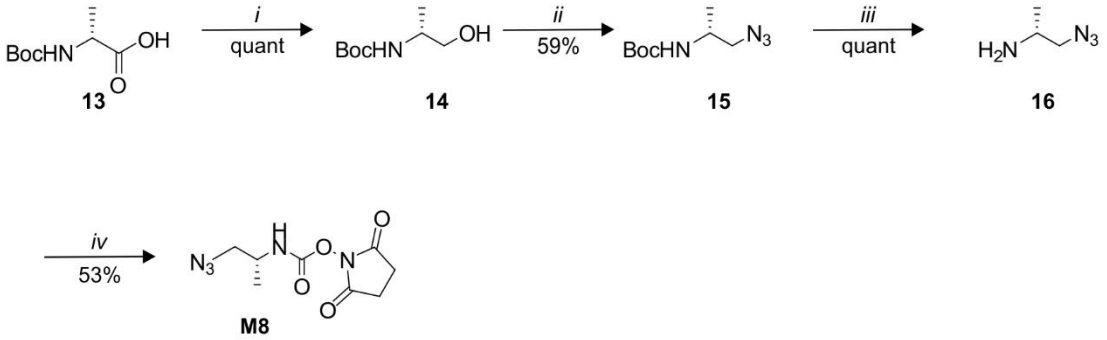


i : 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii* : DIAD, PPh_3, phtalimide, THF, TA ; *iii* : TFA, TA ; *iv* : $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4.5H_2O$, CH_3CN, H_2O, TA ; *v* : N_2H_4, MeOH, TA ; *vi* : DSC, AcOEt, TA

Schéma 3 : Schéma de synthèse par la voie 2 du monomère azoture/carbamate de succinimide **M7**.

L'azoture/carbamate de succinimide dérivé de la norleucine (**M7**) a été obtenu selon la voie 2 décrite dans le **Schéma 3** à partir de l'acide aminé commercial Boc-L-Nle-OH **7**. Le monomère **M7** a été obtenu avec un rendement global de 23%, ce qui en accord avec les rendements globaux des monomères déjà décrits.

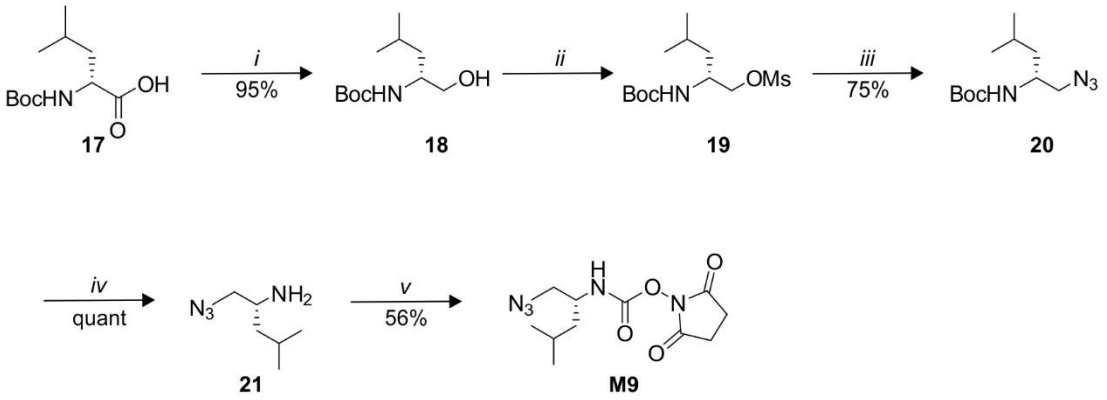
Certaines séquences des chimères nécessitent des monomères activés dont la chaîne latérale est sur le carbone C_α de l'urée. La préparation de ces monomères « inversés » ne suit pas la voie de synthèse décrite précédemment. La synthèse du monomère de l'alanine inversé (**M8**) a déjà été décrite dans la littérature par le laboratoire[10] (**Schéma 4**), et il a été préparé dans les mêmes conditions à partir de l'acide aminé Boc-D-Ala-OH **13** avec un rendement global de 31%.



i : 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii* : DPPA, DBU, NaN_3, DMF, 70°C ; *iii* : TFA, TA ; *iv* : DSC, DIPEA, DCM, TA

Schéma 4 : Schéma de synthèse du monomère azoture/carbamate de succinimide **M8**.

La synthèse du monomère leucine inversé (**M9**) n'a pas été rapportée dans la littérature. La synthèse d'un monomère dérivé de la (*S*)-phénylalanine dont la chaîne latérale est sur le carbone C_α avait été décrite[10]. Nous nous en sommes inspirés pour préparer **M9** à partir de la Boc-D-Leu-OH **17** (**Schéma 5**). La première étape est commune aux précédentes synthèses de monomères, et consiste à la réduction de l'acide carboxylique en alcool. Cet alcool est transformé en mésylate puis converti en azoture en utilisant de l'azoture de sodium. L'amine primaire **21** est ensuite générée par déprotection du groupement protecteur tert-butoxycarbonyle en milieu acide. L'étape finale consiste à activer le dérivé azido-amine sous forme de carbamate de succinimide par réaction avec la DSC. Le monomère **M9** a été obtenu avec un rendement global de 25%.



i : 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii* : MsCl, Et_3N, DCM, 0°C, Ar ; *iii* : NaN_3, DMF, 70°C ; iv :TFA, TA ; *v* : DSC, DIPEA, DCM, TA

Schéma 5 : Schéma de synthèse du monomère azoture/carbamate de succinimide **M9**.

3. Synthèse sur support solide des chimères **C2-C11**

a. Principe de la synthèse sur support solide

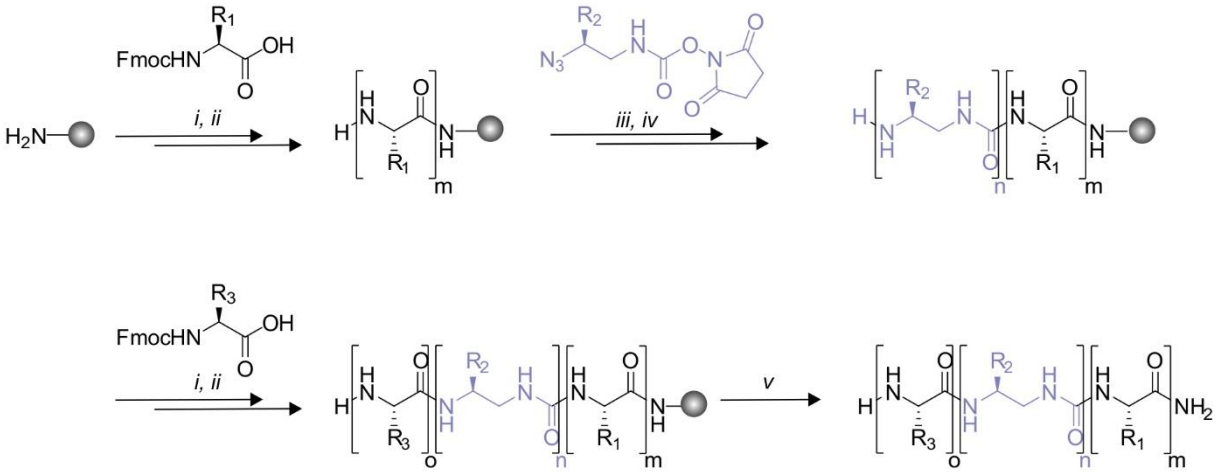
La synthèse sur support solide (SPS), introduite par Merrifield[11] en 1963, a permis de révolutionner les synthèses peptidique et polypeptidique. Le concept général de cette méthodologie repose sur un principe simple : l'attachement du premier acide aminé de la séquence à synthétiser sur une résine polymérique via une liaison covalente. L'addition des acides aminés se fait un par un jusqu'à obtention de la séquence désirée puis le clivage de la résine permet de récupérer le peptide en solution. L'intérêt de la synthèse supportée est de faciliter l'élimination des réactifs en fin de réaction par simple filtration et permet donc l'emploi de large excès de réactifs permettant d'atteindre des rendements élevés (quasi-quantitatifs dans le cas de la synthèse peptidique) à chaque étape. Cette technologie a permis la simplification des manipulations et ainsi la réduction des temps de synthèse des peptides. Le polymère utilisé comme résine doit être insoluble dans les solvants utilisés, être stable et doit contenir un groupe fonctionnel auquel le premier acide aminé protégé pourra bien se greffer via une liaison covalente.

Les oligourées aliphatiques développées dans notre équipe sont synthétisées en solution[12] et sur support solide[6]. Leur synthèse repose sur un couplage séquentiel de monomères dérivés d'éthylène diamine activés sous forme de carbamate de succinimide. Ces monomères permettent d'accéder à des séquences courtes (6-12 résidus) avec de bons rendements et de bonnes puretés. L'irradiation micro-onde permet d'améliorer la synthèse sur support solide : le temps de réaction est réduit et il y a moins de risque d'agrégation. Cependant, il a été montré que des monomères activés contenant un groupement Fmoc pour

protéger l'amine sont incompatibles avec l'utilisation de champ micro-onde : ils se dégradent et un phénomène d'oligomérisation non contrôlée a lieu sur la résine. Afin de pouvoir utiliser une résine labile à l'acide trifluoroacétique et d'éviter l'utilisation de groupement Fmoc, une stratégie dans laquelle l'amine est masquée sous forme d'azoture a été développée au laboratoire. La réduction de l'azoture en amine se fait en présence de triméthylphosphine[13]. L'introduction des blocs urées repose donc sur cette stratégie.

b. Méthodologie de synthèse sur support solide de foldamères mixtes oligoamide/oligourée

La synthèse sur support solide des oligomères chimères a été privilégiée en raison de sa flexibilité et de sa rapidité (**Schéma 6**). Cette méthodologie de synthèse utilise le couplage sous champ micro-onde de monomères azotures/carbamate de succinimide. L'amine est ensuite régénérée par réduction de l'azoture via une réaction de Staudinger sous irradiation micro-onde. Cette étape de déprotection permet l'introduction du monomère suivant. Les résines utilisées pour la synthèse sur support sont la résine polyéthylène glycol NovaPEG Rink Amide et la résine polystyrène Rink Amide MBHA de MerckMillipore. La substitution initiale des résines varie de 0,45 mmol/g à 0,78 mmol/g et la synthèse est réalisée sur une échelle de 75 μmol. La résine polystyrène Rink Amide MBHA est protégée par un groupement Fmoc qui nécessite d'être retiré en présence d'un mélange pipéridine/DMF (2:8) avant de réaliser la synthèse. Ces résines sont des supports solides hydrophiles compatibles avec des solvants organiques et aqueux, ce qui est idéal au vu des solvants utilisés dans la réduction de Staudinger[13]. La première étape de la synthèse des foldamères mixtes consiste au couplage des acides aminés α protégés par un groupement Fmoc sous irradiation micro-ondes (via l'appareil CEM Discovery) en présence de diisopropylcarbodiimide (DIC) et d'hydroxybenzotriazole (HOBt). L'utilisation d'HOBt est essentielle car le fort encombrement stérique de ce composé empêche la racémisation de l'acide aminé couplé. Le couplage des monomères se fait en présence de N,N'-diisopropyléthylamine, et la déprotection de ces monomères est effectuée par une réaction de Staudinger en présence de triméthylphosphine[13]. Une fois la chimère obtenue, celle-ci est clivée de la résine par un mélange acide trifluoroacétique (TFA), eau et triisopropylsilane (TIS) (95:2,5:2,5). Le TFA permet de séparer le peptide de la résine, mais il permet également d'enlever les groupements protecteurs des chaînes latérales sensibles en milieu acide. Le TIS est utilisé pour piéger les carbocations générés par le clivage de la résine et des chaînes latérales. Une fois clivées de leur support, les chimères sont purifiées par HPLC semi-préparative et caractérisées par une analyse HPLC en phase inverse et spectrométrie de masse (*electrospray*).



i :Fmoc-Xaa-OH (5 equiv), HOBt (5 equiv), DIC (5 equiv), (25 W, 75°C, 5 min) ; *ii* : 20% pipéridine, DMF, (50 W, 50°C, 8 min) x2; *iii* : monomère azoture protégé activé **MX^{U}** (1,5 equiv), DIPEA (3 equiv), (50 W, 50°C, 20 min) x2 ; *iv* : PMe_3 (1M) dans le THF (10 equiv), dioxane/H_2O (7:3), (50 W, 50°C, 15 min) x2; *v* : TFA/TIS/H_2O (95:2,5:2,5).

Schéma 6 : Schéma de synthèse sur support solide des oligomères chimères **C2**-**C11**.

*c. Résultats des synthèses sur support solide des chimères **C2-C11***

Les chimères **C2-C11** ont été synthétisées selon la méthode décrite précédemment. Le chromatogramme HPLC de **C2** avant purification montre la présence d'un pic très majoritaire à 4,03 min (**Fig. 43a**). Des analyses par spectrométrie de masse ont révélé que ce pic majoritaire correspondait bien à la chimère **C2** attendue qui a été obtenue avec un rendement de 35% après purification par HPLC semi-préparative.

L'ensemble des chimères synthétisées (**C3-C10**) présente les mêmes caractéristiques de pureté des bruts réactionnels et de rendement (**Tableau 7**).

chimères	séquences	pureté brut réactionnel	ESI-MS $[M+H]^+$	HPLC t_R (min)[a]	rendement[b]
C2	RHKIL^{U}K^{U}L^{U}LQEGS	86%	1508,80	4,03	35%
C3	RHKILK^{U}L^{U}L^{U}QEGS	86%	1508,86	4,63	20%
C4	RHKILK^{U}L^{U}L$_{\alpha}$^{U}QEGS	69%	1508,97	4,35	21%
C5	RHKILK^{U}L$_{\alpha}$^{U}L^{U}QEGS	75%	1508,93	4,90	25%
C6	RHKILK^{U}L$_{\alpha}$^{U}L$_{\alpha}$^{U}QEGS	98%	1508,86	4,91	36%
C7	RHKILR^{U}L$_{\alpha}$^{U}L^{U}QEGS	88%	1536,86	4,53	49%
C8	RHKILR^{U}L$_{\alpha}$^{U}A^{U}QEGS	97%	1494,90	4,47	56%
C9	RHKA$_{\alpha}$^{U}L^{U}K^{U}V^{U}Nle^{U}EGS	61%	1381,80	3,95	5%
C10	KHKA$_{\alpha}$^{U}L^{U}K^{U}V^{U}Nle^{U}DSS	64%	1369,80	4,77	5%
C11	ENAA$_{\alpha}$^{U}L^{U}Y^{U}V^{U}Nle^{U}KDD	59%	1394,60	4,90	7%

Tableau 7 : Séquences, temps de rétention HPLC et rendements de synthèse des chimères **C2-C11**. [a] système RP-HPLC : gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement après purification HPLC.

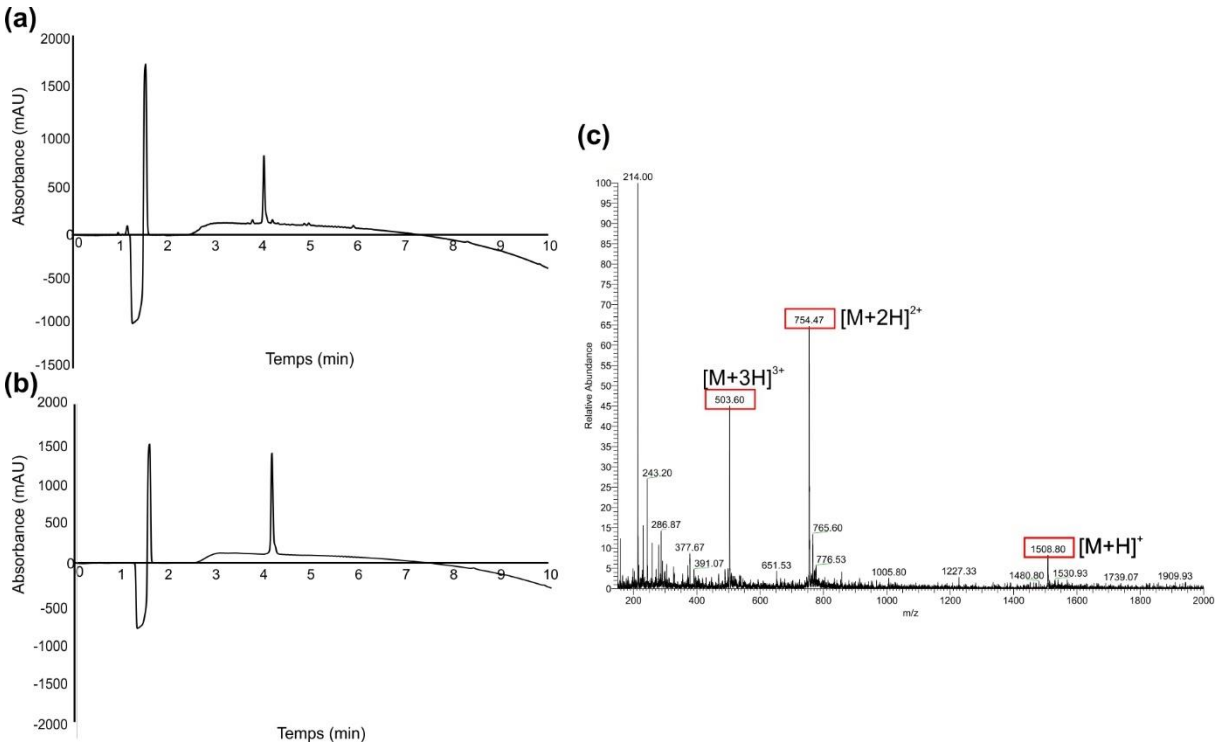


Figure 43 : Caractérisation de la chimère **C2** : **(a)** chromatogramme HPLC de **C2** avant purification ; **(b)** chromatogramme HPLC de **C2** après purification par RP-HPLC semi-préparative sur colonne C18 (gradient linéaire de 20 à 40% de B en 20 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) ; **(c)** spectre de masse de **C2** après purification par RP-HPLC semi-préparative.

Ces résultats permettent d'affirmer que l'insertion d'un monomère urée inversé dans ces séquences ne pose pas de difficulté de synthèse.

La synthèse de la chimère **C11** a été réalisée en suivant le même procédé. Une fois clivée de son support, la chimère obtenue est analysée par HPLC et spectrométrie de masse. Le chromatogramme HPLC indique la présence de plusieurs impuretés (**Fig. 42**), corroborées par l'analyse par spectrométrie de masse, correspondant à des délétions. Le composé **C11** a été purifié par RP-HPLC semi-préparative avec un rendement de 7%. Les pics de délétions identifiés par spectrométrie de masse semblent correspondre à des séquences où les réductions des fonctions azotures des résidus urées n'auraient pas eu lieu, et permettent ainsi d'expliquer le faible rendement de synthèse.

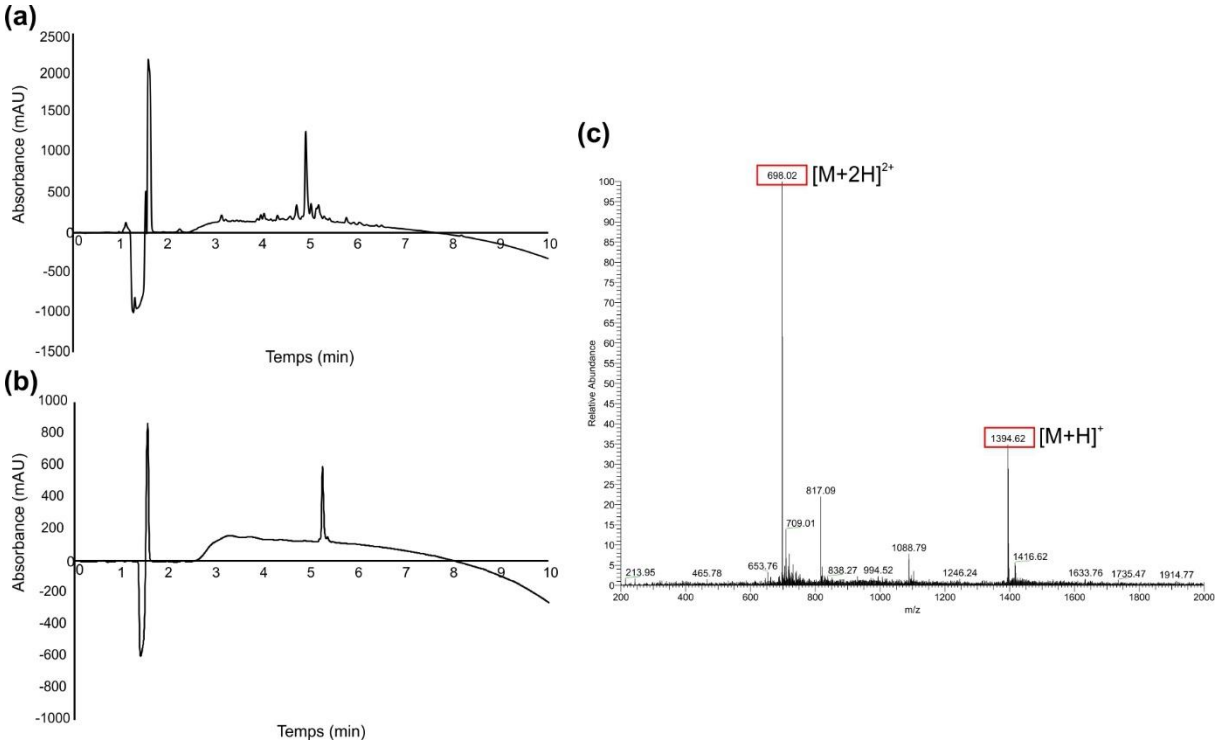


Figure 44 : Caractérisation de la chimère **C11** : **(a)** chromatogramme HPLC de **C11** avant purification ; **(b)** chromatogramme HPLC de **C11** après purification par RP-HPLC semi-préparative sur colonne C18 (gradient linéaire de 20 à 60% de B en 20 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) ; **(c)** spectre de masse de **C11** après purification par RP-HPLC semi-préparative.

4. Conclusion sur la génération de foldamères mixtes oligoamide/oligourée

Une étude de la zone d'interaction entre le peptide SRC1-2 et la protéine VDR, ainsi qu'une superposition manuelle d'une séquence chimère **C1** sur SRC1-2 a permis de concevoir une dizaine de séquences de foldamères mixtes oligoamide/oligourée. La synthèse sur support solide des chimères a été réalisée avec de bons rendements. Ces synthèses ont également permis de révéler que l'insertion de nouveaux résidus urées dont la chaine latérale est sur le C_α n'a pas d'impact sur le bon déroulement de la synthèse.

II. Caractérisation bio-structurale des foldamères mixtes **C2-C11**

Une étude de type relation structure-affinité (*structure-affinity relationship*, SAR) a été réalisée sur les chimères **C2**-**C11**, en collaboration avec le Dr. Natacha Rochel-Guiberteau de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) de Strasbourg. Les caractérisations structurales (DRX) et biophysiques (affinité, stabilité) ont été réalisées par son équipe. Les différents résultats obtenus ont permis d'étudier les foldamères mixtes oligoamide/oligourée comme potentiels inhibiteurs de l'interaction VDR/SRC1-2, et de

comparer l'effet des modifications ponctuelles de ces séquences sur l'affinité et la stabilité des complexes formés.

1. Caractéristiques et statistiques d'affinements des données cristallographiques

La cristallisation des complexes VDR/chimère a été réalisée au sein de l'équipe du Dr. Natacha Rochel-Guiberteau. Les chimères ont été co-cristallisées avec la protéine zVDR issue du poisson-zèbre (*Danio rerio*) et la vitamine D_3 en utilisant la méthode de la goutte suspendue. Les conditions de cristallisation des chimères sont répertoriées dans le **Tableau 8**. Les structures cristallines des chimères en complexe avec VDR ont été obtenues en haute résolution par diffraction des rayons X. Les données ont été collectées aux synchrotrons SOLEIL (Gif-sur-Yvette) et ESRF (Grenoble). Seules trois chimères n'ont pas pu être étudiées par cristallographie : **C4**, **C9** et **C10,** car aucun cristal n'a été obtenu avec zVDR.

chimères	protéine et ligand	condition de cristallisation
C2	zVDR-LBD + $VitD_3$	Sodium acétate 2,5 M, HEPES 0,1 M pH 7,5
C3	zVDR-LBD + $VitD_3$	Sulfate de lithium 1,8 M, HEPES 0,1 M pH 7,5
C5	zVDR-LBD + $VitD_3$	Ammonium acétate 3,5 M, TRIS 0,1 M pH 8,0
C6	zVDR-LBD + $VitD_3$	Sodium acétate 3,0 M, MES 0,1 M pH 6,0
C7	zVDR-LBD + $VitD_3$	Sodium acétate 3,0 M, TRIS 0,1 M pH 8,0
C8	zVDR-LBD + $VitD_3$	Sodium acétate 3,0 M, TRIS 0,1 M pH 8,0
C11	zVDR-LBD + $VitD_3$	Sulfate de lithium 1,8 M, MES 0,1 M pH 6,0

Tableau 8 : Conditions de cristallisation des chimères **C2**-**C11** ayant cristallisées avec la protéine zVDR.

La résolution des structures, effectuée par le Dr.Natacha Rochel-Guiberteau, a été réalisée par remplacement moléculaire avec le logiciel Phaser en utilisant la structure naturelle de VDR (PDB **2HC4)** comme modèle de départ. L'affinement et la construction des modèles atomiques des chimères ont respectivement été effectués par mes soins grâce aux logiciels Phenix et Coot. Les jeux de données ainsi que les statistiques d'affinement sont donnés dans le **Tableau 9** et permettent de valider les modèles obtenus. L'affinement à haute résolution de la structure du complexe VDR/**C3** n'a pas pu être réalisé dû à la qualité des données collectées. Par conséquent les jeux de données ainsi que les statistiques d'affinement n'ont pas pu être validés, et l'étude structurale de ce complexe sera effectuée avec le modèle construit de **C3**.

	VDR/SRC1-2 (2HC4)	C2	C3*	C5	C6	C7	C8	C11
Résolution (Å)	20,0-2,20 (2,28-2,20)	20,4-2,37 (2,55-2,37)	43,0-1,98 (2,05-1,98)	22,0-1,95 (2,02-1,95)	32,96-2,35 (2,53-2,35)	38,0-2,30 (2,47-2,30)	32,9-2,00 (2.09-2.00)	14,9-2,00 (2,09-2,00)
Groupe d 'espace	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P6_522$
Paramètre de maille (Å)	a = b = 65,84; c = 262,84 α = β = 90 γ = 120	a = b = 66,09; c = 260,14 α = β = 90 γ = 120	a = b = 66,18; c = 261,37 α = β = 90 γ = 120	a = b = 66,35; c = 262,37 α = β = 90 γ = 120	a = b = 65,91; c = 262,42 α = β = 90 γ = 120	a = b = 66,20; c = 254,22 α = β = 90 γ = 120	a = b = 66,37; c = 253,72 α = β = 90 γ = 120	a = b = 66,11; c = 262,48 α = β = 90 γ = 120
Nombre de réflexions uniques	18382	14409	24519	25904	14937	15313	23248	23172
Redondance	10,2	6,0	*nd*	7,3	9,1	13,7	9,4	1,6
R_{sym} (%)[a]	4,8	8,4	*nd*	9,3	8,7	7,4	5,9	1,2
Complétude (%)	99,0 (99,5)	99,0 (98,0)	99,1 (99,0)	99,4 (99,0)	99,6 (99,0)	98,2 (97,0)	99,1 (99,8)	96.2 (80,0)
Moyenne *I*/σ (%)	29,1 (7,3)	12,3 (1,34)	*nd*	12,3 (1,33)	17,6(1,37)	21,2(1,36)	19,7 (1,36)	16,0 (1,34)
RMSD.[b] sur les liaisons (Å)	0,006	0,007	0,013	0,008	0,007	0,006	0,006	0,006
sur les angles (°)	1,13	1,10	1,81	1,15	1,13	1,00	0,98	0,87
R_{work}/R_{free} [c] (%)	22,0/25,1	21,3/24,6 (35,4/41,4)	26,0/29,0 (51,1/55,7)	19,2/21,7 (38,55/42,91)	20,3/22,7 (29,9/34,3)	19,2/23,9 (26,6/34,6)	18,3/21,7 (26,0/28,4)	19,9/22,1 (24,3/26,7)
Facteur B								
Protéine		85,21	*nd*	57,89	70,86	66,04	56,71	49,58
Ligand		71,01	*nd*	52,11	70,27	61,76	58,01	45,96
Eau		92,98	*nd*	67,65	77,32	70,37	66,81	55,26

Tableau 9 : Statistiques des jeux de données et de l'affinement des complexes avec la séquence naturelle et avec les séquences chimères **C2**-**C11**. Les valeurs entre parenthèses correspondent aux très hautes résolutions. [a] $R_{sym}(I) = \sum_{h\,k\,l}\sum_i |I_{h\,k\,l,i} - \langle I_{h\,k\,l}\rangle| / \sum_{h\,k\,l}\sum_i |I_{h\,k\,l,i}|$ avec $\langle I_{h\,k\,l}\rangle$, l'intensité moyenne des multiples $I_{h\,k\,l,i}$ observée par les réflexions liés aux éléments de symétrie. [b] Les déviations de la racine des moyenne des carrés (RMSD) sont données à partir de valeurs idéales. [c] $R_{work} = \sum_{h\,k\,l}|F_{obs} - F_{calc}| / \sum_{h\,k\,l} F_{obs}$, F_{obs} et F_{calc} correspondent aux facteurs de structures observés et calculés. R_{free} est similaire à R_{work}, mais calculé sur les 10% des données exclues de l'affinement.

* Statistiques des jeux de données et de l'affinement du modèle construit de **C3**, les valeurs de redondance, R_{sym} et des facteurs d'agitation thermique B ne sont pas disponibles avec le modèle.

Ces structures sont comparées à la structure de la protéine VDR utilisée pour la conception des chimères. Tous les complexes cristallisent dans une maille hexagonale (a = b ≠ c ; α = β = 90°, γ = 120°) avec le même groupe d'espace $P6_522$. Ce premier constat permet d'affirmer que les chimères, quelles que soient leurs modifications (longueur et position de la

portion urée, nature des résidus urée, positions des chaînes latérales des résidus urée), n'affectent pas le groupe d'espace dans lequel les complexes cristallisent.

Nous nous sommes ensuite intéressés à l'étude des effets des diverses modifications des séquences sur l'interaction entre les chimères et VDR. Pour ce faire, nous avons comparé les mesures d'affinité mesurées par thermophorèse à micro-échelle (*microscale thermophoresis* MST), et les mesures de stabilité obtenues par nano fluorimétrie différentielle à balayage (*nano differential scanning fluorimetry* nanoDSF) réalisées au sein de l'équipe du Dr. Natacha Rochel-Guiberteau, aux structures cristallines que nous avons résolues et affinées. La technique de MST est utilisée pour détecter les interactions protéine-protéine en se basant sur un gradient microscopique de température généré par un laser infra-rouge. Tout changement dans la couche d'hydratation, la charge et la forme d'une biomolécule affecte son mouvement thermophorétique et la constante de dissociation est calculée à partir des variations de fluorescence induites par ces mouvements des particules suite au chauffage localisé produit par le laser infra-rouge. La nanoDSF est une méthode qui permet de mesurer à haute résolution la stabilité des protéines en mesurant la fluorescence émise par les acides aminés tryptophanes ou tyrosines. Le suivi de ces changements de fluorescence des résidus acides aminés permet d'évaluer la stabilité de la protéine en présence du ligand. A haute température, les protéines ont tendance à se déplier en exposant ainsi leurs résidus hydrophobes au solvant. Le dépliement de la structure de la protéine est observé en fonction de la température et permet de déterminer la température de fusion (T_m) des protéines. Cette température de fusion caractérise la température au point d'inflexion du changement de structuration repliée-dénaturée de la protéine.

Les modèles des structures cristallines ont été construits par mes soins à partir de la densité électronique. En l'absence de cette dernière, les atomes n'ont pas été construits. Les portions C-terminales des chimères ainsi que celles de la séquence naturelle ne sont pas visibles dans les cartes de densité et n'ont donc pas été construites (**Fig. 45**). L'absence de densité pour cette portion peut signifier une flexibilité de celle-ci et/ou une absence d'interaction avec la surface de la protéine.

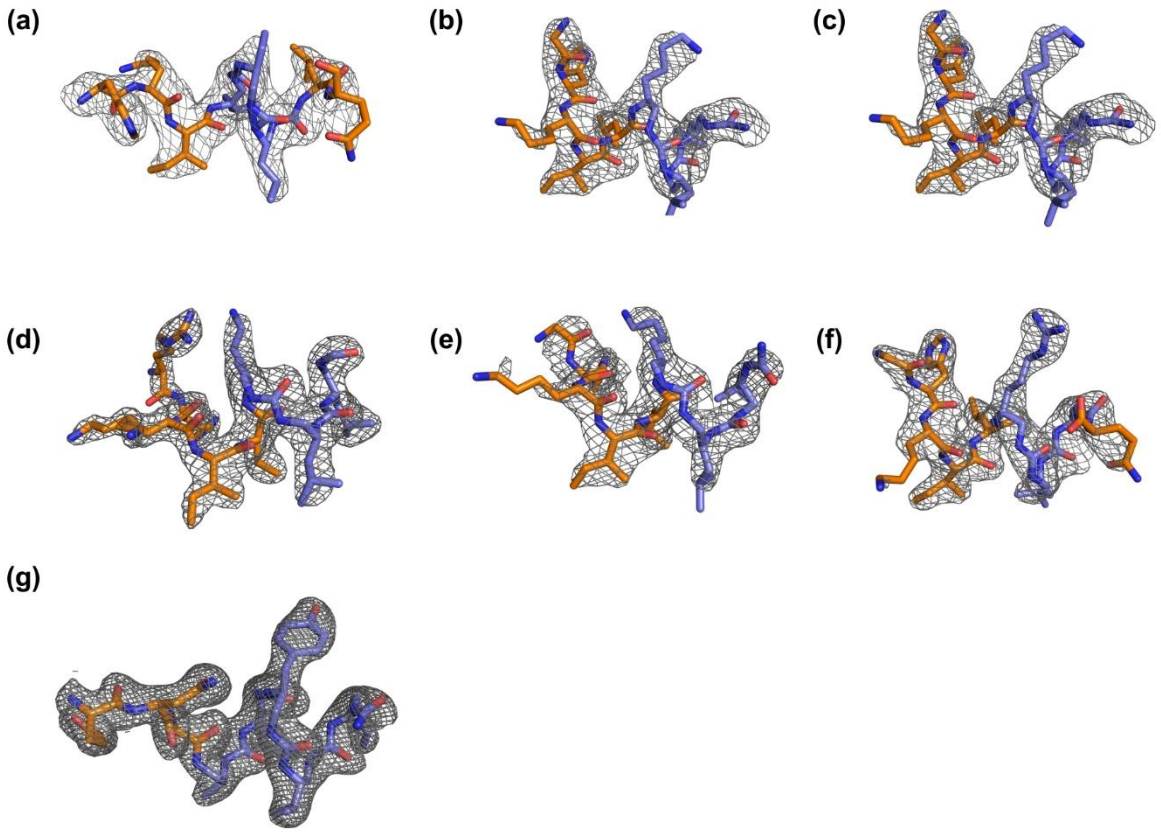


Figure 45 : Cartes de densité électronique $2F_o$-$2F_c$ des foldamères mixtes oligourée/oligoamide, σ = 1,6 Å **(a) C2, (b) C3, (c) C5, (d) C6, (e) C7, (f) C8** et **(g) C11**. Observation de la conformation hélicoïdale des chimères ; les cartes de densité de la protéine n'ont pas été représentées pour visualiser au mieux la topologie en hélice des chimères.

Toutes les chimères se replient en hélice et sont bien ancrées dans la poche hydrophobe de VDR grâce à la présence de liaisons hydrogène entre les résidus aux positions -1, 1 et 3 de la chaîne principale de la chimère et les résidus acides aminés E446 et K274 formant la pince chargée. Le mode de liaison des chimères est donc identique à celui du ligand naturel (**Fig. 46**). D'après les distances mesurées entre les atomes, la pince chargée est conservée dans les structures cristallines pour une partie des chimères (**C2**, **C5**-**C7** et **C11**, **Tableau 10**). La structure du complexe avec la séquence **C8** révèle la présence d'une molécule d'eau entre le carbonyle du résidu urée leucine en position 4 et l'acide aminé K274. La pince chargée peut être conservée grâce à la formation d'un réseau de liaison hydrogène avec cette molécule d'eau. En ce qui concerne la chimère **C3**, les distances interatomiques, mesurées sur le modèle non-affiné à haute résolution, mettent en évidence une distance de plus de 4 Å entre le carbonyle du résidu L_3^U de la chaîne principale et le NH de l'amine de la chaîne latérale de la lysine K274 de VDR. Cette distance est trop importante pour laisser supposer l'existence d'une liaison hydrogène et donc la conservation de la pince chargée. Le modèle utilisé pour la mesure de ces distances est incomplet (les molécules d'eau ne sont pas

construites), à ce jour les mesures ne permettent pas de conclure sur le maintien partiel de la pince chargée entre les acides aminés E446 et K274 avec la chimère **C3**.

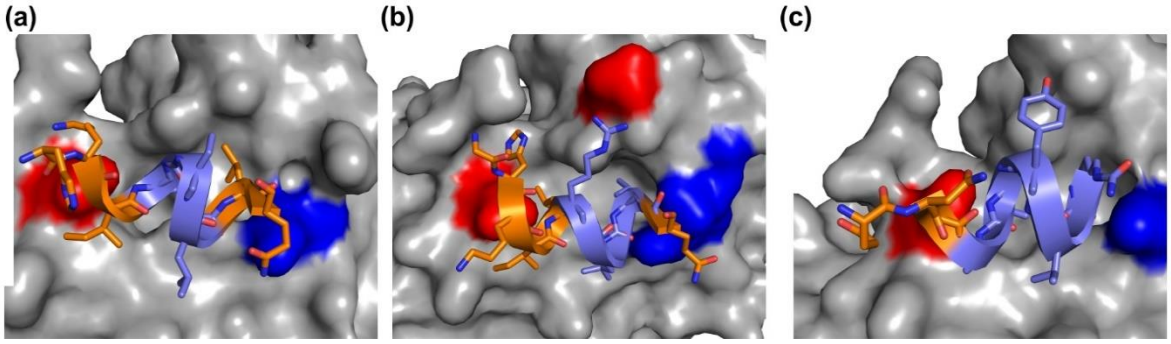


Figure 46 : Schéma avec trois chimères en hélice avec le réseau de liaison hydrogène formant la pince chargée entre la chaîne principale des chimères **C2 (a) C5 (b)** et **C11 (c)** avec les acides aminés E446 (en bleu) et K274 (en rouge) de VDR.

ligands	$d_{CO---NH}$K274 (Å)	$d_{CO--NH1}$E446 (Å)	$d_{CO--NH2}$E446 (Å)
SRC1-2	2,9	2,8	2,8
C2	4,2	3,1	2,6
C3	4,9*	2,9	2,9
C5	2,8	2,5	3,0
C6	3,3	3,0	3,1
C7	2,7	2,4	2,1
C8	nd[a]	2,7	2,9
C11	2,7	2,8	3,0

Tableau 10 : Distances mesurées des liaisons hydrogène entre la chaîne principale du ligand et les résidus K274 et E446 de zVDR formant la pince chargée. [a] La longueur de la liaison hydrogène entre la chaîne principale de **C8** et l'acide aminé K274 n'est pas disponible car une molécule d'eau est placée entre le carbonyle de **C8** et l'amine primaire du résidu K274, la pince chargée est conservée grâce au réseau de liaisons hydrogène présent entre **C8** et l'acide aminé K274. * Mesure de la longueur de la liaison issu du modèle qui ne contient pas de molécules d'eau.

2. Effet du remplacement de la portion LXXL de SRC1-2 par trois urées : chimère **C2**

La première approche de remplacement de quatre résidus d'acides aminés par trois urées concerne le remplacement de la portion LXXL pour laquelle la chimère **C2** a été conçue.

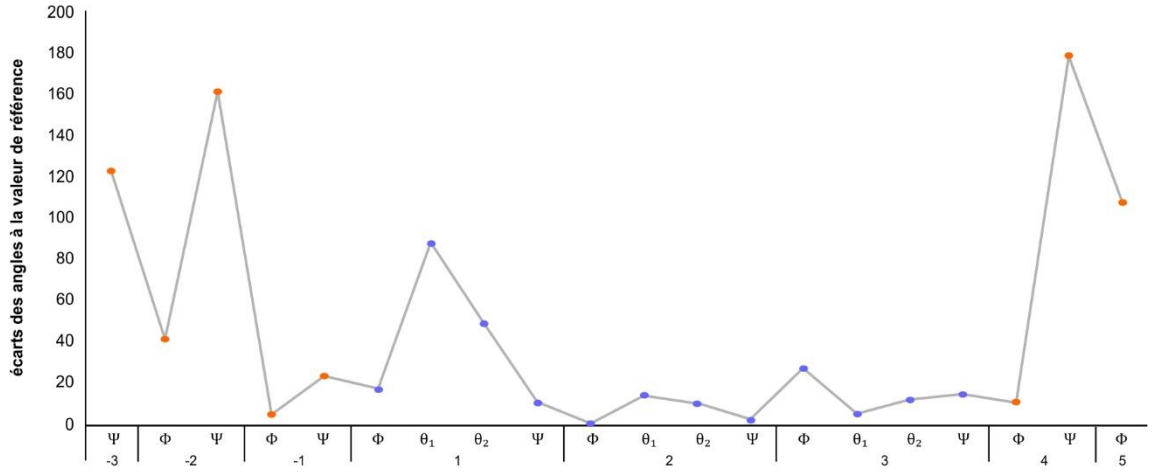
L'affinité de **C2** mesurée par thermophorèse est de 802,48 ± 94,30 µM, ce qui est mille fois supérieur à la valeur du peptide **P1** utilisée comme référence pour lequel le K_d est de 0,90 ± 0,27 µM. La perte de la liaison hydrogène entre la chaîne principale de **C2** et l'acide aminé K274 est un premier élément qui peut expliquer cette perte d'affinité. Les mesures de stabilité

effectuées par nanoDSF avec la protéine zVDR issue du poisson-zèbre mais aussi avec hVDR (issue de l'homme) montrent que le peptide naturel est plus efficace pour stabiliser le complexe protéine-ligand que la chimère **C2**. En effet, les températures de fusion T_m sont de 54,4°C et 57,7°C pour zVDR et hVDR en présence de la séquence **C2**, alors que les valeurs de référence de la protéine en complexe avec **P1** sont de 56,2°C et 60,9°C. Néanmoins, les T_m sont supérieures à celles des protéines seules (51,9°C et 54,6°C), cela suggère que les protéines sont tout de même stabilisées par le composé **C2** (**Tableau 11**).

ligand	séquence	K_d (µM)[a]	zVDR-LBD T_m (°C)[b]	hVDR-LBD T_m (°C)[b]
sans ligand			51,9	54,6
P1[c]	$R_{-4}H_{-3}K_{-2}I_{-1}L_1H_2R_3L_4L_5Q_6E_7G_8S_9P_{10}S_{11}$	0,90 ± 0,27	56,2	60,9
C2	$R_{-4}H_{-3}K_{-2}I_{-1}L_1{}^{U}K_2{}^{U}L_3{}^{U}L_4Q_5E_6G_7S_8$	802,43 ± 94,30	54,4	57,7

Tableau 11 : Tableau récapitulatif des mesures d'affinité et de stabilité pour la chimère **C2** et le peptide **P1**. [a]mesure réalisée par thermophorèse à micro-échelle (MST, T = 25°C, puissance : 80%). [b]mesure réalisée par nanoDSF, valeurs correspondants à la moyenne des températures de fusion à 330 et 350 nm. [c]séquence du peptide naturel utilisée comme référence pour les études d'affinité et stabilité.

La caractérisation de l'hélicité de la chimère a été réalisée en calculant l'écart des angles dièdres par rapport à leurs valeurs de référence pour chaque résidu acide aminé et oligourée. L'écart de l'angle (a) est calculé à partir de la formule suivante : $\Delta a = 180 - \left|\left|a_i - a_{ref}\right| - 180\right|$. Les angles Φ = -58° et Ψ = -47° sont utilisés comme références pour les acides aminés et ceux utilisés pour les résidus urées sont Φ = -100°, Θ_1 = 55°, Θ_2 = 85° et Ψ = -170° (**Graph. 1**). Un faible écart d'angle entre les résidus aux positions -1 et 3 indique que la structure en hélice canonique est partiellement conservée par rapport à celle du peptide. Les angles dièdres du résidu acide aminé isoleucine à la position -1 sont proches de ceux d'un acide aminé d'une hélice α canonique et les angles dièdres des résidus urées aux positions 2 et 3 sont similaires à ceux d'une hélice d'oligourée canonique. Seuls les angles Θ_1 et Θ_2 du résidu urée leucine en position 1 diffèrent des angles de l'hélice canonique d'oligourée. La différence d'angle à cette position ne permet pas de mimer correctement le caractère canonique de l'hélice du motif LXXLL de SRC1-2 et serait un premier indice structural permettant d'expliquer la faible affinité de la chimère **C2** à VDR.



Graphique 1 : Caractérisation de l'hélicité de la chimère **C2**. L'écart des angles dièdres à leurs valeurs de référence est donné pour chaque résidu.

La structure cristalline de la chimère **C2** montre que les chaînes latérales des deux résidus leucine urées semblent être bien projetées dans la poche hydrophobe (**Fig. 47**). La chaîne latérale de la lysine urée n'est pas complète puisque les atomes ne sont pas visibles dans la carte de densité, ce qui ne permet pas de reconstruire la partie distale de celle-ci. Une superposition avec la structure de la séquence naturelle a été réalisée afin d'essayer d'identifier d'autres éléments structuraux responsables de la faible stabilité et affinité de la chimère **C2**.

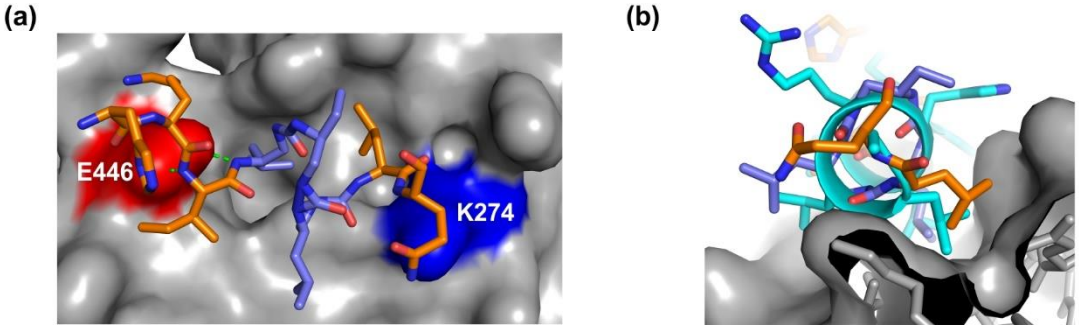


Figure 47 : Structures cristallines de l'interaction entre **C2** (les résidus urées sont en violet et les résidus d'acides aminés en orange) et zVDR (en gris). **(a)** Représentation de l'ancrage de **C2** dans la poche hydrophobe avec la pince chargée formée par les résidus E446 (en rouge) et K274 (en bleu). **(b)** Vue de profil de la superposition de **C2** avec SRC1-2 (en cyan), révélant la projection des chaînes latérales dans la poche hydrophobe.

La superposition des deux séquences montre que la chimère **C2** mime correctement les résidus clés du peptide naturel, en effet le résidu urée L_1^U et le résidu d'acide aminé L_4 ont des valeurs de RMSD de 1,31 Å et 0,80 Å (**Tableau 12**). Les valeurs de RMSD ont été calculées pour chaque résidu pour la superposition de **C2** à SRC1-2 à partir de la formule

suivante : $\mathrm{RMSD} = \sqrt{\frac{\sum d^2}{n}}$ où d est la distance mesurée en angström et n le nombre de mesure par résidu. Les RMSD des acides aminés ont été réalisés sur les distances entre chaînes principales, et ceux des urées ont été faits sur les chaînes latérales des résidus. Les valeurs de RMSD permettent de comparer les deux structures cristallines. Des valeurs proches ou inférieures à 1 Å indiquent une bonne superposition et donc une similarité structurale entre la chimère et le peptide naturel. Le résidu leucine urée L_3^U a un RMSD de 2,88 Å, cette urée n'est pas un bon mime de la leucine naturelle, la chaine latérale de l'urée est orientée vers le solvant et non vers la poche hydrophobe. Cette différence d'orientation de la chaîne latérale est un second élément en faveur de la faible affinité et stabilité de la chimère **C2** par rapport au ligand naturel. Les valeurs de RMSD des acides aminés du composé **C2** indiquent que la chimère commence à bien mimer la portion peptidique de SRC1-2 entre les acides aminés K_{-2} et I_{-1}.

	nature du résidu et position / RMSD (Å)							
chimère	-3	-2	-1	1	2	3	4	5
C2	H 5,37	K 1,51	I 0,73	L^U 1,31	K^U 3,10	L^U 2,88	L 0,80	Q 0,94

Tableau 12 : Tableau récapitulatif des valeurs de RMSD calculées pour la superposition de **C2** à SRC1-2 Les résidus en jaune correspondent aux résidus projetés dans la poche hydrophobe.

3. <u>Effet du remplacement de la portion XXLL de SRC1-2 par trois urées : chimères **C3** à **C8**</u>

La seconde approche de remplacement d'une portion du motif LXXLL concerne le remplacement de la portion XXLL. Cette seconde approche a conduit à la conception de six séquences mixtes oligoamide/oligourée. Les quatre premières séquences **C3**-**C6** sont des séquences parentes où seule la position de la chaîne latérale des résidus urées leucine 3 et/ou 4 a été déplacée en C_α.

La séquence **C7** est similaire à la séquence **C5** mais avec un résidu arginine urée en remplacement du résidu lysine urée. Une modification de la longueur de la chaine latérale de l'urée en position 4 de **C7** a conduit à la séquence **C8**, en effet le résidu leucine urée de **C7** a été remplacé par un résidu alanine urée. Ces diverses modifications ponctuelles des chimères ont permis de réaliser une étude de l'impact de ces changements sur l'affinité et la structure des complexes chimères/VDR.

a. Impact de la modification de la position de la chaîne latérale en Cα

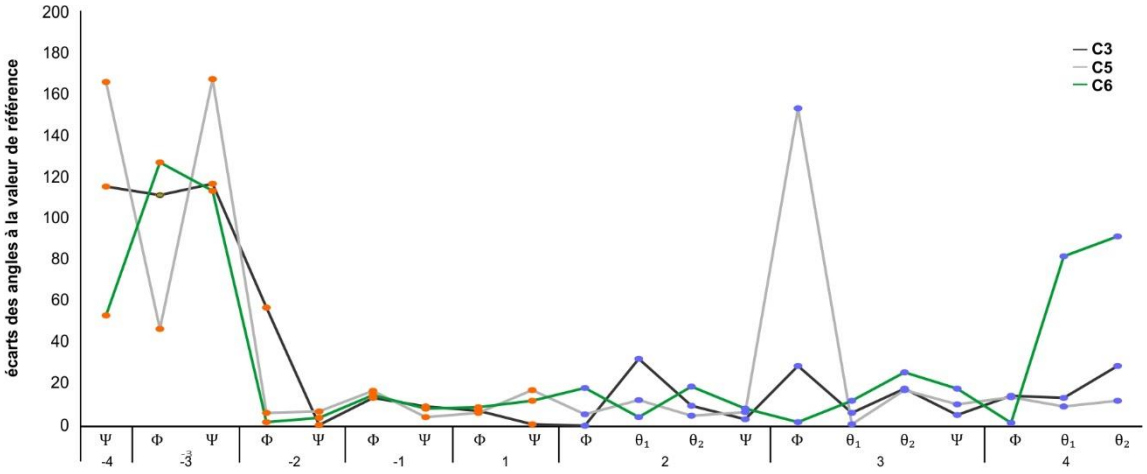
Parmi les quatre séquences mixtes conçues pour étudier l'impact de la modification de la chaîne latérale, seule la séquence **C4** n'a pas donné de structure cristalline. Les mesures d'affinité réalisées (**Tableau 13**) montrent que l'affinité de ces quatre foldamères mixtes est très inférieure à la valeur de référence **P1**. En effet, l'affinité mesurée pour la chimère **C4** est de 21,86 ± 16,2 mM, et celle de la chimère **C6** est de 1124,25 ± 445,39 µM. Ceci est très intéressant car ces deux structures ont toutes les deux la chaîne latérale du résidu leucine urée sur le carbone C_α en position 4 et/ou 3, ce qui suggère que cette modification n'est pas favorable pour l'interaction avec la protéine. L'affinité de **C3** est de 195,82 ± 43,93 µM. Cette affinité est bien meilleure que celle de **C4** et **C6**, bien qu'elle soit encore très inférieure à celle du ligand peptidique **P1**. Il est possible que le déplacement de la chaîne latérale en C_α pour le résidu urée L_4^U diminue l'affinité. Les mesures de stabilité de **C4** avec zVDR et hVDR confirment l'absence de liaison de la chimère à la protéine, en effet ses valeurs de T_m sont proches (52,3°C avec zVDR et 54,8°C avec hVDR) de celles de la protéine seule. Les températures de fusion avec **C3** et **C6** sont également proches de celle des protéines seules (**Tableau 13**), ce qui implique que les chimères **C3** et **C6** aient un très faible effet de stabilisation des complexes qu'elles forment avec les protéines zVDR et hVDR. Ces constats sont en accord avec les observations faites avec les mesures d'affinité pour **C6**, mais pas pour **C3** où l'affinité est améliorée malgré une faible stabilisation du complexe avec la protéine VDR.

ligand	séquence	K_d (µM)[a]	zVDR-LBD T_m (°C)[b]	hVDR-LBD T_m (°C)[b]
sans ligand			51,9	54,6
P1[c]	$R_{-4}H_{-3}K_{-2}I_{-1}L_1H_2R_3L_4L_5Q_6E_7G_8S_9P_{10}S_{11}$	0,90 ± 0,27	56,2	60,9
C3	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2^UL_3^UL_4^UQ_5E_6G_7S_8$	195,82 ± 43,93	52,8	55,2
C4	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2^UL_3^UL_{4\alpha}^UQ_5E_6G_7S_8$	21863,45 ± 162620,25	52,3	54,8
C5	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2^UL_{3\alpha}^UL_4^UQ_5E_6G_7S_8$	215,04 ± 70,36	55,2	59,0
C6	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2^UL_{3\alpha}^UL_{4\alpha}^UQ_5E_6G_7S_8$	1124,25 ± 445,39	52,4	54,8

Tableau 13 : Tableau récapitulatif des mesures d'affinité et de stabilité pour les chimères **C3**-**C6** et le peptide **P1**. [a]mesures réalisées par thermophorèse à micro-échelle (MST, T = 25°C, puissance : 80%). [b]mesures réalisées par nanoDSF, valeurs correspondantes à la moyenne des températures de fusion à 330 et 350 nm. [c]séquence du peptide naturel utilisée comme référence pour les études d'affinité et stabilité.

La caractérisation de l'hélicité des chimères indique que la structure en hélice α canonique est conservée pour les acides aminés entre les positions -2 et 1 (**Graph. 2**). L'hélice canonique d'oligourée est conservée pour **C3**, mais pas pour **C5** où la présence du résidu urée

inversé à la position 3 rompt le caractère canonique de l'hélice avec une différence d'angle Φ de 160°. L'écart d'angle de ce même résidu dans la séquence **C6** est de 2°. La structure en hélice canonique de **C6** est rompue à partir du résidu en position 4 où les écarts d'angle Θ_1 et Θ_2 du résidu urée inversé sont supérieurs à ceux des résidus urées « normaux » à cette position. Ces différences angulaires ne permettent pas de mimer correctement l'hélice canonique du motif LXXLL de SRC1-2 et sont donc les premiers indices structuraux permettant d'expliquer les différences d'affinité des chimères **C3**, **C5** et **C6**.



Graphique 2 : Caractérisation de l'hélicité des chimères **C3-C6**. L'écart des angles dièdres à leurs valeurs de références est donné pour chaque résidu.

Les structures cristallines de **C3** et **C6** sont similaires et les portions peptidiques en N-terminale des chimères sont orientées dans la même direction que le peptide naturel, et sont bien mimées avec des RMSD proches de 1 Å (**Tableau 14**). Les chaînes latérales des résidus urée K_2^U sont bien projetées et miment correctement l'histidine de la séquence naturelle (RMSD de 1,31 Å pour **C3** et 0,97 Å pour **C6**). La valeur de RMSD pour le résidu L_4^U de **C3** est un peu haute avec 1,57 Å et celle-ci est encore plus élevée pour **C6** avec 2,25 Å, ce qui confirme que l'orientation des chaînes latérales dans la poche hydrophobe n'est pas parfaite. L'interaction clé au niveau de la position 3 de la chimère n'est plus conservée avec **C3** : la chaîne latérale du résidu L_3^U est orientée vers le solvant, le RMSD de 6,51 Å pour ce résidu confirme la perte de l'interaction clé (**Fig. 48b**). Cette perte d'interaction est un élément confirmant la faible stabilité du complexe VDR/**C3**, mais ne permet pas d'expliquer la forte affinité de cette chimère avec la protéine. Le modèle actuel de la chimère ne permet pas d'observer l'ensemble des informations structurales expliquant la meilleure affinité de **C3**. Concernant **C6**, l'orientation du résidu urée L_3^U vers la poche hydrophobe est conservée. La structure cristalline n'apporte pas d'éléments de réponse suffisants pour expliquer la faible affinité et stabilité de cette chimère avec VDR.

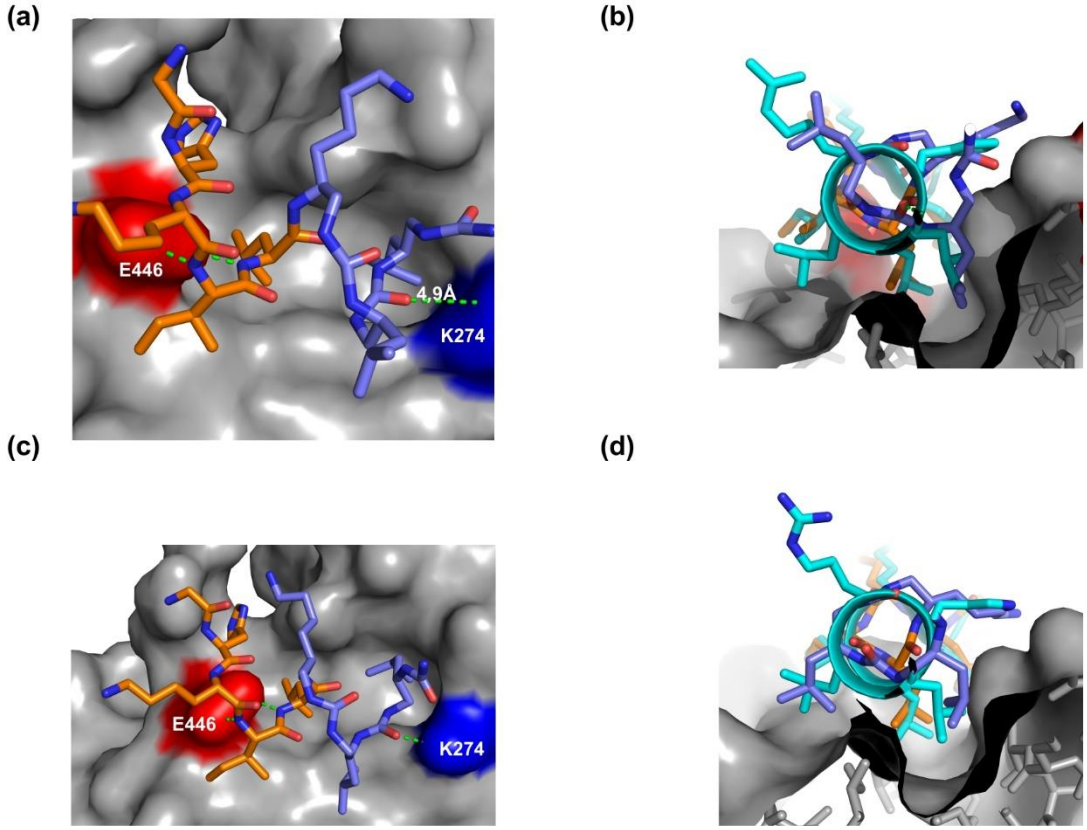


Figure 48 : Structures cristallines de l'interaction entre zVDR en gris) et les chimères **C3** et **C6** (les résidus urées sont en violet et les acides aminés en orange). **(a)** Représentation de l'ancrage de **C3** dans la poche hydrophobe avec la pince chargée formée par E446 (en rouge) et K274 (en bleu). **(b)** Vue de profil de la superposition de **C3** avec SRC1-2 (en cyan), révélant la projection des chaînes latérales dans la poche hydrophobe. **(c)** Représentation de l'ancrage de **C6** dans la poche hydrophobe avec la pince chargée formée. **(b)** Vue de profil de la superposition de **C6** avec SRC1-2, indiquant la projection des chaînes latérales dans la poche hydrophobe.

	nature du résidu et position / RMSD (Å)							
chimères	-4	-3	-2	-1	1	2	3	4
C3	R 1,29	H 0,97	K 1,10	I 0,88	L 1,43	K^U 1,31	L^U 6,51	L^U 1,57
C5	R 2,30	H 1,31	K 1,22	I 0,88	L 1,22	K^U 0,94	L_α^U 1,06	L^U 1,71
C6	R 0,80	H 0,80	K 0,77	I 0,50	L 1,10	K^U 0,97	L_α^U 1,66	L_α^U 2,04

Tableau 14 : Tableau récapitulatif des valeurs de RMSD calculées pour la superposition de **C3**-**C6** à SRC1-2. Les résidus en jaune correspondent aux résidus projetés dans la poche hydrophobe.

La quatrième séquence étudiée est la chimère **C5**. Cette séquence, dont seule la position de la chaîne latérale du résidu urée L_3^U a été déplacée sur le carbone C_α est aussi affine que son analogue **C3** avec un K_d de 215,04 ± 70,36 µM. Cependant la stabilité du complexe VDR/**C5** est améliorée d'après les températures de fusion 55,2°C et 59,0°C (zVDR

et hVDR). Parmi toutes les séquences décrites jusqu'à présent, le composé **C5** a les valeurs de T_m les plus proches de la séquence naturelle **P1**, elle est donc la chimère stabilisant le mieux VDR. Une simple modification de la position de la chaîne latérale du résidu L_3^U en C_α permet de gagner considérablement en stabilité mais n'a pas d'effet sur l'affinité. La structure cristalline de la séquence **C5** montre une liaison hydrogène entre la chaîne latérale du résidu urée K_2^U de **C5** et l'acide aminé E285 de zVDR (**Fig. 49a**). Les atomes de la chaîne latérale du résidu urée lysine à la position 2 sont bien visibles dans la carte de densité et ont été construits. Cette liaison hydrogène permet d'ajouter un point d'ancrage supplémentaire de la chimère à la surface de la protéine. Un troisième point d'ancrage entre **C5** et zVDR a lieu via la formation de deux liaisons hydrogène entre le carbonyle du résidu urée en C-terminale L_4^U et les NH de la guanidine de l'acide aminé R280 de VDR. Les chaînes latérales des leucines urées L_3^U et L_4^U sont mieux projetées dans la poche hydrophobe que **C6** et **C3**. Les valeurs de RMSD sont de 1,06 Å et 1,71 Å (**Tableau 14**), ce qui permet de confirmer le bon mimétisme des chaînes latérales des résidus de la séquence naturelle. Les valeurs des RMSD des acides aminés N-terminaux de **C5** sont plus hautes que celles de **C3** pour les mêmes résidus. Ce bon mimétisme de projection des chaînes latérales, ainsi que la présence des points d'ancrages supplémentaires entre les résidus urées K_2^U et L_4^U et les acides aminés E285 et R280 apportent de la stabilité au complexe formé par zVDR /**C5** et contribue à conserver l'affinité. La différence d'affinité entre les chimères **C3** et **C5** pourrait être liée à cette différence de mimétisme de la portion N-terminale des séquences hybrides au peptide naturel.

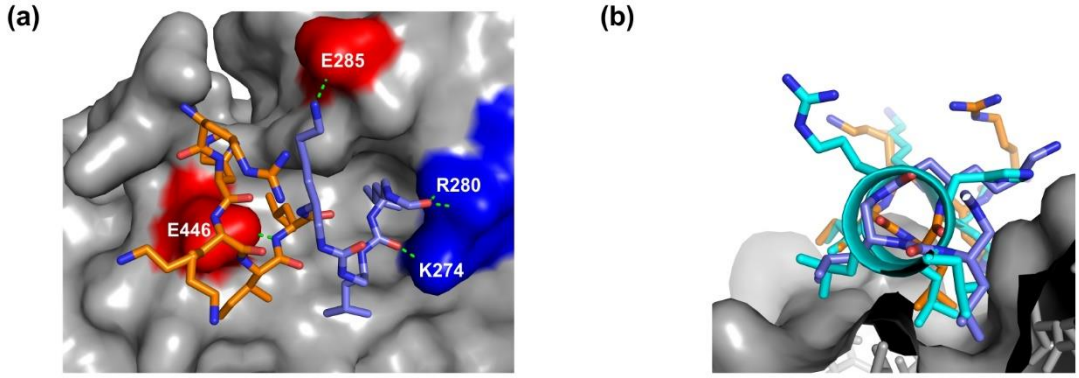


Figure 49 : Structure cristalline de l'interaction entre **C5** (les résidus urées sont en violet et les acides aminés en orange) et zVDR (en gris). **(a)** Représentation de l'ancrage de **C5** dans la poche hydrophobe avec la pince chargée formée par les résidus E446 (en rouge) et K274 (en bleu) ; formation des liaisons hydrogène entre le résidu urée K_2^U et l'acide aminé E285 (en rouge) et le résidu urée L_4^U et l'acide aminé R280 (en bleu). **(b)** Vue de profil de la superposition de **C5** avec SRC1-2 (en cyan), révélant la projection des chaînes latérales dans la poche hydrophobe.

Cette étude de l'impact de la position de la chaîne latérale en C_α permet de souligner son importance sur l'affinité et la stabilité des chimères avec le récepteur nucléaire VDR.

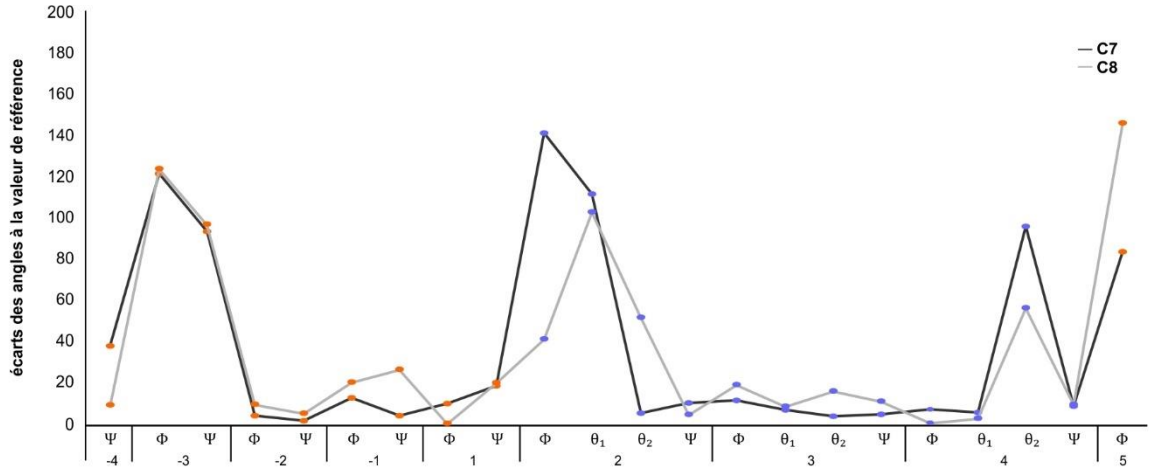
b. Impact de la modification sur deux résidus urée aux positions 2 et 4

Les séquences **C7** et **C5** diffèrent par la nature du résidu urée en position 2 de la chimère. Le remplacement du résidu lysine urée par une arginine urée en position 2 de **C7** semble avoir un impact sur l'affinité de la chimère. Les mesures d'affinité montrent que **C7** a une affinité de 64,25 ± 13,23 µM (**Tableau 15**). Cette amélioration d'affinité suggère que modifier une résidu lysine urée par un dérivé d'arginine urée à la position 2 est plus favorable pour l'interaction avec la protéine VDR. Les températures de fusion obtenues pour **C7** (T_m = 55,9°C et 59,2 °C pour zVDR et hVDR) vont également dans ce sens, puisqu'elles sont très proches de celles obtenues avec **P1** : **C7** est ainsi la séquence qui stabilise le mieux le complexe formé avec VDR.

ligand	séquence	K_d (µM)[a]	zVDR-LBD T_m (°C)[b]	hVDR-LBD T_m (°C)[b]
sans ligand			51,9	54,6
P1[c]	$R_{-4}H_{-3}K_{-2}I_{-1}L_1H_2R_3L_4L_5Q_6E_7G_8S_9P_{10}S_{11}$	0,90 ± 0,27	56,2	60,9
C5	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2{}^{U}L_{3\alpha}{}^{U}L_4{}^{U}Q_5E_6G_7S_8$	215,04 ± 70,36	55,2	59,0
C7	$R_{-4}H_{-3}K_{-2}I_{-1}L_1R_2{}^{U}L_{3\alpha}{}^{U}L_4{}^{U}Q_5E_6G_7S_8$	64,25 ± 13,23	55,9	59,2
C8	$R_{-4}H_{-3}K_{-2}I_{-1}L_1R_2{}^{U}L_{3\alpha}{}^{U}A_4{}^{U}Q_5E_6G_7S_8$	214,33 ± 38,84	53,0	56,2

Tableau 15 : Tableau récapitulatif des résultats des mesures d'affinité et de stabilité pour les chimères **C7**, **C8** et le peptide **P1**. [a]mesure réalisée par thermophorèse à micro-échelle (MST, T = 25°C, puissance : 80%). [b]mesure réalisée par nanoDSF, valeurs correspondantes à la moyenne des températures de fusion à 330 et 350 nm. [c]séquence du peptide naturel utilisée comme référence pour les études d'affinité et stabilité.

La caractérisation de l'hélicité réalisée pour les chimères **C7**-**C8** indique que la structure en hélice α canonique est conservée pour les acides aminés aux positions -2, -1 et 1 (**Graph. 3**). Cependant, seuls les résidus urées en position 3 ont des angles dièdres proches de ceux de l'hélice canonique d'oligourée. Les écarts d'angles dièdres pour les résidus urées aux positions 2 et 4 ne permettent pas d'avoir une hélice canonique d'oligourée dans les séquences **C7** et **C8**.



Graphique 3 : Caractérisation de l'hélicité des chimères **C7** et **C8**. L'écart des angles dièdres à leurs valeurs de référence est donné pour chaque résidu.

La structure cristalline de la chimère **C7** indique la présence de liaisons hydrogène entre les résidus urées aux positions 2 et 4 de la chimère avec les acides aminés E285 et R280 de la protéine (**Fig. 50a**). La conservation des liaisons hydrogène dans la structure de **C7** a été observée lors de l'étude du composé **C5** et ne permettent pas de donner une explication structurale à la meilleure affinité et stabilité de la chimère **C7** par rapport aux autres séquences décrites jusqu'à présent.

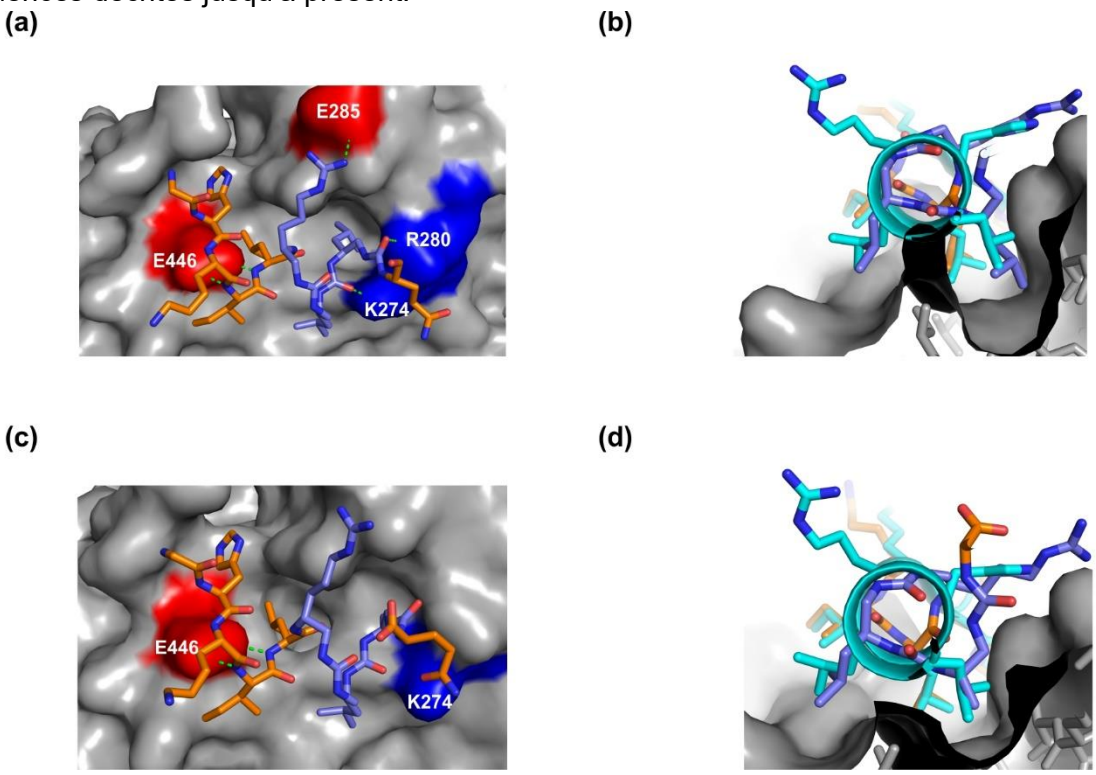


Figure 50 : Structures cristallines de l'interaction entre **C7** et **C8** (les résidus urées sont en violet et les acides aminés en orange) et zVDR (en gris). **(a)** Représentation de l'ancrage de **C7** dans la poche hydrophobe avec la pince chargée formée par E446 (en rouge) et K274 (en bleu) et des liaisons hydrogène entre les résidus R_2^U – E285 et L_4^U et R280. **(b)** Vue de profil de la superposition de **C7** avec

SRC1-2 (en cyan), révélant la projection des chaînes latérales dans la poche hydrophobe. **(c)** Représentation de l'ancrage de **C8** dans la poche hydrophobe avec la pince chargée formée. **(b)** Vue de profil de la superposition de **C8** avec SRC1-2, indiquant la projection des chaînes latérales dans la poche hydrophobe.

La séquence **C8** a une chaîne latérale plus courte pour le résidu urée A_4^U que dans le composé **C7** (L_4^U). Cette modification de longueur de chaîne latérale a entrainé une diminution de l'affinité et de la stabilité du complexe. En effet, la chimère **C8** a une affinité de 214,33 ± 38,84 µM, ce qui est plus faible que celle de la séquence **C7** mais proche de celle du composé **C5**. Les valeurs de T_m ont également diminué à 53,0°C et 56,2°C pour zVDR et hVDR. La structure cristalline de la séquence **C8** révèle que le résidu urée alanine A_4^U mime mieux le résidu naturel que le résidu leucine urée L_4^U de la chimère **C7** : le RMSD de **C8** pour ce résidu est de 1,18 Å contre 1,83 Å pour **C7** (**Tableau 16**). La modification de la chaîne latérale en position 4 ne semble donc pas défavorable pour l'interaction. Les liaisons hydrogène supplémentaires observées pour **C7** ne sont pas présentes pour le complexe avec **C8**. L'absence de ces points d'ancrages avec les acides aminés E285 et R280 pourrait être responsable de la perte d'affinité et de stabilité du complexe avec **C8**.

	nature du résidu et position / RMSD (Å)								
chimères	-4	-3	-2	-1	1	2	3	4	5
C7	R 1,55	H 1,21	K 1,37	I 1,05	L 1,05	R^U 1,07	L_α^U 1,94	L^U 1,83	Q 2,50
C8	R 1,98	H 1,15	K 1,32	I 0,87	L 1,22	R^U 0,93	L_α^U 1,90	A^U 1,18	Q 3,70

Tableau 16 : Tableau récapitulatif des valeurs de RMSD calculées pour la superposition de **C7** et **C8** à SRC-1-2. Les résidus en jaune correspondent aux résidus projetés dans la poche hydrophobe mimant les leucines en i, i+3 et i+4 du peptide naturel.

4. <u>Effet du remplacement de l'intégralité du motif d'interaction LXXLL de SRC1-2 par cinq urées : chimères **C9** à **C11**</u>

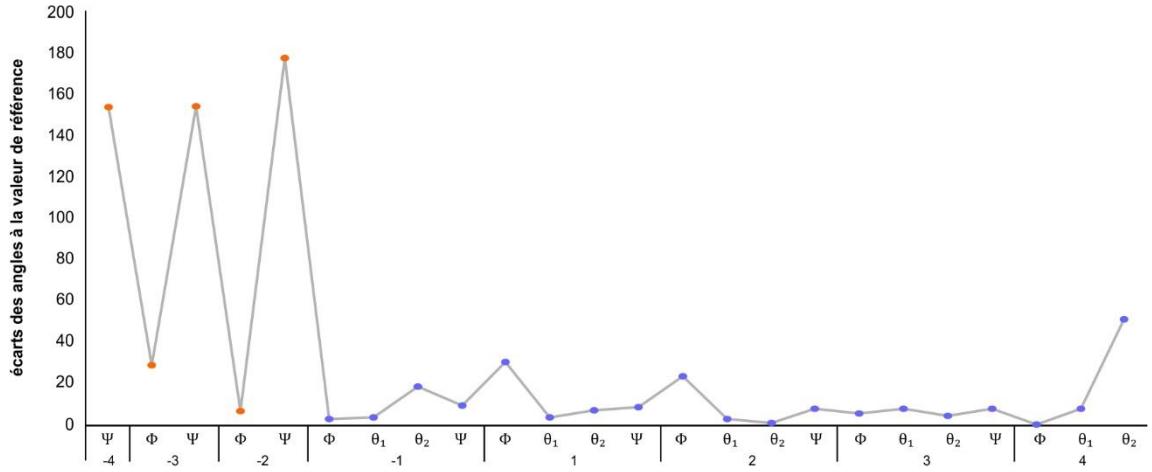
Le remplacement de l'intégralité du motif d'interaction par cinq urées a conduit à la conception de trois chimères **C9**, **C10** et **C11**. Cette dernière est la seule à avoir cristallisé et dont la structure cristallographique a été résolue.

ligand	séquence	K_d (µM)[a]	zVDR-LBD T_m (°C)[b]	hVDR-LBD T_m (°C)[b]
sans ligand			51,9	54,6
P1[c]	$R_{-4}H_{-3}K_{-2}I_{-1}L_1H_2R_3L_4L_5Q_6E_7G_8S_9P_{10}S_{11}$	0,90 ± 0,27	56,2	60,9
C9	$R_{-4}H_{-3}K_{-2}A_{-1\alpha}{}^{U}L_1{}^{U}K_2{}^{U}V_3{}^{U}Nle_4{}^{U}E_5G_6S_7$	69,11 ± 6,11	54,1	57,3
C10	$K_{-4}H_{-3}K_{-2}A_{-1\alpha}{}^{U}L_1{}^{U}K_2{}^{U}V_3{}^{U}Nle_4{}^{U}D_5S_6S_7$	38,91 ± 7,21	53,6	57,3
C11	$E_{-4}N_{-3}A_{-2}A_{-1\alpha}{}^{U}L_1{}^{U}Y_2{}^{U}V_3{}^{U}Nle_4{}^{U}K_5D_6D_7$	1,58 ± 0,43	54,9	58,6

Tableau 17 : Tableau récapitulatif des résultats des mesures d'affinité et de stabilité pour les chimères **C9**-**C11** et le peptide **P1**. [a]mesure réalisée par thermophorèse à micro-échelle (MST, T = 25°C, puissance : 80%). [b]mesure réalisée par nanoDSF, valeurs correspondants à la moyenne des températures de fusion à 330 et 350 nm. [c]séquence du peptide naturel utilisée comme référence pour les études d'affinité et stabilité.

Les mesures d'affinité réalisées avec ces trois séquences révèlent des affinités qui sont similaires voire meilleures que pour les chimères à trois urées (**Tableau 11, 13, 15 et 17**). La chimère ayant la meilleure affinité se trouve parmi elles, il s'agit de la séquence **C11** qui a une affinité de 1,58 ± 0,43 µM. Cette valeur est très proche de celle du peptide **P1**, bien que **C11** soit le mime du peptide SRC2-3. Mimer l'intégralité du motif hélicoïdal LXXLL par une séquence d'oligourée semble donc être la meilleure approche pour concevoir des chimères à visée inhibitrice de l'interaction VDR/SRC1-2. Les mesures de stabilité réalisées pour les trois chimères montrent que **C11** a également les meilleures températures de fusion T_m pour zVDR et hVDR avec 54,9°C et 58,6°C, qui sont très proches des valeurs de **P1**. Cela sous-entend que le complexe qu'elle forme avec VDR est aussi stable que celui formé avec le peptide naturel. Ce résultat est en accord avec les constats faits pour l'affinité : le remplacement du motif naturel par cinq urées est plus favorable pour avoir une bonne interaction avec la protéine VDR.

La caractérisation de l'hélicité de la chimère **C11** a été réalisée et montre que les écarts d'angles dièdres des résidus urée sont faibles et que l'hélice formée par ces résidus est bien une hélice canonique d'oligourée (**Graph. 4**). La chimère **C11** est la seule chimère qui conserve le caractère canonique de la portion en hélice du peptide naturel, ce qui lui permet d'être un bon mime de SRC1-2 et donc d'expliquer la meilleure affinité de cette séquence par rapport à l'ensemble des chimères de la série.



Graphique 4 : Caractérisation de l'hélicité de la chimère **C11**. L'écart des angles dièdres à leurs valeurs de références est donné pour chaque résidu.

	nature du résidu et position / RMSD (Å)							
chimère	-4	-3	-2	-1	1	2	3	4
C11	E 10,08	N 8,03	A 2,83	A_α^U 1,82	L^U 0,79	Y^U 2,26	V^U 2,36	Nle^U 1,22

Tableau 18 : Tableau récapitulatif des valeurs de RMSD calculées pour la superposition de **C11** à SRC1-2. Les résidus en jaune correspondent aux résidus projetés dans la poche hydrophobe mimant les leucines en i, i+3 et i+4 du peptide naturel.

La structure cristalline de **C11** montre un bon mimétisme de la structure du peptide naturel **P1** grâce aux résidus urées L_1^U, et Nle_4^U dont les valeurs de RMSD sont proches de 1 Å (**Tableau 18**). La portion N-terminale est moins bien mimée et s'oriente vers le solvant (**Fig. 51**). Cette observation est confirmée par la valeur des RMSD (10,08 Å et 8,03 Å) des acides aminés aux positions -4 et -3.

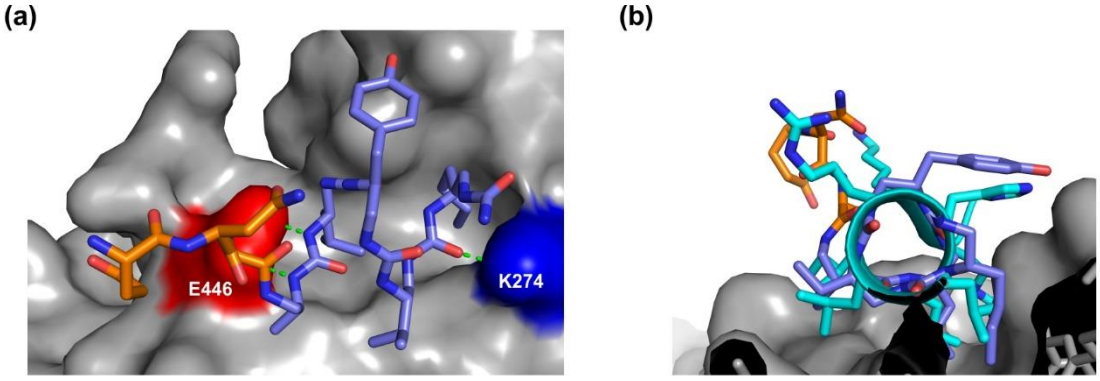


Figure 51 : Structure cristalline de l'interaction entre **C11** (les résidus urées sont en violet et les acides aminés en orange) et zVDR (en gris). **(a)** Représentation de l'ancrage de **C11** dans la poche hydrophobe avec la pince chargée formée par les résidus E446 (en rouge) et K274 (en bleu). **(b)** Vue de profil de la superposition de **C11** avec SRC1-2 (en cyan), révélant la projection des chaînes latérales dans la poche hydrophobe.

5. Conclusion de l'étude de relation structure-affinité sur les chimères **C2** - **C11**

L'ensemble des structures cristallines obtenues confirme les hypothèses réalisées lors de la conception des chimères à partir de la superposition de l'oligourée **C1** à la structure de SRC1-2 en complexe avec VDR. Le remplacement de sept résidus d'acides aminé contenant le motif LXXLL par cinq urées est la meilleure approche pour concevoir des foldamères mixtes oligoamide/oligourée capables d'interagir avec VDR. L'étude de relation structure-affinité a permis de mettre en évidence une chimère tête de série : la chimère **C11**, possédant une bonne affinité et stabilité**.**

Le remplacement d'une portion du motif d'interaction par trois urées a permis de mettre en évidence l'importance des modifications ponctuelles de la séquence sur sa stabilité et affinité avec la protéine. Un résidu arginine urée en position 2 a été identifié comme étant le meilleur résidu à cet endroit pour avoir une interaction stable et favorable. Cette interaction est renforcée par la présence de liaisons hydrogène supplémentaires entre la chimère et les résidus E285 et R280 de VDR, liaisons hydrogène non présentes dans le complexe naturel.

III. Génération d'une nouvelle série de chimères mixtes

1. Optimisation de la chimère tête de série **C11**

chimères	-4	-3	-2	-1	1	2	3	4	5	6	7
C11	E	N	A	A_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}	K	D	D
C12				A_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}			
C13				A^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}			
C14	E	N	A	A_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}			
C15				A_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}	K	D	D
C16				L_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}	K	D	D
C17				Nle_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}	K	D	D
C18				Cha_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}	K	D	D
C19			A^{U}	A_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}	K	D	D
C20			Cin	A_{α}^{U}	L^{U}	Y^{U}	V^{U}	Nle^{U}	K	D	D
C21	E	N	A	A_{α}^{U}	L^{U}	R^{U}	V^{U}	Nle^{U}	K	D	D

Tableau 19 : Séquences des chimères **C12-C21** conçues par l'optimisation de **C11**. Les résidus en violet correspondent aux résidus urée.

L'étude SAR a mis en évidence **C11** comme chimère « tête de série ». Cette séquence a servi de point de départ pour la génération d'une nouvelle série de composés. Nous nous sommes intéressés à l'étude de trois paramètres : l'impact du motif oligourée, celui de la

portion peptidique C-terminale, et l'effet de la portion N-terminale sur l'affinité. L'ensemble des séquences conçues est répertorié dans le **Tableau 19**.

a. Suppression des portions peptidiques

La première approche d'optimisation de **C11** a été d'isoler la portion oligourée pour étudier l'impact des résidus d'acides aminés aux extrémités N- et C-terminales de la séquence sur l'affinité et la stabilité. Cela a conduit à la séquence **C12**. Puis, nous avons voulu étudier l'effet de la position de la chaîne latérale en C_α du résidu urée alanine A_{-1}^{U} de **C12** sur l'interaction. Une séquence **C13** analogue à **C12** a été conçue sans la chaîne latérale sur l'atome C_α de l'alanine urée en position -1.

La portion C-terminale n'est pas visible dans les cartes de densité obtenues par diffraction des rayons X car ces résidus d'acide aminé C-terminaux peuvent être flexibles et/ou ne pas interagir avec la surface de la protéine. La seconde approche d'optimisation a donc été de supprimer la portion peptidique en position C-terminale de **C11,** ainsi une seule séquence a été conçue : **C14**.

La troisième approche d'optimisation concerne la portion N-terminale de **C11**. Dans un premier temps, cette portion a été supprimée donnant la séquence **C15**. Puis une modification de la chaîne latérale en position -1 a conduit à la conception de trois séquences : **C16**, **C17** et **C18**. L'objectif de ces modifications a été de voir l'impact de la longueur de la chaîne latérale du résidu sur l'interaction avec VDR.

Une autre stratégie considérée a été de tenter de prolonger l'hélice jusqu'au résidu -2 par le remplacement de l'acide aminé A_{-2} par une urée (**C19**) ou un groupement cinnamoyle (**C20**). L'objectif est de pouvoir accéder à de nouvelles poches hydrophobes (**Fig. 52**).

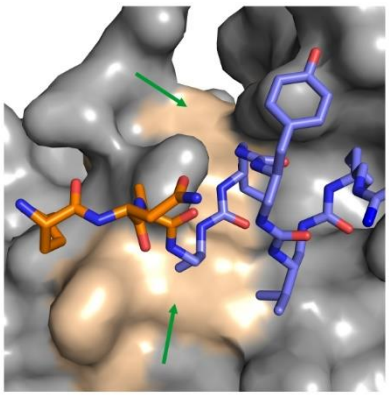


Figure 52 : Représentation des zones hydrophobes ciblées (en beige).

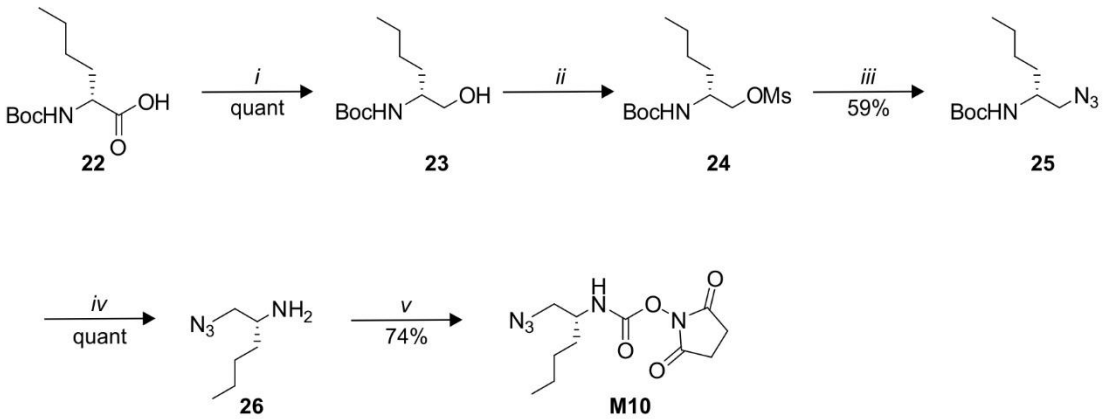
b. Modification du résidu urée en position 2

Les séquences **C5** et **C7** montrent la présence d'une liaison hydrogène supplémentaire entre les résidus urée K_2^U et aminé E285 de VDR. Les résultats de l'étude SAR sur la série des chimères à trois urées (II-3) a montré que l'arginine urée était le meilleur résidu pour obtenir une bonne interaction en position 2. En prenant compte de ce résultat, il a été décidé de concevoir la séquence chimère **C21** possédant un résidu arginine urée en position 2 au lieu du dérivé urée de la tyrosine. Cette modification a pour but de favoriser la liaison hydrogène entre les résidus R_2^U et E285 de VDR en diminuant la distance entre la chaîne latérale de la chimère et le récepteur pour créer deux points d'ancrage supplémentaires au ligand.

2. Synthèse des séquences optimisées **C12-C21**

a. Préparation de monomères activés

Les séquences des chimères conçues en III-1 nécessitent des monomères urées dont la chaîne latérale est sur le carbone C_α de l'urée. La synthèse du monomère urée norleucine inversé (**M10**) n'a pas été rapportée dans la littérature. La même stratégie de synthèse du monomère **M9** à partir de l'acide aminé commercial Boc-D-Nle-OH **22** a été considérée et **M10** a été obtenu avec un rendement global de 44% (**Schéma 7**).

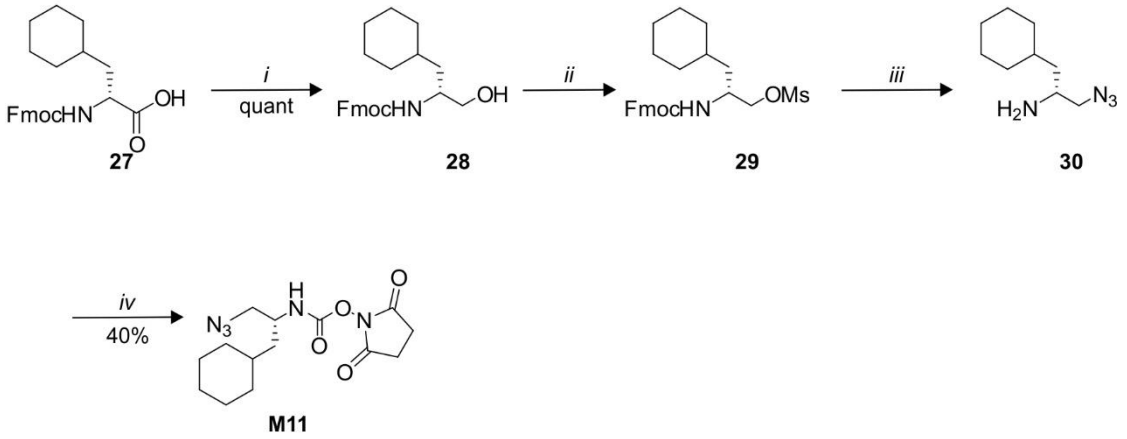


i : 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii* : MsCl, Et_3N, DCM, 0°C, Ar_{atm} ; *iii* : NaN_3, DMF, 70°C ; iv :TFA, TA ; *v* : DSC, DIPEA, DCM, TA

Schéma 7 : Schéma de synthèse du monomère azoture/carbamate de succinimide **M10**

La synthèse du monomère urée cyclohexylalanine inversé (**M11**) n'a également pas été rapportée dans la littérature, et nous nous sommes inspirés de la synthèse de **M9** à partir de la Fmoc-D-Cha-OH **27** (**Schéma 8**). Lors de la substitution nucléophile 2 avec l'azoture de sodium sur le mésylate **29**, l'azido amine **30** a été obtenu en raison d'une réduction du

groupement Fmoc effectuée par l'azoture de sodium. Après activation par la DSC et purification par précipitation dans l'éther froid, le monomère **M11** a été obtenu avec un rendement global de 40%.



i : 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii* : MsCl, Et_3N, DCM, 0°C, Ar_{atm} ; *iii* : NaN_3, DMF, 70°C ; iv : DSC, DIPEA, DCM, TA

Schéma 8 : Schéma de synthèse du monomère azoture/carbamate de succinimide **M11**

*b. Synthèse sur support solide des chimères **C12**-**C21***

Les synthèses des chimères **C12**-**C21** ont été réalisées sur de la résine polystyrène Rink Amide MBHA (loading de 0.52 mmol/g) à une échelle de 50 µmol.

i. Synthèse des chimères **C12** et **C13**

Comme pour les autres chimères, la synthèse des chimères **C12** et **C13** a été réalisée suivant un couplage et une déprotection séquentiels des monomères azotures activés sous forme de carbamate de succinimide.

La synthèse des chimères **C12** et **C13** n'a pas posé de difficulté particulière. Le chromatogramme RP-HPLC de la chimère **C12** (**Fig. 53a**) avant purification par RP-HPLC semi-préparative indique la présence d'un produit largement majoritaire avec une pureté de 78% à 4,86 min. Le chromatogramme RP-HPLC de **C13** indique également la présence d'un produit majoritaire avec une pureté de 80% à 4,72 min. Après purification par HPLC préparative, les chimères **C12** et **C13** sont obtenues avec un rendement de 13% et 79% respectivement.

chimères	séquences	pureté brut réactionnel	ESI-MS $[M+H]^+$	HPLC t_R (min)[a]	rendement[b]
C12	$A_\alpha{}^{U}L^{U}Y^{U}V^{U}Nle^{U}$	78%	722,33	4,86	13%
C13	$A^{U}L^{U}Y^{U}V^{U}Nle^{U}$	80%	722,63	4,72	79%

Tableau 20 : Séquences, temps de rétention HPLC et rendements de synthèses des chimères **C12-C13**. [a] système RP-HPLC : gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement après purification HPLC.

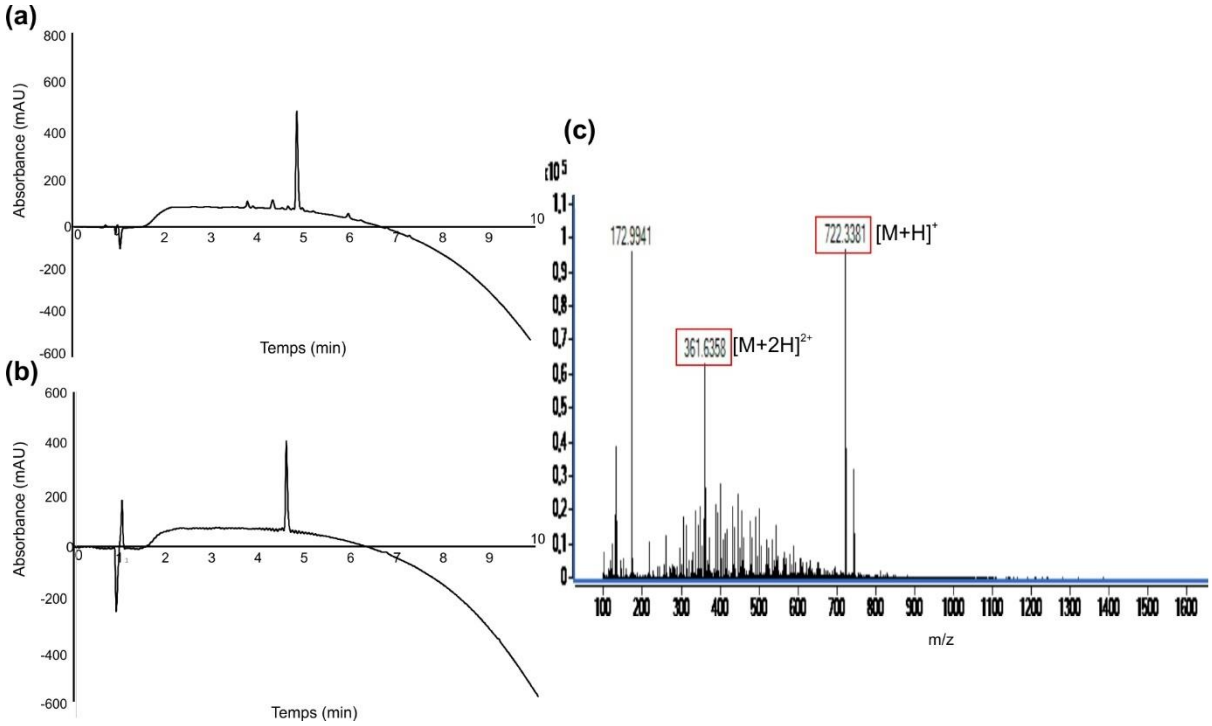


Figure 53 : Caractérisation de la chimère **C12** : **(a)** chromatogramme HPLC de **C12** avant purification ; **(b)** chromatogramme HPLC de **C12** après purification par RP-HPLC préparative sur colonne C18 (gradient linéaire de 20 à 60% de B en 20 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) ; **(c)** spectre de masse de **C12** après purification par RP-HPLC préparative.

ii. Synthèse de la chimère **C14**

La synthèse de la chimère **C14** a été réalisée suivant une série de couplages et de déprotections séquentiels des monomères azotures activés sous forme de carbamate de succinimide suivie d'un couplage séquentiel d'acide aminé α commerciaux protégés par un groupement Fmoc. Cette synthèse n'a pas posé de problème particulier.

chimère	séquence	pureté brut réactionnel	ESI-MS $[M+H]^+$	HPLC t_R (min)[a]	rendement[b]
C14	$ENAA_{\alpha}{}^{U}L^{U}Y^{U}V^{U}Nle^{U}$	79%	1036,43	4,92	32%

Tableau 21 : Séquence, temps de rétention HPLC et rendement de synthèse de la chimère **C14**. [a] système RP-HPLC : gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement après purification

Le chromatogramme RP-HPLC de la chimère **C14** (**Fig. 54a**) avant purification par RP-HPLC semi-préparative indique la présence d'un produit largement majoritaire avec une pureté de 79% à 4,92 min. Après purification par HPLC semi-préparative **C14** est obtenue avec un rendement de 32%.

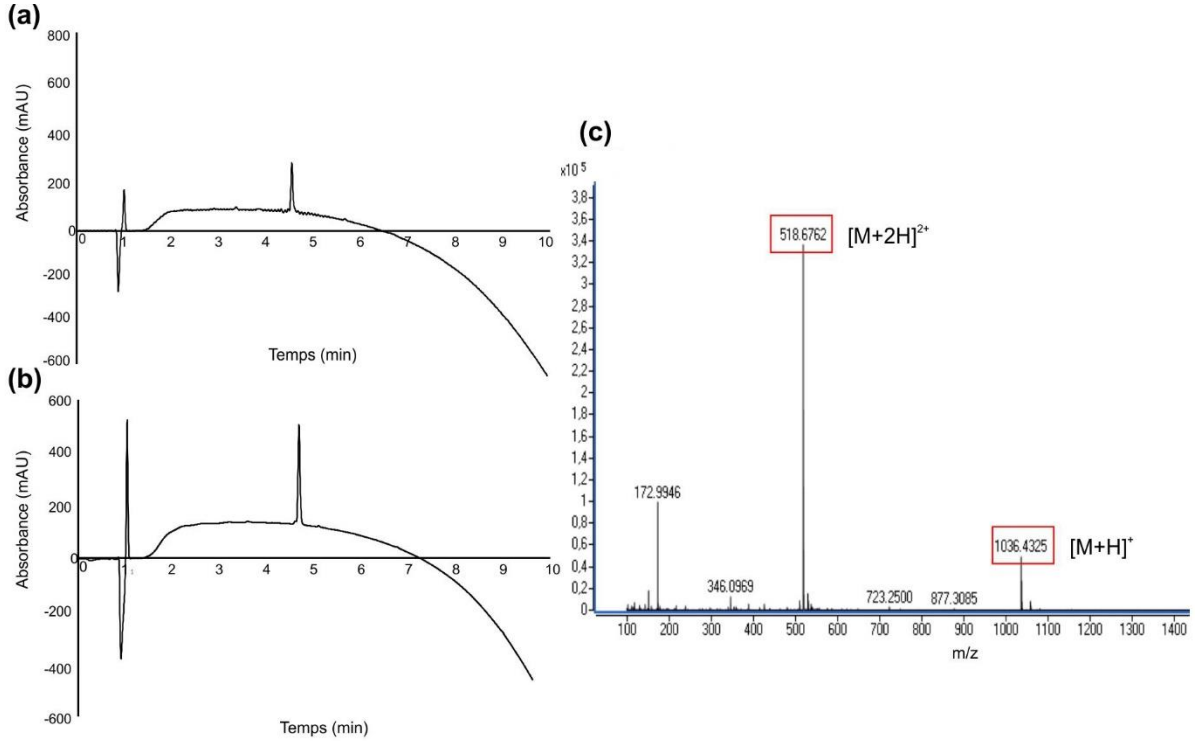


Figure 54 : Caractérisation de la chimère **C14** : **(a)** chromatogramme HPLC de **C14** avant purification ; **(b)** chromatogramme HPLC de **C14** après purification par RP-HPLC préparative sur colonne C18 (gradient linéaire de 20 à 60% de B en 20 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) ; **(c)** spectre de masse de **C14** après purification par RP-HPLC préparative

iii. Synthèse des chimères **C15**-**C21**

Les chimères **C15** à **C19** ont été synthétisées suivant des couplages et déprotections séquentiels d'acides aminés commerciaux protégés par un groupement Fmoc, suivi du couplage et déprotection séquentiels de monomères azoture/carbamate de succinimide.

L'obtention des composés **C15**, **C16**, **C17** et **C18** n'a pas été aussi aisée que dans le cas des chimères précédentes. Les chromatogrammes HPLC des chimères (**Fig. 53a**) avant purification par RP-HPLC préparative indiquent la présence de plusieurs produits correspondant à des délétions de résidus urées. L'analyse par spectrométrie de masse réalisée avant la purification a permis d'identifier la présence des chimères **C15**, **C16**, **C17** et **C18** aux temps de rétention de 4,17 min (**C15**), 4,77 min (**C16**), 4,42 min (**C17**) et 4,80 min (**C18**) avec des puretés de 20% et 15% (**Tableau 22**). Après purification par HPLC préparative **C15**, **C16**, **C17** et **C18** sont obtenues avec des rendements de 8%, 6%, 20% et 10%.

chimères	séquences	pureté brut réactionnel	ESI-MS $[M+H]^+$	HPLC t_R (min)[a]	rendement[b]
C15	$A_\alpha{}^{U}L^{U}Y^{U}V^{U}Nle^{U}KDD$	20%	1080,45	4,17	8%
C16	$L_\alpha{}^{U}L^{U}Y^{U}V^{U}Nle^{U}KDD$	20%	1122,49	4,77	6%
C17	$Nle_\alpha{}^{U}L^{U}Y^{U}V^{U}Nle^{U}KDD$	15%	1122,86	4,42	20%
C18	$Cha_\alpha{}^{U}L^{U}Y^{U}V^{U}Nle^{U}KDD$	20%	1162,89	4,80	10%
C19	$A^{U}A_\alpha{}^{U}L^{U}Y^{U}V^{U}Nle^{U}KDD$	-	-	-	-
C20	$CinA_\alpha{}^{U}L^{U}Y^{U}V^{U}Nle^{U}KDD$	-	-	-	-
C21	$ENAA_\alpha{}^{U}L^{U}R^{U}V^{U}Nle^{U}KDD$	-	-	-	-

Tableau 22 : Séquences, temps de rétention HPLC et rendements de synthèses des chimères **C12-C21**. [a] système RP-HPLC : gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement après purification par RP-HPLC préparative.

Les chromatogrammes RP-HPLC des chimères **C19**, **C20** et **C21** avant purification (**Fig. 56**) révèlent la présence de plusieurs pics. Les analyses par spectrométrie de masse de ces trois chimères avant purification n'ont pas révélé la présence des produits désirés. Certains pics identifiés par spectrométrie de masse semblent correspondre à des séquences où la réduction de l'azoture n'aurait pas fonctionné. Les faibles puretés des chimères **C15**-**C18** avant purification ainsi que le non fonctionnement des synthèses des chimères **C19**-**C21** posent problème. Nous avons décidé d'investiguer afin d'identifier et de comprendre le, ou, les paramètres responsables de ces problèmes de synthèse sur support solide.

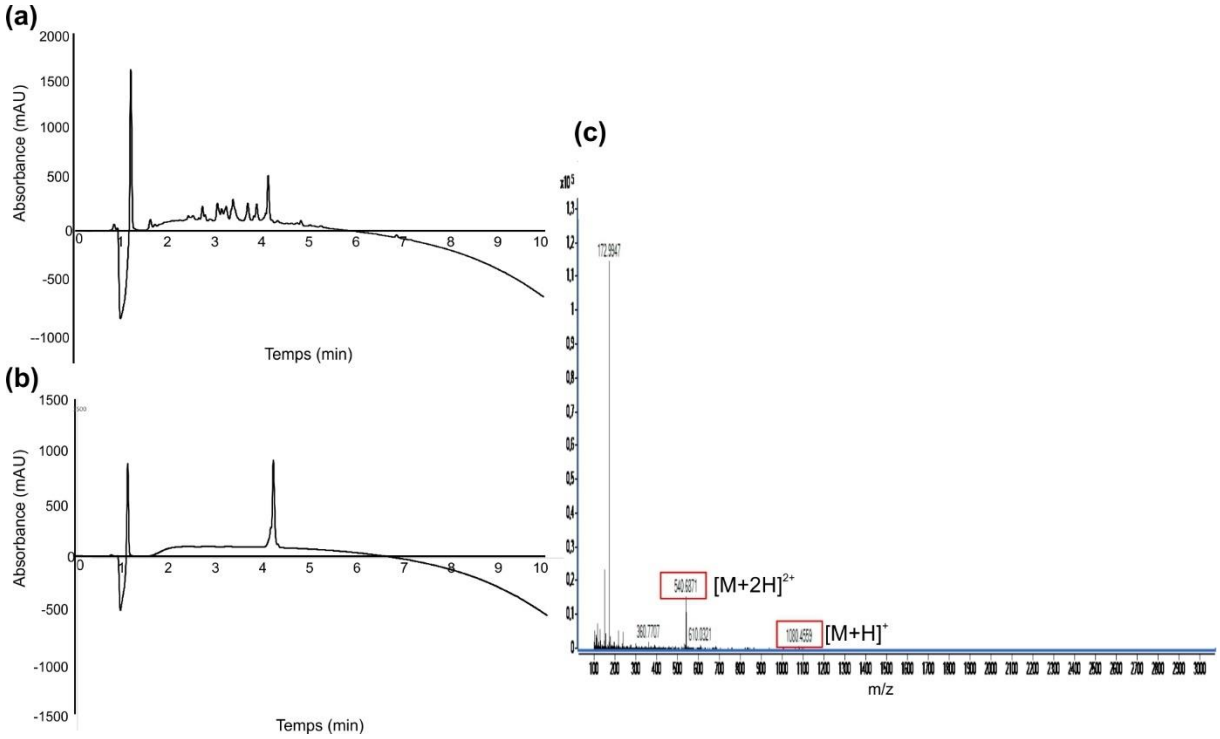


Figure 55 : Caractérisation de la chimère **C15** : **(a)** chromatogramme HPLC de **C15** avant purification ; **(b)** chromatogramme HPLC de **C15** après purification par RP-HPLC préparative sur colonne C18 (gradient linéaire de 20 à 70% de B en 20 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) ; **(c)** spectre de masse de **C15** après purification par RP-HPLC préparative.

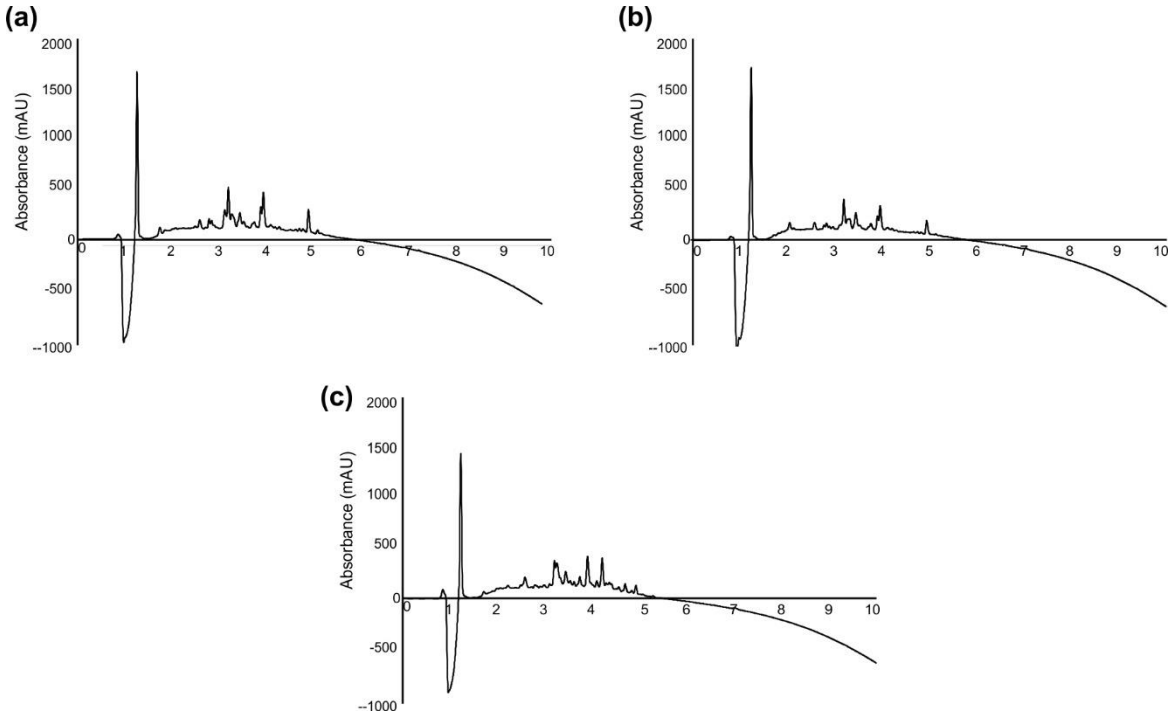


Figure 56 : Chromatogrammes HPLC des chimères après clivage de la résine. **(a)** Chromatogrammes de **C19**, **(b) C20** et **(c) C21**.

c. Problèmes de synthèses identifiés et solutions envisagées

i. Pureté des monomères urées et impact sur la synthèse supportée

Le premier paramètre étudié a été la pureté des lots de monomères carbamate utilisés. Ces monomères ont été synthétisés plusieurs fois, et différents lots ont donc été utilisés au cours des synthèses des chimères **C12**-**C13**, **C15**-**C16** et **C19-C21**. Une analyse RMN du proton des monomères concentrés à 10 mM a révélé des traces résiduelles d'impuretés pour trois d'entre eux : **M5**, **M7** et **M8**. Ces monomères ont été repurifiés par précipitation dans l'éther à froid et les nouvelles analyses RMN ont confirmé la bonne pureté de ces monomères (**Fig. 57**).

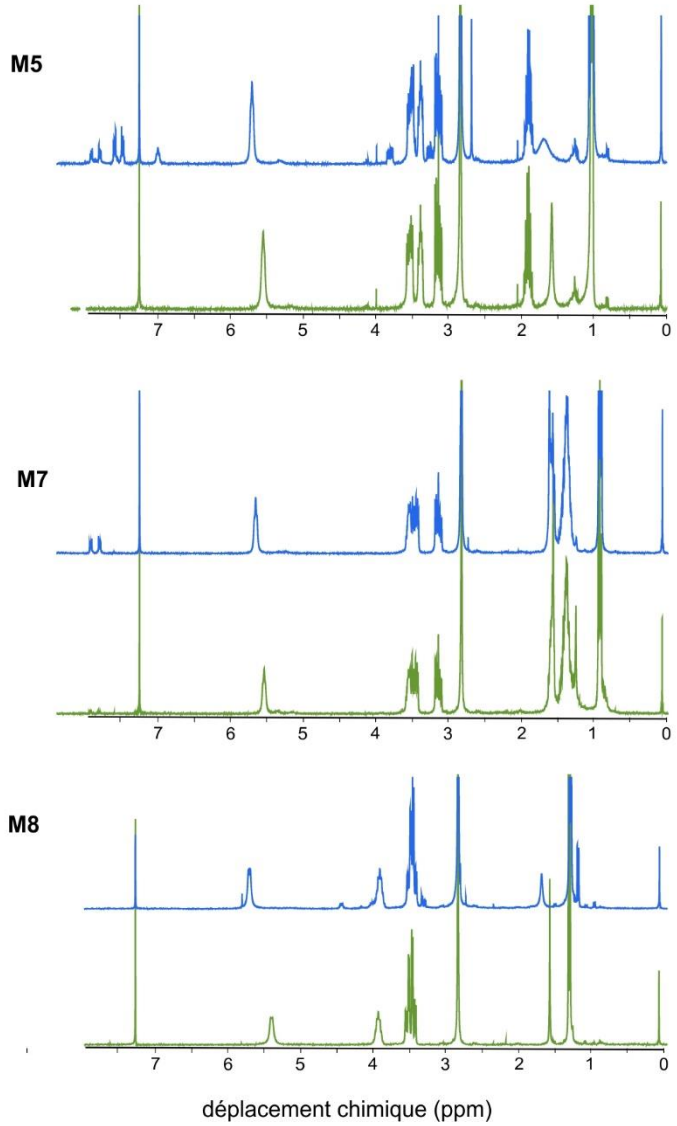


Figure 57 : Spectres RMN ^{1}H (300 MHz, $CDCl_3$) des monomères **M5**, **M7** et **M8** avant (en bleu) et après purification (en vert).

Afin de vérifier que la pureté des monomères n'est pas responsable des problèmes de synthèse rencontrés, nous avons choisi de refaire la synthèse de la chimère **C12** qui avait bien fonctionné. Le chromatogramme HPLC avant purification révèle le bon déroulement de la synthèse avec un pic majoritaire à 4,94 min et une pureté de 87% (**Tableau 23**). Après purification par HPLC semi-préparative **C12-2** est obtenue avec un rendement de 87%. La synthèse a bien fonctionné et les monomères activés ne devraient plus poser de problème lors des synthèses.

Les chimères **C15**, **C16**, **C19**, **C20** et **C21** ont donc été resynthétisées en utilisant les mêmes conditions standard décrites au paragraphe II-3-*b*. Les synthèses des chimères ont mieux fonctionné qu'auparavant, mais les chromatogrammes HPLC de ces chimères avant purification (**Fig. 58** et **Fig. 59**) indiquent toujours la présence de plusieurs produits. Les analyses par spectrométrie de masse réalisées avant la purification ont permis d'identifier la présence des chimères **C15-2** et **C16-2** ainsi que des chimères **C19-2 C20-2** et **C21-2** mais également la présence des composés où la réduction de l'azoture n'aurait pas été totale. L'ensemble des temps de rétention, puretés et rendements de ces composés sont répertoriés dans le **Tableau 23.**

chimères	séquences	pureté brut réactionnel	ESI-MS $[M+H]^+$	HPLC t_R (min)[a]	rendement[b]
C12-2	$A_\alpha{}^UL^UY^UV^UNle^U$	87%	722,50	4,94	87%
C15-2	$A_\alpha{}^UL^UY^UV^UNle^UKDD$	27%	1080,81	4,25	6%
C16-2	$L_\alpha{}^UL^UY^UV^UNle^UKDD$	22%	1122,85	4,78	5%
C19-2	$A^UA_\alpha{}^UL^UY^UV^UNle^UKDD$	7%	1180,87	4,23	3%
C20-2	$CinA_\alpha{}^UL^UY^UV^UNle^UKDD$	2%	1210,85	5,92	0,3%
C21-2	$ENAA_\alpha{}^UL^UR^UV^UNle^UKDD$	23%	1387,99	3,92	2%

Tableau 23 : Séquences, temps de rétention HPLC et rendements de synthèses des chimères **C12-2 – C21-2**. [a] système RP-HPLC :gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement obtenu après purification HPLC

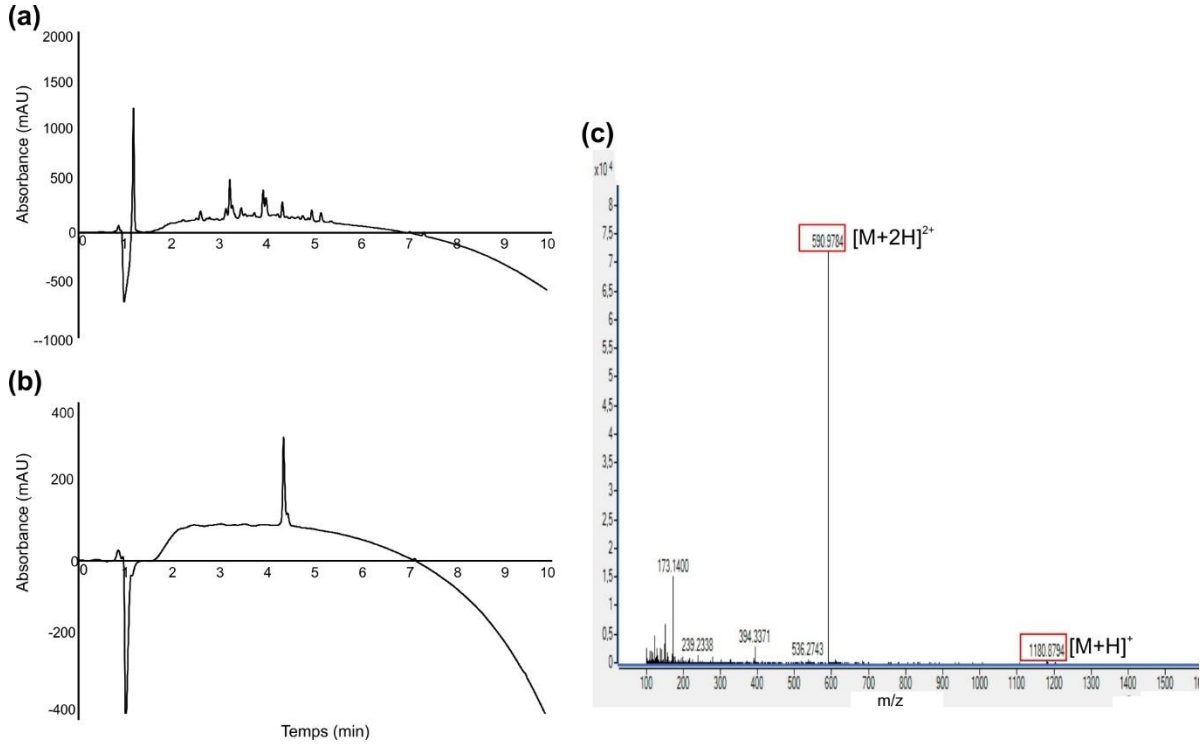


Figure 58 : Caractérisation de la chimère **C19-2** : **(a)** chromatogramme HPLC de **C19-2** avant purification ; **(b)** chromatogramme HPLC de **C19-2** après purification par RP-HPLC préparative sur colonne C18 (gradient linéaire de 20 à 60% de B en 20 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) ; **(c)** spectre de masse de **C19-2** après purification par RP-HPLC préparative.

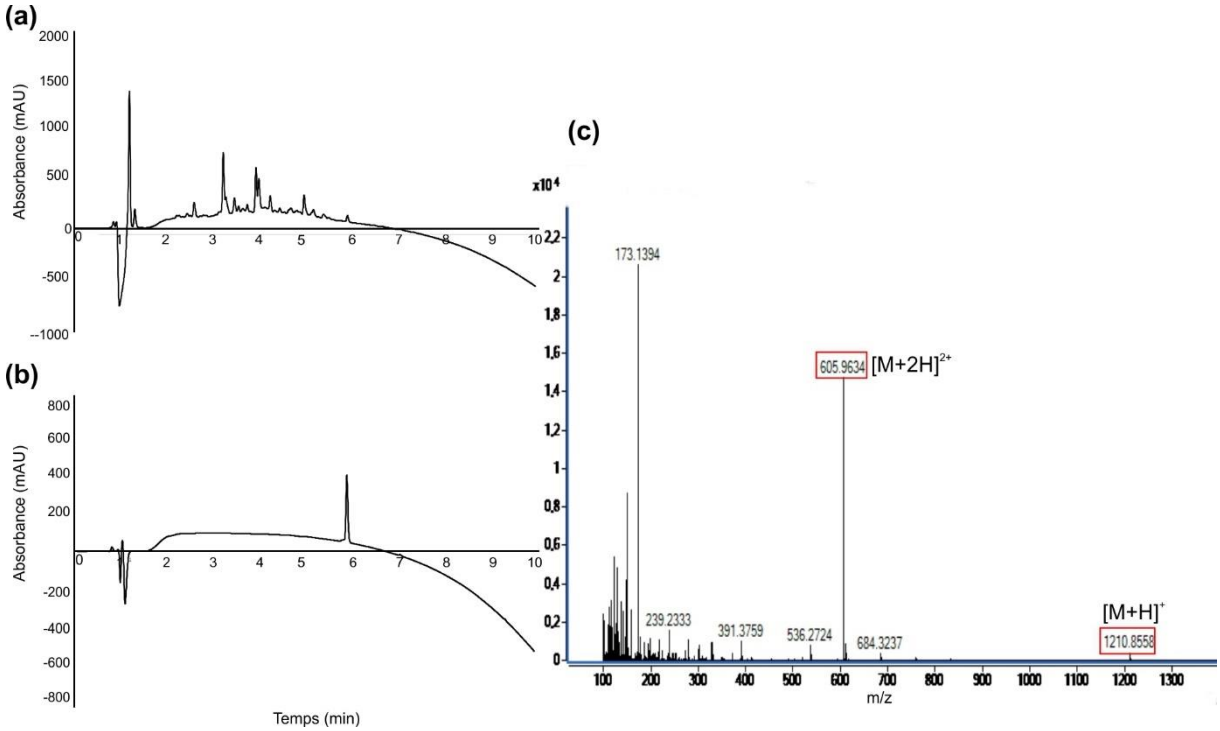


Figure 59 : Caractérisation de la chimère **C20-2** : **(a)** chromatogramme HPLC de **C20-2** avant purification ; **(b)** chromatogramme HPLC de **C20-2** après purification par RP-HPLC préparative sur colonne C18 (gradient linéaire de 20 à 60% de B en 20 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) ; **(c)** spectre de masse de **C20-2** après purification par RP-HPLC préparative

ii. Suivi de réaction par micro-clivage

La synthèse des chimères **C19-2** et **C20-2** pose toujours problème (**Tableau 23**). Ces deux composés sont les seuls dont la synthèse est réellement problématique : très faibles puretés avant purification (7 et 2%) et rendements (3 et 0,3%). Afin de comprendre et identifier l'origine du problème, un suivi de la réaction a été réalisé via des étapes de micro-clivages après chaque étape de couplage et analysé par HPLC et spectrométrie de masse (**Fig. 60**).

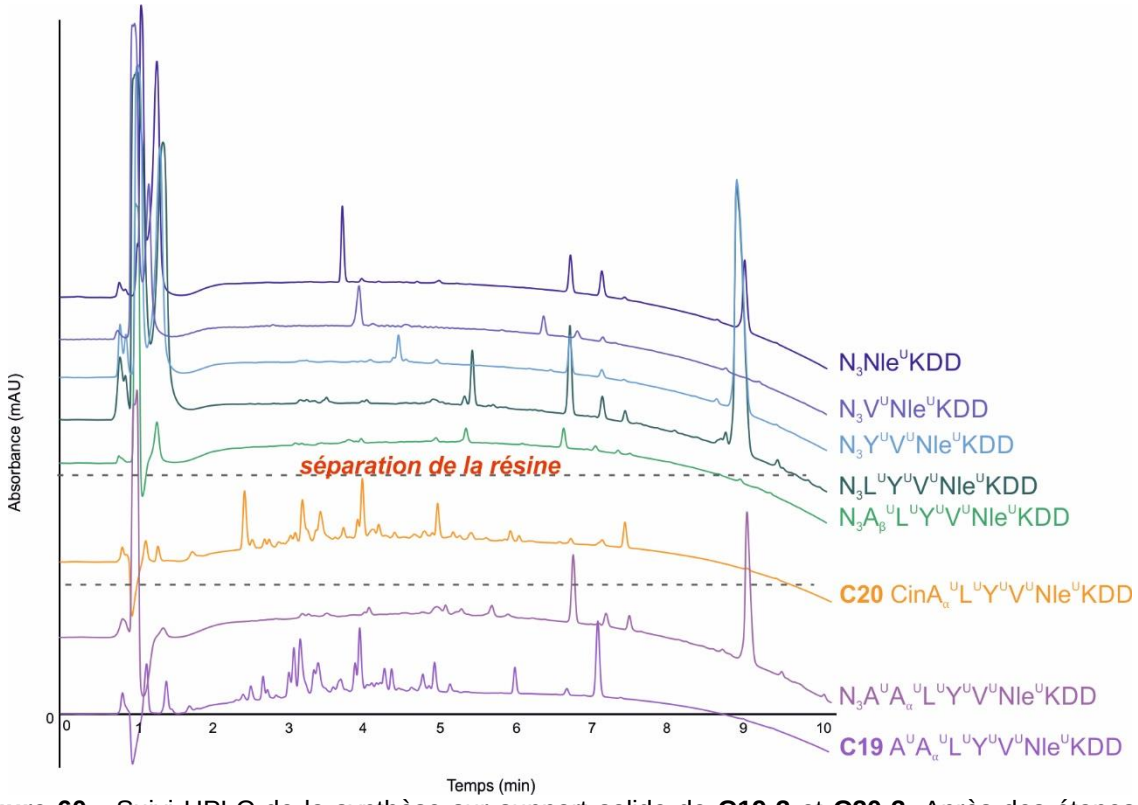


Figure 60 : Suivi HPLC de la synthèse sur support solide de **C19-3** et **C20-3**. Après des étapes de micro-clivages à chaque étape de couplage des monomères activés. La synthèse a été réalisée sur une échelle de 100 µmol et la résine a été séparée en deux lors de la réduction du résidu A_{α}^{U} pour permettre de préparer **C19-3** et **C20-3**. Les pics après 6,5 min ne sont pas visibles par spectrométrie de masse et correspondent aux résidus de clivage. Les micro-clivages ont été réalisés dans un mélange de TFA/TIS/H_2O (95:2,5:2,5) pendant 1h30 à température ambiante, puis le TFA a été évaporé et le composé a été dilué dans un mélange eau/acétonitrile (8:2) puis analysé par HPLC (gradient linéaire de 10 à 100% de B en 10 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) et spectrométrie de masse.

Les chromatogrammes RP-HPLC ainsi que les spectres de masse indiquent le bon déroulement de la synthèse jusqu'au couplage du monomère alanine inversé **M8** où de petits pics résiduels apparaissent. Néanmoins la chimère correspond au pic majoritaire. Lors de la dernière étape de synthèse, les chromatogrammes HPLC ainsi que les spectres de masse indiquent la présence des chimères **C19-3** et **C20-3** comme produits minoritaires (**Fig. 60a**)

avec une pureté de 4% et 2% (**Tableau 24**). Les autres pics semblent correspondre à des délétions où la réduction de l'azoture n'aurait pas fonctionné. Ces résultats sont identiques à ceux observés pour **C19-2** et **C20-2**.

chimères	séquences	pureté brut réactionnel	ESI-MS $[M+H]^+$	HPLC t_R (min)[a]	rendement[b]
C19-3	$A^UA_\alpha{}^UL^UY^UV^UNle^UKDD$	4%	1180,87	4,23	1,5%
C20-3	$CinA_\alpha{}^UL^UY^UV^UNle^UKDD$	2%	1210,85	5,92	2,3%

Tableau 24 : Séquences, temps de rétention HPLC et rendements de synthèses des chimères **C19-3 – C20-3**. [a] système RP-HPLC :gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement obtenu après purification par RP-HPLC préparative.

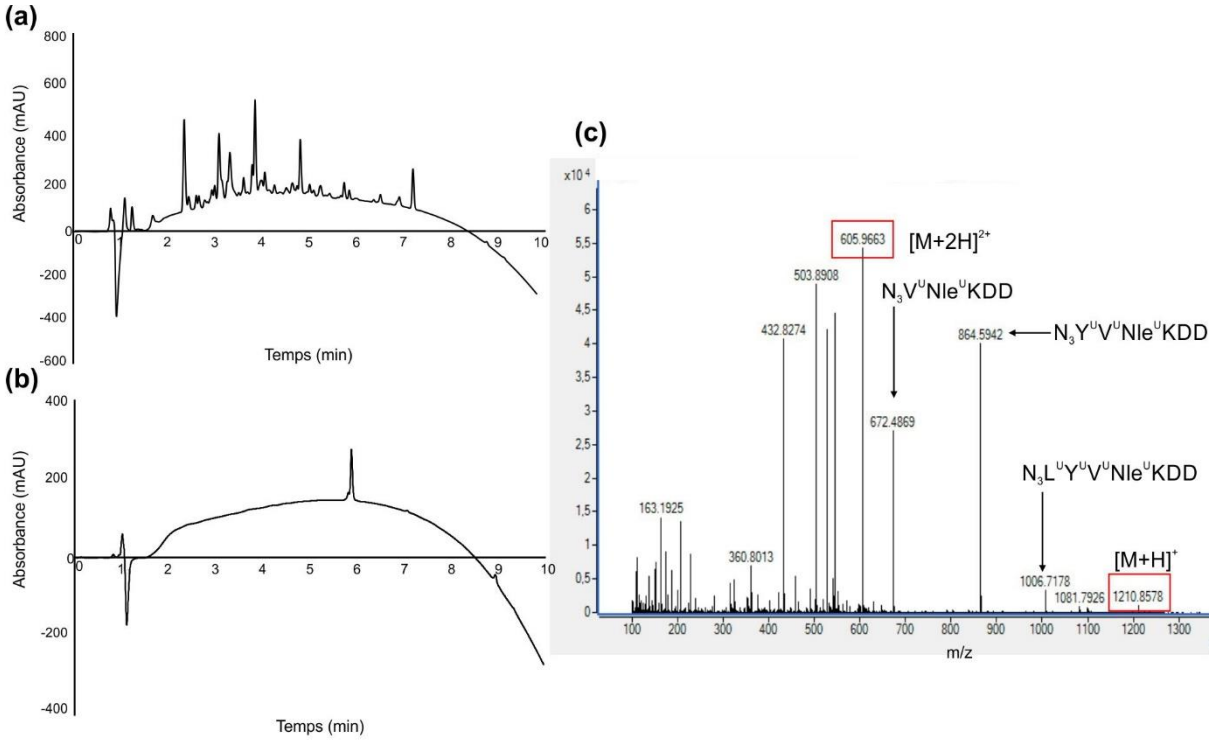


Figure 61 : Caractérisation de la chimère **C20-3** : **(a)** chromatogramme HPLC de **C20-3** avant purification ; **(b)** chromatogramme HPLC de **C20-3** après purification par RP-HPLC préparative sur colonne C18 (gradient linéaire de 20 à 60% de B en 20 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) ; **(c)** spectre de masse de **C20-3** avant purification par HPLC préparative.

iii. Simulation de réaction sur support solide

L'observation de ces profils HPLC semblent également montrer une diminution de l'intensité du signal correspondant à la chimère au cours de la synthèse. Il semble possible qu'une partie de la chimère se séparerait de la résine au cours de la synthèse. Cela serait en faveur du mauvais fonctionnement des synthèses. Afin de vérifier cette hypothèse, une « simulation » de la synthèse a été envisagée, la séquence **N_3NIle^{U}KDD** a été choisie comme modèle et a été soumise aux conditions de champs micro-ondes, températures et temps identiques à ceux de la synthèse des chimères **C19** et **C20**. C'est-à-dire du solvant (DMF) a été ajouté en lieu et place du monomère activé et la résine est chauffée dans les conditions de couplage (50 W, 50°C, 20 min), puis de réduction (50 W, 50°C, 15 min). Après chaque étape de couplage fictif, de la résine est prélevée pour réaliser un micro-clivage. Les produits obtenus dans le filtrat sont analysés par HPLC et spectrométrie de masse. Le suivi HPLC montre qu'il n'y a pas de diminution du signal, ni perte de produit au cours de la simulation de synthèse (**Fig. 62**).

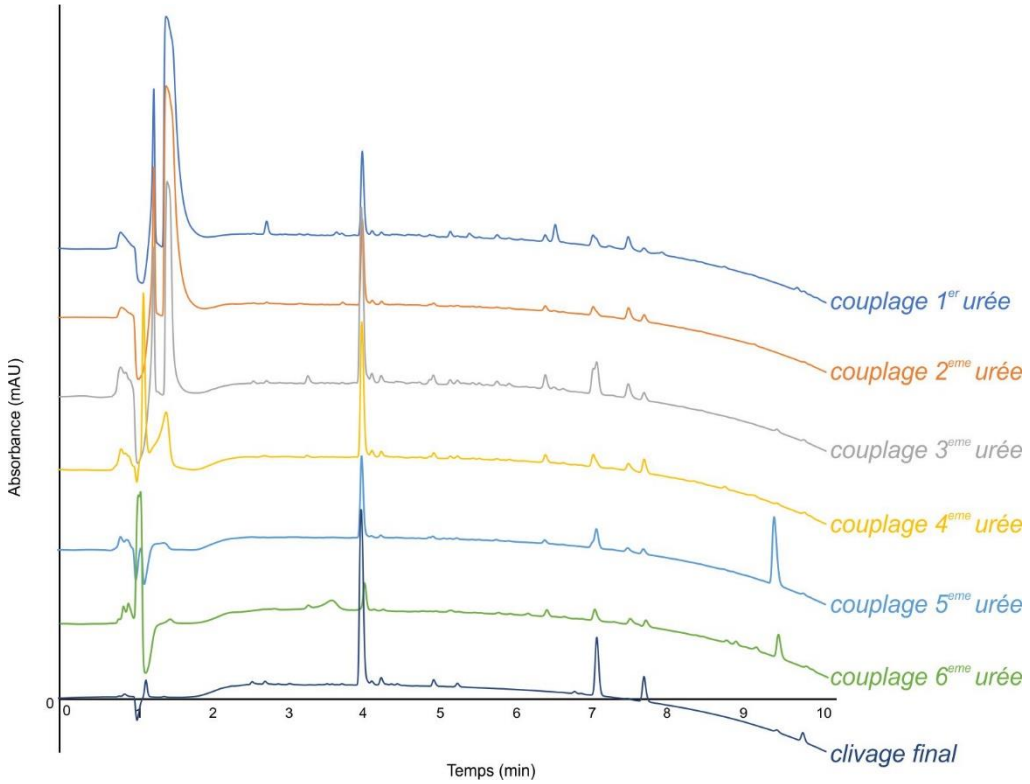


Figure 62 : Suivi HPLC des micro-clivages de la simulation de la synthèse sur support solide. Les pics après 6,5 min ne sont pas visibles par spectrométrie de masse et correspondent aux résidus de clivage. Les micro-clivages ont été réalisés dans un mélange de TFA/TIS/H_2O (95:2,5:2,5) pendant 1h30 à température ambiante, puis le TFA a été évaporé et le composé a été dilué dans un mélange eau/acétonitrile (8:2) puis analysée par HPLC (gradient linéaire de 10 à 100% de B en 10 minutes avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN)) et spectrométrie de masse.

Les analyses HPLC et de spectrométrie de masse des solvants de réaction ne révèlent pas la présence des sous-produits observés lors des synthèses initiales de **C19** et **C20**. Aucune analyse ne confirme la dégradation de la chimère sur la résine.

Au sein du laboratoire, il a été constaté que le bon fonctionnement de la synthèse sur support solide de chimères et d'oligourées est parfois dépendant de la séquence. Il est alors possible d'imaginer que la synthèse des chimères **C19** et **C20** soit concernée par ce phénomène. Afin de vérifier cette hypothèse, une modification de la position des résidus dans la séquence et/ou de leur nature a été considérée. La synthèse n'a pas été réalisée par manque de temps, mais est envisagée dans le futur.

3. Caractérisation bio-structurale des foldamères mixtes de nouvelle génération

La cristallisation des complexes VDR/chimère a été réalisée au sein de l'équipe du Dr. Natacha Rochel-Guiberteau. Les chimères ont été co-cristallisées avec la protéine zVDR issue du poisson-zèbre (*Danio rerio*) et la vitamine D_3 en utilisant la méthode de la goutte suspendue. Les conditions de cristallisation des chimères sont répertoriées dans le **Tableau 25**. Les structures cristallines des chimères **C14** et **C19** en complexe avec VDR ont été obtenues en haute résolution par diffraction des rayons X. Les données ont été collectées aux synchrotrons SOLEIL (Gif-sur-Yvette) et ESRF (Grenoble). Les modèles des chimères ont été construits par mes soins comme décrit au paragraphe II-1, les jeux de données ainsi que les statistiques d'affinement sont donnés dans le **Tableau 26** et permettent de valider les modèles obtenus. A ce jour les autres séquences hybrides n'ont pas cristallisé avec zVDR.

chimères	protéine et ligand	condition de cristallisation
C14	zVDR-LBD + $VitD_3$	Ammonium sulfate 1,5 M, BisTrisPropane 0,1 M pH 7,0
C19	zVDR-LBD + $VitD_3$	Sulfate de lithium 1,5 M

Tableau 25 : Conditions de cristallisation des chimères **C14** et **C19** ayant cristallisées avec VDR.

Ces structures sont comparées à la structure de la protéine VDR utilisée pour la conception des chimères et à **C11**, qui est la séquence « tête de série » utilisée pour la génération des nouveaux composés. Tous les complexes cristallisent dans une maille hexagonale ($a = b \neq c$; $\alpha = \beta = 90°$, $\gamma = 120°$) avec le même groupe d'espace $P6_522$. Cette observation permet d'affirmer que les chimères, quelle que soient leurs modifications (longueur de la portion urée, nature des résidus urées, et positions des chaînes latérales des résidus urées), n'affectent pas le groupe d'espace dans lequel les complexes cristallisent.

Les chimères **C14** et **C19** conservent le mode de liaison canonique des chimères de la première génération par rapport au ligand naturel. En effet, la pince chargée est conservée dans les structures cristallines pour les deux chimères (**Tableau 27**). Les portions C-terminales de la chimère **C19**, tout comme celle de la séquence naturelle, ne sont pas visibles dans les cartes de densité et n'ont donc pas été construites.

	VDR/SRC1-2 (2HC4)	C14	C19
Résolution (Å)	20,0-2,20 (2,28-2,20)	48,0-2,20 (2,26-2,20)	19,9-2,30 (2,38-2,30)
Groupe d 'espace	$P6_522$	$P6_522$	$P6_522$
Paramètre de maille (Å)	a = b = 65,84; c = 262,84 α = β = 90 γ = 120	a = b = 66,47; c = 261,96 α = β = 90 γ = 120	a = b = 66,24; c = 252,52 α = β = 90 γ = 120
Nombre de réflexions uniques	18382	18379	15412
Redondance	10,2	2,0	37,1
R_{syn} (%)[a]	4,8	1,5	10,9
Complétude (%)	99,0 (99,5)	99,7 (99,0)	99,3 (99,8)
Moyenne *I*/σ (%)	29,1 (7,3)	21,6 (1,35)	25,5 (1,35)
RMSD.[b] sur les liaisons (Å)	0,006	0,009	0,006
sur les angles (°)	1,13	1,10	0,93
R_{work} / R_{free} [c] (%)	22,0/25,1	18,2/23,1 (27,8/31,4)	19,9/23,9 (29,1/35,0)
Facteur B			
Protéine		65,41	78,85
Ligand		59,21	78,47
Eau		66,96	75,59

Tableau 26 : Statistiques des jeux de données et de l'affinement des complexes avec la séquence naturelle et avec les séquences chimères **C14**-**C19**. Les valeurs entre parenthèse correspondent aux très hautes résolutions. [a] $R_{sym}(I) = \sum_{h\,k\,l}\sum_i |I_{h\,k\,l,i} - \langle I_{h\,k\,l}\rangle| / \sum_{h\,k\,l}\sum_i |I_{h\,k\,l,i}|$ avec $\langle I_{h\,k\,l}\rangle$, l'intensité moyenne des multiples $I_{h\,k\,l,i}$ observée par les réflexions liées aux éléments de symétrie. [b] Les déviations de la racine des moyenne des carrés (RMSD) sont données à partir de valeurs idéales. [c] $R_{work} = \sum_{h\,k\,l}|F_{obs} - F_{calc}| / \sum_{h\,k\,l} F_{obs}$, F_{obs} et F_{calc} correspondent aux facteurs de structures observés et calculés. R_{free} est similaire à R_{work}, mais calculé sur les 10% des données exclues de l'affinement.

ligands	$d_{CO---NH}$K274 (Å)	$d_{CO--NH1}$E446 (Å)	$d_{CO--NH2}$E446 (Å)
SRC1-2	2,9	2,8	2,8
C14	2,5	3,1	2,7
C19	3,1	2,9	2,7

Tableau 27 : Distances mesurées des liaisons hydrogène entre la chaîne principale du ligand et les résidus K274 et E446 de zVDR formant la pince chargée.

Les mesures d'affinité par MST et de stabilité par nanoDSF sont en cours et ne sont pas encore disponibles pour la totalité des séquences des chimères de seconde génération. Cependant, des premières mesures de stabilité par nanoDSF ont été réalisées avec les séquences des chimères **C12**, **C13**, **C14**, **C15**, **C16** et **C19** (**Tableau 28**).

ligands	séquences	hVDR-LBD T_m (°C)
sans ligand		57,7
P1[a]	$R_{-4}H_{-3}K_{-2}I_{-1}L_1H_2R_3L_4L_5Q_6E_7G_8S_9P_{10}S_{11}$	63,6
C12	$A_{-1\alpha}{}^{U}L_1{}^{U}K_2{}^{U}V_3{}^{U}Nle_4{}^{U}$	60,1
C13	$A_{-1}{}^{U}L_1{}^{U}Y_2{}^{U}V_3{}^{U}Nle_4{}^{U}$	59,4
C14	$E_{-4}N_{-3}A_{-2}A_{-1\alpha}{}^{U}L_1{}^{U}K_2{}^{U}V_3{}^{U}Nle_4{}^{U}$	59,7
C15	$A_{-1\alpha}{}^{U}L_1{}^{U}Y_2{}^{U}V_3{}^{U}Nle_4{}^{U}K_5D_6D_7$	60,8
C16	$L_{-1\alpha}{}^{U}L_1{}^{U}Y_2{}^{U}V_3{}^{U}Nle_4{}^{U}K_5D_6D_7$	61,8
C19	$A_{-2}{}^{U}A_{-1\alpha}{}^{U}L_1{}^{U}Y_2{}^{U}V_3{}^{U}Nle_4{}^{U}K_5D_6D_7$	60,0

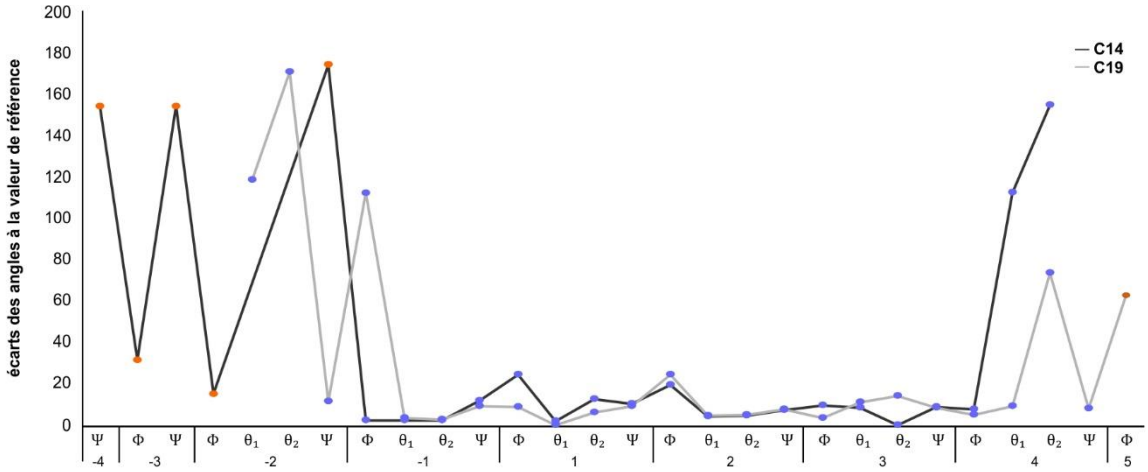
Tableau 28 : Tableau récapitulatif des résultats des mesures de stabilité pour les chimères **C12**-**C19** et le peptide **P1**. Les mesures sont réalisées par nanoDSF et les valeurs correspondent à la moyenne des températures de fusion à 330 et 350 nm. [a] séquence du peptide utilisée comme références pour les études d'affinité et de stabilité

Ces résultats préliminaires semblent montrer l'importance de la position de la chaîne latérale du résidu urée inversé sur le carbone C_α pour stabiliser le complexe (**C12** vs **C13**). La présence de la séquence peptidique en C-terminale (**C15**) semble également plus importante pour stabiliser le complexe hVDR/chimère par rapport à la séquence peptidique située à l'extrémité N-terminale (**C14**). Ces premiers résultats sont confirmés par l'étude des structures cristallines disponibles à ce jour.

	nature du résidu et position / RMSD (Å)								
chimères	-4	-3	-2	-1	1	2	3	4	5
C14	E 9,81	N 7,86	A 2,64	$A_\alpha{}^U$ 1,75	L^U 1,67	Y^U 1,71	V^U 1,67	Nle^U 1,56	
C19			A^U 3,80	$A_\alpha{}^U$ 1,45	L^U 1,01	Y^U 1,64	V^U 1,11	Nle^U 1,18	K 2,23

Tableau 29 : Tableau récapitulatif des valeurs de RMSD calculées pour la superposition de **C14** et **C19** à SRC1-2. Les résidus urée projetés dans le sillon hydrophobe de VDR sont surlignés en jaune.

La caractérisation de l'hélicité des chimères **C14** et **C19** (**Graph. 5**) montre que pour **C14** l'hélice d'oligourée canonique est conservée entre les résidus -1 et 3. Dans le cas de **C19**, l'hélice d'oligourée commence à être canonique à partir du résidu -1 et se termine lors du dernier résidu urée en position 4. La présence du résidu urée à la position N-terminale -2 ne semble pas initier une pré-organisation de la structure en hélice canonique d'oligourée de **C19**. Les hélices canoniques d'oligourée sont plus courtes pour ces deux chimères comparées à celle de **C11**.



Graphique 5 : Caractérisation de l'hélicité des chimères **C14** et **C19**. L'écart des angles dièdres à leurs valeurs de références est donné pour chaque résidu.

Les structures cristallines de **C14** et **C19** ne miment pas correctement la portion en N-terminale, les valeurs des RMSD sont supérieures à 7 Å pour les résidus en positions -4 et -3. Le résidu $A_{-1\alpha}^{U}$ de **C19** a un RMSD de 1,45 Å, qui est inférieur à celui de **C11** et **C14** (**Tableau 29**) : la position -1 est donc mieux mimée par **C19**. Les résidus urées L_1^U, V_3^U et Nle_4^U de **C19** ont des meilleures valeurs RMSD que **C14** et **C11**, leurs projections dans la poche hydrophobe est donc meilleure et plus proche de la séquence naturelle. Les résidus urées tyrosine en position 2 de **C14** et **C19** ont des valeurs de RMSD améliorées par rapport à **C11**. L'optimisation du mimétisme du résidu à cette position a été réalisée avec succès.

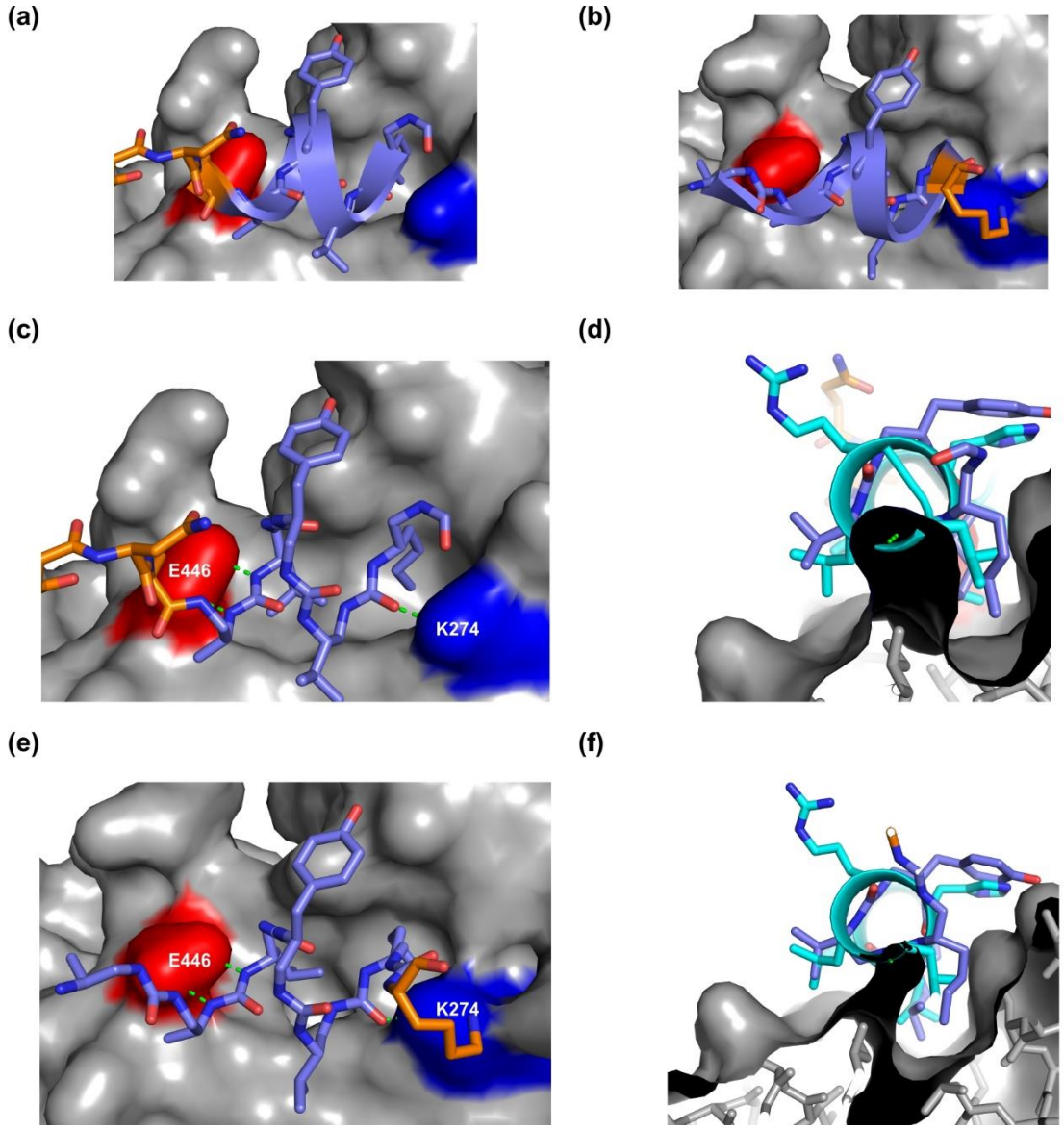


Figure 63 : Structures cristallines de l'interaction entre zVDR (en gris) et les chimères **C14** et **C19** (les résidus urées sont en violet et les acides aminés en orange). Représentation des hélices sous forme de schéma des chimères **(a) C14** et **(b) C19**. **(c)** Représentation de l'ancrage de **C14** dans la poche hydrophobe avec la pince chargée formée par les acides aminés E446 (en rouge) et K274 (en bleu). **(d)** Vue de profil de la superposition de **C14** avec SRC1-2 (en cyan), révélant la projection des chaînes latérales dans la poche hydrophobe. **(e)** Représentation de l'ancrage de **C19** dans la poche hydrophobe avec la pince chargée formée. **(f)** Vue de profil de la superposition de **C19** avec SRC1-2, indiquant la projection des chaînes latérales dans la poche hydrophobe.

L'acide aminé C-terminal lysine à la position 5 de **C19** est visible dans la carte de densité, et a donc été reconstruit. La superposition avec SRC1-2 montre que ce résidu mime difficilement la partie en C-terminale du ligand naturel, il semble être orienté vers les chaînes latérales des résidus aminés K274 et R280 de VDR. La valeur RMSD de 2,23 Å confirme le mauvais mimétisme de cette acide aminé à cet emplacement de la séquence (**Fig. 63**). L'ajout

d'une urée supplémentaire en N-terminal dans la séquence **C19** suggère la formation d'une liaison hydrogène supplémentaire entre l'acide aminé E446 et le dérivé alanine urée en position -2.

4. Conclusion sur la génération d'une seconde série de foldamères mixtes

L'étude des données cristallographiques de la chimère « tête de série » **C11** a permis de concevoir huit nouveaux composés. La préparation de deux nouveaux monomères urée nécessaire à la synthèse de ces nouveaux composés a été réalisée avec de bons rendements. Les synthèses des chimères **C12** à **C18** et **C21** n'ont pas posé de difficulté contrairement aux chimères **C19** et **C20**. Un suivi de réaction sur support solide par micro-clivage et une simulation de synthèse sur support de ces synthèses difficiles n'ont pas permis d'identifier le, ou les paramètres responsables de ces problèmes de synthèse. Cette étude n'a pas pu être amenée plus loin faute de temps, mais elle sera développée dans le futur via l'étude de la synthèse séquence-dépendante pour les chimères **C19** et **C20**. Les données structurales obtenues seront comparées et mises en relation avec les mesures biophysiques lorsqu'elles seront disponibles. Cette nouvelle étude SAR permettra la génération d'une troisième série de foldamères mixtes oligoamide/oligourée ciblant l'interaction VDR/SRC1-2.

IV. Application à d'autres récepteurs nucléaires

D'autres interactions cibles entre des peptides corégulateurs et récepteurs nucléaires ont également pu être investiguées pour le développement de nouvelles séquences chimères oligoamide/oligourée. Une étude de relation structure-affinité avec les chimères **C2**-**C11** a été initiée par le Dr. Natacha Rochel-Guiberteau avec deux autres récepteurs nucléaires : ERα et RARα. Les mesures de stabilités été réalisés (**Tableau 30**) avec les LBD de ces deux protéines indiquent que les chimères **C7 C9** et **C11** stabilisent le mieux le complexe formé avec RARα. Cet effet de stabilisation est très léger pour les oligomères **C5**, **C7** et **C9** avec la protéine ERα. Récemment, une structure cristalline du complexe **C11** avec RARα a été obtenue, résolue et affinée (**Fig. 64**). Lorsque l'ensemble des informations bio-structurales seront disponibles, l'étude de structure-affinité permettra la génération de nouvelles séries de foldamères hybrides oligoamide/oligourée ciblant les interactions entre les récepteurs ERα et RARα avec les protéines SRC.

ligands	séquences	ERα-LBD T_m (°C)	RARα-LBD T_m (°C)
sans ligand		47,1	47,1
P1[a]	$R_{-4}H_{-3}K_{-2}I_{-1}L_1H_2R_3L_4L_5Q_6E_7G_8S_9P_{10}S_{11}$	49,0	49,0
C2	$R_{-4}H_{-3}K_{-2}I_{-1}L_1{}^{U}K_2{}^{U}L_3{}^{U}L_4Q_5E_6G_7S_8$	47,3	46,8
C3	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2{}^{U}L_3{}^{U}L_4{}^{U}Q_5E_6G_7S_8$	47,3	47,1
C4	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2{}^{U}L_3{}^{U}L_{4\alpha}{}^{U}Q_5E_6G_7S_8$	47,4	47,2
C5	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2{}^{U}L_{3\alpha}{}^{U}L_4{}^{U}Q_5E_6G_7S_8$	47,5	47,6
C6	$R_{-4}H_{-3}K_{-2}I_{-1}L_1K_2{}^{U}L_{3\alpha}{}^{U}L_{4\alpha}{}^{U}Q_5E_6G_7S_8$	47,5	46,8
C7	$R_{-4}H_{-3}K_{-2}I_{-1}L_1R_2{}^{U}L_{3\alpha}{}^{U}L_4{}^{U}Q_5E_6G_7S_8$	47,4	47,8
C8	$R_{-4}H_{-3}K_{-2}I_{-1}L_1R_2{}^{U}L_{3\alpha}{}^{U}A_4{}^{U}Q_5E_6G_7S_8$	47,3	46,9
C9	$R_{-4}H_{-3}K_{-2}A_{-1\alpha}{}^{U}L_1{}^{U}K_2{}^{U}V_3{}^{U}Nle_4{}^{U}E_5G_6S_7$	47,4	48,0
C10	$K_{-4}H_{-3}K_{-2}A_{-1\alpha}{}^{U}L_1{}^{U}K_2{}^{U}V_3{}^{U}Nle_4{}^{U}D_5S_6S_7$	47,3	47,7
C11	$E_{-4}N_{-3}A_{-2}A_{-1\alpha}{}^{U}L_1{}^{U}Y_2{}^{U}V_3{}^{U}Nle_4{}^{U}K_5D_6D_7$	47,2	48,0

Tableau 30 : Tableau récapitulatif des résultats des mesures de stabilité avec les protéines ERα et RARα pour les chimères **C2**-**C11** et le peptide **P1**. Les mesures sont réalisées par nanoDSF et les valeurs correspondent à la moyenne des températures de fusion à 330 et 350 nm. [a] séquence du peptide utilisée comme références pour les études d'affinité et de stabilité

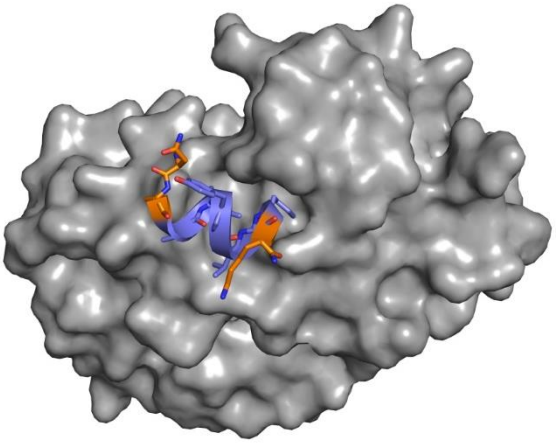


Figure 64 : Structure cristalline de la chimère **C11** (les résidus urées sont en bleu et les résidus d'acides aminés en orange) avec RARα (en gris)

V. Conclusions du chapitre et perspectives

La conception de chimères par superposition d'un foldamère mixte à la séquence du peptide SRC1-2 en interaction avec VDR a permis de générer une première série de foldamères hybrides **C2**-**C11**. L'un d'entre eux s'est révélé être très prometteur, et sa séquence a été optimisée en vue d'améliorer son interaction avec la protéine et inhiber l'action de VDR/SRC1-2. L'identification d'une tête de série dans la seconde génération de foldamères mixtes oligoamide/oligourée permettrait la génération d'une nouvelle série de chimères. L'approche utilisée pour la conception de la première série sera répétée jusqu'à l'obtention

d'un foldamère mixte mimant parfaitement le peptide naturel et possédant de bonnes activités d'inhibition de l'interaction VDR/SRC1-2.

Une autre stratégie de conception de foldamères mixtes a été considérée, mais pas réalisée faute de temps. L'objectif était de développer des foldamères hybrides issus de **C11** contenant un motif de type stapled peptide. Nous nous sommes inspirés des travaux de Speltz et al[14] sur le développement de peptides réticulés inhibiteurs de l'interaction entre ERα et SRC2-2. La conception des foldamères mixtes oligoamide/oligourée réticulés s'est basée sur la superposition des séquences développées par Speltz à la structure de **C11**. La superposition des quatre peptides stapled de Spletz montre que la réticulation des chaînes latérales pourrait se faire aux positions -1 et 3 de **C11**. La distance entre les deux chaînes latérales étant de 4,4 Å, la réticulation des chaînes latérales par la formation d'un pont disulfure entre deux résidus urées cystéine et homocystéine a été considérée (**Fig. 65**) et a conduit à la conception de séquences foldamères chimères réticulées **FCR1-FCR4** (**Tableau 31**).

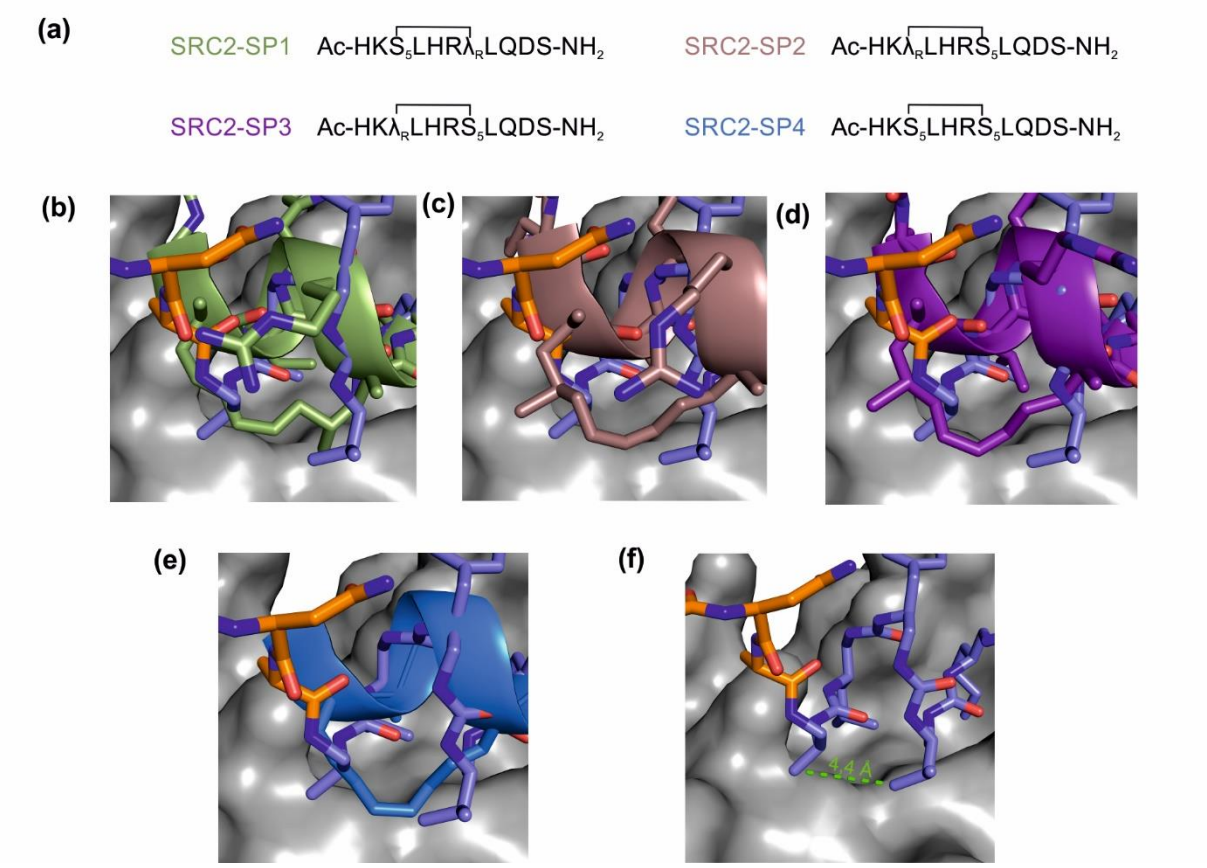


Figure 65 : Stratégie de conception de foldamères chimères réticulés. **(a)** Séquences des peptides réticulés développés par Speltz *et al.* Superposition de **C11** aux peptides réticulés **(b)** SRC2-CP1, **(c)** SRC2-SP2, **(d)** SRC2-SP3, **(e)** SC2-SP4. **(f)** Distance mesurée (4,4 Å) entre les deux chaînes latérales en positions -1 et 3 où la réticulation pourrait avoir lieu via la formation d'un pont disulfure entre deux résidus urées de type cystéine.

chimères	-4	-3	-2	-1	1	2	3	4	5	6	7
FCR1	E	N	A	C^{U}	L^{U}	Y^{U}	C^{U}	Nle^{U}	K	D	D
FCR2	E	N	A	hC^{U}	L^{U}	Y^{U}	C^{U}	Nle^{U}	K	D	D
FCR3	E	N	A	C^{U}	L^{U}	Y^{U}	hC^{U}	Nle^{U}	K	D	D
FCR4	E	N	A	hC^{U}	L^{U}	Y^{U}	hC^{U}	Nle^{U}	K	D	D

Tableau 31 : Séquences des chimères réticulés **FCR1-FCR4.** Les résidus en jaune correspondent aux résidus projetés dans la poche hydrophobe mimant les leucines en i, i+3 et i+4 du peptide naturel.

L'ensemble de ces travaux de thèse démontrent qu'il existe différentes approches pour cibler les interactions protéine-protéine et qu'il est possible de les cibler avec des foldamères mixtes oligoamide/oligourée pour mimer la petite portion peptidique en hélice α responsable de l'interaction.

Références bibliographiques

1. (a) P. R. Housley, E. R. Sanchez, M. Danielsen, G. M. Ringod, W. Pratt *J. Biol. Chem.* **1990**, *265*, 12778-12781 ; (b) D. M. Heery, E. Kalkhoven, S. Hoare, M. G. Parker *Nature* **1997**, *387*, 733-736
2. N. McKenna, R. B. Lanz, B. W. O'Malley *Endocr. Rev.* **1999**, *20*, 321-344
3. A. Teichert, L. A. Arnold, S. Otieno, Y. Oda, I. Augustinaite, T. T. Greistlinger, R. W. Kriwacki, R. K. Guy, D. D. Bikle *Biochemistry* **2009**, *48*, 1454-1461
4. (a) Y Demizu, S. Nagoya, M. Shirakawa, M. Kawamura, N. Yamagata, Y. Sato, M. Doi, M. Kurihara *Bioorg. Med. Chem. Lett.* **2013**, *23*, 4292-4296 (b) T. Misawa, Y. Demizu, M. Kawamura, N. Yamagata, M. Kurihara *Bioorg. Med. Chem.* **2015**, *23*, 1055-1061
5. (a) J. Fremaux, L. Mauran, K. Pulka-Ziach, B. Kauffmann, B. Odaert, G. Guichard *Angew. Chem. Int. Ed.* **2015**, *54*, 9816-9820 (b) L. Mauran, B.Kauffmann, B. Odaert, G. Guichard *C. R. Chimie* **2016**, *19*, 123-131
6. C. Douat-Casassus, K. Pulka, P. Claudon, G. Guichard *Org. Lett.* **2012**, *14*, 3130-3133
7. P. B. Alper, S.-C. Hung, G.-H. Wong *Tetrahedron Lett.* **1996**, *34*, 6029-6032
8. (a) E. D. Goddard-Borger, R. V. Stick *Org. Lett.* **2007**, *19*, 3797-3800 (b) N. Fischer, E. D. Goddard-Borger, R. Greiner, T. M. Klapötke, B. W. Skelton, J. Stierstorfer *Org. Chem.* **2012**, *4*, 1760-1764
9. (a) Z. Zhang, E. Fan *Org. Chem.* **2005**, *22*, 8801-8810 (b) K. Burgess *J. Am. Chem.* **1997**, *7*, 1556-1564
10. N. Pendem, C. Douat, P. Claudon, M. Laguerre, S. Castano, B. Desbat, D. Cavagnat, E. Ennifar, B. Kauffmann, G. Guichard *J. Am. Chem.* **2013**, *288*, 10914-10922
11. R. B. Merrifield *J. Am. Chem.* **1963**, *85*, 2149-2154
12. (a) L. Fischer, V. Semetey, J-M. Lozano, A-P. Schaffner, J-P. Briand, C. Didierjean, G. Guichard *Eur. J. Org. Chem.* **2007**, 2511-2525 (b) V. Diemer, L. Fischer, B. Kauffmann, G. Guichard *Chem. Eur. J.* **2016**, *22*, 15684-15692
13. (a) J. T. Lundquist, J. C. Pelletier *Org. Lett.* **2011**, *3*, 781-783 (b) F. Debaene, N. Wissinger *Org. Lett.* **2003**, *5*, 4445-4447
14. T. E. Speltz, S. W. Fanning, C. G. Mayne, C. Fowler, E. Tajkhorshid, G. L. Greene, T. W. Moore *Angew. Chem. Int. Ed.* **2016**, *55*, 4252-4255

Chapitre 3

Développement de mimes de structures tertiaires hélicoïdales à base de foldamères d'oligourée

Les structures secondaires en hélice sont fréquemment présentes et impliquées dans l'interaction entre protéines[1]. Toutefois, des modes d'interaction plus complexes mettant en jeu des motifs hélicoïdaux de structures tertiaires et quaternaires ont été identifiées[2] sous forme de dimères d'hélices. Le développement d'une base de données répertoriant toutes les cibles d'IPP impliquant des dimères d'hélices aux interfaces protéine-protéine a ouvert la voie à la conception et au développement de mimes de structures tertiaires et quaternaires comme candidats d'une nouvelle classe d'inhibiteurs d'IPP[3]. Parmi les structures tertiaires et quaternaires identifiées sous forme de dimère d'hélices, le motif hélice-coude-hélice (HCH ou *hairpin*) est le système le plus simple. Ce motif consiste en deux hélices α connectées par un coude contenant moins de sept résidus d'acides aminés (**Fig. 66a**). Le développement de dimères d'hélices non naturels pour effectuer la reconnaissance de surface de protéines est assez récent, les premiers résultats présents dans la littérature et décrits dans le Chapitre 1 sont très encourageants[4].

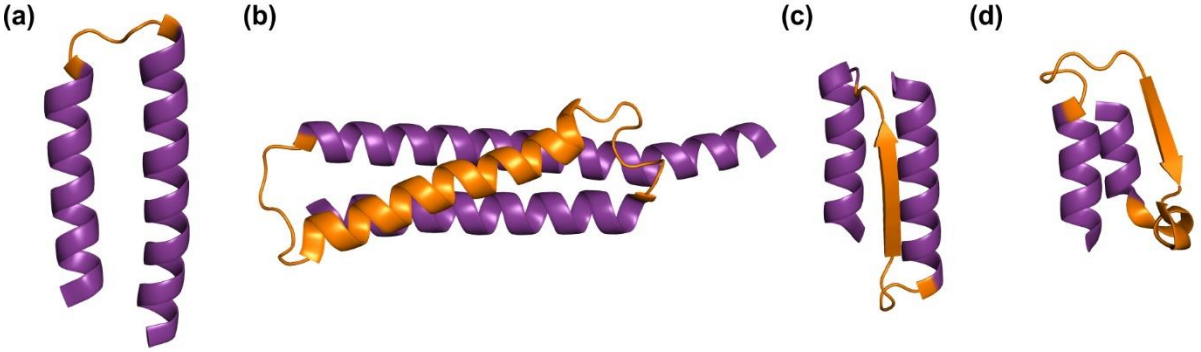


Figure 66 : Structures de dimères d'hélices α reliées par **(a)** un coude (structure en *hairpin*) (PDB :**1I4Y**), **(b)** hélice-boucle-hélice (PDB : **1OAH**), **(c)** hélice-feuillet-hélice (PDB : **2YCB**) et **(d)** hélice-boucle-feuillet-hélice-hélice (PDB : **2UYG**). L'ensemble des séquences reliant les deux hélices (en violet) est représenté en orange.

Au sein du laboratoire, nous souhaitons développer des foldamères d'oligourées comme peptidomimes pour cibler de larges surfaces d'interactions entre protéines. La compréhension du repliement de séquences oligourée sous forme d'hairpin hélicoïdal serait un premier pas dans cette direction, et c'est dans ce contexte que mon projet de thèse s'inscrit. La conception de séquences hairpin en solvant organique a permis de réaliser une étude séquence-structure et d'identifier un principe gouvernant le repliement de ces architectures artificielles en vue d'une application au développement de foldamères amphiphiles repliés en hairpin.

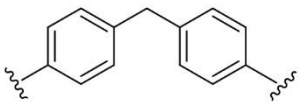
I. Conception et caractérisation des oligourées capable de se replier en structure hélice-coude-hélice en milieu organique.

Les foldamères d'oligourées développés au sein du laboratoire adoptent une structure hélicoïdale très stable en solution et à l'état solide. La conception de la structure en hélice-coude-hélice à base de foldamères s'est basée sur une approche ascendante dans laquelle deux séquences de foldamères d'oligourées connues pour adopter une structure hélicoïdale stable[5] sont connectées entre elles par un espaceur. La conception des premières séquences oligourées **OL1**, **OL2** et **OL3** (**Tableau 32**) capables de se replier en structure hélice-coude-hélice en milieu organique a été initiée au laboratoire avant le début de mon projet de thèse par deux post-doctorants, Yella-Reddy Nelli et Nagendar Pendem.

Ces oligourées hairpin **OL1**, **OL2** et **OL3** (**Fig. 67**) sont constituées de deux hélices d'oligourées identiques de huit et six résidus respectivement, connectées à leurs extrémités N-terminales par un espaceur aromatique et rigide dérivé du méthylène diphényl diisocyanate (MDI). Les composés oligourées **OL1** et **OL2** sont tous deux constitués de résidus urées aromatiques dérivés de l'acide aminé phénylalanine. La séquence de l'oligomère **OL3** est composée d'une portion de résidus urées aliphatiques dérivés de la valine, l'alanine et la leucine répétée deux fois.

oligourées		Espaceur	1	2	3	4	5	6	7	8	C-term
OL1	*Hélice 1*	MDI	F^U	F^U	F^U	F^U	F^U	F^U	F^U	F^U	NHMe
	Hélice 2		F^U	F^U	F^U	F^U	F^U	F^U	F^U	F^U	NHMe
OL2	*Hélice 1*	MDI	F^U	F^U	F^U	F^U	F^U	F^U			NHMe
	Hélice 2		F^U	F^U	F^U	F^U	F^U	F^U			NHMe
OL3	*Hélice 1*	MDI	V^U	A^U	L^U	V^U	A^U	L^U			NHMe
	Hélice 2		V^U	A^U	L^U	V^U	A^U	L^U			NHMe

Espaceur :



MDI

Tableau 32 : Séquences des oligourées hairpin **OL1**, **OL2** et **OL3** avec l'espaceur aromatique MDI

A l'état cristallin, les oligourées adoptent une structure repliée de type hairpin. Les hélices droite 2,5 des séquences oligourées sont bien définies, parallèles l'une à l'autre quelque soit la longueur et la nature des séquences d'oligourée et respectivement séparées par des distances de 11,9 Å, 11,0 Å et 13,9 Å entre les NH terminaux. Fort de ces premiers résultats, je me suis ensuite intéressée à l'effet de la nature de l'espaceur sur le repliement en structure tertiaire sous forme d'hairpin des séquences d'oligourées.

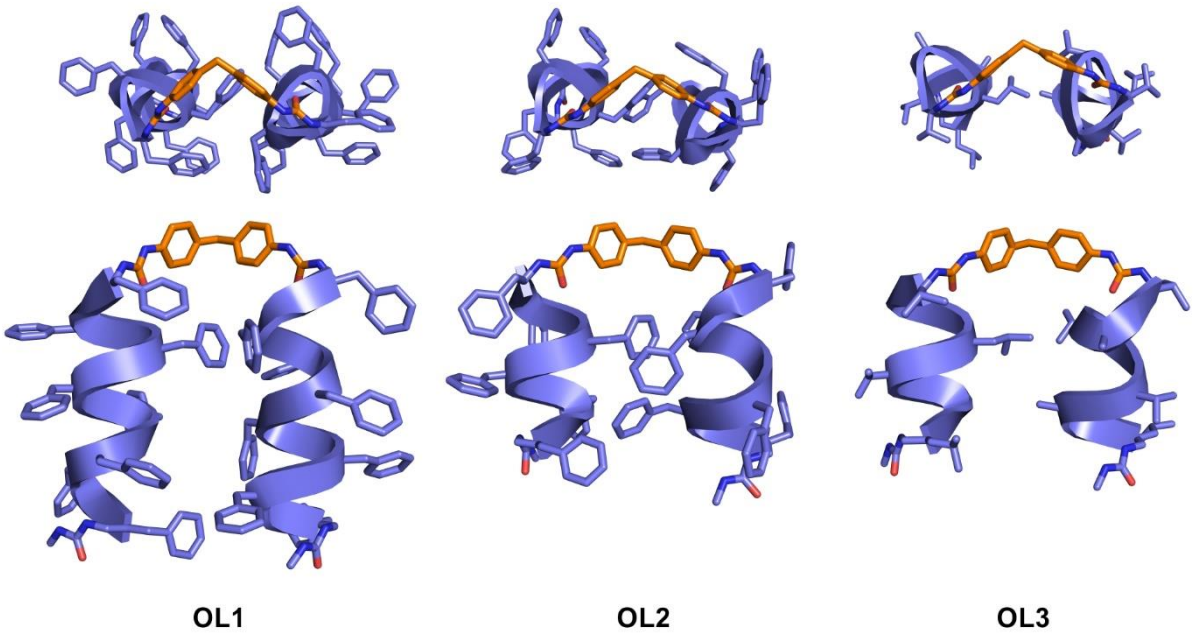


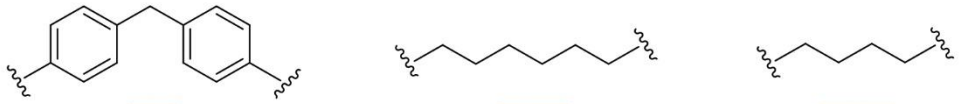
Figure 67 : Structures cristallines des oligourées **OL1**, **OL2** et **OL3** développées au laboratoire qui adoptent une structure repliée hélice-coude-hélice, aussi appelé harpin d'hélices. Les séquences d'oligourées sont représentées en bleu et l'espaceur MDI est en orange.

1. <u>Modification de la nature de l'espaceur d'une séquence d'oligourée modèle soluble en milieu organique</u>

Pour mieux comprendre les principes de repliement des structures tertiaires d'oligourées sous forme d'hairpin, l'espaceur rigide aromatique MDI a été modifié par des espaceurs aliphatiques plus flexibles, et a conduit à la conception de deux nouvelles séquences d'oligourées. L'oligourée **OL3** a été choisie comme séquence modèle pour cette étude, et un premier espaceur aliphatique à six atomes de carbone, l'hexaméthylène dérivée de l'hexaméthylene diisocyanate (HMDI), a été considéré et a conduit à la séquence **OL4**. Un espaceur plus court dérivé de la tétraméthylène diisocyanate (TtMDI) a conduit à la séquence **OL5** (**Tableau 33**).

oligourées		Espaceur	1	2	3	4	5	6	C-term
OL3	*Hélice 1*	MDI	V^U	A^U	L^U	V^U	A^U	L^U	NHMe
	Hélice 2		V^U	A^U	L^U	V^U	A^U	L^U	NHMe
OL4	*Hélice 1*	HMDI	V^U	A^U	L^U	V^U	A^U	L^U	NHMe
	Hélice 2		V^U	A^U	L^U	V^U	A^U	L^U	NHMe
OL5	*Hélice 1*	TtMDI	V^U	A^U	L^U	V^U	A^U	L^U	NHMe
	Hélice 2		V^U	A^U	L^U	V^U	A^U	L^U	NHMe

Espaceurs :

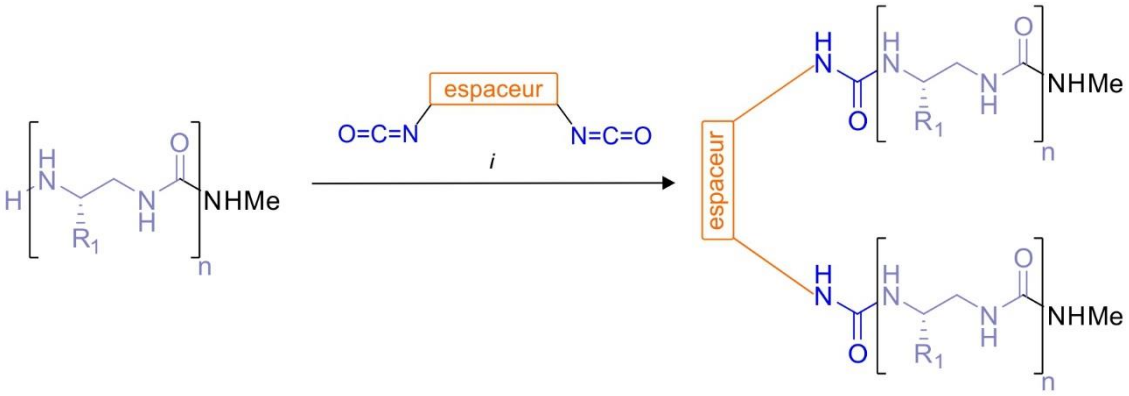


MDI HMDI TtMDI

Tableau 33 : Séquences des hairpins d'oligourées avec des espaceurs aliphatiques **OL4** et **OL5** dérivées de la séquence de l'oligourée **OL3**.

2. <u>Synthèse en solution des oligourées</u>

La synthèse des molécules de type **OL** a consisté en la réaction des extrémités N-terminales des séquences sur les fonctions diisocyanates de l'espaceur formant ainsi un lien urée entre ce dernier et les brins oligourées. La réaction de couplage avec l'espaceur se fait en présence d'une base et du précurseur de l'espaceur en défaut.



i : espaceur diisocyanate (0,5 equiv), DIPEA, 12 h, TA

Schéma 9 : Méthode générale de synthèse et dimérisation des oligourées **OL1**-**OL5** en solution

Les séquences d'hexaoligourée N-Boc protégées ainsi que le composé **OL4** ont été synthétisés par Yella-Reddy Nelli. **OL5** a été synthétisée au cours de ce projet de thèse et a été obtenue avec un rendement de couplage de 34%.

oligourées	séquences	MALDI-MS $[M+Na]^+$	rdt[a]
OL4	$L^uA^uV^uL^uA^uV^u$-HMDI-$V^uA^uL^uV^uA^uL^u$	1734,42	94%
OL5	$L^uA^uV^uL^uA^uV^u$-TtMDI-$V^uA^uL^uV^uA^uL^u$	1706,22	34%

Tableau 34 : Résultats de la dimérisation des hélices d'oligourée **OL4** et **OL5** en solution avec les espaceurs aliphatiques HMDI et TtmDI. [a] rendement de l'étape de dimérisation.

3. Caractérisation structurale des oligourées **OL4** et **OL5**

Au sein du laboratoire, la cristallisation des oligourées se fait classiquement en solvant organique. Des cristaux du composé **OL1** avaient été obtenus mais la présence de six molécules indépendantes dans l'unité asymétrique a compliqué la résolution et l'affinement de la structure pendant de nombreuses années malgré tous nos efforts et collaborations. La cristallisation racémique avait alors été envisagée pour résoudre cette problématique et a été appliquée pour l'ensemble des séquences d'oligourées **OL1** à **OL4** par Nagendar Pendem et Yella-Reddy Nelli, qui ont dû synthétiser au préalable les oligomères énantiomères **OL*ent*-1** - **OL*ent*-4** à partir d'acides aminés D.

La cristallisation racémique de l'oligourée **OL4** dans du méthanol a été réalisée par Yella-Reddy Nelli et l'affinement de la structure a été terminé au cours de mon projet de thèse. L'oligomère **OL5** a cristallisé seul dans du DMSO. Les structures cristallines des oligourées **OL4** et **OL5** ont été obtenues à haute résolution par diffraction des rayons X (DRX). La résolution des structures et l'affinement ont été réalisés par méthode directe avec les logiciels SHELXD et SHELXL par Brice Kauffmann, cristallographe au sein de l'Institut Européen de Chimie et Biologie (IECB, UMS 3033).

a. Cristallisation racémique comme outil de cristallisation et résolution de structure

La cristallisation racémique consiste à mélanger deux énantiomères dans un solvant de cristallisation pour obtenir des cristaux contenant les molécules de deux énantiomères du mélange racémique. La présence des deux énantiomères dans le mélange favorise la formation de centrosymétries via des miroirs et des centres d'inversion. La présence de ces éléments de symétries va ordonner les molécules du cristal, augmenter le nombre d'éléments de symétrie, réduire la taille de l'unité asymétrique et ainsi faciliter la cristallisation, la diffraction

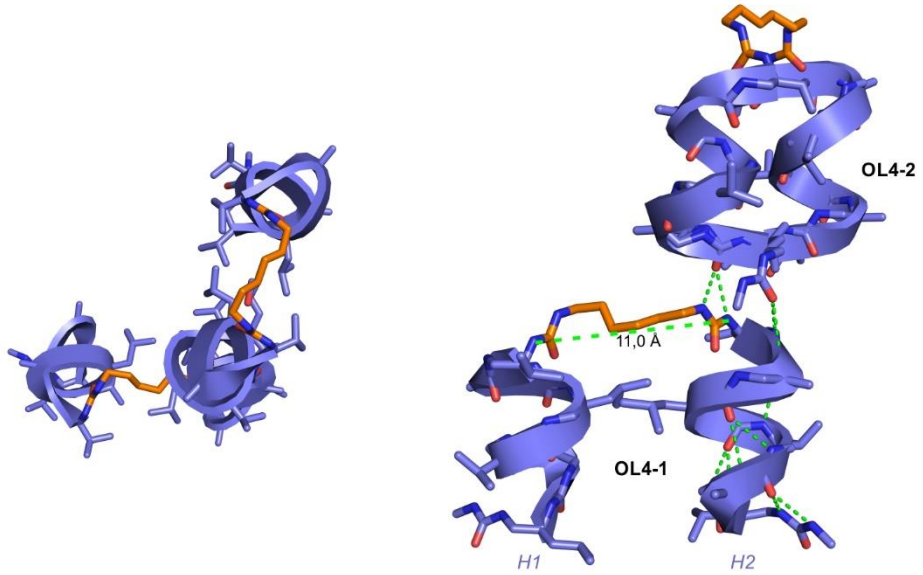
des cristaux et la résolution de la structure. La présence des éléments de centrosymétrie dans les cristaux obtenus par cristallisation racémique simplifie le phasage, car seules les amplitudes des phases à 0° et 180° sont mesurées à haute résolution. Dans le cas de cristaux d'un énantiomère, les combinaisons de phasage sont plus nombreuses[6] : de 0° à 360°. La résolution des structures obtenues par cristallisation racémique peut avoir lieu soit par remplacement moléculaire soit par méthode directe. La cristallisation racémique a d'abord été réalisée pour les petites molécules[7], puis adaptée et largement employée pour la cristallisation et résolution de structures des protéines[8] puis adaptées à la résolution d'acides nucléiques[9] et de foldamères[10]. Les inconvénients majeurs de cette méthode de cristallisation sont le coût des acides aminés D nécessaires à la synthèse de l'énantiomère et l'effort de synthèse requis.

*b. Structure cristalline de l'oligourée **OL4***

L'oligourée **OL4** a cristallisé dans un système cristallin monoclinique dans le groupe d'espace $P2_1/c$. L'unité asymétrique est ici composée de deux molécules indépendantes et parallèles (**Fig. 68a**). Chaque hélice d'oligourée est stabilisée par un réseau hydrogène à trois centres et la caractérisation de l'hélicité des angles dièdres (Φ, Ψ, Θ_1 et Θ_2) des résidus des hélices confirment la canonicité des hélices d'oligourée (**Fig. 68c**).

Une superposition des hélices dans chaque molécule indépendante (**OL4-1** et **OL4-2**), ainsi qu'une superposition des deux molécules indépendantes ont été réalisées avec les atomes de la chaîne principale des hélices d'oligourées (**OL4-H1** et **OL4-H2**). Les valeurs de RMSD calculées pour chaque molécule indépendante sont toutes inférieures ou égales à 1 Å, les hélices de chaque molécule indépendante sont donc similaires. Les valeurs de RMSD obtenues pour les résidus des deux hélices sont similaires exceptés ceux de l'espaceur et des résidus L_6^U aux extrémités C-terminales (**Fig. 68b**). Les structures cristallines des deux molécules de **OL4** sont similaires à celles observées pour les oligourées **OL1**-**OL3**. L'oligomère **OL4** se replie sous forme d'une structure en hélice-coude-hélice (hairpin). Les hélices droite de chaque molécule sont parallèles l'une à l'autre et séparées par 11 Å (**Fig. 68a**). La distance mesurée entre les hélices se fait entre les NH terminaux de chaque hélice. Les deux molécules de l'oligourée **OL4** contenues dans l'unité asymétrique interagissent ensemble via des liaisons hydrogène entre les C=O terminaux des résidus urée A_5^U et L_6^U d'une hélice et les NHs des amides des premiers résidus V_1^U et A_2^U d'une hélice de la seconde molécule. La structure cristalline du composé **OL4** avec un espaceur flexible aliphatique montre qu'il est possible d'obtenir une structure repliée sous forme d'hélice-coude-hélice similaire à celle observée pour les oligourées **OL1** – **OL3**.

(a)



(b)

oligourées	nature des résidus et position / RMSD (Å)							
	Espaceur	1	2	3	4	5	6	N-term
OL4-1	HMDI 0,77	V^U 0,30	A^U 0,20	L^U 0,17	V^U 0,19	A^U 0,16	L^U 0,22	NHMe 0,38
OL4-2	HMDI 0,80	V^U 0,48	A^U 1,19	L^U 1,11	V^U 1,58	A^U 0,81	L^U 1,10	NHMe 1,51
H1	HMDI 1,88	V^U 1,78	A^U 1,76	L^U 0,77	V^U 0,95	A^U 0,82	L^U 2,09	NHMe 2,81
H2		V^U 1,67	A^U 1,42	L^U 0,92	V^U 1,11	A^U 1,10	L^U 3,20	NHMe 4,02

(c)

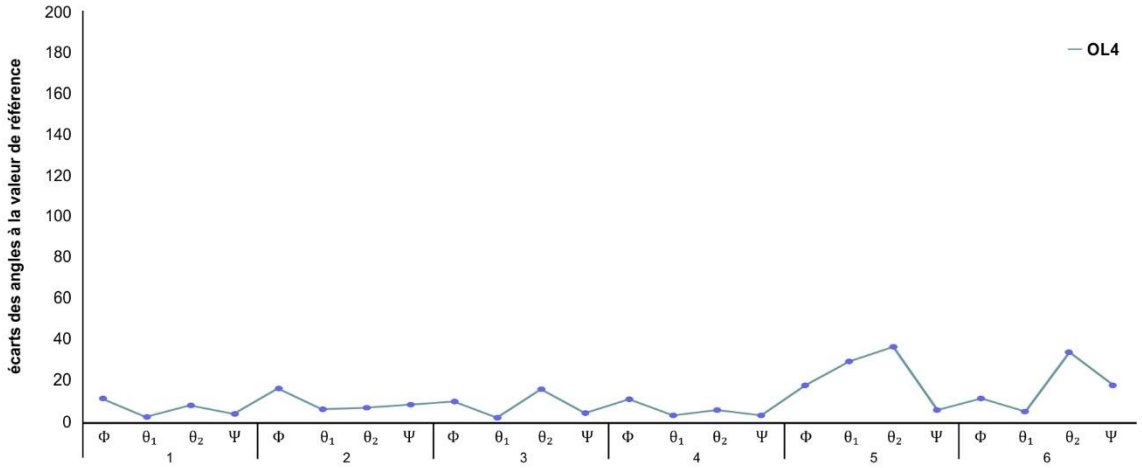
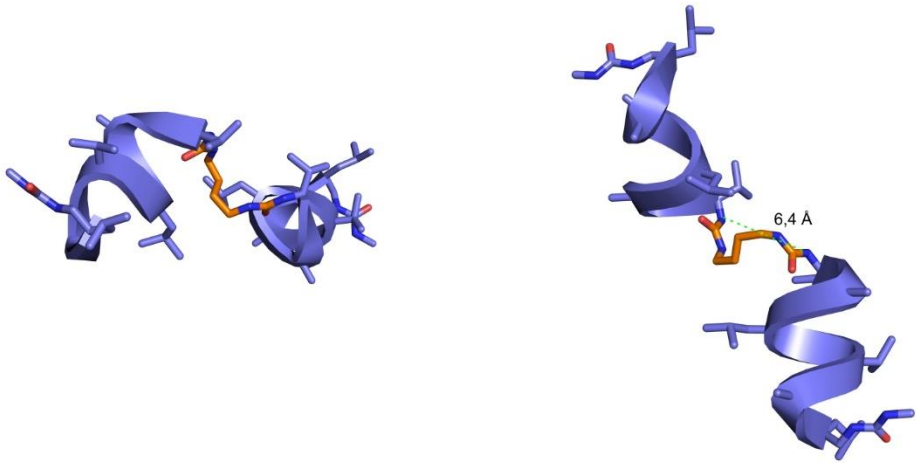


Figure 68 : **(a)** Structure cristalline de l'oligourée **OL4**. **(b)** Tableau récapitulatif des valeurs de RMSD calculées pour les deux molécules indépendantes (**OL4-1** et **OL4-2**) et les hélices (**H1** et **H2**) de l'oligomère **OL4. (c)** Ecart des angles dièdres à leur valeur de référence pour chaque résidu de **OL4**

c. *Structure cristalline de l'oligomère **OL5***

Des cristaux de l'oligomère **OL5** ont été obtenus dans du DMSO, et l'oligourée a cristallisé dans le groupe d'espace $P2_12_12_1$. L'unité asymétrique est composée d'une seule molécule indépendante. Chaque hélice d'oligourée de l'oligomère est stabilisée par un réseau hydrogène à trois centres.

(a)



(b)

oligourée	nature des résidus et position / RMSD (Å)							
	Espaceur	1	2	3	4	5	6	Nter
OL5	TtMDI 0,75	V^U 0,42	A^U 0,39	L^U 0,29	V^U 0,31	A^U 0,22	L^U 0,65	NHMe 1,24

(c)

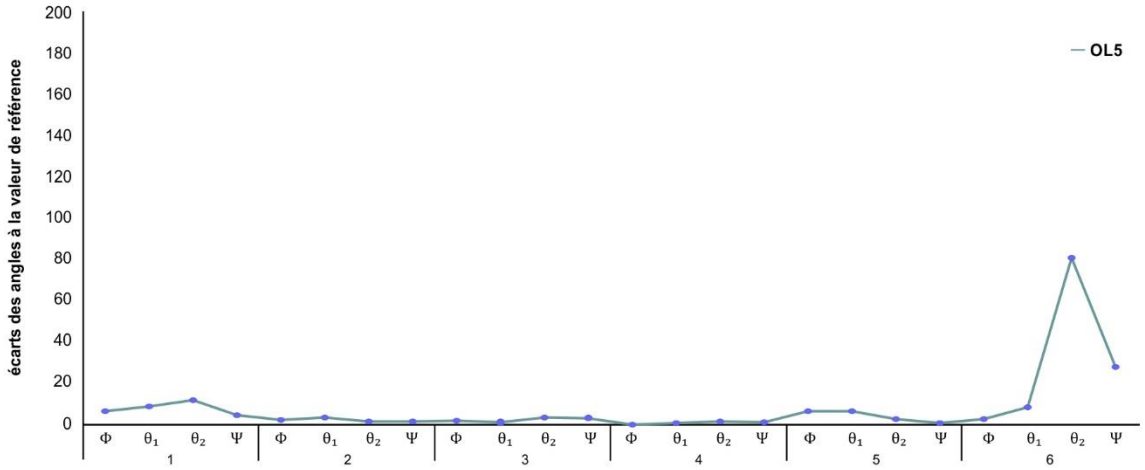


Figure 69 : **(a)** Structure cristalline de l'oligourée **OL5**. **(b)** Table récapitulative des valeurs de RMSD calculées par superposition des deux hélices de l'oligourée. **(c)** Ecart des angles dièdres à leurs valeurs de référence pour chaque résidu de l'oligomère **OL5**.

La structure cristalline obtenue montre que l'oligomère **OL5** adopte une structure « ouverte », différente de celles observées jusqu'à présent pour l'ensemble des oligourées (**Fig. 69**). Les hélices ne sont pas parallèles, orientées dans des sens opposés et séparées

par 6,4 Å. Cette distance n'est pas suffisante pour que les hélices soient côte à côte. La structure cristalline de l'oligourée **OL5** révèle qu'une modification de la longueur de l'espaceur aliphatique a un effet sur le repliement de la séquence d'oligourée modèle en une structure de type hélice-coude-hélice, « fermé », ou ouverte.

d. Détermination d'une règle de conception des foldamères d'oligourées pour former des structures de type Hélice coude hélice

L'ensemble des informations obtenues par l'étude des structures des oligourées **OL1** à **OL5** révèle que la nature du bras espaceur a un effet prédominant sur le repliement (**Fig. 70**). L'espaceur est le paramètre clé pour permettre la conception de foldamères d'oligourée capable de se replier en structure hélice coude hélice. Un espaceur aliphatique composé d'une chaîne à six atomes de carbone permet d'obtenir une structure repliée en hairpin similaire à celle obtenue avec un espaceur rigide aromatique. Cependant, une structure préliminaire de l'oligourée **OL4** qui a cristallisée seule lors de la cristallisation racémique, indique la présence d'une conformation « ouverte » de l'oligomère avec l'espaceur aliphatique HMDI. Un espaceur rigide aromatique permet donc de préorganiser au mieux les hélices dans une orientation parallèle sans tenir compte de la nature des chaînes latérales et du nombre de résidu par tour. Fort de ces résultats, l'objectif est maintenant de concevoir des oligourées hydrosolubles capables de se replier en structure HCH en milieu aqueux.

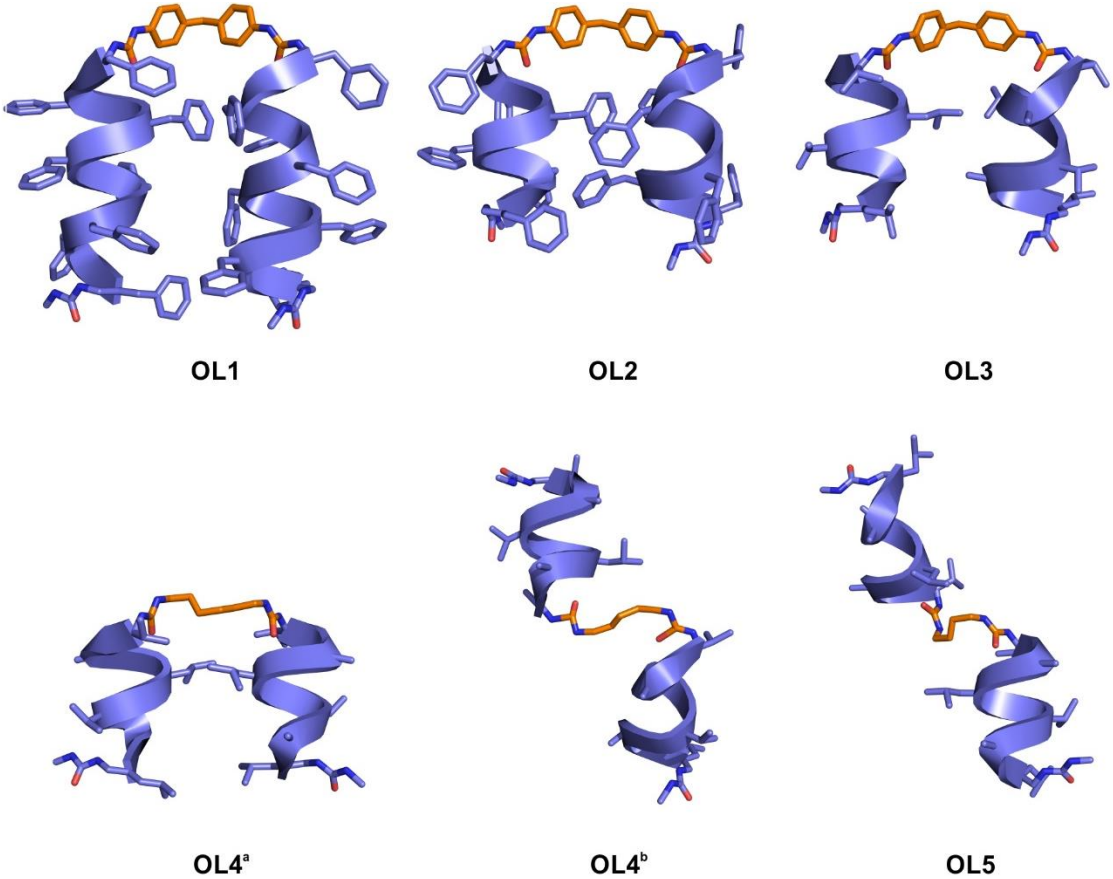


Figure 70 : Ensemble des structures cristallines obtenues pour les oligourées en solvant organique (DMF pour **OL1** et **OL2**, méthanol pour **OL3** et **OL4** et DMSO pour **OL5**).

II. Conception de structure repliée en hairpin à base de foldamères hydrosolubles

Le laboratoire a récemment développé une série de foldamères amphiphiles capables de former des assemblages quaternaires en condition aqueuse[11]. Deux topologies d'assemblages ont été obtenues : une capsule et un canal (**Fig. 71**). Ces assemblages formés par les séquences **OL6** et **OL7** sont dirigés par les chaînes latérales hydrophobes des résidus urées dérivés de la leucine (L^U). Chaque assemblage obtenu est constitué d'un cœur hydrophobe stabilisé de part et d'autre par des ponts salins formés par les résidus urées E^U et K^U. La répartition des chaînes latérales chargées/non-chargées des résidus urées permet de guider la topologie de l'assemblage. La structure cristallographique des hélices d'oligourées de l'oligomère **OL6** révèle un faisceau constitué de six hélices d'oligourées canoniques bien définies avec un cœur hydrophobe constitué de résidus urée leucine ainsi qu'une surface hydrophile chargée. Les hélices de l'oligourée **OL7** s'organisent dans le cristal pour créer un assemblage de type canal étendu avec un pore hydrophile chargé pourvu d'un diamètre interne de 17 Å.

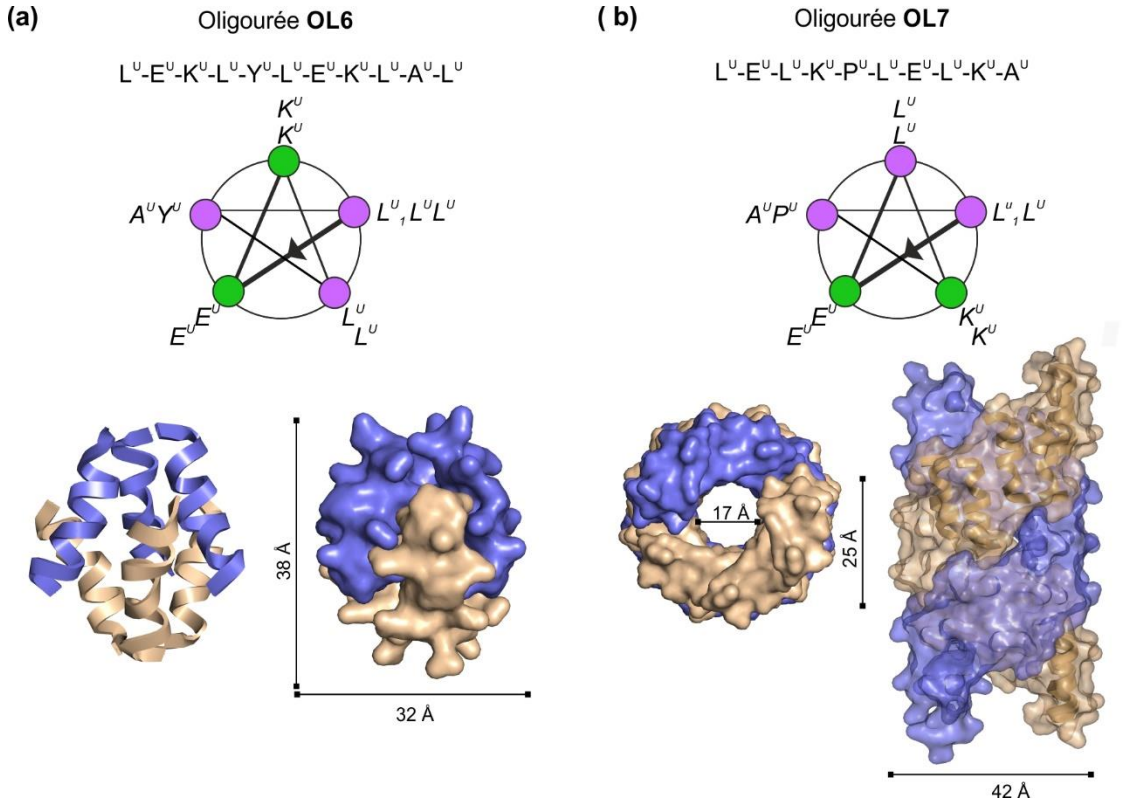


Figure 71 : Foldamères d'oligourée amphiphiles capables de former des assemblages quaternaires en solution aqueuse : **(a)** une capsule (**CCDC : 1030455**) avec le composé **OL6**, et **(b)** un canal (**CCDC : 1030456**) pour l'oligomère **OL7**. Les hélices sont colorées en bleu et en blanc selon leurs orientations dans le cristal, et les positions des chaînes latérales de séquence d'oligourée sont représentées dans les diagrammes en hélice où les résidus chargés sont en vert, et les résidus non chargés en violet.

1. Conception d'oligourées amphiphiles

Fort de ces connaissances, les séquences des résidus urées chargés/non chargés des oligourées amphiphiles **OL6**-**OL7** ont été conservées pour concevoir des foldamères hydrosolubles structurés en hairpin. La conception repose sur l'étude des structures cristallographiques obtenues en milieu organique. Le bras espaceur est, en effet, le paramètre permettant de contrôler l'orientation des séquences d'oligourées en hélice parallèles repliées en structure tertiaire HCH. Afin de concevoir une première séquence amphiphile à base de foldamère d'oligourée adoptant une structure HCH en milieu aqueux, **OL8**, nous avons choisi comme base de travail la structure cristallographique de la molécule **OL3** et la séquence de l'oligomère **OL6**. Cet oligomère est connu pour réaliser un assemblage quaternaire spontané en milieu aqueux, il a donc été supposé que ce phénomène d'auto-assemblage se produirait avec les oligourées sous forme de dimère d'hélice.

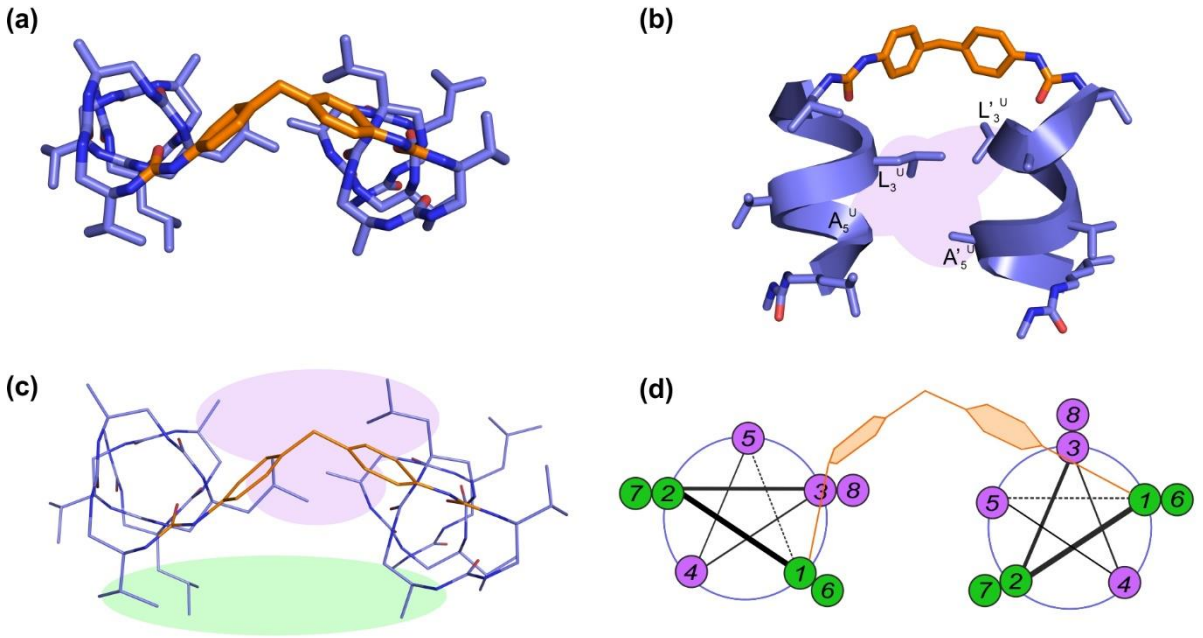


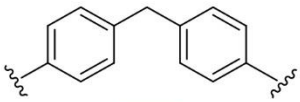
Figure 72 : **(a)** Vue axiale de l'oligourée **OL3** utilisée pour la conception du hairpin et **(b)** identification de la zone d'interaction (en violet) entre les résidus A^U et L^U des deux hélices de **OL3**. **(c)** Identification des zones d'interactions possibles entre les résidus de chaque hélice. Les zones en violet correspondent au cœur hydrophobe et la zone verte correspond aux interactions électrostatiques. **(d)** Diagramme en hélice représentant les positions des résidus des deux hélices reliées par un espaceur. Les positions en violet correspondent aux résidus nécessaires pour réaliser des effets hydrophobes, celles en vert correspondent aux résidus hydrophiles.

L'observation de la structure de l'oligomère **OL3** indique que des interactions inter hélices seraient susceptibles d'avoir lieu entre les résidus aux positions 3 et 5 de chaque hélice (**Fig. 72c**). Nous avons décidé de concevoir des hélices composées de huit résidus oligourée

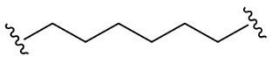
afin de renforcer le cœur hydrophobe pour stabiliser le repliement de la structure. La formation d'un cœur hydrophobe entre les hélices au niveau des résidus 3, 5 et 8 est réalisée via le positionnement des résidus urées non chargés. Des résidus hydrophobes dérivés de leucine sont donc positionnés en 3, 5 et 8 pour former le cœur hydrophobe, des résidus chargées en position 1, 2, 6 et 7 pour la solubilité et un dérivé de la tyrosine est en position 4. La structure du composé **OL3** suggère que les chaînes latérales du résidu situé à cette position de l'hélice n'effectuent pas d'interactions, et des travaux menés au sein du laboratoire par Gavin Collie sur la résolution de structure d'oligourée par iodation des résidus tyrosyle, ont motivé le choix de positionner ce dérivé urée en position 4 comme outil pour résoudre la structure[20]. L'ensemble de ces observations a conduit à la conception de l'oligourée **OL8** avec MDI comme espaceur et une seconde séquence **OL9** a été conçue avec l'espaceur plus flexible HMDI (**Tableau 35**). Afin d'étudier l'effet de l'espaceur sur le phénomène d'auto-assemblage, trois séquences « contrôle » correspondant à un brin des dimères d'hélices **OL8** et **OL9** ont été imaginés. Une première synthèse de la portion en hélice d'oligourée seule sur support solide a été envisagée avant de réaliser l'étape de dimérisation, l'objectif étant de pouvoir identifier les étapes où des problèmes de synthèses des oligourées **OL8** et **OL9** seraient rencontrés. L'oligomère **OL10** correspondant à cette portion oligourée a donc été imaginée et l'introduction de cap tolyle et propyle à l'extrémité N-terminale de l'oligomère **OL10** a conduit aux séquences **OL11** et **OL12** (**Tableau 35**).

oligourées		Espaceur/Cap	1	2	3	4	5	6	7	8
OL8	*Hélice 1*	MDI	E^U	K^U	L^U	Y^U	L^U	E^U	K^U	L^U
	Hélice 2		E^U	K^U	L^U	Y^U	L^U	E^U	K^U	L^U
OL9	*Hélice 1*	HMDI	E^U	K^U	L^U	Y^U	L^U	E^U	K^U	L^U
	Hélice 2		E^U	K^U	L^U	Y^U	L^U	E^U	K^U	L^U
OL10			E^U	K^U	L^U	Y^U	L^U	E^U	K^U	L^U
OL11		*p*-tolyl	E^U	K^U	L^U	Y^U	L^U	E^U	K^U	L^U
OL12		Pr	E^U	K^U	L^U	Y^U	L^U	E^U	K^U	L^U

Espaceurs :



MDI

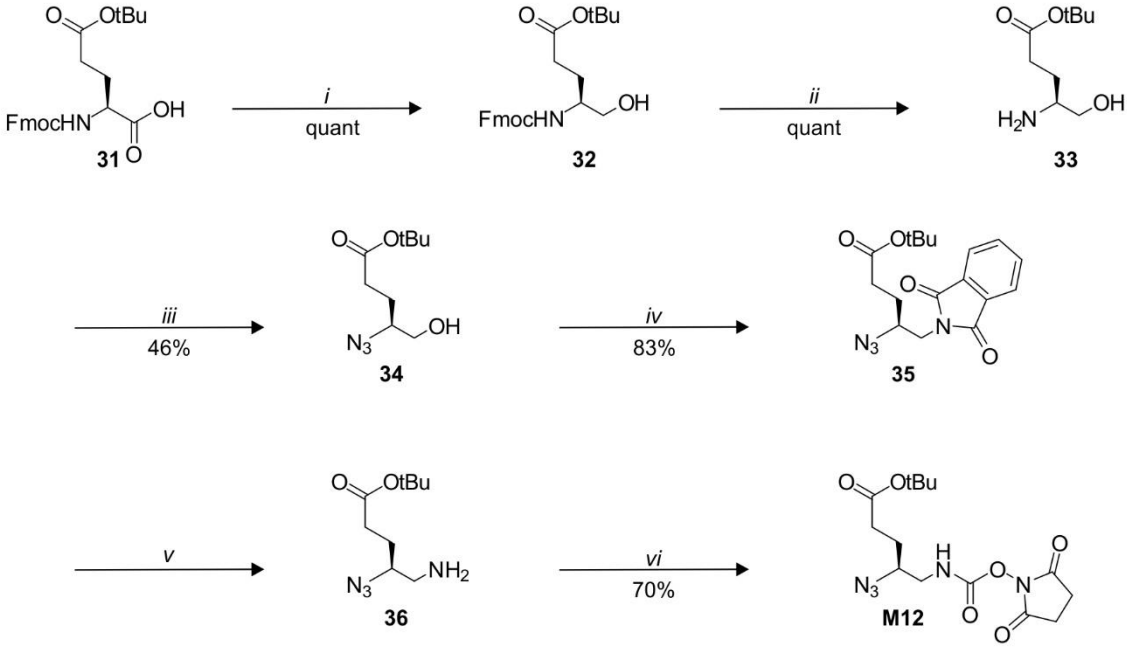


HMDI

Tableau 35 : Séquences des hairpins d'oligourées hydrosolubles **OL8** et **OL9** et des séquences « contrôle » d'hélices d'oligourée hydrosolubles **OL10-OL12.**

2. Préparation du monomère activé dérivé de l'acide glutamique **M12**

L'ensemble des séquences d'oligourée requiert un résidu urée dérivé de l'acide glutamique dont le monomère activé sous forme d'azoture carbamate de succinimide doit être préparé. Le monomère activé **M12** a été obtenu selon la voie 2 de synthèse décrite dans le **Schéma 10** à partir de l'acide aminé commercial Fmoc-L-Glu(OtBu)-OH **31**. Le monomère **M12** a été obtenu avec un rendement global de 27% ce qui en accord avec les rendements globaux des monomères déjà décrits.



i : 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii* : DBU, THF, TA ; *iii* : $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4.5H_2O$, CH_3CN, H_2O, TA ; *iv* : DIAD, PPh_3, phtalimide, THF, TA ; *v* : N_2H_4, MeOH, TA ; *vi* : DSC, DCM, TA

Schéma 10 : Schéma de synthèse du monomère **M12**

3. Synthèse sur support solide

a. Dimérisation sur support solide

La synthèse sur support solide de foldamères d'oligourée hydrosolubles est bien connue au laboratoire. Afin de diminuer le nombre d'étapes de purification et de gagner en temps de synthèse nous avons choisi de développer une stratégie de dimérisation sur support solide compatible avec la stratégie de synthèse des oligomères d'urée.

Le premier exemple de dimérisation sur support solide de peptide a été décrit en 2008 par Kodadek *et al.* qui se sont intéressés à l'introduction d'un lien covalent entre deux séquences de peptoïdes[13]. Leur approche de dimérisation sur support solide consiste à coupler un premier résidu lysine sur la résine. Les fonctions amine de ce résidu sont libres afin

de permettre les couplages séquentiels simultané d'unités aminés et peptoïdes sur chacune de ces fonctions. Cette approche est originale car elle permet de générer un dimère dont les séquences sont reliées par leurs extrémités C-terminales. La dimérisation des deux séquences peptoïdes antagonistes à VEGFR2 a conduit à une amélioration de l'affinité d'un facteur 100 par rapport aux affinités obtenues pour les deux séquences de peptoïdes seules. Cette méthodologie de dimérisation par les extrémités C-terminales (**Fig. 73a**) a été appliquée par Hooks *et al.* pour le développement d'une mini-librairie de séquences mixtes peptides/peptoïdes spécifiques au cancer des poumons sous forme de dimères et de trimères[14].

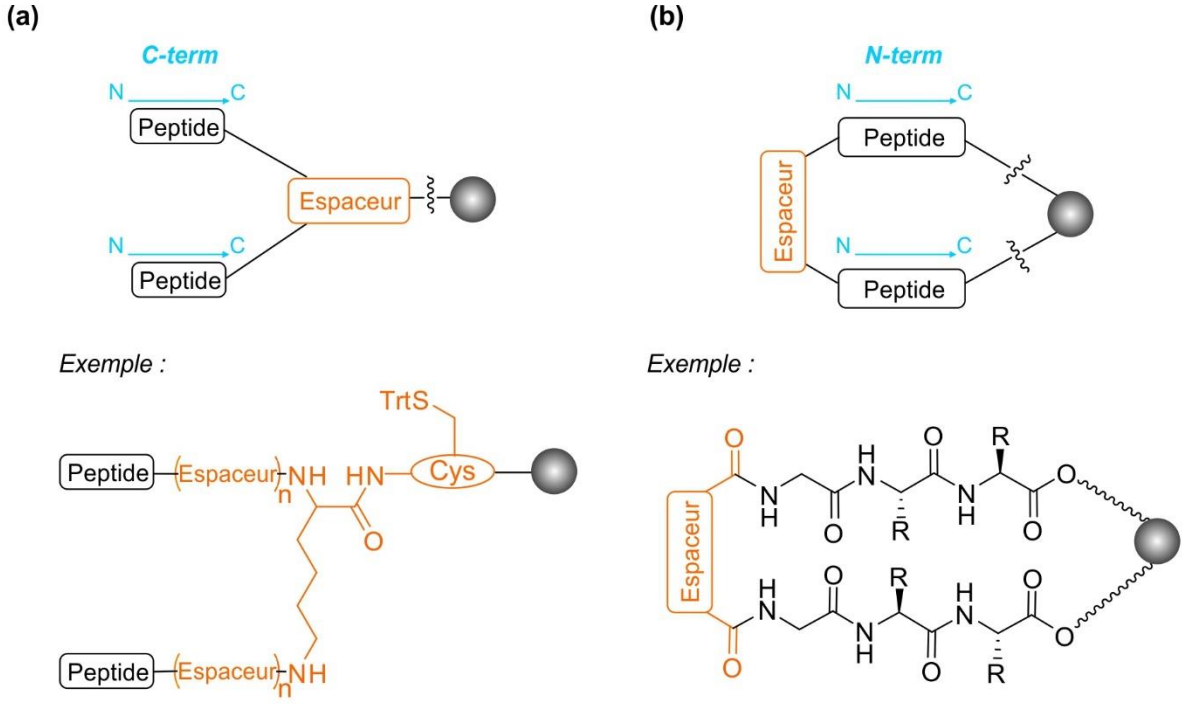
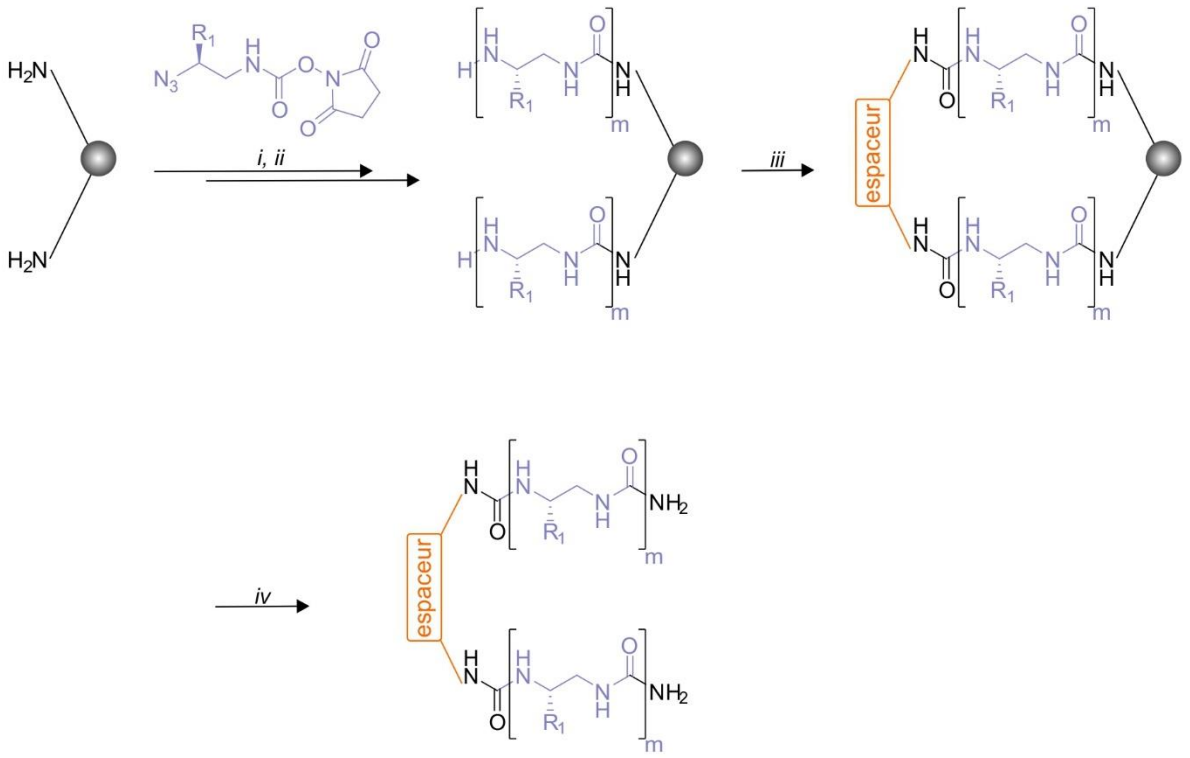


Figure 73 : Approches de dimérisation. Dimérisation par **(a)** l'extrémité C-terminale[14] et **(b)** N-terminale[15].

Une seconde approche de dimérisation sur support solide a été proposée en 2010 avec l'équipe de John Tovar[15]. Leur stratégie de dimérisation repose sur la connexion, par un composé diélectrophile, des extrémités N-terminales de deux séquences peptidiques immobilisées sur la résine (**Fig. 73b**). Les oligopeptides sont synthétisés sur support solide en stratégie Fmoc. Le réactif diélectrophile utilisé est ajouté en léger défaut par rapport au chargement de la résine, pour effectuer la dimérisation sur support solide et le dimère est obtenu après clivage de la résine.

*b. Protocole de dimérisation sur support solide des oligourées **OL8** et **OL9***

Le couplage des hélices d'oligourée aux espaceurs MDI et HMDI a été réalisé sur support solide en s'inspirant de l'approche de dimérisation sur support solide par Tovar *et al.*[15].



i : monomère azoture protégé activé **MX** (3 equiv), DIPEA (3 equiv), (50 W, 50°C, 20 min) ; *ii* : PMe_3 (1M) dans le THF (10 equiv), dioxane/H_2O (7:3), (50 W, 50°C, 15 min) ; iii : 1) espaceur sous forme de diisocyanate (0,3 equiv), DIPEA (3 equiv), DCM, TA, 24 h , 2) dioxane/H_2O (7:3), TA, 2 h ; i*v* : TFA/TIS/H_2O (95:2,5:2,5).

Schéma 11 : Schéma de synthèse sur support solide des dimères **OL8** et **OL9**.

La stratégie de dimérisation sur support solide des oligourées hydrosolubles est basée sur la synthèse, classiquement utilisée au laboratoire, d'oligourée sur support solide avec une résine pourvu d'un loading suffisamment important pour pouvoir avoir les hélices proches les unes des autres et ainsi promouvoir la dimérisation. La dimérisation sur support solide a été réalisée dans les mêmes conditions que la dimérisation en solution décrite précédemment dans ce chapitre. La réaction a été réalisée à température ambiante durant 24h sous agitation mécanique, la résine a ensuite été rincée avec du dichlorométhane, puis un test chloranil a été réalisé pour vérifier le bon fonctionnement du couplage. Un mélange de dioxane/eau à 30% a ensuite été ajouté à la résine pour quencher la réaction et laissé à réagir pendant deux heures à température ambiante sous agitation mécanique pour piéger les fonctions isocyanate qui n'auraient pas réagis. Les billes de résine sont ensuite rincées plusieurs fois avant l'étape

de clivage de l'oligourée de la résine réalisé avec la solution de clivage TFA/TIS/H_2O (95:2,5:2,5).

c. Résultat des synthèses sur support solide

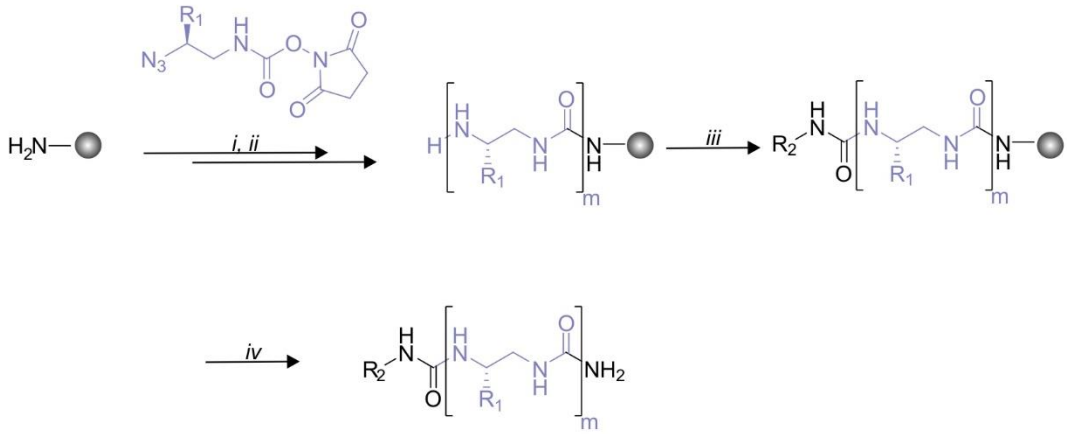
La synthèse de la séquence d'oligourée **OL10** a été réalisée en premier pour contrôler le bon déroulement de la synthèse sur support solide de la portion oligourée sur une échelle de 75 µmol. Les segments en hélice d'oligourées des oligomères **OL8**, **OL9**, **OL11** et **OL12** ont été synthétisés en même temps sur une échelle de 100 µmol. La résine a été partagée en quatre pour réaliser chaque étape de capping N-terminal des brins d'hélices (**OL11** et **OL12**) et de dimérisation sur support solide (**OL8** et **OL9**) sur une échelle de 25 µmol. La synthèse des composés **OL8**-**OL12** a été réalisée sur une résine polystyrène Rink Amide MBHA (loading de 0,87 mmol/g).

i. Synthèse des séquences « contrôle » **OL10**, **OL11** et **OL12**

La synthèse de l'oligourée **OL10** a été réalisée suivant la procédure décrite dans le Chapitre 2, et les oligourées **OL11** et **OL12** ont été synthétisés selon le protocole décrit au paragraphe *b* (**Schéma 12**). Les synthèses n'ont posé aucun problème particulier. Les chromatogramme RP-HPLC des chimères **OL10** (**Fig. 74a**), **OL11** (**Fig. 75a**) et **OL12** avant purification par RP-HPLC préparative indique la présence des produit largements majoritaires avec une pureté de 88%, 68% et 56% aux temps de rétentions de 5,14 min, 6,05 min et 5,89 min. Après purification par HPLC préparative en phase inverse **OL10**, **OL11** et **OL12** ont été obtenus avec des rendements globaux respectifs de 43%, 30% et 26%.

oligourées	séquences	pureté brut réactionnel	ESI-MS $[M+H]^+$	HPLC t_R (min)[a]	rdt[b]
OL10	$E^UK^UL^UY^UL^UE^UK^UL^U$	88%	1266,59	5,14	43%
OL11	*p-tolyl*-$E^UK^UL^UY^UL^UE^UK^UL^U$	68%	1399,74	6,05	30%
OL12	*Pr*-$E^UK^UL^UY^UL^UE^UK^UL^U$	56%	1351,73	5,89	26%

Tableau 36 : Séquences, temps de rétention HPLC et rendements de synthèses des chimères **OL10-OL12**. [a] système RP-HPLC : gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement après purification HPLC.



i : monomère azoture protégé activé **MX** (3 equiv), DIPEA (3 equiv), (50 W, 50°C, 20 min) ; *ii* : PMe_3 (1M) dans le THF (10 equiv), dioxane/H_2O (7:3), (50 W, 50°C, 15 min) ; iii : 1) « cap » sous forme d'iocyanate (1 equiv), DIPEA (3 equiv), DCM, TA, 24 h , 2) dioxane/H_2O (7:3), TA, 2 h ; i*v* : TFA/TIS/H_2O (95:2,5:2,5).

Schéma 12 : Schéma de synthèse sur support solide des oligomères **OL11** et **OL12**.

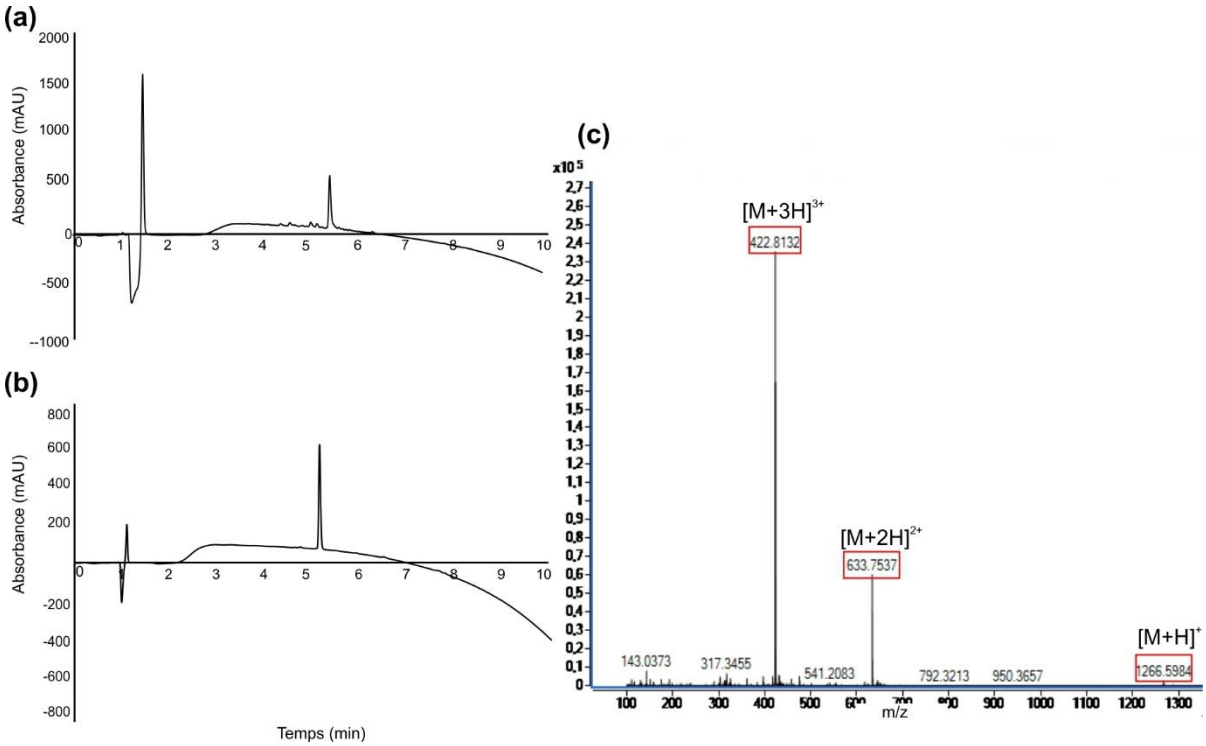


Figure 74 : Caractérisation de l'oligourée **OL10** : chromatogramme HPLC de **OL10 (a)** avant purification et **(b)** après purification par RP-HPLC préparative sur colonne C18; **(c)** spectre de masse de **OL10** après purification.

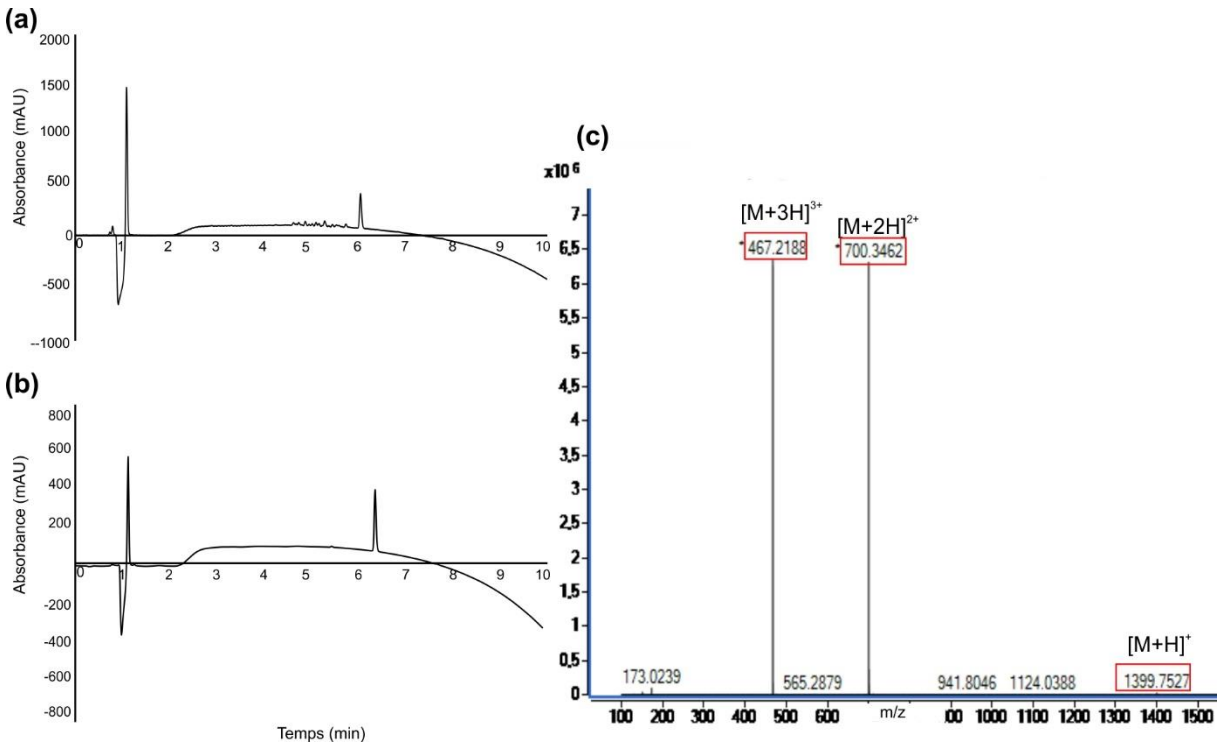


Figure 75 : Caractérisation de l'oligourée **OL11** : chromatogramme HPLC de **OL11 (a)** avant purification et **(b)** chromatogramme HPLC de **OL11** après purification par RP-HPLC préparative sur colonne C18 ; **(c)** spectre de masse de **OL11** après purification.

ii. Synthèse des dimères d'hélices **OL8** et **OL9**

La synthèse des dimères d'hélices d'oligourées **OL8** et **OL9** a été réalisée sur support solide suivant la procédure précédemment décrite. La synthèse des oligourées ne pose aucun problème particulier. Les chromatogrammes RP-HPLC avant purification des oligourées **OL8** et **OL9** révèlent la présence d'un produit largement majoritaire avec une pureté de 71% et 51% aux temps de rétention de 6,05 min et 7,21 min. Des analyses par spectrométrie de masse des bruts de réaction confirment la présence des dimères d'hélices **OL8** et **OL9** qui ont été obtenus, après purification sur HPLC préparative, avec des rendements globaux de 29% et 15%.

oligourées	séquences	pureté brut réactionnel	ESI-MS $[M+2H]^{2+}$	HPLC t_R (min)[a]	rdt[b]
OL8	$L^UK^UE^UL^UY^UL^UK^UE^U$-MDI-$E^UK^UL^UY^UL^UE^UK^UL^U$	71%	1392,22	6,05	29%
OL9	$L^UK^UE^UL^UY^UL^UK^UE^U$-HMDI-$E^UK^UL^UY^UL^UE^UK^UL^U$	51%	1351,23	7,21	15%

Tableau 37 : Séquences, temps de rétention HPLC et rendements de synthèses des oligourées **OL8-OL9**. [a] système RP-HPLC : gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement après purification HPLC.

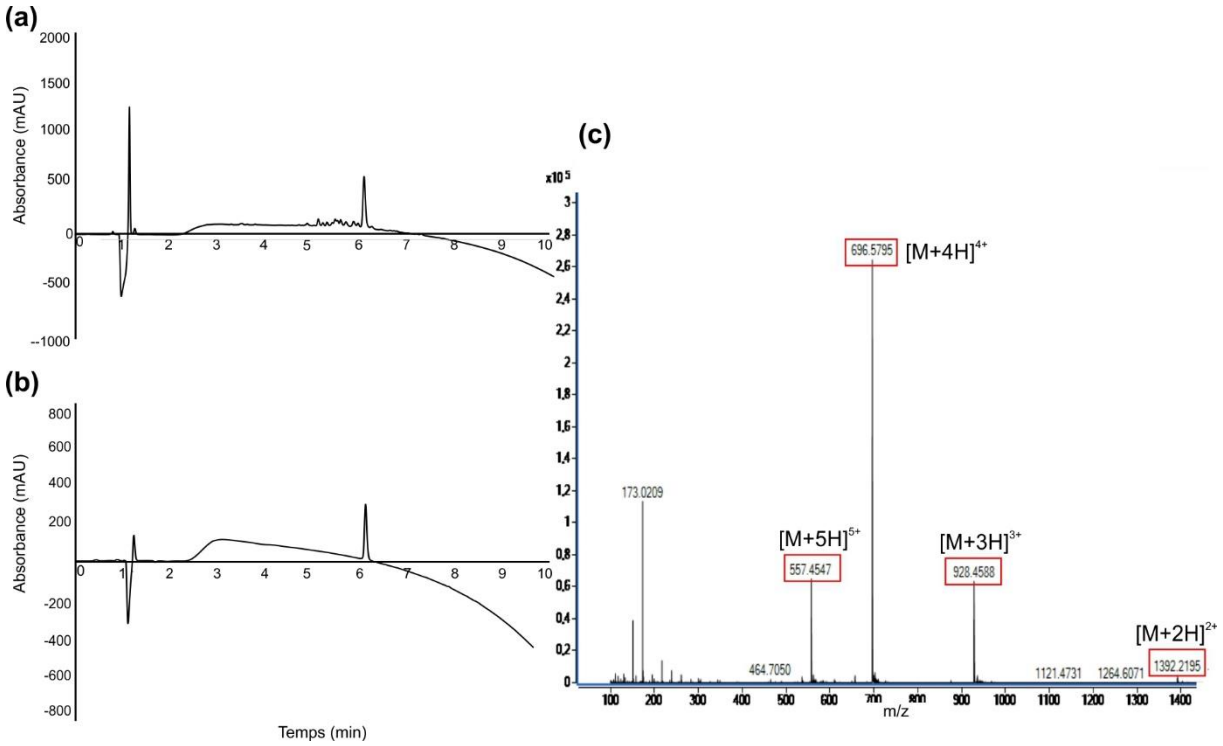


Figure 76 : Caractérisation de l'oligourée **OL8** : chromatogramme HPLC de **OL8 (a)** avant purification et **(b)** chromatogramme HPLC de **OL8** après purification par RP-HPLC préparative sur colonne C18 ; **(c)** spectre de masse de **OL8** après purification.

Le bon fonctionnement de ces réactions de dimérisation réalisées pour la première fois a été très surprenant car aucune optimisation des conditions de réaction n'a été nécessaire pour améliorer la pureté des bruts réactionnels. Cette méthode de synthèse sur support solide est une approche innovante pour le développement de la dimérisation.

4. Caractérisation structurale des oligourées hydrosolubles

a. Caractérisation à l'état solide par diffraction des rayons X

La cristallisation des oligourées **OL8**-**OL12** en condition aqueuse par diffusion en goutte suspendue a été réalisée par mes soins, et 96 conditions de cristallisation ont été criblées par oligourée. Un phénomène de précipitation a été observée à de nombreuses reprises pour les composés **OL8** et **OL9** lors de la cristallogenèse, et certains de ces précipités ont évolué vers la formation de cristaux. Des problèmes de diffraction et de phasage n'ont pas permis d'obtenir de structures cristallines pour ces deux oligourées. Des cristaux du composé **OL9** ont été finalement obtenus dans les conditions suivantes : 30% méthyl-2,4-pentanediol, 100 mM acétate de sodium, 200 mM de chlorure de calcium, pH 4,6, et en quantité suffisante pour permettre une collecte au synchrotron. Bien qu'un jeu de données du dimère **OL9** ait été obtenu par DRX, la résolution par méthode directe n'a pas pu aboutir à l'obtention d'une structure cristalline.

Récemment, des petits cristaux de qualité suffisante pour permettre une collecte des données du brin d'hélice **OL12** à l'IECB ont été obtenus dans les conditions suivantes : 28% PEG 400, 100 mM HEPES sels de sodium, 200 mM chlorure de calcium, pH 7,5. La structure cristalline de l'oligourée a été résolue et affinée à haute résolution par méthode direct avec les logiciels SHELXD et SHELXL par Brice Kauffmann. Le composé **OL12** correspond à un brin d'hélice du dimère **OL9** pour lequel un jeu de données cristallographiques était disponible. Une résolution de la structure de ce dimère par remplacement moléculaire avec la structure du brin d'hélice **OL12** comme modèle a été réalisée et s'est révélée infructueuse. A ce jour, seule la structure cristalline de l'oligourée **OL12** est disponible.

L'oligourée **OL12** cristallise dans une maille monoclinique centré avec un groupe d'espace C2. Une seule molécule indépendante de l'oligourée **OL12** est présente par unité asymétrique (**Fig. 77a**). La molécule adopte une structure en hélice d'oligourée canonique stable. L'hélice amphiphile présente une face hydrophobe constituée de trois résidus urées portant des chaînes latérales isobutyle ainsi que d'une face hydrophile avec les résidus urées chargés. La structure cristallographique de l'oligourée **OL12** révèle la présence de quatre molécules dans la maille (**Fig. 77b**). Les interactions entre deux hélices antiparallèles le long de l'axe transversal s'articulent autour d'un cœur hydrophobe entre les chaînes latérales des résidus urées de la leucine. La présence de molécules d'eau permet de relier les deux hélices parallèles le long de l'axe longitudinal entre elles.

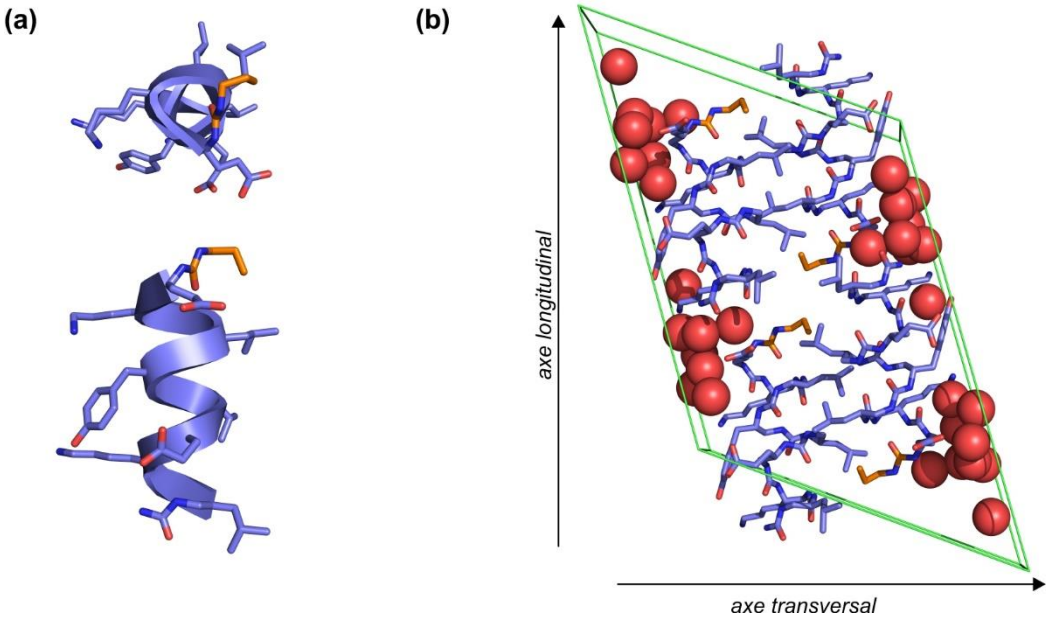


Figure 77 : Structure cristalline de l'oligourée **OL12 (a)** dans l'unité asymétrique et **(b)** dans la maille. Les boules rouges correspondent aux molécules d'eau présentes dans la maille.

Bien que la séquence des résidus d'oligourée du composé **OL12** soit inspirée des séquences **OL6** et **OL7**, les systèmes cristallins et les groupes d'espace de ces oligourées sont différents[11]. Une comparaison des empilements cristallins n'est pas possible entre celui de l'oligourée **OL12** à ceux des oligourées **OL6** et **OL7** qui s'auto-assemblent pour former des capsules ou et des canaux en milieux aqueux.

b. Caractérisation en solution par dichroïsme circulaire

Une caractérisation conformationnelle en solution par dichroïsme circulaire a été réalisée pour les oligourées **OL8**-**OL12** en parallèle des tentatives de caractérisation structurale. L'objectif de cette étude a été de voir si des phénomènes de structuration quaternaire semblable aux assemblages des oligourées amphiphiles **OL6** étaient obtenus. Les phénomènes d'assemblage et de structuration quaternaire sont caractérisés par CD via une augmentation de l'intensité de l'ellipticité molaire en fonction de la concentration[11]. Ainsi l'absence de relation de dépendance entre l'ellipticité molaire et la concentration indique qu'aucun phénomène de structuration tertiaire ou quaternaire n'a lieu.

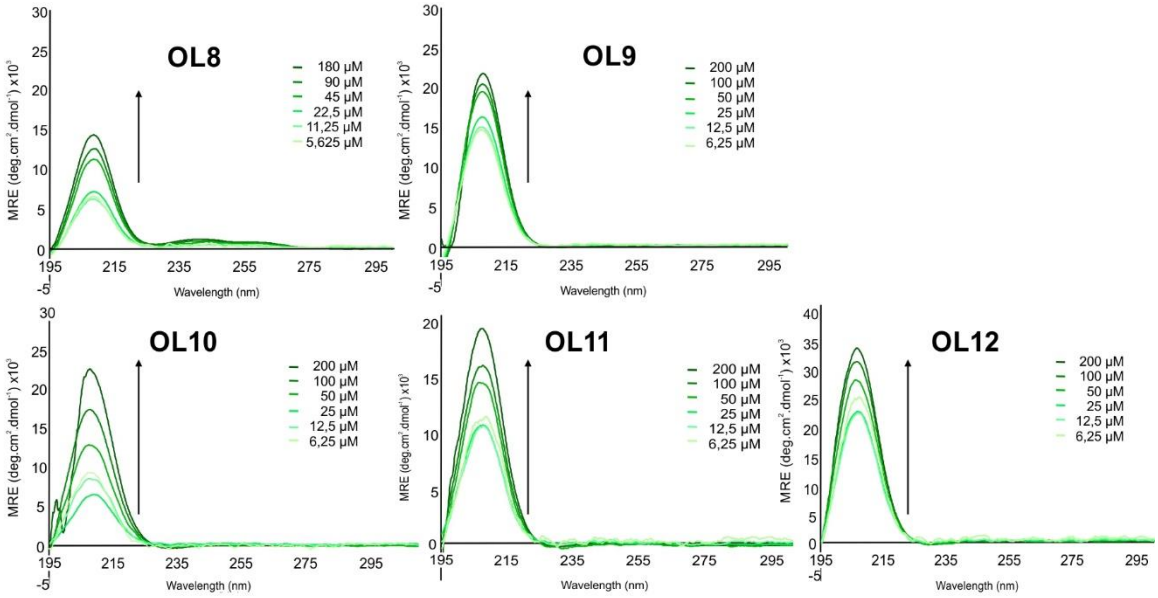


Figure 78 : Spectres CD en concentration des oligourées **OL8-OL12** en solution aqueuse non-tamponnée. Le sens de la flèche indique l'effet de l'augmentation de la concentration sur l'ellipticité.

Les premières mesures réalisées dans une solution aqueuse non-tamponnée semblent indiquer une dépendance de l'ellipicité molaire en fonction de la concentration, ce qui suggère la présence d'un phénomène de structuration tertiaire ou quaternaire (**Fig. 78**). Une étude en concentration de ces assemblages en condition aqueuse dans une solution tamponnée de sodium phosphate à 20 mM à différents pH (5,8 et 8,2) a été réalisée (**Fig. 79**). L'ensemble des résultats obtenus en condition acide (pH 5,8) et basique (pH 8,2) semble indiquer la présence de phénomènes d'assemblage pour les oligourées **OL12** à pH 5,8 et pour **OL10** à pH 8,2 où les valeurs d'ellipticités molaires augmentent en fonction de la concentration. Le spectre CD du dimère **OL8** obtenu en condition basique révèle la présence d'une relation différente et inexpliquée entre la concentration et l'ellipticité molaire : celle-ci diminue lorsque la concentration augmente.

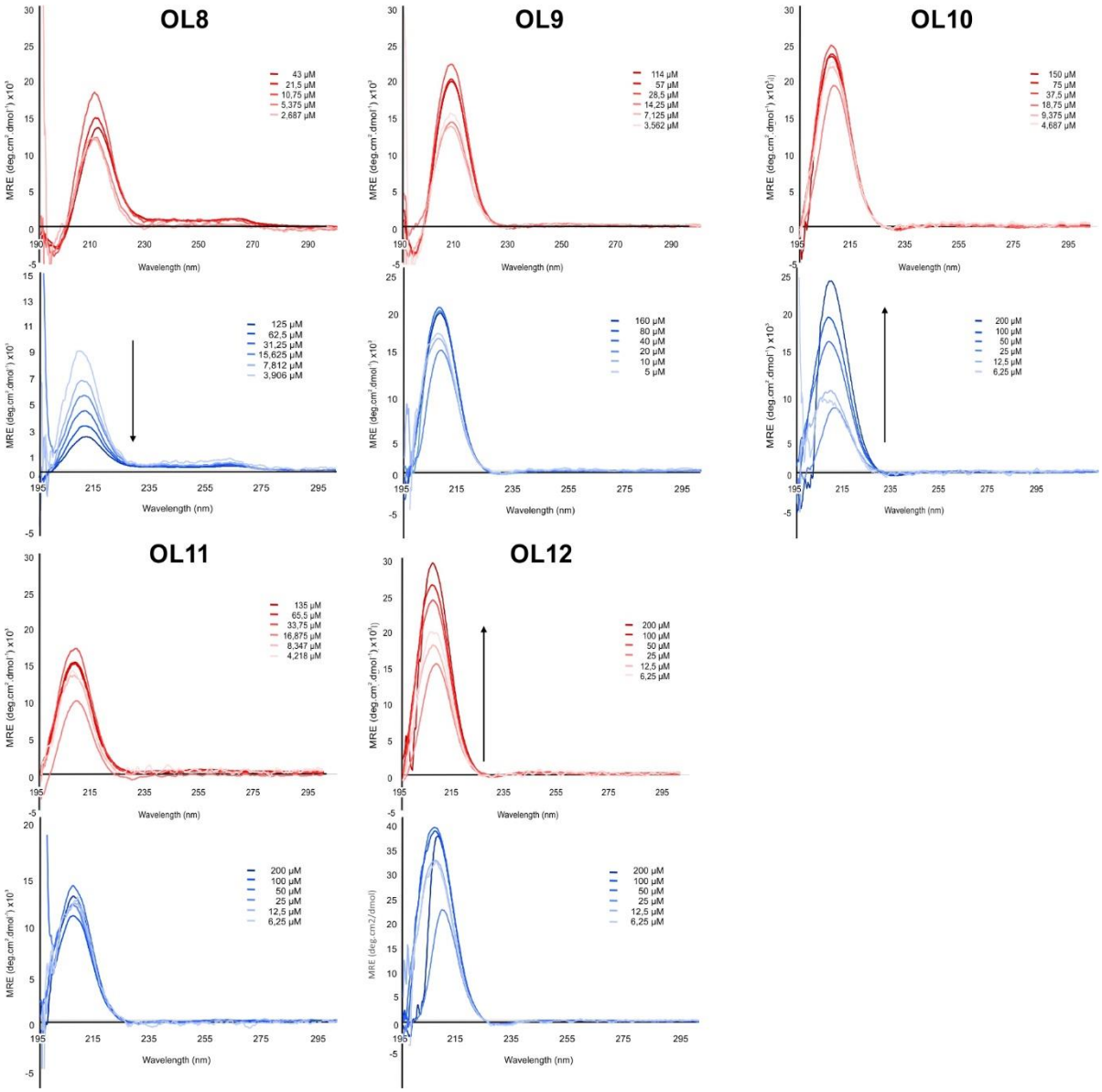


Figure 79 : : Spectres CD en concentration des oligourées **OL8-OL12** en tampon Sodium phosphate 20 mM à pH 5,8 (rouge) et pH 8,2 (bleu). Le sens des flèches indique l'effet de l'augmentation de la concentration sur l'ellipticité.

III. Utilisation de l'iodation des résidus tyrosyles pour résoudre les problèmes de cristallisation et de phase

A ce jour, aucune structure cristalline des oligourées **OL8** à **OL11** n'est disponible dû à des difficultés de cristallisation (formation instantanée de précipité), de diffraction et de résolution de phases. Des données RX ont été mesurées par Jérémie Buratto au synchrotron SOLEIL pour l'oligourée **OL9**, et un remplacement moléculaire avec la structure cristalline de brin d'hélice N-capée **OL12** a été tentée par Jérémie Buratto, en vain. L'obtention de structures cristallines pour l'ensemble des oligourée est nécessaire pour la poursuite de l'étude structurale du projet. Un résidu urée dérivé de la tyrosine a été considérée lors de la conception des oligourées pour permettre l'introduction d'iode sur celui-ci, la présence de cet atome dans la séquence peut aider à la cristallisation, la diffraction et la résolution de la phase.

1. Iodation des résidus tyrosyles comme outil pour résoudre les structures RX

L'iodation des résidus tyrosyles a été décrite dans la littérature par voie enzymatique[16] et par voie chimique[17-22]. Celle-ci consiste en une substitution électrophile aromatique des résidus tyrosyles d'une protéine en présence d'agents iodants[17]. L'iodation de la chaîne latérale du résidu aminé de la tyrosine peut se faire en *ortho* du groupe hydroxyle aux positions 3 et/ou 5 pour conduire soit à la mono-, soit à la di-iodation du résidu aromatique. Cette approche a permis de résoudre les problèmes de phase dans la résolution des structures cristallines de nombreuses protéines[18]. Bien que le développement de méthodes d'iodation innovantes pour la résolution de structures RX des protéines ont été rapportées en nombre dans la littérature[19], très peu d'exemples appliqués à l'iodation de foldamères ont été décrits[20].

L'iodation de la chaîne latérale tyrosyle peut se faire en amont[18,21], pendant la cristallogenèse[18e] ou immédiatement après en plongeant les cristaux dans une solution contenant des agents iodants[18d, 19,22]. La possibilité d'obtenir des dérivés de foldamères ayant des atomes lourds par la simple addition d'agent iodant (N-iodosuccinimide, diiode ou iodure de potassium) pendant la cristallogenèse permet d'éviter la préparation, la purification, la caractérisation de molécules supplémentaires et la synthèse de monomères activés adéquats.

a. Iodation chimique in situ de foldamères oligourées amphiphiles

Gilles Guichard et Gavin Collie ont décrit une méthode d'iodation chimique *in situ*[20] des résidus tyrosyles du foldamère oligourée amphiphile **OL6**. Ce foldamère a la capacité de s'auto-assembler sous forme de capsule en milieu aqueux où les chaînes latérales des résidus

urées dérivés de la tyrosine sont orientées vers le solvant[11,20], rendant les résidus tyrosyles accessibles pour réaliser une réaction d'iodation. La substitution électrophile aromatique a lieu pendant la cristallogenèse, ce qui permet de qualifier cette méthode de *in situ*. L'iodation du composé **OL6** en **OL6-I_2** a été réalisée par Gavin Collie où une solution de diiode à 0,75 mM est ajoutée à la condition de cristallisation (15% d'isopropanol, acétate de sodium à 100 mM, pH 4,6 et chlorure de calcium à 200 mM). La cristallisation en goutte suspendue de l'oligourée **OL6** en présence de la condition de cristallisation contenant l'agent d'iodation a permis d'obtenir des cristaux de **OL6-I_2**. Ces derniers étaient similaires à ceux obtenus pour l'oligourée non iodée. Les paramètres de maille et le groupe d'espace du composé iodée **OL6-I_2** obtenus par DRX sont également identiques à ceux du foldamère non iodé **OL6**. La structure cristalline obtenue est très proche de celle de son analogue non iodé (**Fig. 80**), l'insertion de deux iodes dans la chaîne latérale tyrosyle du dérivé urée n'a pas d'effet sur le phénomène d'auto-assemblage de l'oligourée.

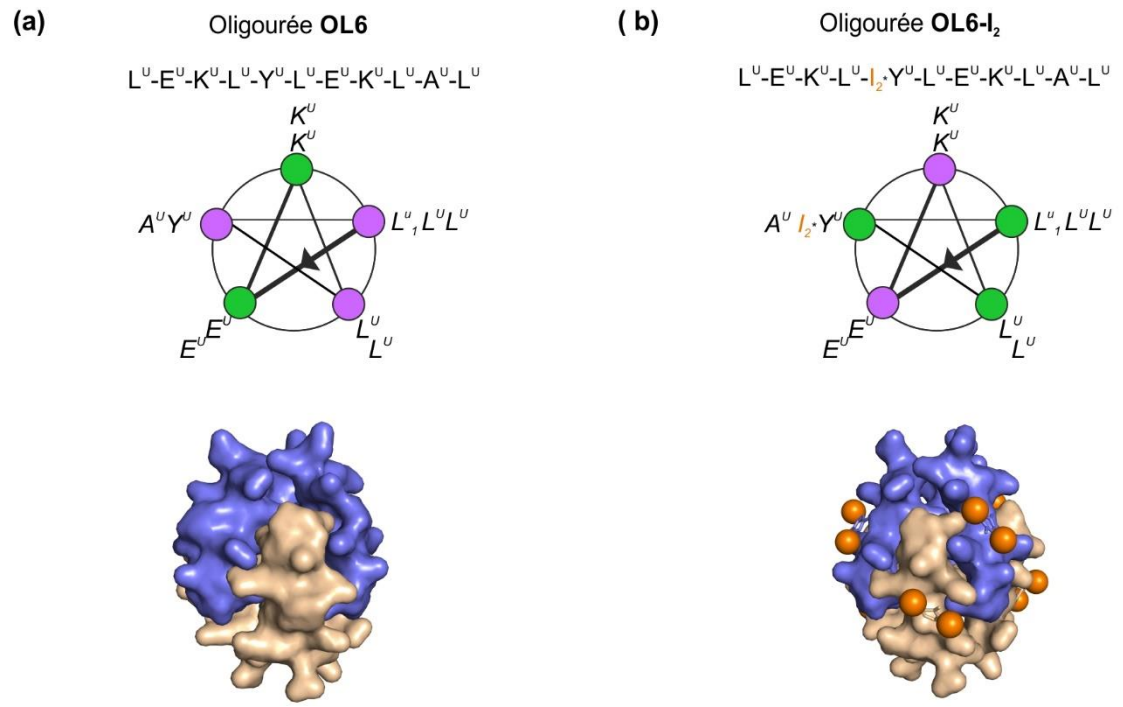


Figure 80 : Représentation de l'oligourée **OL6** (CCDC : **1030455**) et de son analogue di-iodé **OL6-I_2** (**CCDC : 1038892)** sous forme de diagramme en hélice où les résidus non-chargés sont en vert et les résidus chargés en violet. Les boules orange correspondent aux atomes d'iode.

*b. Résultats de l'iodation in situ et post cristallogenèse des oligourées **OL8** et **OL9***

Fort de ces constats, l'iodation *in situ* des composés **OL8** et **OL9** a été considérée et réalisée par Jérémie Buratto sans succès. Des précipités ont été observés lors de la cristallogenèse et n'ont pas évolué en cristaux. Une autre approche réalisée par Brice

Kauffmann a été d'ajouter une solution de triiode (KI/I_2) dans la solution réservoir des conditions de cristallisation où des cristaux avait été obtenus, la résolution de ces cristaux avait posé des problèmes de phase. La diffusion de l'iode dans ces cristaux a conduit à leur perte de pouvoir de diffraction.

2. Iodation d'oligourée en amont de la cristallogenèse

Les méthodes d'iodation pendant et après la cristallogenèse des oligourées **OL8** et **OL9** n'ont donc pas été concluantes. L'iodation des noyaux aromatiques des résidus tyrosyles des oligourées a donc été envisagée avant de procéder à la cristallogenèse.

a. Iodation des résidus tyrosines des protéines et peptides décrits dans la littérature

En 2000 Jia *et al.* ont décrit une méthode d'iodation simple[18a] des résidus tyrosines d'une protéine antigel du scarabée *Tenebrio molitor* (TmAFP). La protéine TmAFT purifiée par HLPC est incubée pendant deux heures à température ambiante avec une solution de N-iodosuccinimide (NIS) à 10 mM dans une solution d'acétate de sodium à 50 mM à pH 3,7. La protéine iodée est ensuite purifiée par RP-HPLC semi-préparative et les analyses par spectrométrie de masse MALDI confirme la présence de deux iodes liés à la protéine et la structure cristalline de la protéine a pu être résolue.

Plus récemment, Schottelius *et al.* ont rapporté une méthode d'iodation directe de résidus tyrosine de peptides pour l'imagerie in vivo de récepteurs ciblés par ces peptides[23] (**Schéma 13a**). La mono-iodation directe d'un pentapeptide cyclique radiomarqué **FC131**, affin au récepteur de la chémokine CXCR4, a été réalisée en présence de NIS (1 equiv.) dans un mélange eau-acétonitrile (v/v). La réaction se fait en quelque minutes (5-10 min) à température ambiante et le peptide iodé est directement purifié sur colonne RP-HPLC semi-préparative. Cette méthode est inspirée des réactions d'iodation des dérivés phénoliques[24] réalisées en présence d'acide para-toluènosulfonique ou de TFA (**Schéma 13b et c**).

L'iode du NIS est un électrophile faible. En présence d'acide tel que le TFA, même en quantité catalytique, le NIS va réagir pour former un iodonium trifluoroacétate[25] qui est un électrophile très fort (**Schéma 14**). Il a été montré que plus l'acide est fort, plus le caractère électrophile du composé halogéné sera fort[26]. Ainsi les noyaux aromatiques les moins activés et réactifs pourront réagir et réaliser une réaction de substitution électrophile aromatique avec des agents d'halogénation sous forme de N-imide/amide. Puisque le noyau aromatique de la chaîne latérale du résidu tyrosine est très activé, la substitution électrophile aromatique se fait rapidement en quelques minutes.

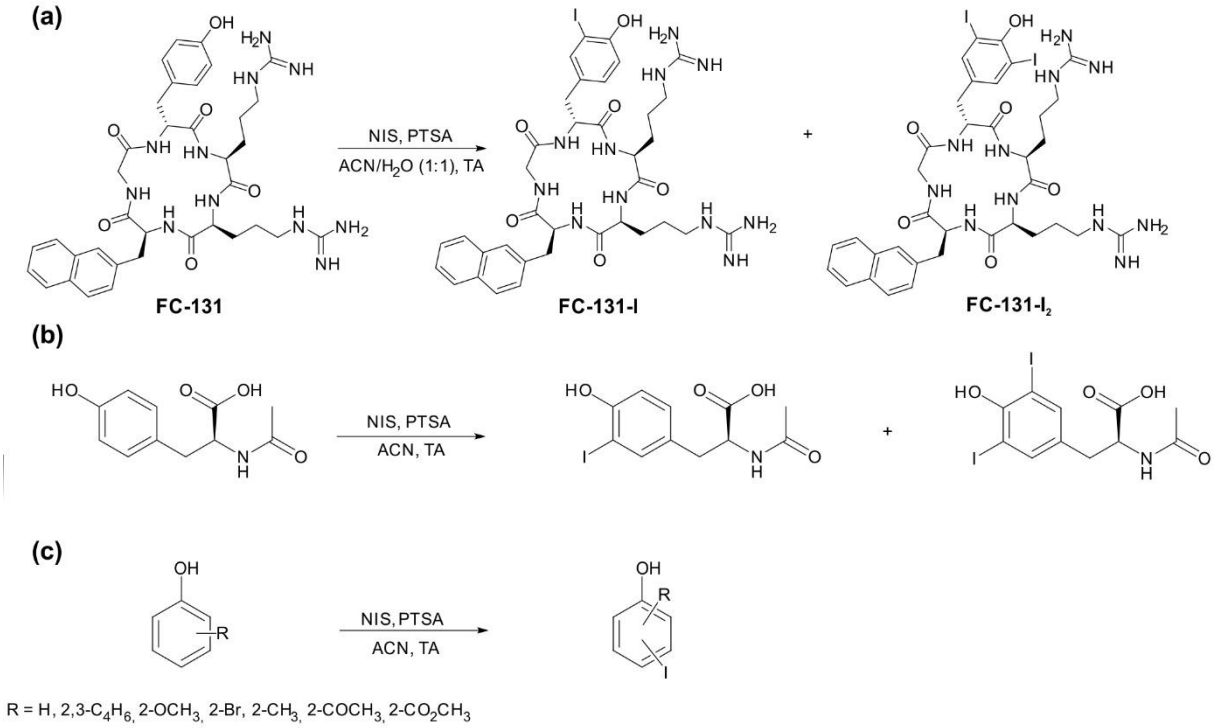


Schéma 13 : Réactions d'iodation in situ en en présence de N-iodosuccinimide et d'acide para-toluènosulfonique **(a)** de FC131[23] **(b)** et d'analogues de N-acétyl-tyrosine [24a] et **(c)** de phénol[24b].

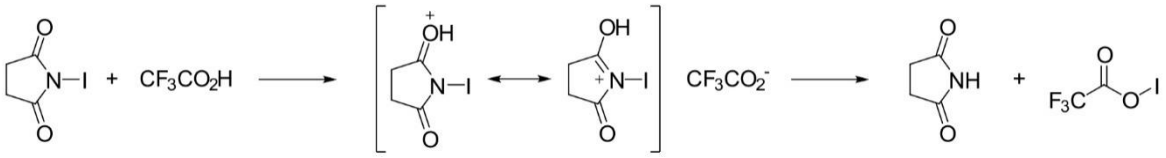
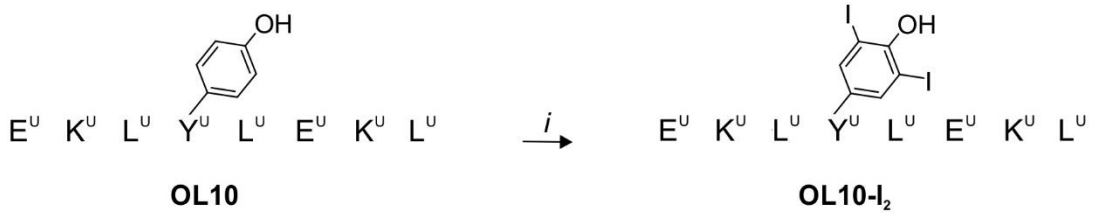


Schéma 14 : Réaction proposée pour la génération de l'iodonium trifluoroacétate[25]

b. *Premières applications de cette méthode d'iodation sur les foldamères d'oligourées **OL9** et **OL10***

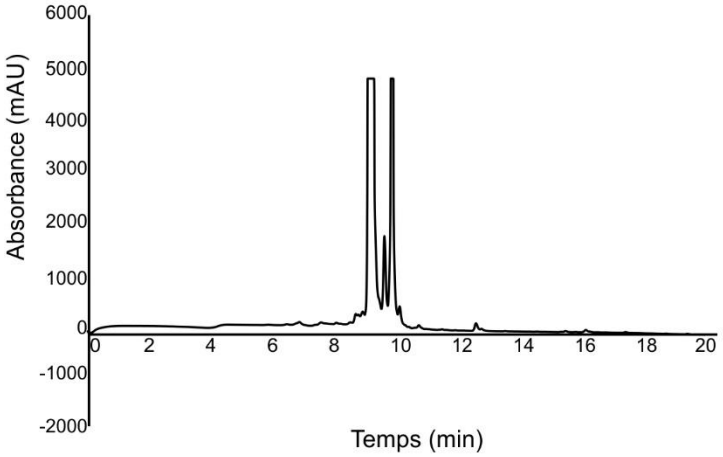
Fort de ces résultats et inspirée de l'ensemble de ces travaux[23,25,26a], une approche d'iodation des chaînes latérales des résidus urées dérivés de la tyrosine des oligomères **OL8** et **OL9** a été envisagée. L'oligourée **OL9** a été le dimère d'hélice sélectionné pour initier l'étude d'iodation car une réaction de substitution électrophile aromatique supplémentaire du NIS sur les noyaux aromatiques de l'espaceur du composé **OL8** pourrait se produire et complexifier l'étude. Un premier test sur la séquence contrôle **OL10** a été réalisé (**Schéma 15**). L'iodation a été réalisée sur 1 mg d'oligourée sous forme de sel de TFA (0,789 µmol) à une concentration de 4,5 mM dans un mélange eau/acétonitrile (1:1) en présence d'agent iodant à température ambiante, puis le mélange réactionnel a été directement injecté sur RP-HPLC semi-préparative (gradient linéaire de 10 à 100 % de B en 20 min avec A (0,1% TFA dans H_2O) et

B (0,1% TFA dans ACN)) (**Fig. 81**). Trois produits sont obtenus à 9,25 min, 9,64 min et 9,90 min, ils ont été analysés par spectrométrie de masse à basse résolution et l'oligourée **OL10-I_2** correspond au pic observé à 9,90 min en HPLC semi-préparative. Des analyses par spectrométrie de masse à haute résolution sont en cours pour identifier les produits formés à 9,25 min et 9,64 min.



i : NIS (2,5 equiv), ACN(0,1% TFA)/H_2O (0,1% TFA), TA, 10 min

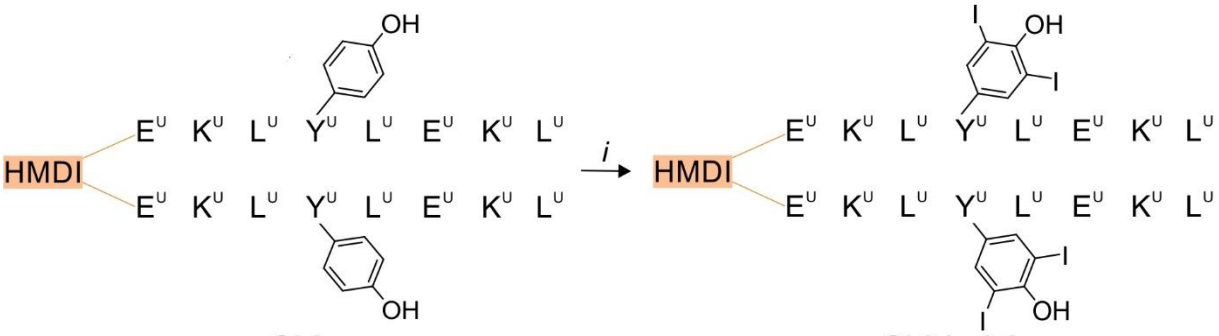
Schéma 15 : Iodation des résidus tyrosyle de l'oligourée **OL10**



t_r(min)	pureté	ESI-MS $[M+H]^+$
9,25	55%	1516,79
9,64	18%	1516,79
9,90	26%	1518,80

Figure 81 : Chromatogramme de HPLC semi-préparative avec l'observation des produits formées lors de l'iodation réalisée pour **OL10** avec 2,5 equiv. NIS, 10 min

Une première iodation du dimère d'hélice **OL9** sous forme de sel de TFA a été réalisée dans les mêmes conditions sur 1 mg (0,37 µmol, **Schéma 16**), et trois autres conditions de réactions d'iodation ont ensuite été testées (**Tableau 38**).



OL9 **OL9-I$_4$, *1-4***

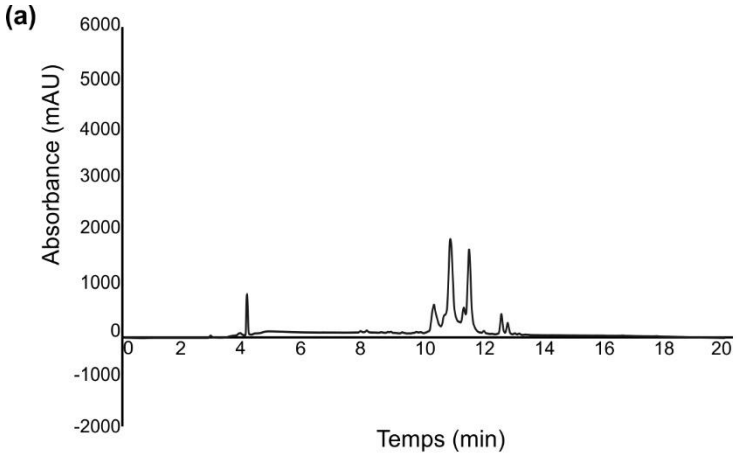
i : NIS (3,5 – 5 equiv), ACN(0,1% TFA)/H$_2$O (0,1% TFA), TA, 2,5 à 7,5 min

Schéma 16: Iodation des résidus tyrosyle de l'oligourée **OL9**

Test	NIS (eq)	temps (min)	pureté brut réactionnel	ESI-MS $[M+2H]^{2+}$	HPLC t_R (min)[a]	rdt[b]
1	5,0	5,0	18%	1603,36	10,89	25%
2	3,5	5,0	76%	1603,27	11,15	34%
3	5,0	2,5	89%	1603,26	11,22	34%
4	5,0	7,5	73%	1603,26	11,21	59%

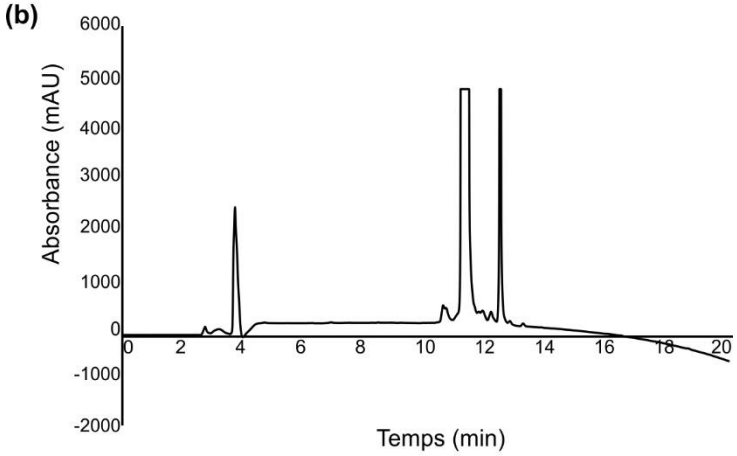
Tableau 38 : Conditions de réaction (équivalent de NIS et temps de réaction), temps de rétention HPLC et rendements de synthèses de tétra-iodation de l'oligourée **OL9**. [a] système RP-HPLC : gradient linéaire 10 à 100% de B en 10 min avec A (0,1% TFA dans H_2O) et B (0,1% TFA dans CH_3CN), détection UV à 200 nm. [b] rendement après injection en HPLC semi-préparative.

Les quatre conditions testées révèlent la présence de plusieurs produits formés (**Fig. 82**). Des analyses par spectrométrie de masse des signaux ont été réalisées à basse résolution par ESI-LCMS. La masse des pics entre 11,15 et 11,49 min correspond à la masse du produit di-chargé ($[M+2H]^{2+}$) de l'oligourée tétra-iodée **OL9-I$_4$**. Des analyses RMN et de spectrométrie de masse à haute résolution sont en cours pour identifier les autres produits formés. La réaction d'iodation à petite échelle des oligourées **OL9** et **OL10** a fonctionné mais l'identification des produits secondaires formés ainsi qu'une optimisation de la réaction devront être réalisées pour pouvoir étendre l'application de cette méthode d'iodation à l'oligourée **OL8**. La cristallisation de l'oligourée **OL9-I$_4$** a été réalisée par diffusion vapeur en goutte suspendue par mes soins dans les mêmes conditions de cristallisation que l'analogue non iodé, dont la résolution des données de diffraction des cristaux obtenus avait rencontré des difficultés. A ce jour, seul un précipité est observé dans la goutte de cristallisation.



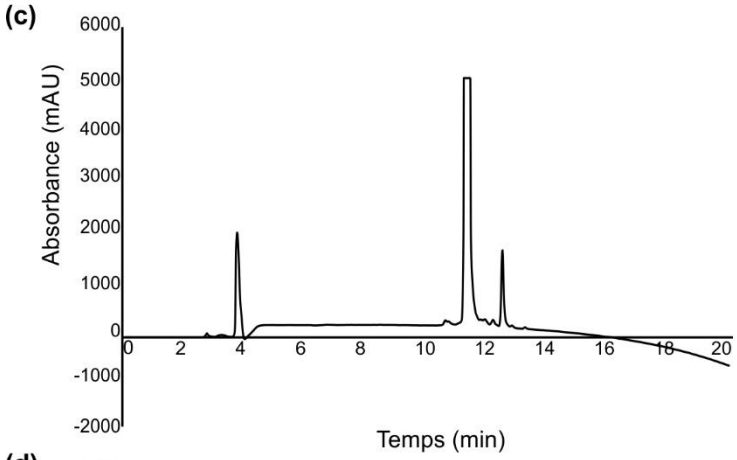
OL9-I$_4$ *Test 1*

t_r(min)	pureté	ESI-MS $[M+2H]^{2+}$
10,36	15%	1601,34
10,89	55%	1602,35
11,49	18%	1603,36
12,59	12%	1229,53



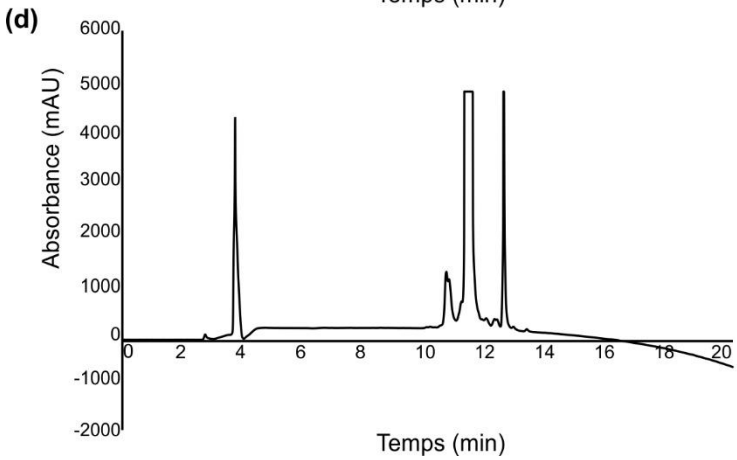
OL9-I$_4$ *Test 2*

t_r(min)	pureté	ESI-MS $[M+2H]^{2+}$
10,58	4%	1602,32
11,15	76%	1603,27
12,44	20%	1229,55



OL9-I$_4$ *Test 3*

t_r(min)	pureté	ESI-MS $[M+2H]^{2+}$
11,22	89%	1603,26
12,50	11%	1229,49



OL9-I$_4$ *Test 4*

t_r(min)	pureté	ESI-MS $[M+2H]^{2+}$
10,62	12%	1602,26
11,21	73%	1603,26
12,49	15%	1229,50

Figure 82 : Chromatogrammes de HPLC semi-préparative avec l'observation des produits formées lors des réactions d'iodation réalisées pour **OL9** dans quatre conditions : **(a)** 5 equiv. NIS, 5 min, **(b)** 3,5 equiv. NIS, 5 min, **(c)** 5 equiv. NIS, 2,5 min et **(d)** 5 equiv. NIS, NIS, 7,5 min.

L'ensemble des méthodes d'iodation décrites dans la littérature, ainsi que le développement préliminaire d'une méthode d'iodation de l'oligourée **OL9**, n'a pas permis de résoudre les problèmes de cristallisation et de phase rencontrés. La cristallisation racémique de l'oligourée **OL9** pour résoudre ces problèmes de cristallisation et de résolution de phase serait la prochaine approche considérée pour obtenir une structure cristalline de l'oligomère **OL9** et des autres composés **OL8**-**OL11** en cas d'échec de cristallisation de leurs analogues iodés.

IV. Conclusion du chapitre et perspectives

La cristallisation racémique a permis d'obtenir des structures cristallines de foldamères d'oligourées sous forme de dimère d'hélices repliés en hairpin. L'étude des structures a permis d'identifier l'espaceur MDI. Un espaceur rigide aromatique s'est révélé être le meilleur pour préorganiser les hélices dans une orientation parallèle l'une à l'autre. Cette étude structurale a été le point de départ pour concevoir des hairpins d'hélices solubles dans l'eau à base de foldamères d'oligourée. La dimérisation sur support solide des hélices d'oligourée a été réalisée avec succès. Cette voie de synthèse supportée est une approche rapide et originale pour dimériser des séquences d'oligomères.

La compréhension des phénomènes de repliement et d'assemblage nécessite maintenant des données structurales. Malgré tous nos efforts d'iodation des chaînes latérales tyrosyle des oligourées, la cristallisation et la résolution des structures des oligomères restent problématiques. Le succès de l'approche de cristallisation racémique utilisée pour la résolution des structures en solvant organique semble donc être la stratégie à adopter pour obtenir des données structurales. Cette approche nécessite la synthèse des énantiomères de chaque monomère pour ensuite réaliser la synthèse sur support solide des hélices énantiomères.

A ce jour, le manque de données structurales de ces oligourées ne permet pas de confirmer la présence de tels assemblages, ni d'expliquer le phénomène observé par CD pour l'oligourée **OL8** en condition basique. Une caractérisation biostructurale par spectrométrie de masse native et RMN à très haut champs (dans de l'oxyde de deutérium) permettrait de mettre en évidence la présence d'assemblages.

L'obtention des structures cristallines permettrait de mieux connaitre les règles de conception de structures artificiels avec l'objectif de pouvoir développer des mimes de structure tertiaire comme inhibiteurs d'interactions protéine-protéine.

Références bibliographiques

1. A. L. Jochim, P. S. Arora *ACS Chem. Biol.* **2010**, *5*, 919-923
2. A. M. Watkins, M. G Wuo, P. S. Arora *J. Am. Chem. Soc.* **2015**, *137*, 11622-11630
3. M. Raj, B. N. Bullock, P. S. Arora *Bioorg. Med. Chem*. **2013**, *32*, 4051-4057
4. (a) N. Sawyer, A. M. Watkins, P S. Arora *Acc. Chem. Res*. **2017**, *50*, 1313-1322 (b) J. W. Checco, D. F. Kreitler, N. C. Thomas, D. G. Belair, N. J. Rettko, W. L. Murphy, K. T. Forest, S. H. Gellman *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 4552-4557
5. L. Fischer, P. Claudon, N. Pendem, E. Miclet, C. Didierjean, E. Ennifar, G. Guichard *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 1067-1070
6. J.M. Berg *Proteins* **1993**, *16*, 301-305
7. (a) Wallach *Liebigs. Ann. Chem.* **1895**, *286*, 90-113 (b) J. D. Dunitz *J. Am. Chem. Soc.* **1991**, *113*, 9811-9820
8. (a) J.M. Berg *J. Am. Chem. Soc.* **1992**, *114*, 4002-4003 (b) T. O. Yeates *Nat. Struct. Mol. Biol.* **1995**, *2*, 1062-1067 (c) T.O. Yeates *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 16230-16235 (d) S. B. H. Kent *Science* **1994**, *266*, 776-779 (e) S. B. H. Kent *J. Am. Chem. Soc.* **2009**, *131*, 1362-1363 (f) S. B. H. Kent *Annu. Rev. Biophys.* **2012**, *41*, 41-61 (g) O. Schneewind *Protein Sci.* **2009**, *18*, 1146-1154 (h) C. Toniolo *Nat. Struct. Mol. Biol.* **1994**, *1*, 908-914
9. (a) Y. Ueda *J. Am. Chem.Soc.* **1993**, *115*, 10432-10433, P. E. Nielsen *Nat. Struct. Biol.* **1997**, *4*, 98-101 (b) I. Huc *Angew. Chem. Int. Ed.* ***2014***, *53*, 14424-14427 (c) I. Huc *Nucleic Acids Res.* **2016**, *44*, 5936-5943
10. (a) I. Huc *J. Am. Chem. Soc.* ***2005***, *127*, 12943-12951 (b) I. Huc, *J. Am. Chem. Soc.* **2013**, *135*, 9628-9631 (c) S. H. Gellman *J. Am. Chem. Soc.* **2016**, *138*, 6498-6505
11. J. Fremaux, L. Mauran, K. Pulka-Ziach, B. Kauffmann, B. Odaert, G. Guichard *Angew. Chem. Int. Ed.* **2015**, *54*, 9816-9820
12. C. Douat-Casassus, K. Pulka, P. Claudon, G. Guichard *Org. Lett.* **2012**, *14*, 3130-3133
13. D. G. Udugamasooriya, S. P. Dineen, R. A. Brekken, T. Kodadek *J. Am. Chem. Soc.* **2008**, *130*, 5744-5752
14. J. C. Hooks, J. P. Matharage, D. G. Udugamasooriya *Biopolymers* **2011**, *96*, 567-577
15. (a) G. S. Vadehra, B. D. Wall, S. R. Diegelmann, J. D. Tovar *Chem. Commun.* **2010**, *46*, 3947-3949 ; (b) A. M. Sanders, T. J. Dawidczyk, H. E. Katz, J. D. Tovar *ACS Macro Lett.* **2012**, *1*, 1326-1329
16. J. Ruf and P. Carayon *Arch. Biochem. Biophys*. **2006**, *445*, 269–277

17. D. Ghosh, M. Erman, M. Sawicki, P. Lala, D. R. Weeks, N. Li, W. Pangborn, D. J. Thiel, H. Jörnvall, R. Gutierrez,J. Eyzaguirre *Acta Crystallogr*. *Sect. D: Biol. Crystallogr.* **1999**, *55*, 779–784
18. (a) Y.-C. Liou, P. L. Davies, Z. Jia *Acta Crystallogr. Sect. D: Biol. Crystallogr***. 2000**, *56*, 354-356 ; (b) E. K. Leinala, P. L. Davies and Z. Jia *Acta Crystallogr. Sect. D: Biol. Crystallogr***. 2002**, *58*, 1081–1083 ; (c) P. H. Zwart, S. Banumathi, M. Dauter and Z. Dauter *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1958–1963 ; (d) P. B. Siegler *Biochemistry***. 1970**, *9*, 3609-3617 ; (e) R. H. Kretsinger *J. Mol. Biol.* **1968**, *31*, 315-318
19. (a) K. Sakamoto, K. Murayama, K. Oki, F. Iraha, M. Kato-Murayama, M. Takahashi, K. Ohtake, T. Kobayashi, S. Kuramitsu, M. Shirouzu, S. Yokoyama *Structure* **2009**, *17*, 335–344 ; (b) H. Miyatake, T. Hasegawa, A. Yamano, *Acta Crystallogr., Sect. D: Biol. Crystallogr*. **2006**, *62*, 280–289
20. G. W. Collie, K. Pulka-Ziach, G. Guichard *Chem. Commun*. **2016**, *52*, 1202-1205
21. L. Brady, A. M. Brzozowski, Z. S. Derewenda, E. Dodson, G. Dodson, S. Tolley, J. P. Turkenburg, L. Christiansen, B. Huge-Jen, L. Norskov, L. Thim, U. Merge *Nature* **1990**, *343*, 767-770
22. G. Evans, G. Bricogne *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003**, *59*, 1923-1929
23. Schottelius, M.; Konrad, M.; Osl, T.; Poschenrieder, A.; Wester, H.-J. *Tetrahedron Letters,* **2015**, *56,* 6602-6605
24. (a) P. Bovonsombat, P. Khanthapura, M. M. Krause, J. Leykajarakul *Tetrahedron Letters,* **2008**, *49,* 7008-7011 ; (b) P. Bovonsombat, J. Leykajarakul, C. Khan, K. Pla-on, M. M. Krause, P. Khanthapura, R. Ali, N. Doowa, *Tetrahedron Letters,* **2009**, *50,* 2664-2667
25. (a) M. C. Carreño, J. Ruano, A. Urbano *Tetrahedron Letters* **1996**, *37,* 4081-4084 ; (b) A.-S. Castanet, F. Colobert, P.-E. Broutin *Tetrahedron Letters* **2002**, *43,* 5047-5048
26. (a) M Bergström, G. Sueresh, V. R. Naidu, C. R. Unelius *Eur. J. Org. Chem.* **2017**, 3234-3239 ; (b) E. B. Merkushev *Synthesis* **1988**, 923-937 ; (c) V. K. Chaikovskii, V. D. Filimonov, V. I. Skorkhodov, V. D. Ogorodnikov *Russ. J. Org. Chem.* **2007**, *43,* 1278-1281

Conclusion et perspectives

Les interactions protéine-protéine sont impliquées dans de nombreux processus biologiques. Des motifs secondaires structurés en hélice α sont généralement impliqués aux interfaces de ces interactions. Ces interfaces peuvent mettre en jeu une petite portion peptidique d'une protéine ou bien des surfaces plus étendues avec des structures tertiaires et quaternaires. Différentes stratégies sont possibles pour cibler les interactions mises en jeu entre protéines en vue de concevoir des molécules capables d'inhiber la reconnaissance avec le partenaire naturel. En termes d'approches chimiques et de design basé sur la structure, il est plus aisé de se focaliser sur des interfaces ne faisant intervenir qu'un fragment d'hélice peptidique Le développement de mimes peptidiques à base de foldamères est en plein essor et permet de palier les problèmes intrinsèques des courtes séquences peptidiques (moins de 10 résidus d'acides aminés) comme le manque de stabilité plasmatique ou l'absence de conformation bien définie. Lorsque l'interaction implique plus qu'une simple hélice, le développement de petites molécules organiques (M_W <500 Da) ou de courts peptides peut se révéler inapproprié car il leur est difficile d'interagir avec des surfaces aussi importantes et se lier aussi fortement que le partenaire ciblé. Ainsi le développement de mimes de structures tertiaires et quaternaires à base de peptides ou de foldamères peut se révéler plus approprié pour cibler des surfaces d'interactions larges. Dans le cas des foldamères, cette approche se heurte cependant aux manques de connaissances structurales sur les principes gouvernant le repliement des foldamères en structure tertiaire et quaternaire. Ces travaux de thèse ont reposé sur la conception de molécules artificielles capables de mimer les structures secondaires et tertiaires présentes aux surfaces d'interactions entre protéines comme les structures de type hélice-coude-hélice (HCH). Au sein du laboratoire, nous nous intéressons aux foldamères à base d'oligourées comme classe de peptidomimes. Ces foldamères d'oligourées sont capables de se replier en structures hélicoïdales 2,5 stables grâce à la répétition d'un réseau de liaisons hydrogène intramoléculaires à longue distance entre fonctions urées.

Des travaux récents ont démontré que la présence d'un segment oligourée inséré au cœur ou située à l'extrémité d'une séquence peptidique en hélice permettait de conserver les propriétés de repliement des deux molécules et de générer une hélice unique bien repliée. Les propriétés structurales de ces foldamères mixtes oligoamide/oligourée (chimères) ont été utilisées pour la reconnaissance de surfaces de protéines entre le récepteur nucléaire de la vitamine D (VDR) et son coactivateur (SRC1). Une série de foldamères chimères (**C2**-**C11**) a été conçue par superposition d'un foldamère mixte à la séquence de la portion peptidique responsable de l'interaction de la protéine coactivatrice SRC1avec VDR. Une étude de relation structure-affinité a été réalisée en collaboration avec le Dr. Natacha Rochel-Guiberteau et a permis d'identifier le foldamère **C11** comme tête de série. Sa séquence a été optimisée en vue

d'améliorer son interaction avec le récepteur nucléaire et inhiber l'interaction entre VDR et SRC1. La cristallisation des nouvelles séquences hybrides a conduit à l'obtention des structures cristallines de deux composés **C14** et **C19**. L'obtention des mesures d'affinité et de l'ensemble des structures cristallines permettra de pouvoir réaliser une seconde étude de relation structure-affinité pour générer une troisième série de foldamères mixtes cible de l'interaction VDR/SRC1. La détermination des valeurs d'IC_{50} de l'ensemble des séquences affines à VDR permettrait une caractérisation plus fine des foldamères mixtes oligoamide/oligourée et de leur aptitude à inhiber l'interaction entre les protéines nucléaire et coactivatrice.

La conception de structures tertiaires capables de mimer des biopolymères (taille, forme et fonction) en connectant des structures secondaires entre elles, reste un défi. La difficulté de synthèse de tels systèmes moléculaires, ainsi qu'une méconnaissance des principes de conception, expliquent le peu d'exemple décrits dans la littérature de structures de foldamères tertiaires et quaternaires. Dans notre équipe, les premières investigations initiées par les Dr. Yella-Reddy Nelli et Dr. Nagendar Pendem se sont portées sur la synthèse de structures tertiaires simples : deux hélices identiques, connectées via un espaceur, sont repliées pour former un motif type HCH, dit hairpin d'hélices.

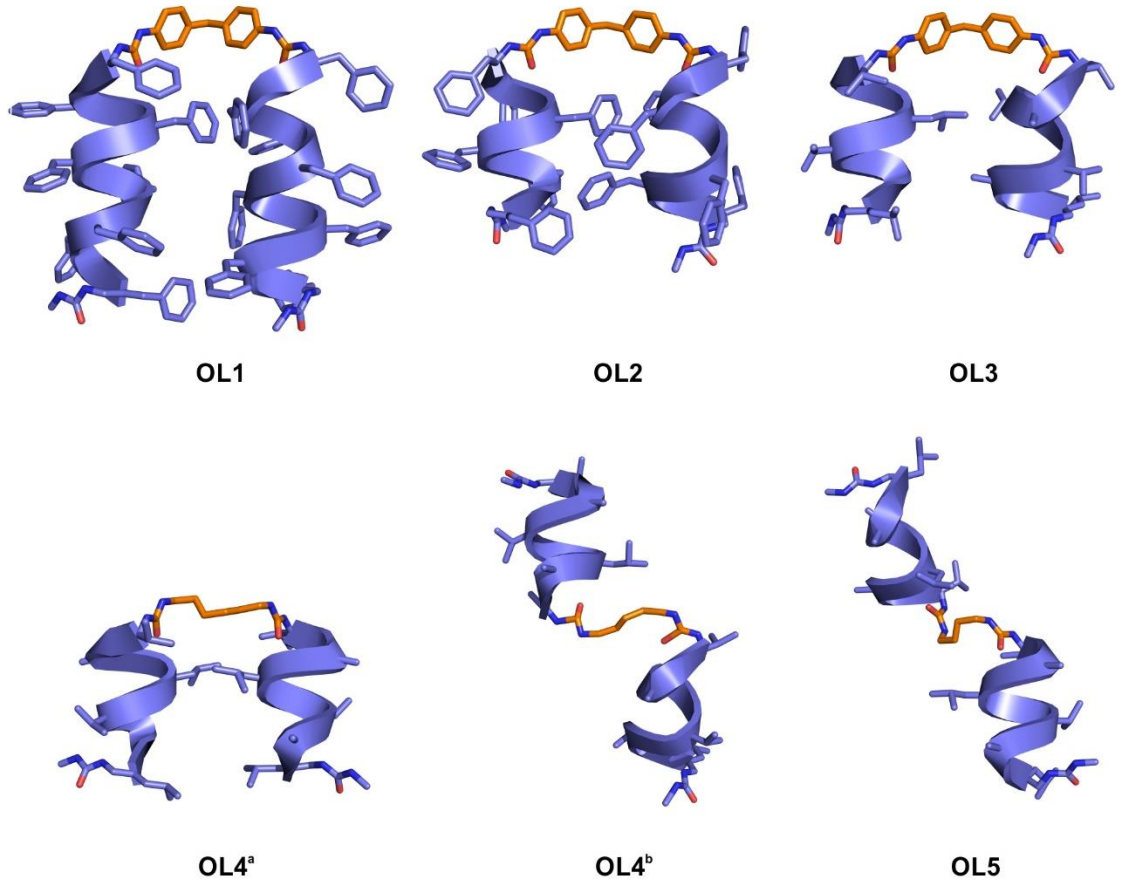


Figure 83 : Ensemble des structures cristallines obtenues pour les oligourées en solvant organique.

Les différentes structures cristallographiques obtenues ont permis de révéler l'importance de la nature de l'espaceur sur le repliement en hairpin du système. Deux types de conformation ont été observées : une dite ouverte, où les hélices sont de part et d'autre de l'espaceur ; une fermée où il y a un repliement de la structure en hairpin, les hélices sont alors orientées parallèlement l'une par rapport l'autre (**Fig. 83**). La nature de l'espaceur est donc un facteur important pour la conception de structures tertiaires en milieu organique avec nos systèmes. Cette étude préliminaire a donc permis d'identifier un espaceur rigide aromatique capable de pré-organiser les hélices dans une orientation parallèle quelles que soient la longueur des hélices ou la nature des chaînes latérales.

Le projet s'est ensuite concentré sur la synthèse de séquences hydrosolubles repliées en hairpin. En milieu aqueux, outre le choix de l'espaceur, la nature et la position des chaînes latérales sont des paramètres importants pour permettre aux deux hélices d'interagir via l'effet hydrophobe. Le choix des chaînes latérales s'est inspiré dans un premier temps des travaux sur l'auto-assemblage d'hélices d'oligourées hydrosolubles développés au laboratoire, les hélices possèdent une face hydrophobe pour induire le repliement par interaction entre les deux hélices, et une face hydrophile pour assurer la solubilité du système dans l'eau. La stratégie de synthèse choisie repose sur une synthèse sur support solide, une approche peu décrite dans la littérature : elle consiste à synthétiser dans un premier temps les hélices sur la bille de résine, puis à dimériser par couplage de l'espaceur entre deux hélices. Un protocole a donc dû être développé et les premières synthèses ont été très encourageantes avec un rendement global de 20% après purification, un rendement très satisfaisant sur support solide. Une caractérisation en solution par RMN et dichroïsme circulaire, et dans le solide par diffraction des rayons X, est en cours pour évaluer le repliement de nos systèmes. Les structures cristallographiques permettront d'obtenir des informations structurales clés en vue de l'optimisation et de la conception de nouvelles séquences. L'optimisation des conditions de cristallisation et la résolution des structures cristallographiques permettra à nouveau de valider l'orientation des hélices de nos structures supersecondaires. Enfin, l'identification d'interactions protéine-protéine sur des surfaces étendues pertinentes sera réalisée afin de pouvoir appliquer les outils développés dans ce projet à une cible biologique.

Partie expérimentale

I. General methods

Commercially available reagents were used throughout without purification. Solvents were purchased from Sigma-Aldrich. THF and DCM were preliminary dried by passing through solvent drying system (SPS 800 Manual from MBraun). TFA was purchased from Sigma-Aldrich. DMF and DCM for solid phase synthesis were purchased from Carlo ERBA. RP-HPLC-quality acetonitrile (ACN) and MilliQ water were used for RP-HPLC analyses and purification. *N*-protected amino acids were purchased from Iris Biotech. Rink amide resins were purchased from Novabiochem laboratories.

Thin layer chromatography (TLC) was performed on silica gel 60 F254 (Merck) with detection by UV light and charring with 1% w/w ninhydrin in ethanol followed by heating. Flash column chromatographies were carried out on silica gel (40-63 µm, Merck).
Solid phase oligourea synthesis with microwave irradiation were carried out on DiscoverBio system from CEM (CEM µWaves S.A.S., Orsay, France).

Analytical RP-HPLC analyses were performed on a Dionex Ultimate U3000 using Macherey-Nagel Nucleodur columns (4.6 x 100 mm, 3µm, λ = 200 nm) at a flow rate of 1mL.min^{-1}. Semi-preparative purifications of oligoureas were performed on a Dionex Ultimate U3000 using a Macherey-Nagel Nucleodur VP250/10 100-16 C18ec column (10 x 250 mm, 16µm) at a flow rate of 4 mL.min^{-1}. Preparative purifications of oligoureas were performed on a Gilson GX-281 using a Macherey-Nagel Nucleodur VP250/21 100-5 C18ec column (21 x 250 mm, 5µm) at a flow rate of 20 mL.min^{-1}. The mobile phase was composed of 0.1% (v/v) TFA-H_2O (Solvent A) and 0.1% (v/v) TFA-ACN (Solvent B).

^{1}H NMR and ^{13}C NMR spectra were recorded on three different NMR spectrometers: (1) an Advance II NMR spectrometer (Bruker Biospin) with a vertical 7.05T narrow-bore/ultrashield magnet operating at 300 MHz for ^{1}H observation and 75 MHz for ^{13}C observation by means of a 5-mm direct BBO 1H/19F_XBB_H probe with Z gradient capabilities; (2) a DPX-400 NMR spectrometer (Bruken Biospin) with avertical 9.4T narrow-bore/ultrashield magnet operating at 400 MHz for ^{1}H observation by means of 5-mm direct QNP ^{1}H/^{13}C/^{31}P/^{19}F probe with gradient capabilities; (3) an avance III NMR spectrometer at 700 MHz for ^{1}H observation by means of 5-mm TXI ^{1}H/^{13}C/^{15}N probe with Z gradient capabilities. Chemical shifts are reported in parts per million (ppm, δ) relative to the ^{1}H or ^{13}C residual signal of the deuterated solvent used. ^{1}H NMR splitting patterns with observed first-order coupling are designated as singlet (s), broad singlet (brs), doublet (d), triplet (t), or quartet (q). Not defined ^{1}H NMR splitting patterns are designated as multiplet (m). Coupling constants (J) are reported in Hertz (Hz).

All MS experiments were performed at the Mass Spectrometry platform from the European Institute of Chemistry and Biology (UMS 3033 – IECB, Pessac, France). ESI-HRMS analyses were carried out on a Thermo Exactive with an ion trap mass analyser. ESI-MS analyses were carried out on a ThermoElectron LCQ Advantage spectrometer equipped with an ion trap mass or on a Tof Agilent 6230 with a time of flight analyser with an UHPLC Agilent 1290 Infinity (agilent Eclipse Plus C18 RRHD 1.8 µm 2.1 x 50 mm). Matrix Assisted Laser Desorption Ionisation Mass Spectrometry (MALDI-MS) analysis were performed on a MALDI-TOF-TOF mass spectrometer (Ultraflex III, Bruker Daltonics, Germany) equipped with a SmartBeam laser (Nd:YAG, 355nm). Solutions of foldamers were prepared as follows: the lyophilized powders were resuspended in dimethyl sulfoxide and then diluted in water/acetonitrile (1/1, vol/vol). The samples at a final concentration of around 100µM were mixed with the matrix α-Cyano-4-hydroxycinnamic acid (5mg/mL in acetonitrile/0.1% aqueous TFA (1/1, vol/vol)). All MALDI-MS measurements were acquired in the reflectron positive mode and a mixture of standard peptides was used for external calibration in the mass range from m/z 962.43 to m/z 3494.65.

All circular dichroism (CD) spectra were recorded on a J-815 Jasco discographe (Jasco France, Nantes, France) using a quartz cell with a path length of 2 mm or 5 mm (Hellma, Paris, France). Samples temperature was regulated at 20 °C. Data were collected in continuous scan mode with a data pitch of 0.1 nm, a scanning speed of 50 nm.min^{-1}, 2 nm bandwith and two accumulations per sample. Samples data were collected as raw ellipticity (Ψ in mdeg) and converted to pear residue mean ellipticity (PRME with [Θ] = ellipticity) in deg.cm^2.dmol^{-1}.residue^{-1} using the following equation:

$$[\Theta] = \frac{\Psi \times 10^{-3}}{res \times l \times c}$$

where res is the number of residues in the oligomer, l is the pathlength in cm, and c is the oligourea concentration in dmol.cm^{-3}.

All Root-Mean-Square-Deviation (RMSD) were calculated for all residues using the following equation:

$$\mathrm{RMSD} = \sqrt{\frac{\sum d^2}{n}}$$

where d is the distance measured in angström and n the number of measurements per residue. Distance measurements were done with Pymol software using measurement wizard. RMSD of amino acids were performed on the distance between atoms of the backbone, and RMSD of urea residues were done on side chains.

For angle deviation (a) studies of chimeras and oligourea, the calculation was done for all residues using the following equation:

$$a = 180 - \left| \left| a_i - a_{ref} \right| - 180 \right|$$

where angles Φ = -58° and Ψ = -47° are used as reference values for amino acids and urea reference values are Φ = -100°, Θ_1 = 55°, Θ_2 = 85° and Ψ = -170°

For X-ray crystallographic studies of chimeras, collections were performed at the European Synchrotron Radiation Facility (ESRF), Grenoble, France or at SOLEIL Synchrotron, Gif-sur-Yvette, France. Data were collected on ID30B, ID30A and ID23 beam at ESRF and on beam PX2 at SOLEIL. Crystals were cryo-protected by addition of either 20% glycerol or 10% glycerol in $LiSO_4$ 2.2 M. Indexation and integration of diffraction data were performed on XDS software and scaling was performed with AIMLESS. All crystals structures were solved using molecular replacement with Phaser software using the crystal structure of VDR (2HC4) as a model. Refinement and construction of atomic models were performed with Phenix and Coot softwares respectively. For X-Ray diffractions studies of oligoureas, collections were performed at the IECB X-ray facility (UMS 3033) on two different high flux microfocus Rigaku rotating anodes at the copper kα wavelength. Data were collected on a micromax MM07 (800W) equipped with osmic Varimax mirrors and semi-cylindrical R-Axis spider IP detector or on a FRX (2.7 kW) equipped with osmic Varimax mirrors and a Dectris Pilatus 200 K hybrid detector. The crystals were mounted on cryo-loops after quick soaking on Paratone-N oil from Hampton research and flash-frozen. Both diffractometers have partial chi geometry goniometer allowing omegascan data collections. The data were processed with the CrystalClear suite version 1.36 and 2.1b25 (CrystalClear (Rigaku/MSC), 2006). All crystal structures were solved using direct methods implemented in SHELXD and were refined using SHELXL 2013 version. Fullmatrix least-squares refinement were performed on F2 for all unique reflections, minimizing w(Fo2- Fc2)2, with anisotropic displacement parameters for non-hydrogen atoms. Hydrogen atoms were positioned in idealized positions and refined with a riding model, with Uiso constrained to 1.2 Ueq value of the parent atom (1.5 Ueq when CH_3). The positions and isotropic displacement parameters of the remaining hydrogen atoms were refined freely. SIMU and DELU commands were used to restrain some side chains as rigid groups and restrain their displacement parameters.

II. General procedure A for solid phase oligouea synthesis under microwave irradiation using azide strategy.

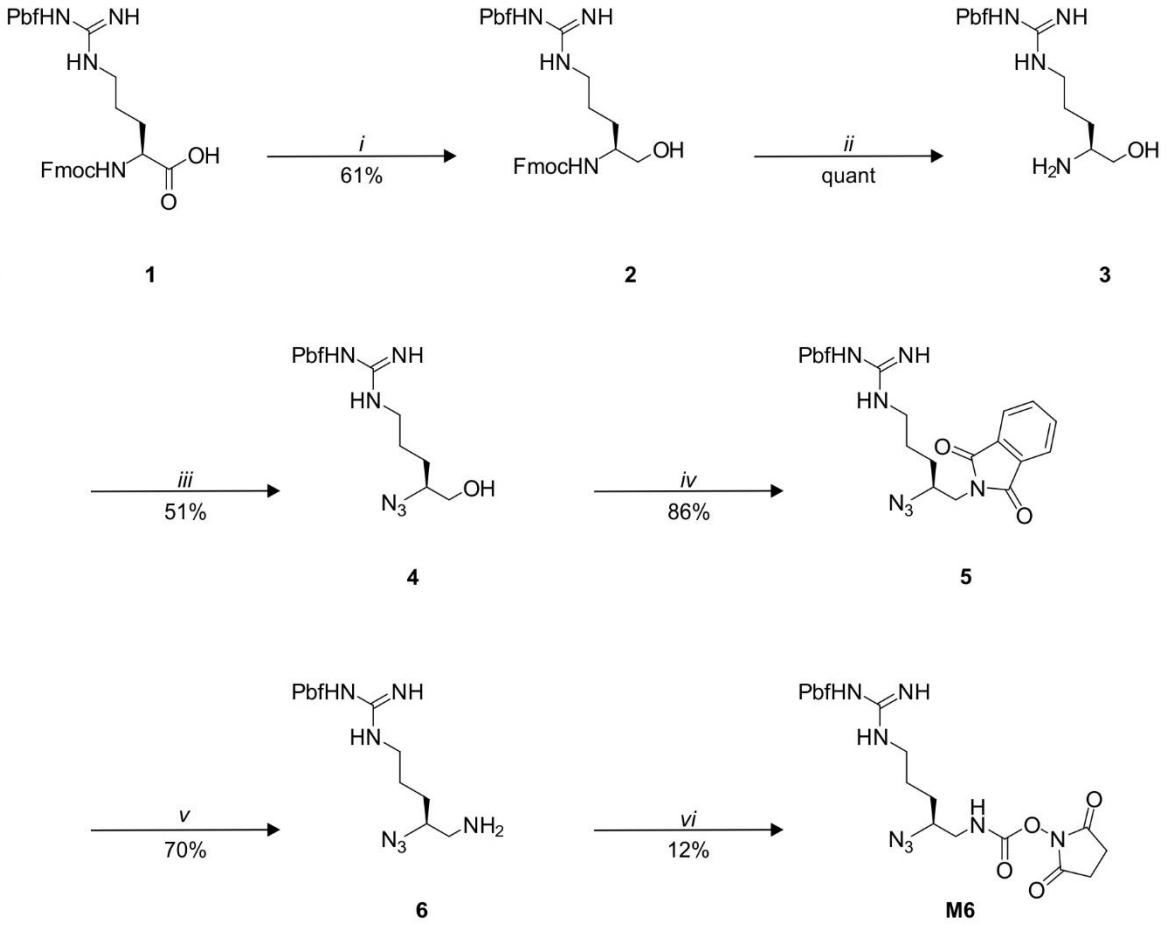
Oligomers were prepared on a 50 - 75 µmol scale. Polyethylene glycol resin (loading 0.45 mmol.g^{-1}) or polystyrene rink amide resin (loading 0.52 mmol.$^{g-1}$), was put in a polypropylene SPE tube (CEM) and preswelled with DMF. All steps were performed under inert atmosphere and microwave irradiation on DiscoverBio (CEM) system. The temperature was maintained by modulation of power and controlled with a fiber optic sensor. Fmoc-Xaa-OH (5 equiv. relative to the resin loading), DIC (5 equiv. relative to resin loading), HOBt (5 equiv. relative to resin loading) were dissolved in DMF. The mixture was added into the reaction vessel (CEM) and placed inside the microwave reactor (CEM DiscoverBio). The vessel was then irradiated (25 W, 75 °C, 5 min). After that, the resin was filtered and washed with DMF (6 x 2 mL). Fmoc was deprotected with 20% of piperidine in DMF, under microwave irradiation (50 W, 50 °C, 8 min), the deprotection steps was performed twice systematically. Azide protected succinimidyl carbamate N_3-Xaa^U-OSu monomer (1.5 equiv. relative to resin loading) and DIPEA (3 equiv. relative to resin loading) were dissolved in DMF (3 mL), added to the reaction vessel, placed inside the microwave reaction and irradiated (50 W, 50 °C, 20 min). The coupling was performed twice systematically. The resin was filtered and washed with DMF (6 x 2 mL). The reduction of the azide group was performed in a mixture of 1,4-dioxane:H_2O (7:3, v/v), so before the reaction, the resin was washed with this mixture of solvents. Then the Staudinger reaction was performed under microwave irradiation (50 W, 50 °C, 15 min) in 1,4-dioxane:H_2O and with 1M PMe_3 solution in THF (10 equiv. relative to resin loading) as the reductive agent. A double deprotection step was performed systematically. After the reaction, the resin was filtered and washed with 1,4-dioxane:H_2O (1 x 2 mL) and DMF (6 x 2 mL). The coupling and reduction efficiency were monitored with the Kaiser test for amino acid and chloranil test for oligourea[1]. Resin was filtered off, washed with DMF (4 x 2 mL), DCM (4 x 2 mL), diethyl ether (2 x 2 mL) and dried. The resin was finally swelled in a mixture of TFA/TIS/H_2O (95:2.5:2.5, v/v/v) and left to react for 240 min. Then, the resin was filtered off, washed with the cleavage mixture and the filtrate was evaporated under reduced pressure. After precipitation with cold diethylether, crude oligomers were purified by preparative or semi-preparative RP-HPLC using the appropriate gradient and the solvent was removed by lyophilization. This procedure provided the desired compounds with a purity ≥90%.

III. Experimental data of Chapter 2

1. Preparation of the azide protected O-succinimidyl carbamates **M1-M11**

Activated monomers **M1, M2, M3**, **M4** and **M5** were prepared using a previously described procedure[2]. Monomer **M8** was prepared using a procedure previously described for D activated monomers[3].

*a. Synthesis of O-succinimidyl carbamate **M6***

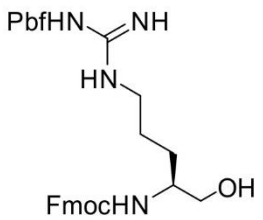


i: 1) IBCF, NMM, THF, -20 °C ; 2) $NaBH_4$, H_2O ; *ii*: DBU, THF, rt, *iii*: $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4.5H_2O$, CH_3CN, H_2O, rt; *iv:* DIAD, PPh_3, phtalimide, THF, rt:; *v*: N_2H_4, MeOH, rt; *vi*: DSC, AcOEt, rt

Scheme S1: Synthesis of activated monomer **M6**.

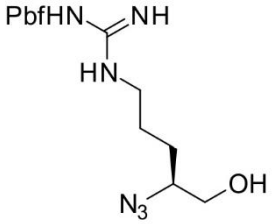
(S)-(9H-fluoren-9-yl)methyl(1-hydroxy-5-(3-((2,2,4,6,7-pentamethyl-2,3 dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentan-2-yl)carbamate 2

Fmoc-L-Arg(Pbf)-OH protected α-amino acid (1 equiv., 5 g, 7.70 mmol) was dissolved in anhydrous THF under N_2 atmosphere and cooled to -10 °C. *N*-methylmorpholine (1.2 equiv., 1.0 mL, 9.2 mmol), isobutylchloroformate (1.2 equiv., 1.2 mL, 9.2 mmol) were successively



added and the reaction mixture was stirred at -10 °C for 30 min. Resulting white suspension was then filtered off and a solution of $NaBH_4$ (2 equiv., 0.58 g,15.4 mmol) in water (10 mL) was added dropwise to the filtrate. The reaction was left to react 1 h and then concentrated under reduced pressure. AcOEt was added and the organic layer was washed with 1M $KHSO_4$ solution (x2) and brine, dried over magnesium sulfate and concentrated under vacuum. N-Fmoc protected amino acid **2** was obtained as a white foam (3.12 g, 61%) after purification by flash chromatography (cyclohexane/AcOEt 70%) **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 6.14 (brs, 3H), 5.45 (brs, 1H), 4.35-4.37 (m, 2H), 4.10-4.17 (m, 2H), 3.50-3.70 (m, 3H), 3.12-3.26 (m, 2H), 2.90 (s, 2H), 2.56 (s, 3H), 2.49 (s, 3H), 2.07 (s, 3H), 2.04 (s, 3H), 1.42 (s, 3H).

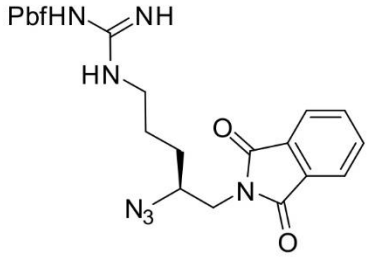
(S)-*N*-(*N*-(4-amino-5-hydroxypentyl)carbamimidoyl)-2,2,4,6,7-pentamethyl-2,3 dihydrobenzofuran-5-sulfonamide 4



2 (1 equiv., 3.12 g, 4.91 mmol) was dissolved in anhydrous THF under N_2 atmosphere. DBU (1.25 equiv., 0.92 mL, 6.13 mmol) was added dropwise and the reaction was stirred 30 min at room temperature. THF was removed under reduced pressure and the amine was directly engaged in the next step without further purification. To a solution of amine in ACN/H_2O mixture (1:1, 50 mL) were successively added K_2CO_3 (n x Acid + 0.05 mmol, 0.881 g), $CuSO_4.5H_2O$ (0.01 equiv., 124 mg, 49,1 µmol) and imidazole-1-sulfonyl azide hydrochloride (1.2 equiv.,1.23 g, 5.89 mmol). The reaction mixture was stirred at room temperature for 3 h then concentrated under reduced pressure and re-dissolved in AcOEt (50 mL). The organic layer was then washed with $KHSO_4$ four times, twice with brine, dried over magnesium sulfate and concentrated under vacuum to afford after silica gel flash chromatography (90% AcOEt in cyclohexane) the azide **4** as a pure white foam with 51% yield (1.06 g). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 6.25 (brs, 3H), 6.06 (brs, 1H), 3.66 (dd, *J* = 4 Hz, 11 Hz, 2H), 3.59 (dd, *J* = 6 Hz, 11 Hz, 2H), 3.42-3.51 (m, 2H), 3.13-3.37 (m, 3H), 2.93 (s, 2H), 2.56 (s, 3H), 2.49 (s, 3H), 2.09 (s, 3H), 2.04 (s, 3H), 1.46 (s, 3H).

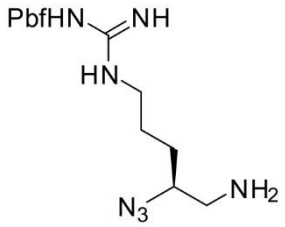
(S)-N-(N-(4-azido-5-(1,3-dioxoisoindolin-2-yl)pentyl)carbamimidoyl)-2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonamide 5.

To a solution of triphenylphosphine (1.2 equiv., 0.76 g, 2.9 mmol) and phtalimide (0.43 g, 2.9 mmol) in anhydrous THF (20 mL) at 0°C and under N_2 atmospheric pressure, was added DIAD



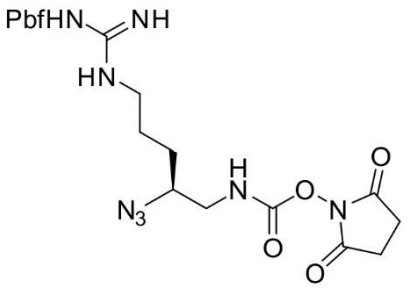
(1.2 equiv., 0.41 mL, 2.9 mmol). A solution of **4** in anhydrous THF was next added dropwise to the reaction mixture and was stirred overnight at room temperature. THF was evaporated under reduced pressure and the crude material was purified by silical gel flash chromatography (DCM 100%, then DCM/MeOH 99:1) to afford **5** as a white foam with 86% yield (1.18 g). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 7.87 (dd, *J* = 3 Hz, 5 Hz, 2H), 7.75 (dd, *J* = 3 Hz, 5 Hz, 2H), 6.02 (brs, 3H), 3.67-3.82 (m, 5H), 3.22-3.82 (m, 5H), 3.22 (dd, *J* = 6 Hz, 12 Hz, 4H), 2.95 (s, 2H), 2.58 (s, 3H), 2.52 (s, 3H), 2.09 (s, 3H), 1.45 (s, 6H).

(S)-*N*-(*N*-(5-amino-4-azidopentyl)carbamimidoyl)-2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-sulfonamide 6.



To a solution of **5** dissolved in methanol (40 mL) was added to a hydrazine hydrate solution at 64% (3 equiv. 0.45 mL, 9.75 mmol). The reaction mixture was left stirring at room temperature overnight and a white precipitate appeared. The reaction mixture was filtered off, washed with ethyl acetate and concentrated under reduced pressure. Then it was dissolved in AcOEt and the organic layer was washed twice with a solution of saturated $NaHCO_3$, once with water, dried over magnesium sulfate and concentrated. **6** was obtained as white foam with 70% yield without any further purification. **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 6.42 (brs, 2H), 6.20-6.36 (m, 3H), 3.62-3.73 (m, 1H), 3.12-3.36 (m, 4H), 2.98-3.08 (m, 1H), 2.94 (s, 2H), 2.56 (s, 3H), 2.49 (s, 3H), 2.09 (s, 3H), 1.58-1.75 (m, 4H), 1.45 (s, 6H).

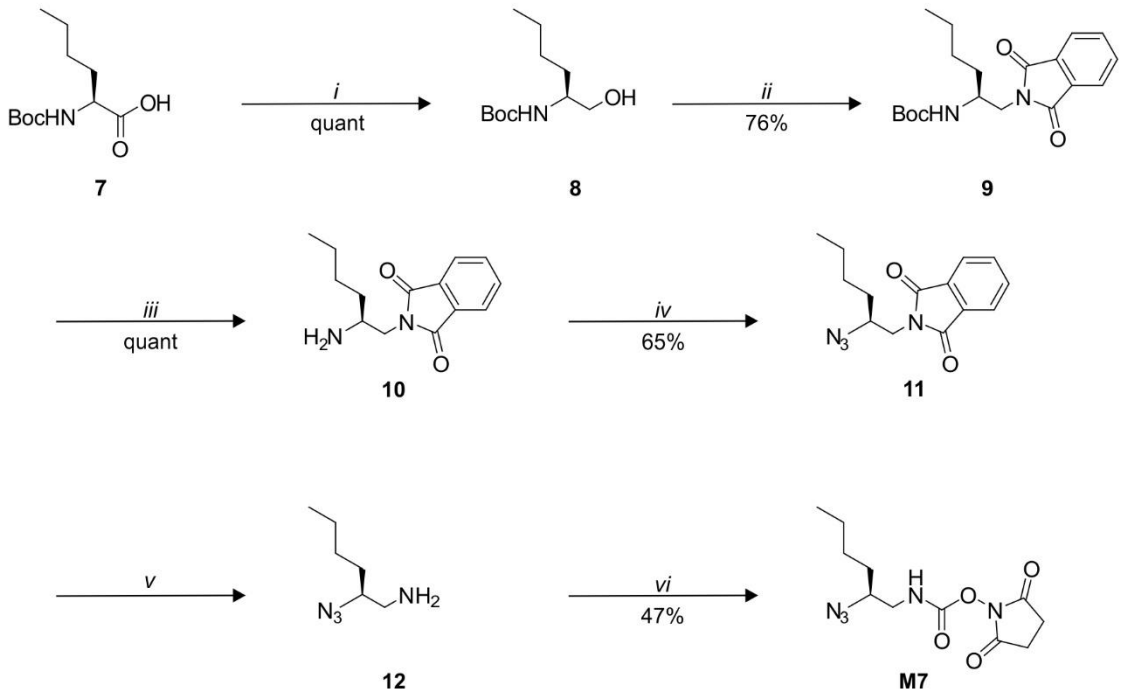
(S)-2,5-dioxopyrrolidin-1-yl(2-azido-5-(3-((2,2,4,6,7-pentamethyl-2,3-dihydrobenzofuran-5-yl)sulfonyl)guanidino)pentyl)carbamate M6.



To a suspension of DSC (0.44 g, 1.74 mmol) in anhydrous DCM (10 mL) was added dropwise a solution of **6** in DCM (10 mL).The reaction mixture was stirred overnight at room temperature under N_2 atmosphere. DCM was evaporated and the crude material was dissolved in AcOEt. The organic layer was washed with $KHSO_4$ and brine, dried over magnesium sulfate and concentrated. **M6** was obtained as white foam with 12% yield after purification by flash chromatography (DCM/AcOEt 10% up to 50%). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 6.33

(brs, 2H), 6.22 (t, *J* = 6 Hz, 2H), 3.58-3.69 (m, 3H), 3.19-3.50 (m, 6H), 2.95 (s, 2H), 2.82 (s, 2H), 2.59 (s, 3H), 2.52 (s, 3H), 2.10 (s, 3H), 1.46 (s, 3H), 1.42 (s, 3H); **^{13}C NMR ($CDCl_3$, 75 MHz)** δ (ppm): 170.5, 159.0, 156.4, 152.4, 138.6, 132.7, 124.9, 117.7, 86.6, 61.1, 60.5, 44.9, 43.3, 40.4, 28.7, 28.5. **HRMS (ESI-TOFMS)** m/z cald for $C_{24}H_{35}N_8O_7S$ $[M+H]^+$, 579.23494, found 579.23348.

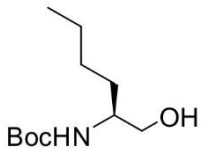
*b. Synthesis of O-succinimidyl carbamate **M7***



i: 1) IBCF, NMM, THF, -20 °C ; 2) $NaBH_4$, H_2O ; *ii*: DIAD, PPh_3, phtalimide, THF, rt; *iii*: TFA, TA ; *iv:* $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4.5H_2O$, CH_3CN, H_2O, rt; *v:* N_2H_4, MeOH, rt; *vi:* DSC, AcOEt, rt

Scheme S2: Synthesis of activated monomer M7

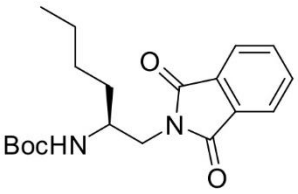
(S)-*tert*-butyl (1-hydroxyhex-2-yl)carbamate 8



Boc-L-Nle-OH protected α-amino acid (1 equiv., 5 g, 21.61 mmol) was dissolved in anhydrous THF under N_2 atmosphere and cooled to -10 °C. *N*-methylmorpholine (1.2 equiv., 2.84 mL, 25.93 mmol), isobutylchloroformate (1.2 equiv., 3.26 mL, 25.93 mmol) were successively added and the reaction mixture was stirred at -10 °C for 30 min. Resulting white suspension was then filtered off and a solution of $NaBH_4$ (2 equiv., 1.70 g, 43.22 mmol) in water (20 mL) was added dropwise to the filtrate. The reaction was left to react overnight and then concentrated under reduced pressure. AcOEt was added and the organic layer was washed with 1M $KHSO_4$ solution (x3), saturated $NaHCO_3$ (x3) and brine (x1), dried over magnesium sulfate and concentrated under vacuum. N-Boc

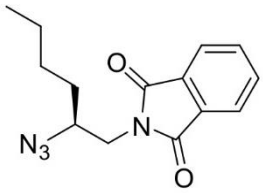
protected amino acid **8** was obtained as an oil (quant.). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 4.68 (brs, 1H), 3.48-3.68 (m, 3H), 2.95 (brs, 1H), 1.43 (s, 9H), 1.27-1.31 (m, 5H), 0.85-0.92 (m, 4H).

(S)-*tert*-butyl (1-(1,3-dioxoisoindolin-2-yl)hexan-2-yl)carbamate 9



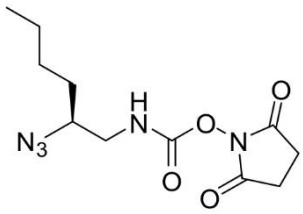
N-Boc protected amino alohol **8** (1 equiv. 21.61 mmol) was dissolved in anhydrous THF under N_2 atmosphere. Triphenylphosphine (1.2 equiv., 6.76 g, 25.78 mmol) and phtalimide (1.2 equiv., 3.79 g, 25.78 mmol) were added, followed by dropwise addition of DIAD (1.2 equiv., 5.09 mL, 25.78 mmol). The reaction mixture was stirred for 30 min at 0 °C and was then allowed to reach room temperature. The reaction was completed after 12 h of stirring. THF was evaporated under reduced pressure and the crude material was purified on silica gel flash chromatography (petroleum ether/AcOEt 10%) and **9** was obtained as pure compound (5.63 g, 76% after two steps).**^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 7.81 (dd, *J* = 3 Hz, 5 Hz, 2H), 7.67 (dd, *J* = 3 Hz, 5 Hz, 2H), 4.53 (d, *J* = 9 Hz, 1H), 3.87-4.03 (m, 1H), 3.59-3.73 (m, 2H), 1.28-1.38 (m, 6H), 1.19 (s, 9H), 0.86-0.89 (m, 3H).

(S)-2-(2-azidohexyl)isoindoline-1,3-dione 11



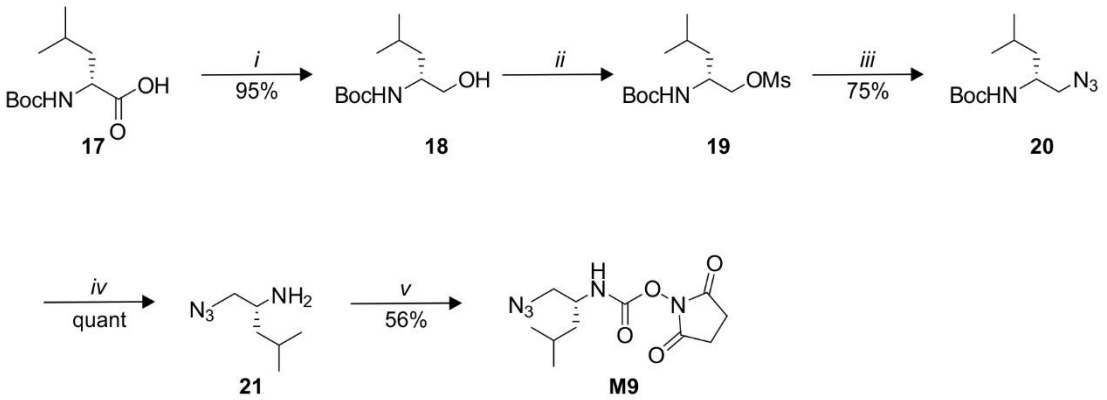
9 was dissolved in pure TFA (10 mL) and left to react without stirring for 15 min. TFA was removed under reduced pressure and the resulting TFA salt was precipitated in cold Et_2O, filtered off and dried on vacuum line. Nor purification, nor isolation of the TFA salt alone were performed, and it was directely engaged in the next step. To a solution of amine in ACN/H_2O mixture (1:1, 50 mL) were successively added K_2CO_3 (n x Acid + 0.05 mmol, 2.76 g), $CuSO_4.5H_2O$ (0.01 equiv., 40.6 mg, 0.162 mmol) and imidazole-1-sulfonyl azide hydrochloride (1.2 equiv.,4.08 g, 19.51 mmol). The reaction mixture was stirred at room temperature overnight, then concentrated under reduced pressure and dissolved in AcOEt (50 mL). The organic layer was then washed with $KHSO_4$ four times, twice with brine, dried over magnesium sulfate and concentrated under vacuum to furnish after silica gel flash chromatography (10% AcOEt in cyclohexane), the azide **11** as an oil (2.89 g, 65% over two steps). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 7.87 (dd, *J* = 3 Hz, 5 Hz, 2H), 7.73 (dd, *J* = 3 Hz, 5 Hz, 2H), 3.65-3.85 (m, 3H), 1.33-1.61 (m, 6H), 0.92 (t, *J* = 7 Hz, 3H).

(S)-2,5-dioxopyrrolidin-1-yl (2-azidohexyl)carbamate M7



To a solution of azido phtalimide derivative **11** (1 equiv., 2.89 g,10.2 mmol) in MeOH (30 mL) was added hydrazine hydrate 55% (2.5 equiv., 2.34 mL, 48.27 mmol). The reaction mixture was left to react overnight at room temperature and a white precipitate appeared. The reaction mixture was filtered off, washed with AcOEt and the filtrate was concentrated under reduced pressure. The crude material was dissolved in AcOEt (30 mL). The organic layer was washed with water twice (2 x 15 mL), dried over magnesium sulfate and resulting azido amine was directely engaged for activation step without further purification. To a stirred suspension of DSC (1.2 equiv., 3.26 g, 12.74 mmol) in AcOEt, was added dropwise a solution of azido amine derivative **12**. The reaction was left to react 4 h and a solution of 1M $KHSO_4$ was added. The organic layer was extracted, washed one time with brine, dried over magnesium sulfate and concentrated under vacuum. Resulting crude was purified by precipitation in Et_2O to activated carbamate **M7** as a white solid with 47% yield after 2 steps (1.42 g). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 5.70 (brs, 1H), 3.53-3.60 (m, 1H), 3.47 (ddd, *J* = 3 Hz, 6 Hz, 13 Hz, 1H), 3.14 (ddd, *J* = 5 Hz, 7 Hz, 13 Hz, 1H), 2.82 (s, 4H), 1.56-1.63 (m, 2H), 1.33-1.45 (m, 4H), 0.92 (t, *J* = 7 Hz, 3H). **^{13}C NMR ($CDCl_3$, 75 MHz)** δ (ppm): 169.9, 151.7, 62.0, 45.5, 31.5, 28.0, 25.7, 22.4, 14.0. **HRMS (ESI-TOFMS)** m/z cald for $C_{11}H_{18}N_5O_4$ $[M+H]^+$, 284.13588, found 284.13537.

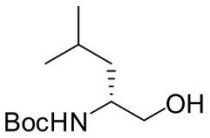
*c. Synthesis of O-succinimidyl carbamate **M9***



i: 1) IBCF, NMM, THF, -20 °C; 2) $NaBH_4$, H_2O; *ii*: MsCl, Et_3N, DCM, 0 °C, Ar_{atm}; *iii*: NaN_3, DMF, 70 °C; iv: TFA, rt; *v:* DSC, DIPEA, DCM, rt

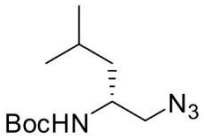
Scheme S3: Synthesis of O-succinimidyl carbamates monomers **M9**

(R) *tert*-butoxycarbonylaminopentanol 18.



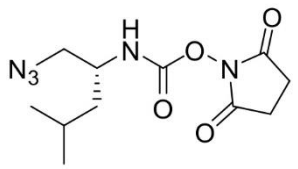
Boc-D-Leu-OH (1 g, 4.0 mmol) was dissolved in 15 mL of anhydrous THF and cooled at -10 °C. *N*-methylmorpholine (0.53 mL, 4.8 mmol), isobutylchloroformate (0.63 mL, 4.8 mmol) were successively added and the reaction mixture was stirred at -10 °C under nitrogen atmosphere for 40 min. The reaction mixture was filtered off and added dropwise to the solution of $NaBH_4$ in water at 0 °C and stirred overnight at room temperature. THF was evaporated and AcOEt was added, the organic layer was washed with $NaHCO_3$ and brine, dried over magnesium sulfate and concentrated under vacuum. **18** was obtained as colourless oil (0.83 g) with 95% yield without purification. **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 4.61 (brs, 1H), 3.67-3.74 (m, 1H), 3.61-3.66 (m, 1H), 3.48 (dd, *J* = 5 Hz, 11 Hz, 1H), 1.58-1.71 (m, 1H), 1.43 (s, 9H), 1.27-1.32 (m, 2H), 0.91-0.92 (m, 3H), 0.89-0.90 (m, 3H).

***tert*-butyl (R)-1-azidopentan-2-ylcarbamate 20**



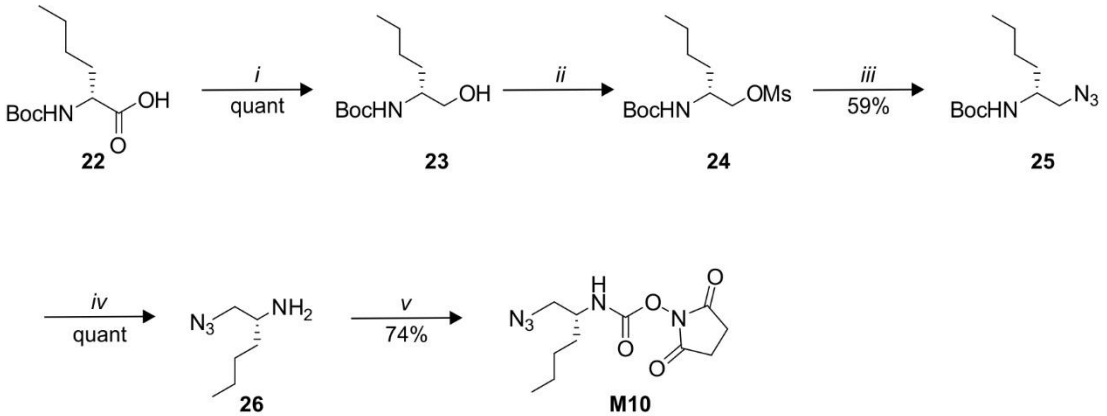
To a stirred solution of **18** (0.83 g, 1 equiv., 3.81 mmol) in dry DCM (30 mL) at 0° C, distilled Et_3N (2.0 equiv., 0.93 mL, 12.03 mmol) was added and after 10 min, mesylchloride (1.2 equiv., 0.72 mL, 9.62 mmol) was added dropwise. Stirring was maintained for 2 h at 0 °C. The reaction mixture was then diluted with brine and the organic layer was dried over $MgSO_4$. Solvent was removed under reduced pressure and the resulting oil was engaged as such in the next step without further purification. Mesylate derivative was then dissolved in DMF (50 mL) and NaN_3 (5 equiv., 2.6 g, 40.1 mmol) was added and the reaction mixture was heated to 70 °C for 5 hours and stirring was maintained overnight at room temperature. The reaction mixture was then diluted with water and extracted several times with diethyl ether. The combined organic layers were then dried over $MgSO_4$ and solvent was removed under reduced pressure. Purification with flash chromatography in EtOAc/cyclohexane (3:7) affored pure azide **20** (0.70 g, 75% over three steps) as a yellow oil. **^{1}H NMR (300 MHz, $CDCl_3$)** δ (ppm): 4.49 (brs, 1H), 3.71-3.86 (m, 1H), 3.43 (dd, *J* = 4 Hz, 12 Hz, 1H), 3.30 (dd, *J* = 4 Hz, 12 Hz, 1H), 1.57-1.71 (m, 1H), 1.44 (s, 9H), 1.28-1.37 (m, 2H), 0.91 (d, *J* = 6 Hz, 6H).

(R)-2,5-dioxopyrrolidin-2-yl (1-azido-4-methylpentyl) carbamate M9



Boc azide **20** (0.70 g, 2.88 mmol) was dissolved in pure TFA (3 mL), and the reaction was left 5 min without stirring. TFA was evaporated and the organic layer was extracted with ether and concentrated under vacuum. Resulting TFA salt was dissolved in DCM (15 mL) at 0 °C and DIPEA (2 equiv., 1.0 mL, 5.76 mmol) was added drop wise. After 15 min, DSC (1.2 equiv., 0.885 g, 3.459 mmol) was added and stirring was maintained overnight. The crude mixture was then washed with 1N $KHSO_4$ and dried over magnesium sulfate. After concentration under reduced pressure and purification by precipitation in cold Et_2O, carbamate **M9** (0.46 g, 56% after two steps) was obtained as white solid. **^{1}H NMR (300 MHz, $CDCl_3$)** δ (ppm): 5.22 (d, *J* = 8 Hz, 1H), 3.81-3.91 (m, 1H), 3.55 (dd, *J* = 4 Hz, 12 Hz, 1H), 3.45 (dd, *J* = 4 Hz, 12 Hz, 1H), 2.82 (s, 4H), 1.62-1.75 (m, 1H), 1.51 (ddd, *J* = 5 Hz, 8 Hz, 13 Hz, 1H), 1.38 (ddd, *J* = 5 Hz, 8 Hz, 14 Hz, 1H), 0.94 (d, *J* = 6 Hz, 6H).**^{13}C NMR (75 MHz, $CDCl_3$)** δ (ppm): 169.7, 151.1, 67.7, 54.6, 50.4, 41.2, 25.6, 24,8, 23.0, 22.1, 20.3.

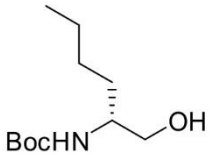
*d. Synthesis of O-succinimidyl carbamate **M10***



i: 1) IBCF, NMM, THF, -20 °C; 2) $NaBH_4$, H_2O ; *ii*: MsCl, Et_3N, DCM, 0 °C, Ar_{atm}; *iii*: NaN_3, DMF, 70 °C; iv:TFA, rt; *v:* DSC, DIPEA, DCM, r.t.

Scheme S4: Synthesis of O-succinimidyl carbamates monomers **M10**

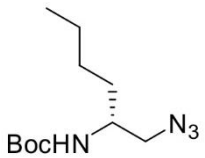
(R)-*tert*-butoxycarbonylaminohexanol 23.



Boc-D-Nle-OH (1 equiv., 2 g, 8.65 mmol) was dissolved in 15 mL of anhydrous THF and cooled at -10 °C. *N*-methylmorpholine (1.2 equiv., 1.14 mL, 10.38 mmol), isobutylchloroformate (1.2 equiv., 1.35 mL, 10.38 mmol) were successively added and the reaction mixture was stirred at -10 °C under nitrogen atmosphere for 30 min. The reaction mixture was filtered off and added

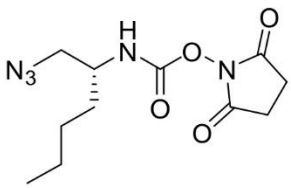
dropwise to the solution of $NaBH_4$ (2 equiv., 0.65 g, 17.30 mmol) in water at 0 °C and stirred overnight at room temperature. THF was evaporated and AcOEt was added, the organic layer was washed with $NaHCO_3$ and brine, dried over magnesium sulfate and concentrated under vacuum. **23** was obtained as colourless oil (0.83 g).**^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 4.65 (brs, 1H), 3.48-3.72 (m, 3H), 2.42 (bs, 1H), 1.42 (s, 9H), 1.27-1.35 (m, 6H), 0.88 (s, 3H).

(R)-*tert*-butyl (1-azidohexan-2-yl)carbamate 25



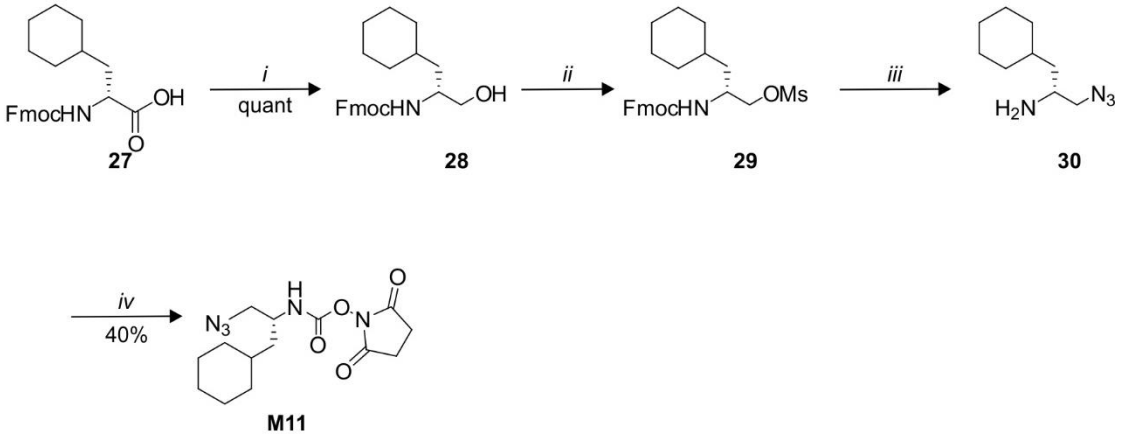
To a stirred solution of **23** (1 equiv., 8.65 mmol) in dry DCM (50 mL) at 0 °C, distilled Et_3N (2.0 equiv., 1.30 mL, 17.30 mmol) was added and after 10 min, mesylchloride (1.5 equiv., 1.0 mL, 12.97 mmol) was added dropwise. Stirring was maintained for 2 h at 0 °C. The reaction mixture was then diluted with brine and the organic layer was dried over $MgSO_4$. Solvent was removed under reduced pressure and the resulting oil was engaged as such in the next step without further purification. Mesylate derivative was then dissolved in DMF (50 mL) and NaN_3 (5 equiv., 2.81 g, 43.25 mmol) was added and the reaction mixture was heated to 70 °C for 5 hours and stirring was maintained overnight at room temperature. The reaction mixture was then diluted with water and extracted several times with diethyl ether. The combined organic layers were then dried over $MgSO_4$ and solvent was removed under reduced pressure. Purification with flash chromatography in AcOEt/Cyclohexane (3:7) was performed to afford pure azide **25** (1.24 g, 59% over three steps) as a yellow oil. **^{1}H NMR (300 MHz, $CDCl_3$)** δ (ppm): 4.53 (brs, 1H), 3.56-3.76 (m, 1H), 3.52 (dd, *J* = 4 Hz, 12 Hz, 1H), 3.41 (dd, *J* = 5 z, 12 Hz, 1H), 1.43 (s, 9H), 1.25-1.37 (m, 6H), 0.89 (s, 3H).

(R)-2,5-dioxopyrrolidin-1-yl (1-azidohexan-2-yl)carbamate M10



Boc azide **25** was dissolved in pure TFA (6 mL), and the reaction was left 5 min without stirring. TFA was evaporated and the organic layer was extracted with ether and concentrated under vacuum. Resulting TFA salt was dissolved in DCM (20 mL) at 0 °C and DIPEA (2 equiv., 1.78 mL, 10.24 mmol) was added drop wise. After 15 min, DSC (1.2 equiv., 1.57 g, 6.14 mmol) was added and stirring was maintained 2 h. The crude mixture was then washed with 1N $KHSO_4$ and dried over magnesium sulfate. After concentration under reduced pressure and purification by precipipation in a solution of cold Et_2O/pentane, carbamate **M10** (1.08 g, 74% after two steps) was obtained as white solid. **^{1}H NMR (300 MHz, $CDCl_3$)** δ (ppm): 5.23 (s, 1H), 3.71-3.81 (m, 1H), 3.55 (dd, *J* = 4 Hz, 12 Hz, 1H), 3.47 (dd, *J* = 4 Hz, 12 Hz, 1H), 2.82 (s, 4H), 11.28-1.40 (m, 6H), 0.91 (s, 3H).**^{13}C NMR (75 MHz, $CDCl_3$)** δ (ppm): 169.7, 151.3, 54.3, 52.2, 32.0, 28.1, 25.9, 22.5, 14.0.

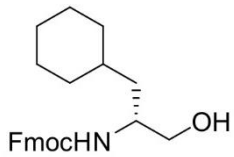
*e. Synthesis of O-succinimidyl carbamate **M11***



i: 1) IBCF, NMM, THF, -20 °C; 2) $NaBH_4$, H_2O; *ii*: MsCl, Et_3N, DCM, 0 °C, Ar_{atm}; *iii*: NaN_3, DMF, 70 °C; iv : DSC, DIPEA, DCM, r.t.

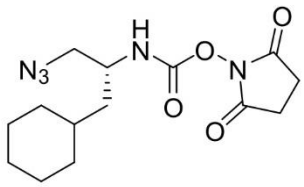
Scheme S5: Synthesis of O-succinimidyl carbamates monomers **M11**

(R)-(9H-fluoren-9-yl)methyl (1-cyclohexyl-3-hydroxypropan-2-yl)carbamate 28



Fmoc-D-Cha-OH (1 equiv., 2.43 g, 6.11 mmol) was dissolved in anhydrous THF under N_2 atmosphere and cooled to -10 °C. *N*-methylmorpholine (1.2 equiv., 0.80 mL, 7.33 mmol), isobutylchloroformate (1.2 equiv., 1.0 mL, 7.33 mmol) were successively added and the reaction mixture was stirred at -10 °C for 30 min. The resulting white suspension was then filtered off and a solution of $NaBH_4$ (2 equiv., 0.46 g, 12.22 mmol) in water (5 mL) was added dropwise to the filtrate. The reaction was left to react 30 min and then concentrated under reduced pressure. AcOEt was added and the organic layer was washed with 1M $KHSO_4$ solution (x3), saturated $NaHCO_3$ (x3), brine, dried over magnesium sulfate and concentrated under vacuum. N-Fmoc protected amino alcool **28** was obtained as a white solid (quant).**^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 7.76 (d, *J* = 7 Hz, 2H), 7.59 (d, *J* = 8 Hz, 2H), 7.42 (dd, *J* = 7 Hz, 14 Hz, 1H), 7.39 (dd, *J* = 7 Hz, 14 Hz, 1H) 7.33 (d, *J* = 7 Hz, 14 Hz 1H), 7.30 (dd, *J* = 9 Hz, 14 Hz, 1H), 4.73 (brs, 1H), 4.44 (brs, 1H), 4.23 (t, *J* = 7 Hz, 2H), 3.48-3.80 (m, 3H), 1.12-1.83 (m, 13 H).

(R)-2,5-dioxopyrrolidin-1-yl (1-azido-3-cyclohexylpropan-2-yl)carbamate M11

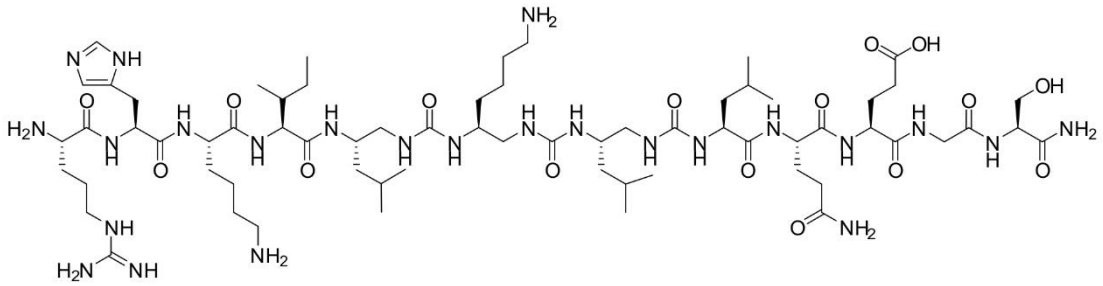


To a stirred solution of **28** (1 equiv., 6.11 mmol) in dry DCM (50 mL) at 0 °C, distilled Et_3N (2.0 equiv., 0.92 mL, 12.22 mmol) was added and after 10 min, mesylchloride (1.5 equiv., 0.70 mL, 9.16 mmol) was added dropwise. The stirring was maintained for 2 h at 0 °C. The reaction mixture was then diluted with brine and the organic layer was dried over $MgSO_4$. The solvent was removed under reduced pressure and the resulting oil was engaged

as such in the next step without further purification. The mesylate derivative was then dissolved in DMF (40 mL) and NaN_3 (5 equiv., 1.98 g, 30.55 mmol) was added and the reaction mixture was heated to 70 °C for 5 hours and the stirring was maintained overnight at room temperature. The reaction mixture was then diluted with water and extracted several times with diethyl ether. The combined organic layers were then treated with 1N HCl twice and the resulting aqueous layer was basified with solid $NaHCO_3$ until pH = 8, and the organic layer was extracted with DCM, dried over $MgSO_4$ and solvent was removed under reduced pressure. Crude material was dissolved in DCM (20 mL) at 0 °C, DSC (1.2 equiv., 1.20 g, 7.33 mmol) was added and stirring was maintained 2 h. The crude mixture was then washed with 1N $KHSO_4$, brine and dried over magnesium sulfate. After concentration under reduced pressure and purification by precipitation in a solution of cold Et_2O/pentane, carbamate **M11** (0.802 g, 40% after four steps) was obtained as white solid. **^{1}H NMR (300 MHz, $CDCl_3$)** δ (ppm): 5.21 (d, *J* = 8 Hz, 1H), 3.84-3.94 (m, 1H), 3.54 (dd, *J* = 4 Hz, 12 Hz, 1H), 3.43 (dd, *J* = 4 Hz, 12 Hz, 1H), 2.82 (s, 4H), 1.63-1.81 (m,5H), 1.10-1.53 (m, 6H) 0.79-1.08 (m, 2H).**^{13}C NMR (75 MHz, $CDCl_3$)** δ (ppm):170.0, 151.1, 66.0, 54.7, 49.8, 40.1, 34.1, 33.6, 32.9,, 26.5, 26.2, 26.1, 25.7, 15.4.

2. Synthesis of oligomers **C2-C21**

Compound **C2**



C2 was synthesized on 75 µmol scale following general procedure A with NovaPEG resin (loading 0.45 mmol.g^{-1}). The product was purified by semi-preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-50%, 20 min, flow = 4 mL.min^{-1}. 39,4 mg was obtained with a global yield of 35%. ESI-MS (M_W 1507.85): m/z 1508,80 $[M+H]^+$, 754.49 $[M+2H]^{2+}$, 503.32 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.03 min.

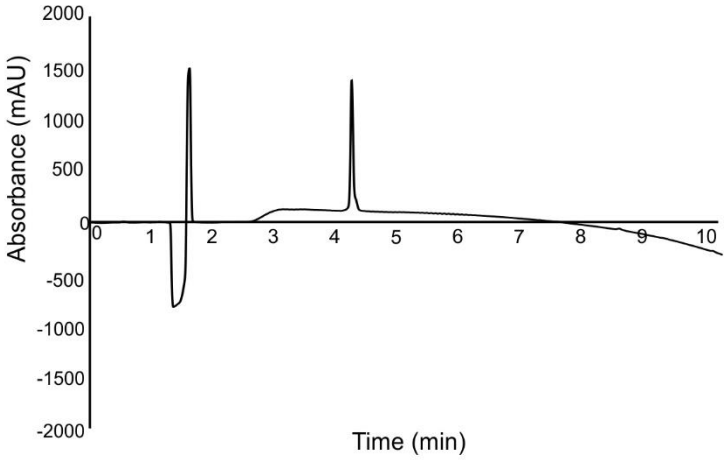
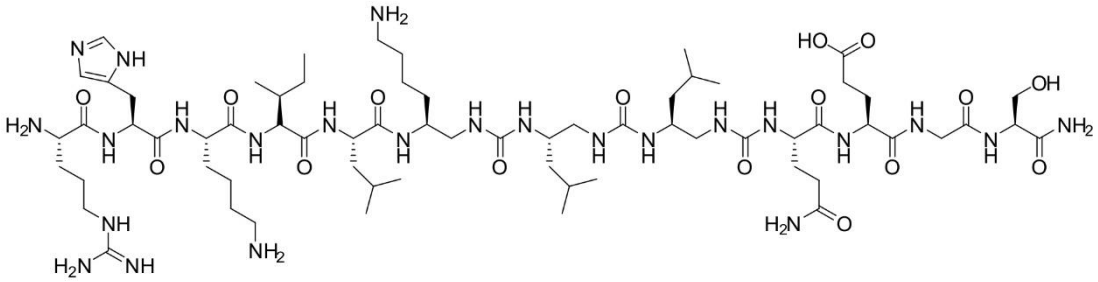


Figure S1: RP-HPLC chromatogram of **C2** after purification (λ = 200 nm).

Compound **C3**



C3 was synthesized on 75 µmol scale following general procedure A with NovaPEG resin (loading 0.45 mmol.g^{-1}). The product was purified by semi-preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-50%, 20 min, flow = 4 mL.min^{-1}. 22,7 mg was obtained with a global yield of 20%. ESI-MS (M_W 1507.85): m/z 1507,86 $[M+H]^+$, 754.48 $[M+2H]^{2+}$, 503.32 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.63 min.

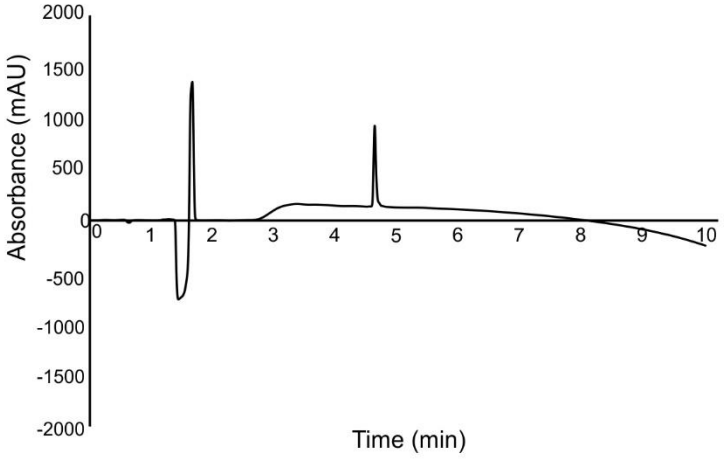
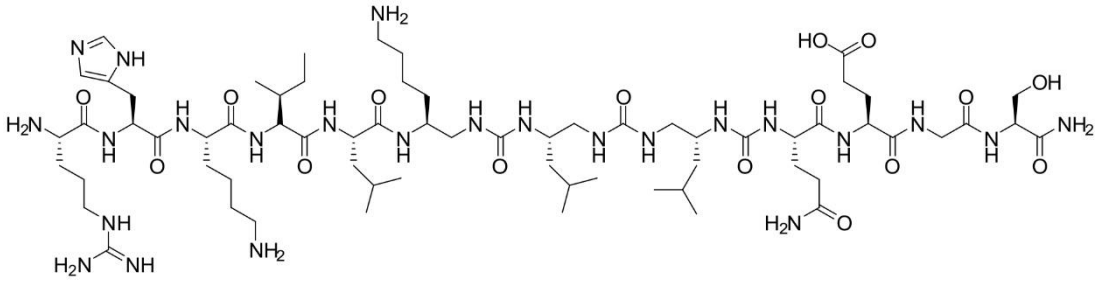


Figure S2: RP-HPLC chromatogram of **C3** after purification(λ = 200 nm).

Compound **C4**



C4 was synthesized on 75 µmol scale following general procedure A with NovaPEG resin (loading 0.45 mmol.g^{-1}). The product was purified by semi-preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-50%, 20 min, flow = 4 mL.min^{-1}. 23.2 mg was obtained with a global yield of 21%. ESI-MS (M_W 1507.85): m/z 1508,97 $[M+H]^+$, 754.61 $[M+2H]^{2+}$, 503.43 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.35 min.

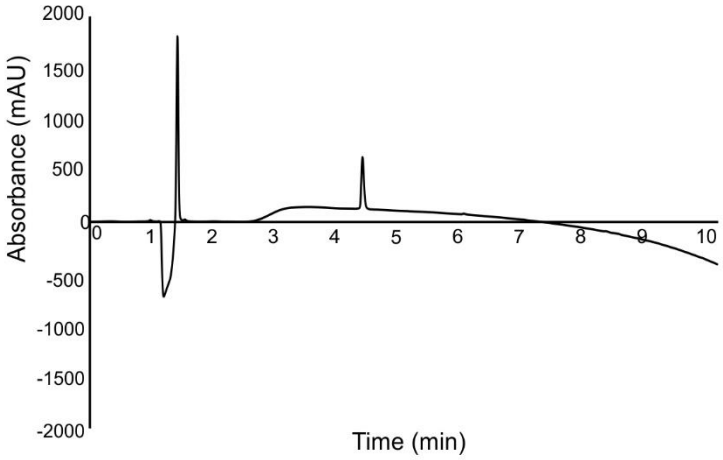
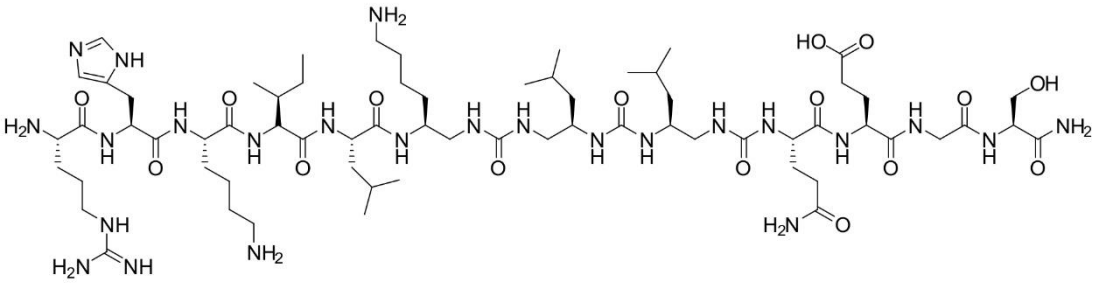


Figure S3: RP-HPLC chromatogram of **C4** after purification (λ = 200 nm).

Compound **C5**



C5 was synthesized on 75 µmol scale following general procedure A with NovaPEG resin (loading 0.45 mmol.g^{-1}). The product was purified by semi-preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-50%, 20 min, flow = 4 mL.min^{-1}. 27.9 mg was obtained with a global yield of 25%. ESI-MS (M_W 1507.85): m/z 1507,93 $[M+H]^+$, 754.93 $[M+2H]^{2+}$, 503.53 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.90 min.

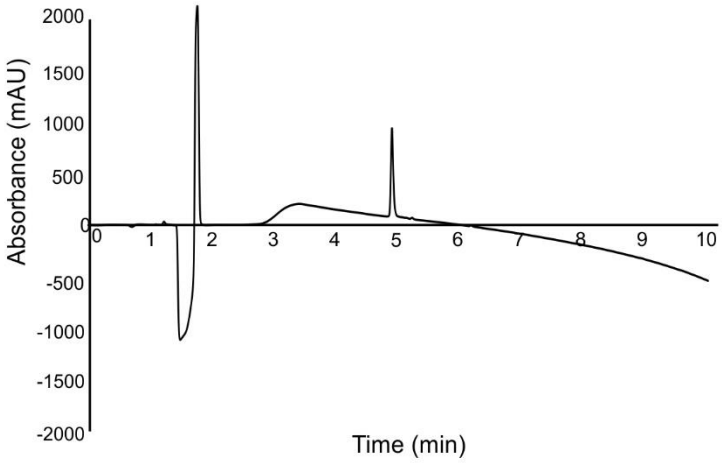
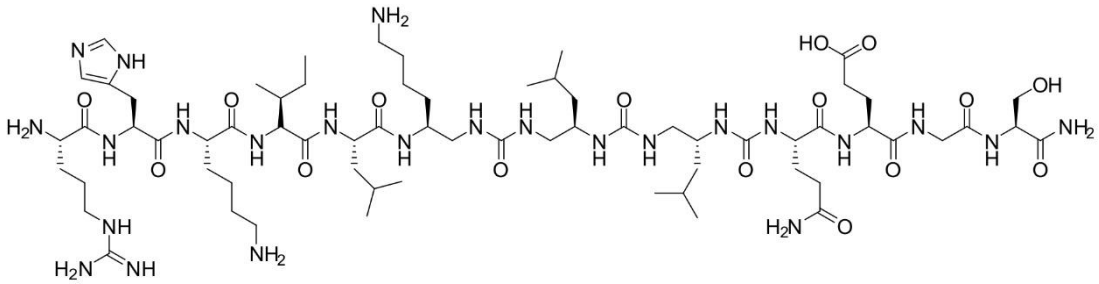


Figure S4: RP-HPLC chromatogram of **C5** after purification (λ = 200 nm).

Compound **C6**



C6 was synthesized on 75 µmol scale following general procedure A with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by semi-preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 4 mL.min^{-1}. 40.4 mg was obtained with a global yield of 36%. ESI-MS (M_W 1507.85): m/z 1508,86 $[M+H]^+$, 754.41 $[M+2H]^{2+}$, 503.27 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10 -100%, 10 min) t_r = 4.91 min.

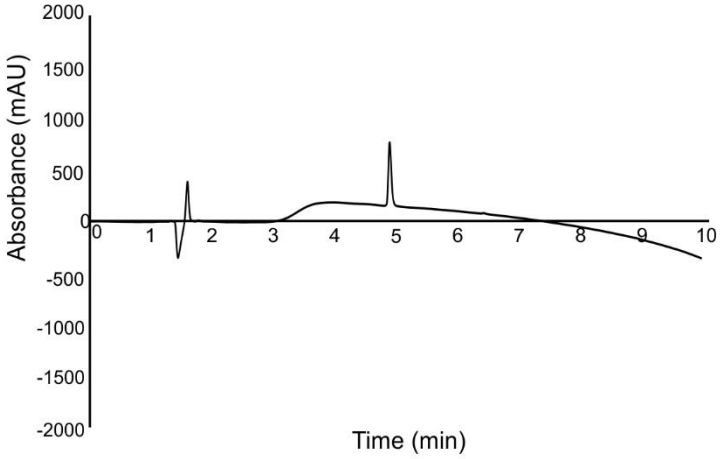
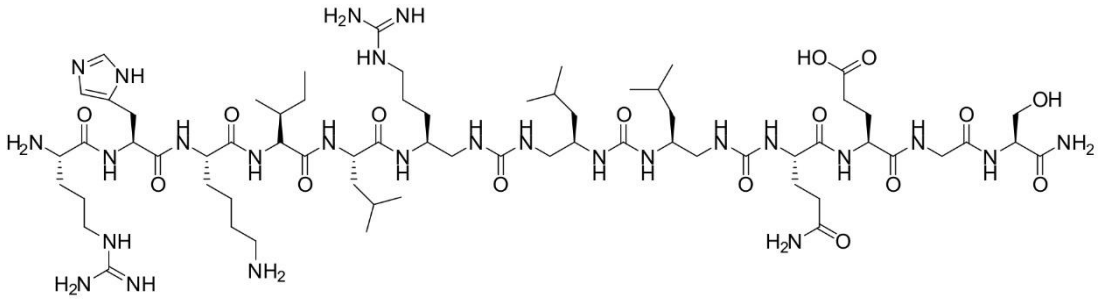


Figure S5: RP-HPLC chromatogram of **C6** after purification (λ = 200 nm).

Compound **C7**



C7 was synthesized on 75 µmol scale following general procedure A with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 56.1 mg was obtained with a global yield of 49%. ESI-MS (M_W 1535.87): m/z 1536,96 $[M+H]^+$, 768.48 $[M+2H]^{2+}$, 512.65 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10 -100%, 10 min) t_r = 4.53 min.

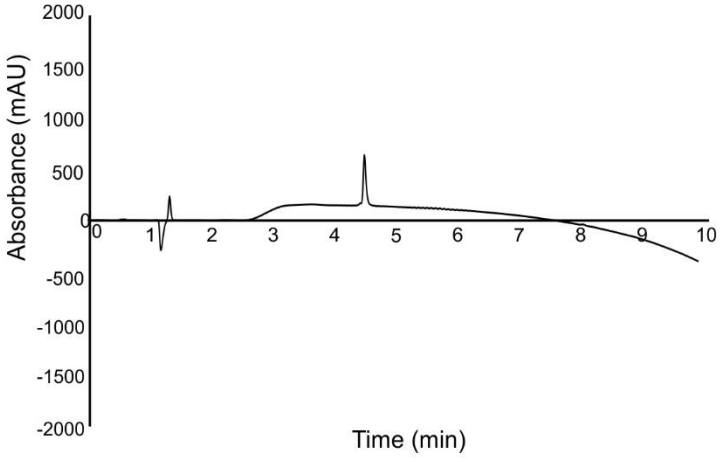
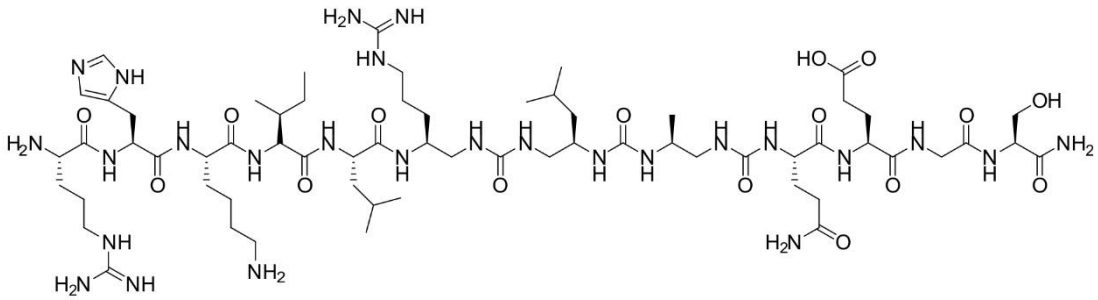


Figure S6: RP-HPLC chromatogram of **C7** after purification (λ = 200 nm).

Compound **C8**



C8 was synthesized on 75 µmol scale following general procedure A with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 63.2 mg was obtained with a global yield of 56%. ESI-MS (M_W 1493.79): m/z 1494,90 $[M+H]^+$, 747.45 $[M+2H]^{2+}$, 498.63 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.47 min.

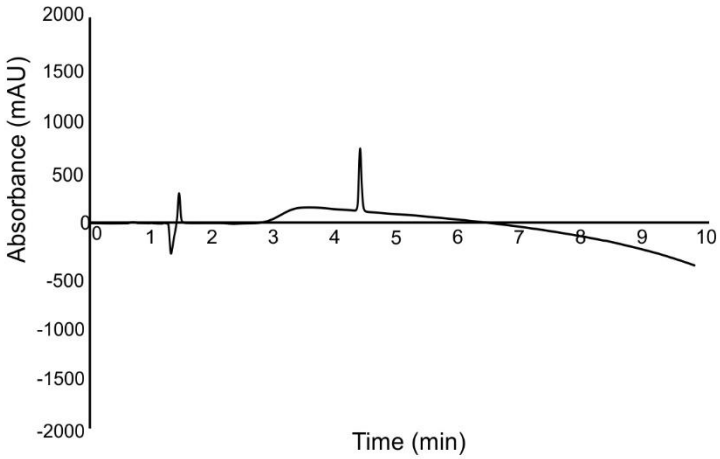
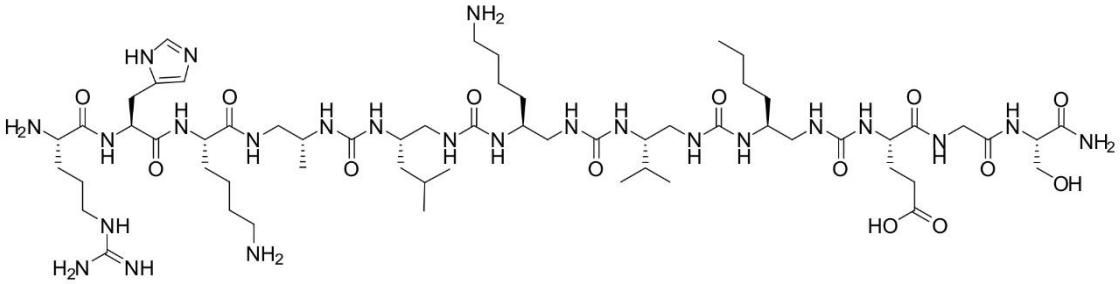


Figure S7: RP-HPLC chromatogram of **C8** after purification (λ = 200 nm).

Compound **C9**



C9 was synthesized on 50 µmol scale following general procedure A with NovaPEG resin (loading 0.45 mmol.g^{-1}). The product was purified by semi-preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-40%, 20 min, flow = 4 mL.min^{-1}. 3.4 mg was obtained with a global yield of 5%. ESI-MS (M_W 1380.88: m/z 1381.80 $[M+H]^+$, 691.47 $[M+2H]^{2+}$, 461.27 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 3.95 min.

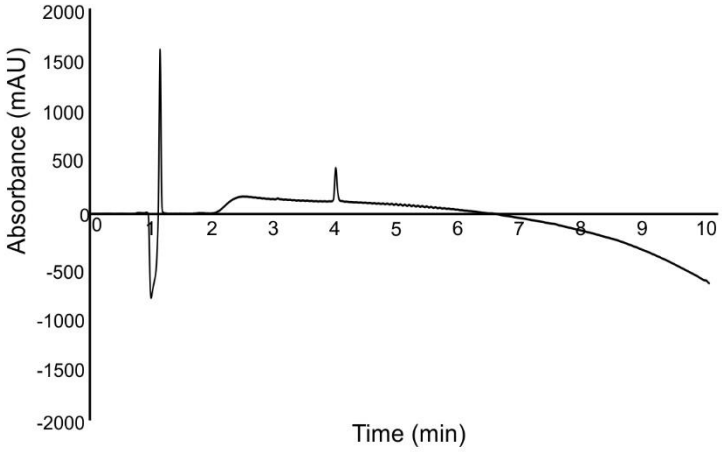
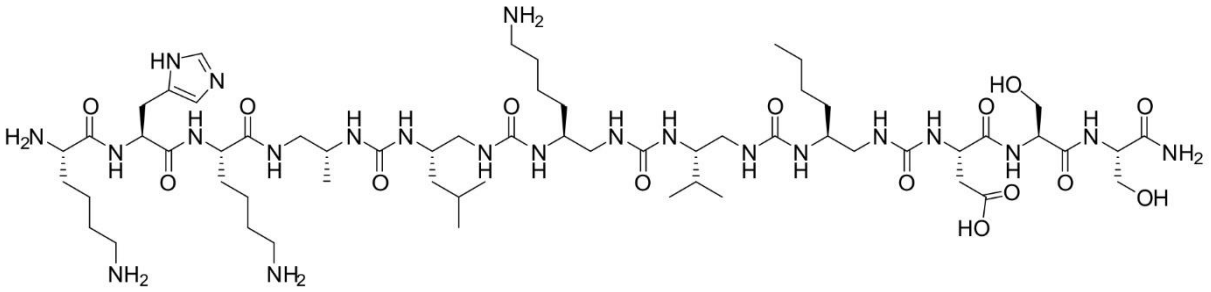


Figure S8: RP-HPLC chromatogram of **C9** after purification (λ = 200 nm).

Compound **C10**



C10 was synthesized on 75 µmol scale following general procedure A with NovaPEG resin (loading 0.45 mmol.g^{-1}). The product was purified by semi-preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 4 mL.min^{-1}. 3.7 mg was obtained with a global yield of 5%. ESI-MS (M_W 1368.87): m/z 1369.80 $[M+H]^+$, 685.4 $[M+2H]^{2+}$, 457.33 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100% ACN, 10 min) t_r = 4.77 min.

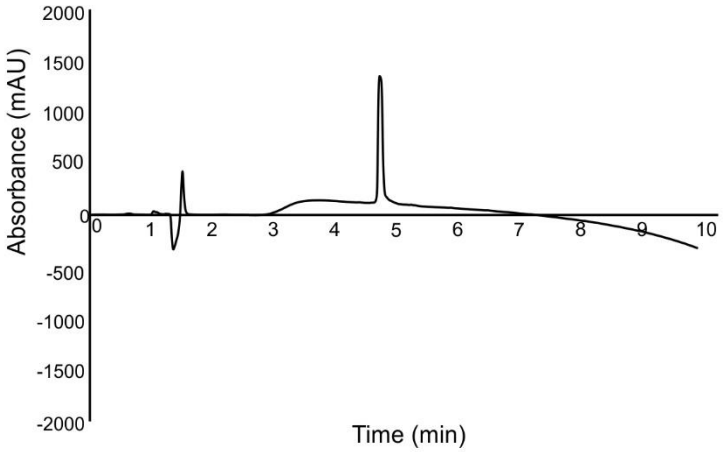
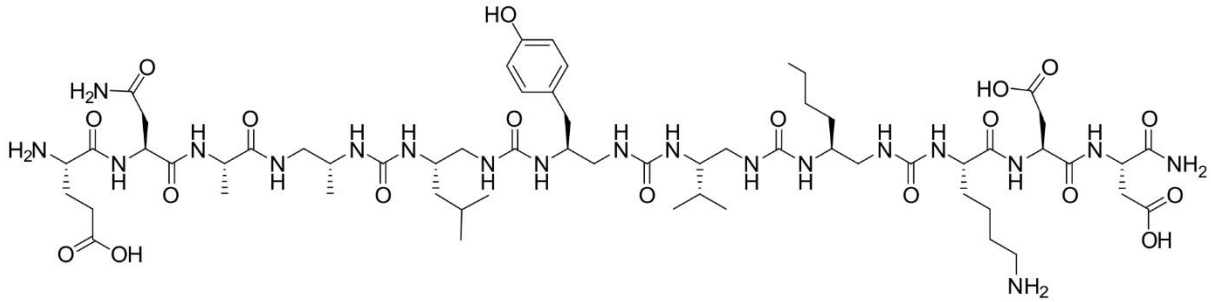


Figure S9: RP-HPLC chromatogram of **C10** after purification (λ = 200 nm).

Compound **C11**



C11 was synthesized on 75 µmol scale following general procedure A with NovaPEG resin (loading 0.45 mmol.g^{-1}). The product was purified by semi-preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 4 mL.min^{-1}. 7.4 mg was obtained with a global yield of 7%. ESI-MS (M_W 1393.77): m/z 1394.60 $[M+H]^+$, 697.93 $[M+2H]^{2+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.90 min.

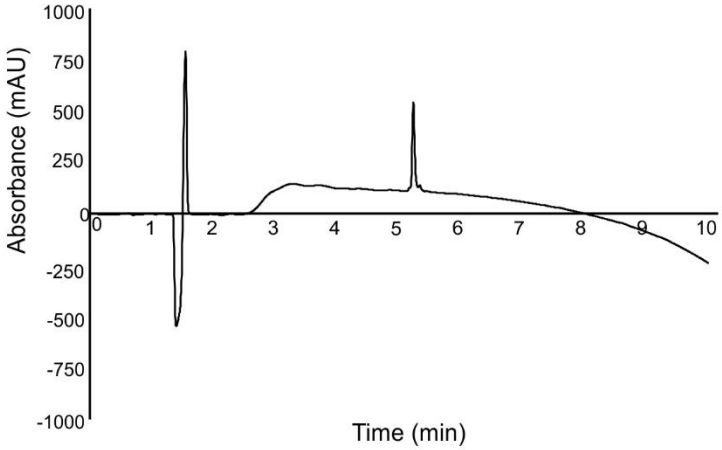
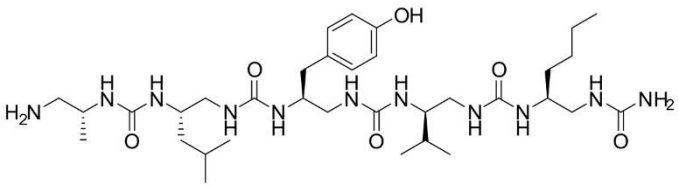


Figure S10: RP-HPLC chromatogram of **C11** after purification (λ = 200 nm).

Compound **C12**



C12 and **C12-2** were synthesized on 50 µmol scale following general procedure A with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 4.9 mg of **C12** was obtained with a global yield of 13%, and 31.5 mg of **C12-2** was obtained with a global yield of 87%. ESI-MS (M_W 721.49): m/z 722.33 $[M+H]^+$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.86 min.

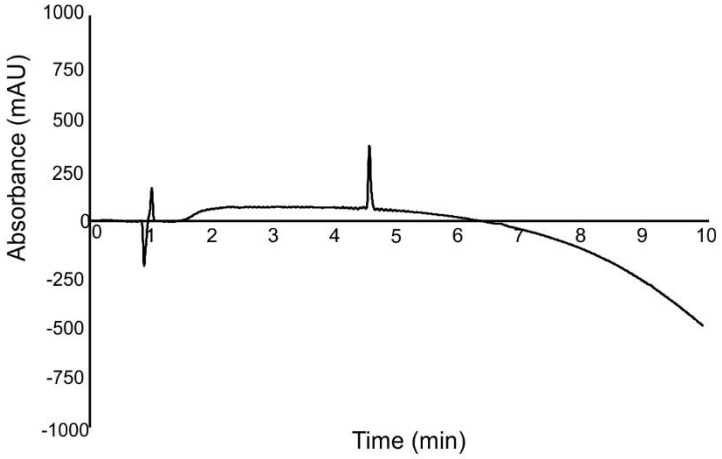


Figure S11: RP-HPLC chromatogram of **C12** after purification (λ = 200 nm).

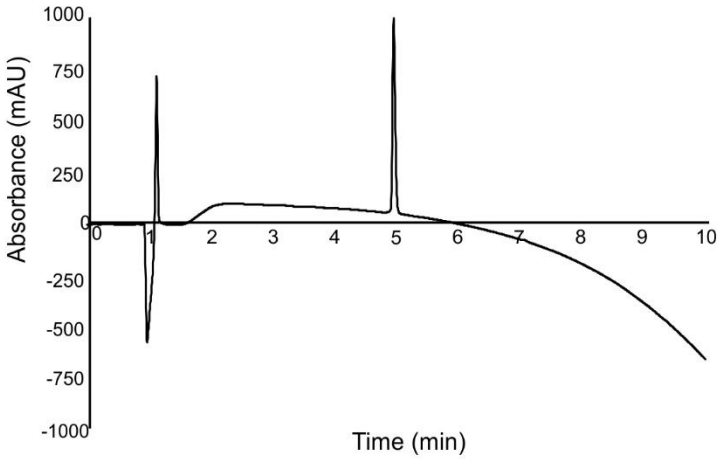
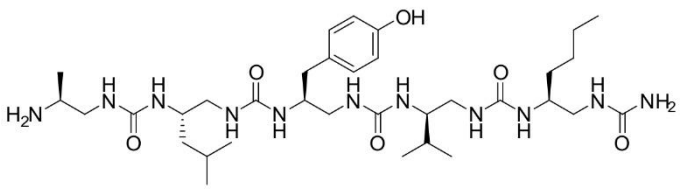


Figure S12: RP-HPLC chromatogram of **C12-2** after purification (λ = 200 nm).

Compound **C13**



C13 was synthesized on 50 µmol scale following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 25-65%, 15 min, flow = 20 mL.min^{-1}. 27.5 mg was obtained with a global yield of 79%. ESI-MS (M_W 721.50): m/z 722.33 $[M+H]^+$, 361.63 $[M+2H]^{2+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.72 min.

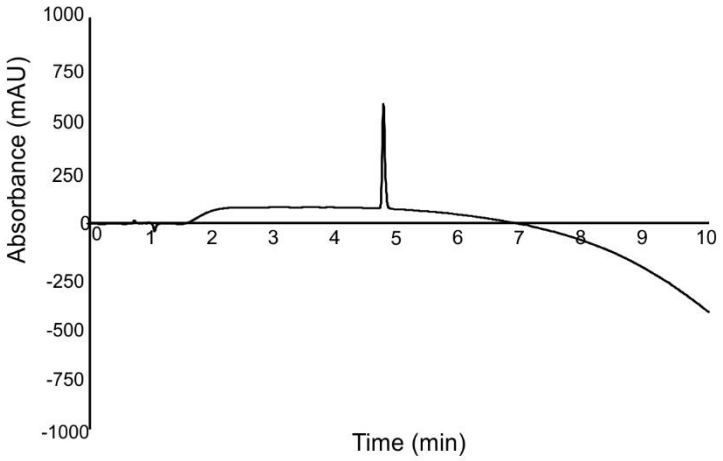
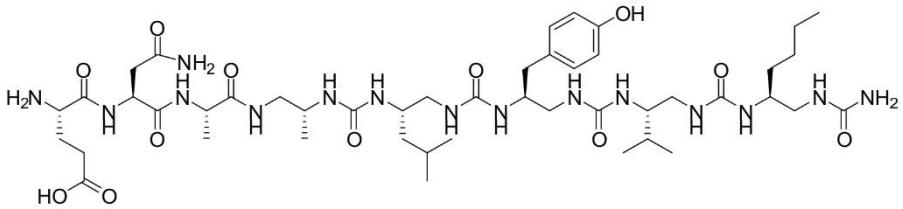


Figure S13: RP-HPLC chromatogram of **C13** after purification (λ = 200 nm).

Compound **C14.**



C14 was synthesized on 50 µmol following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC:H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 16.6 mg was obtained with a global yield of 79%. ESI-MS (M_W 1035.6): m/z 1036.43 $[M+H]^+$, 518.67 $[M+2H]^2$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.92 min.

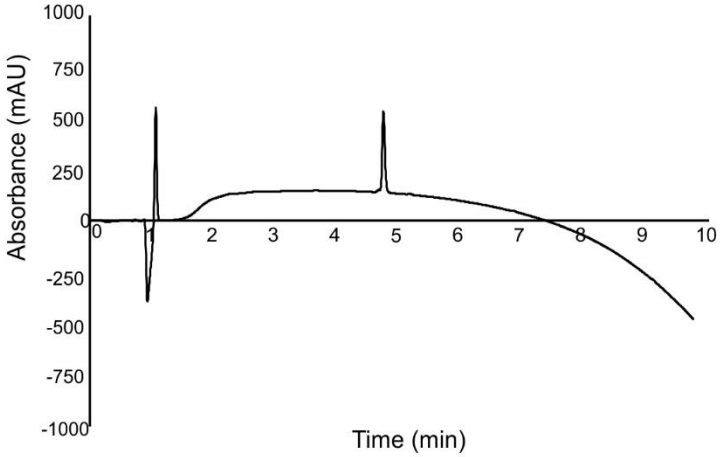
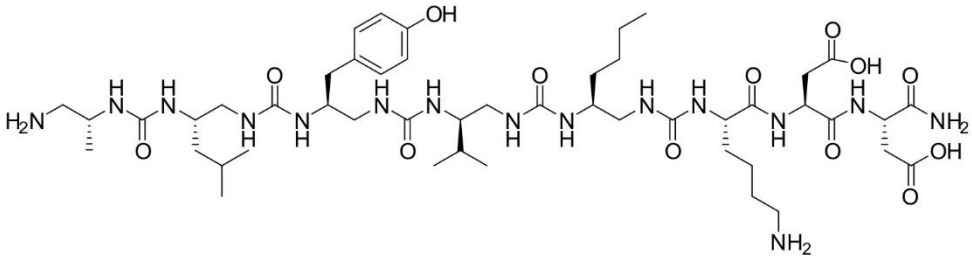


Figure S14: RP-HPLC chromatogram of **C14** after purification (λ = 200 nm).

Compounds **C15**



C15 and **C15-2** were synthesized on 50 µmol scale following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 4.5 mg of **C15** was obtained with a global yield of 8%, and 3.4 mg of **C15-2** was obtained with a global yield of 6% ESI-MS (M_W 1079.6): m/z 1080.45 $[M+H]^+$, 540.68 $[M+2H]^{2+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_R = 4.17 min.

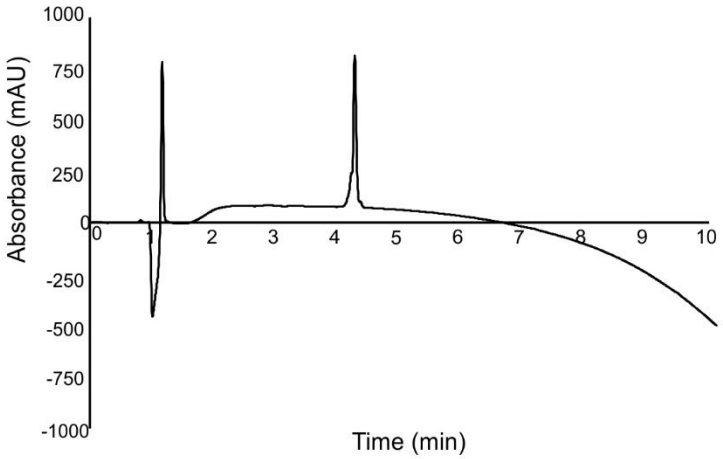


Figure S15: RP-HPLC chromatogram of **C15** after purification (λ = 200 nm).

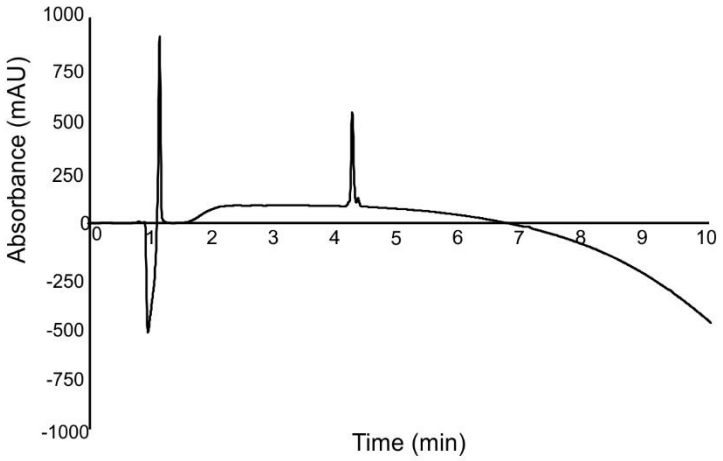
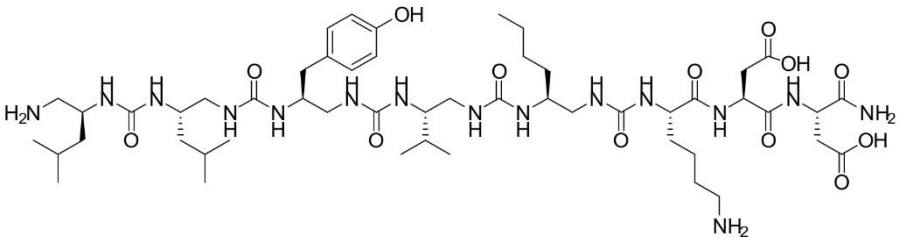


Figure S16: RP-HPLC chromatogram of **C15-2** after purification.

Compounds **C16**



C16 and C16-2 were synthesized on 50 µmol scale following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). Products were purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 3.5 mg of **C16** was obtained with a global yield of 6%, and 2.7 mg of **C16-2** was obtained with a global yield of 5%. ESI-MS (M_W 1121.7): m/z 1122.49 $[M+H]^+$, 561.70 $[M+2H]^{2+}$, 374.78 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.77 min.

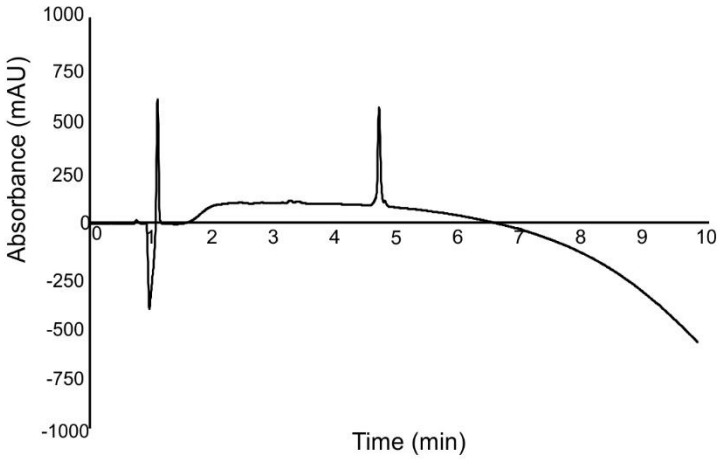


Figure S17: RP-HPLC chromatogram of **C16** after purification (λ = 200 nm).

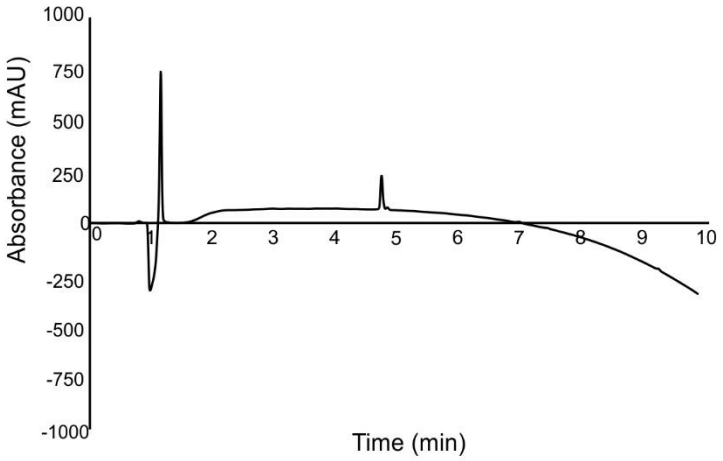
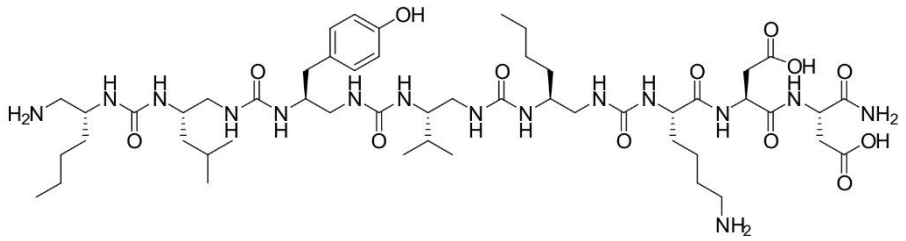


Figure S18: RP-HPLC chromatogram of **C16-2** after purification (λ = 200 nm).

Compound **C17**



C17 was synthesized on 50 µmol scale following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 11.5 mg was obtained with a global yield of 20%. ESI-MS (M_W 1121.7): m/z 1122.86 $[M+H]^+$, 561.97 $[M+2H]^{2+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.42 min.

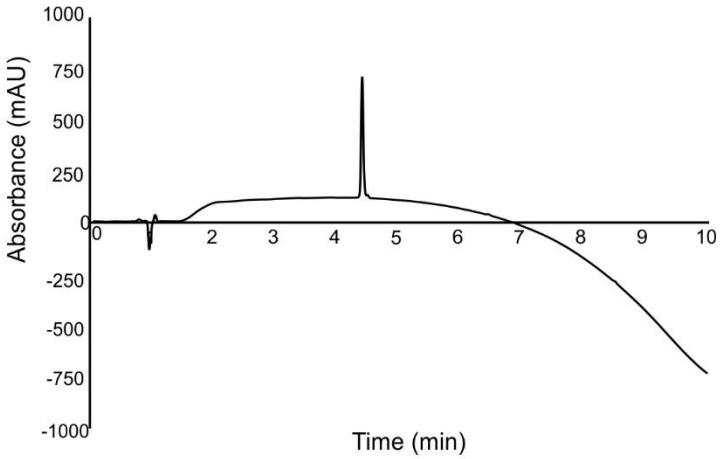
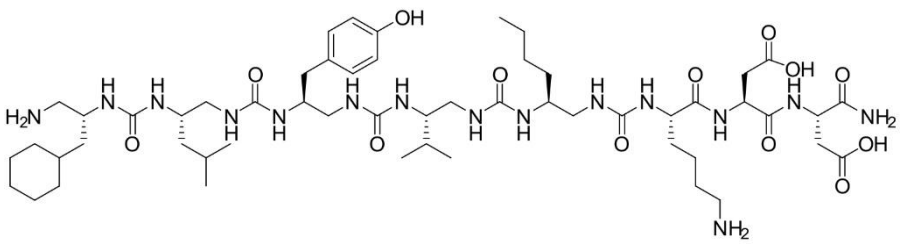


Figure S19: RP-HPLC chromatogram of **C17** after purification (λ = 200 nm).

Compound **C18**



C18 was synthesized on 50 µmol scale following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 5.8 mg was obtained with a global yield of 10%. ESI-MS (M_W 1161.7): m/z 1162.89 $[M+H]^+$, 581.98 $[M+2H]^{2+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_R = 4.80 min.

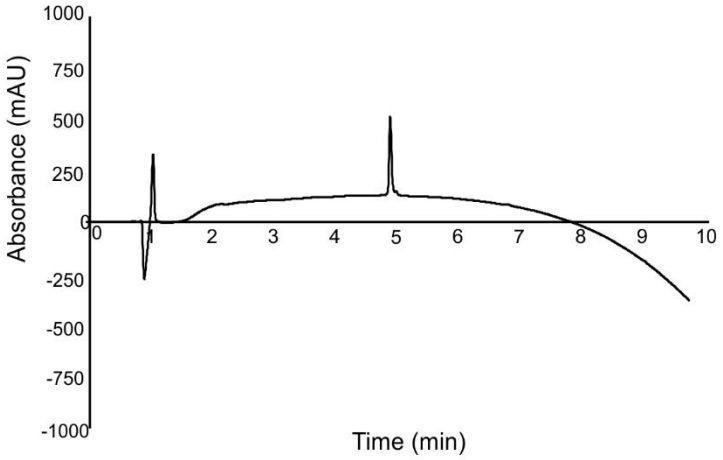
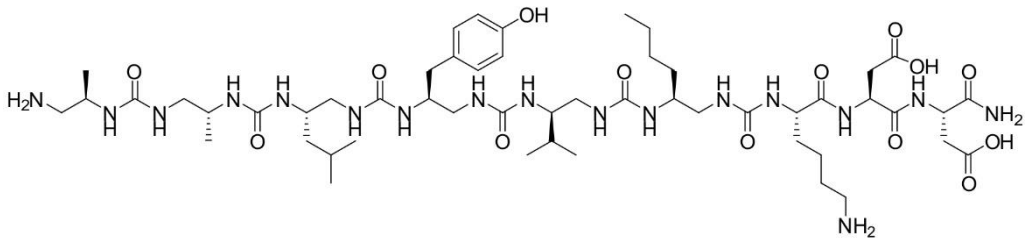


Figure S20: RP-HPLC chromatogram of **C18** after purification (λ = 200 nm).

Compounds **C19**



C19-2 and C19-3 were synthesized on 50 µmol following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). Crude products were purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 1.7 mg of chimera **C19-2** was obtained with a global yield of 9%, and 0.88 mg of **C19-3** was obtained with a global yield of 1.5%. ESI-MS (M_W 1179.7): m/z 1180.87 $[M+H]^+$, 590.97 $[M+2H]^{2+}$, 394.33 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 4.23 min.

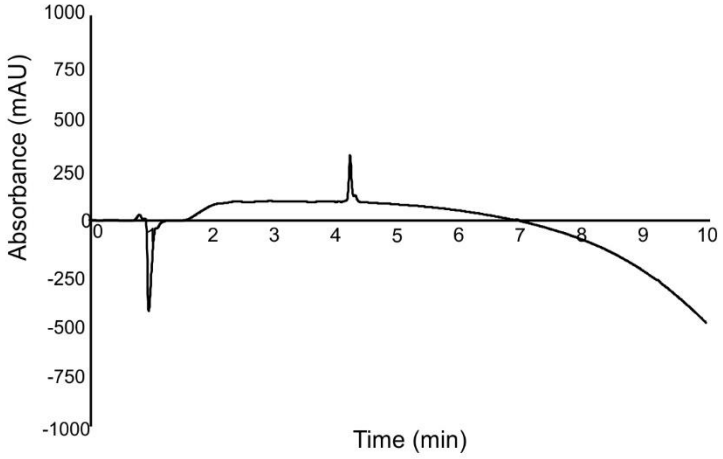


Figure S21: RP-HPLC chromatogram of **C19-2** after purification (λ = 200 nm).

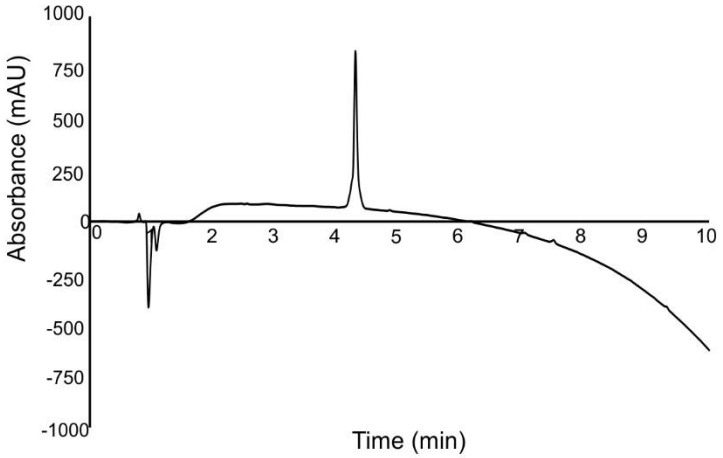
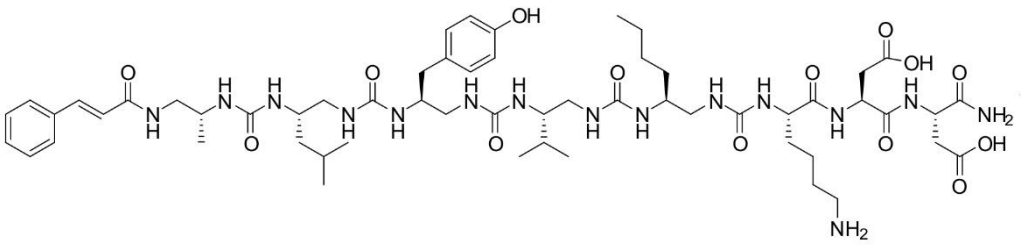


Figure S22: RP-HPLC chromatogram of **C19-3** after purification (λ = 200 nm).

Compounds **C20**



C20-2 and **C20-3** were synthesized on 50 µmol scale following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). Products were purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 0.2 mg of chimera C20-2 was obtained with a global yield of 0.3%, and 1.39 mg of **C20-3** was obtained with a global yield of 2.3%. ESI-MS (M_W 1209.8): m/z 1210.85 $[M+H]^+$, 605.96 $[M+2H]^{2+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 5.92 min.

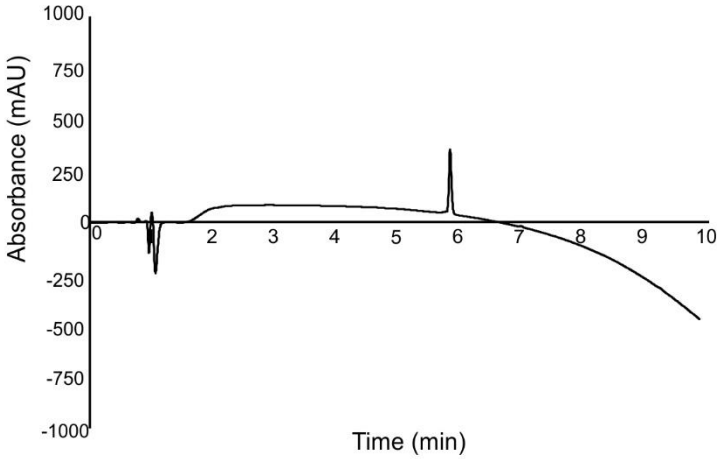


Figure S23: RP-HPLC chromatogram of **C20-2** after purification (λ = 200 nm).

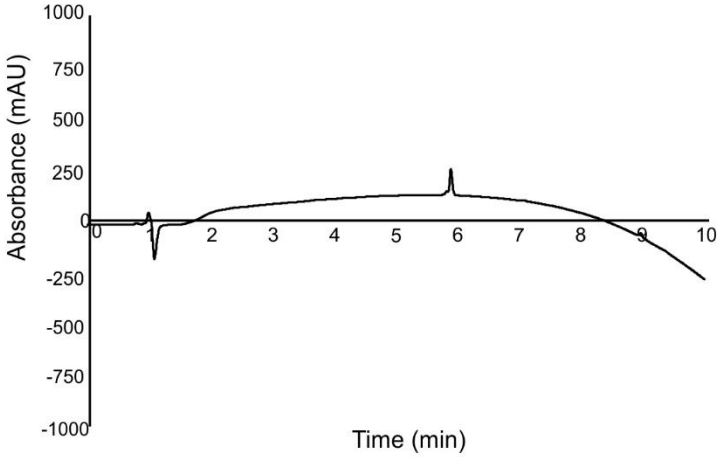
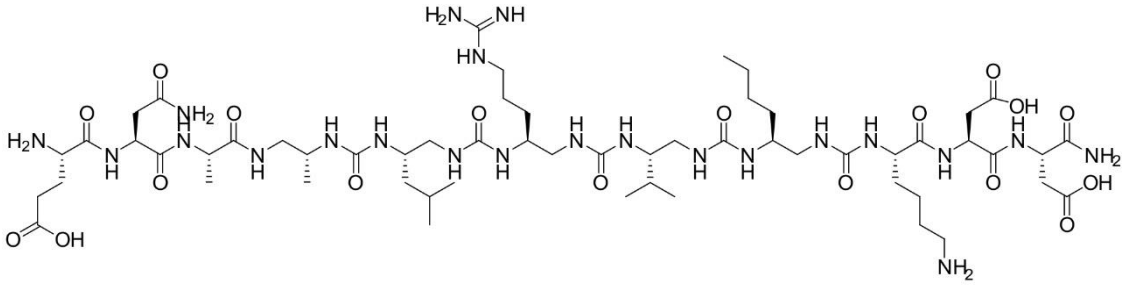


Figure S24: RP-HPLC chromatogram of **C20-3** after purification (λ = 200 nm).

Compound **C21-2**



C21-2 was synthesized on 50 µmol scale following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 1.38 mg was obtained with a global yield of 2%. ESI-MS (M_W 1386.9): m/z 1387.99 $[M+H]^+$, 694.53 $[M+2H]^{2+}$, 463.73 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 3.92 min.

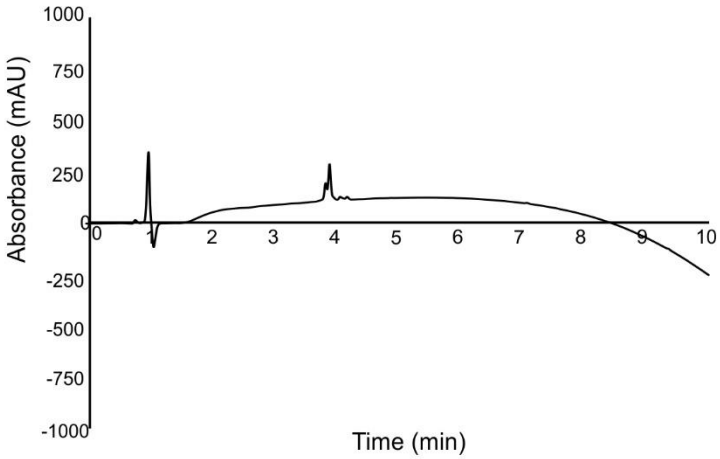


Figure S25: RP-HPLC chromatogram of **C21-2** after purification (λ = 200 nm). There are still some impurities left. Further purification steps were not performed due to a lack of time but will be done in the future

3. Simulation of solid-phase synthesis of chimeras.

The simulation of the solid phase synthesis was performed on a 75 µmol scale following general procedure A with with Rink amide MBHA resin (loading 0.52 mmol.g^{-1}). After coupling **M7** on resin beads, DMF was added and put into the microwave reactor. A micro-cleavage was perfomed with a small amount of resin beads which were swelled in a mixture of TFA/TIS/H_2O (95:2.5:2.5, v/v/v) and left to react for 120 min. After that time, the resin was filtered off, washed with the cleavage mixture and the filtrate was evaporated under reduced pressure and dissolved into a mixture of ACN/H_2O and analyzed by analytical RP-HPLC (HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) and ESI-MS HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 5 min).

4. Biostructural studies

a. General procedure for nanoDSF experiments

Nano Differential Scanning Fluorimetry (nanoDSF) is a modified differential scanning fluorimetry method that determine protein stability at high resolution by employing intrinsic tryptophan or tyrosine fluorescence. NanoDSF is a fast and easy technique to identify ligands (protein, peptide or small molecule) binding and stabilizing a target protein. Only microliters of sample are required, a wide range of concentrations can be measured in solution, and hard-to-measure viscous samples are examined with ease. Protein stability is determined by the change of amino acids (Trp and Tyr) fluorescence in presence and in absence of ligand. A melting temperature (T_m) of the studied protein can be obtained. Stabilization of protein-ligand complex is proportional to the concentration and to the affinity of the ligand[4]. Melting temperature is mesured as a function of protein fluorescence intensity which is plotted against temperature (gradient 20-95 °C, 1.5 °C.min^{-1}). Obtained curve reveals two states of protein folding: a stable state where the protein is folded and an unstable state corresponding to unfolded protein. Proteins are folded at room temperature and have a tendency to unfold at high temperature, and the inflexion point of this transition corresponds to the T_m. Nano-DSF mesures were performed Prometheus NT.48. apparatus (NanoTemper) that mesures intrinsic fluorescence of a protein itself at 330 nm and 350 nm. 10 µL of protein (zVDR and hVDR) were in a buffer solution of HEPES KOH 50 mM pH 8.0, NaCl 50 mM, Glycerol 2%, TCEP 1 mM at a concentration of 4 mg.mL^{-1} for the protein — when measured alone — and with 5 equiv. of ligands (**P1** or chimera). Mesurements were performed three times on three different samples and melting temperatures at each wavelength were deduced from the first derivative of fluorescence ratio 350/330 nm.Data were analyzed by the software package included in the Prometheus NT.48 apparatus.

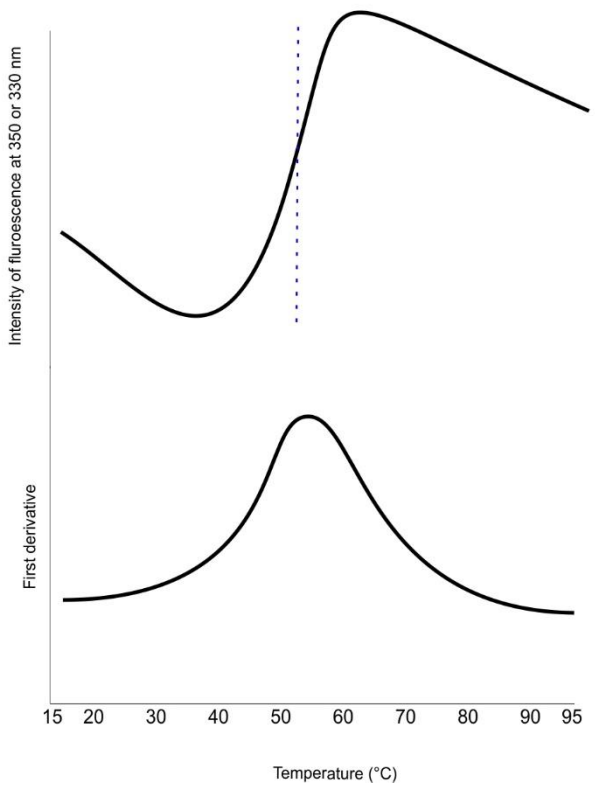


Figure S26: Schematic representation of nanoDSF sigmoidal curve (top) before and after first derivative (down).

b. Results of nanoDSF experiments

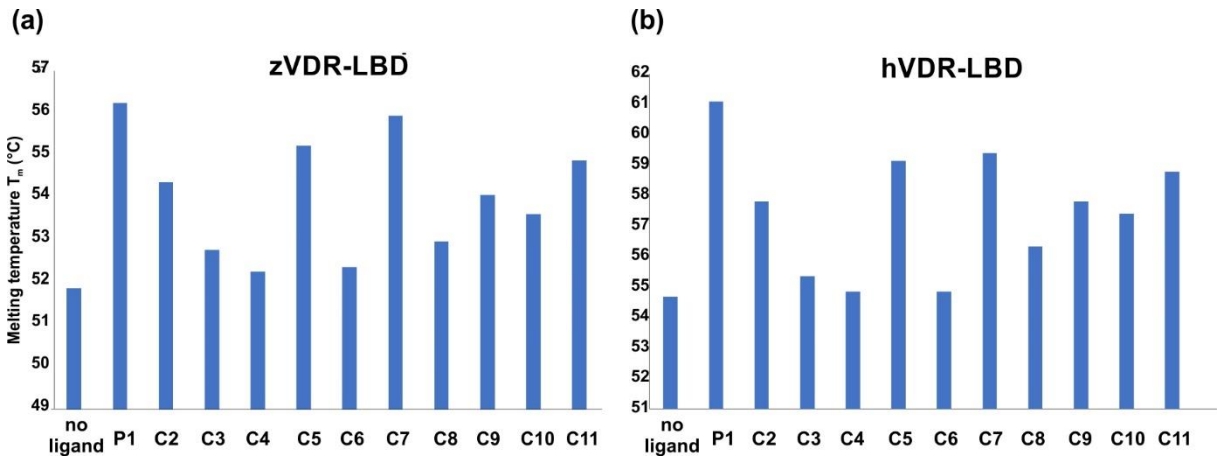


Figure S27: Melting temperature measured by nano-DSF for **(a)** zVDR-LBD and **(b)** hVDR-LBD with peptide **P1** and chimeras.

c. General procedure for MST measurements

Microscale thermophoresis (MST) is an all-optical technic that quantifies biomolecular interactions between proteins under a microscopic gradient of temperature generated by infrared (IR) laser. Any modification in the hydratation layer, charge and shape of a biomolecule affect its thermophoretic movement. Thermophoresis can be described as a molecular flux which is proportional to a gradient of temperature. Modifications in the concentration induce a change of thermophoresis, MST records the fluorescence intensity of an excited protein sample with a IR laser before, during and after the laser has been switched on (**Fig. 28**). MST measurements take 30 s per sample, and as thermophoresis has a limited diffusion, distintives areas are observed and can be separated to study binding event. Fluorescence intensity is collected before heating with the IR laser, this fluorescence intensity needs to be stable for each sample of the serial dilution. Once the IR laser is on, a brutal change in fluorescence is observed with the increase of temperature. This jump of temperature corresponds to the change of fluorescence that depends on temperature[5].

Mesurement of interactions were performed with the Monolith™ NT.115 (Nanotemper technologies GmbH, Munich) and data were analyzed by the software NanoTemper Analysis 1.4.27. One biomolecule is labeled with a fluorochrome and titrated by another biomolecule in order to deduce a dissociation sigmoidal curve[6]. Mesurements were performed by using non-treated capillaries (MO-K022). Dissociation constants at the equilibrium (K_d) were calculated from the variation of fluorescence induced by the thermophoretic movement of particules by IR laser irradiation. Interaction measurements between LBD (Ligand Binding Domain) of hVDR and ligands were performed by hVDR LBD labelled with Alexa Fluor 488 NHS Ester (ThermoFischer Scientific) reagent or with NT-647 dye (Nanothemper catalogue: MO-L001). The buffer condition used of dilution in a 1:1 series is composed of 50 mM HEPES KOH pH 8.0, 50 mM NaCl, 2% Glycerol, 1mM TCEP + 0.1%Tween-20. HEPES buffer was used because VDR LBD labeling reaction can not be done in Tris buffer. VDR labelling an amine-labeling process and the Tris buffer contains lot of free amine groups. The Tween 20 is commonly used as a detergent in MST measurement because it helps to avoid protein to stick on capillary surface. If the protein is sticked on the capillary it can not move freely during the thermophoresis, and the measurement can not be performed correctely. Protein, fluorochrome and ligands are incubated for 15 min and then centrifuged 5 min at room temperature. Then, around 10-12 µL of the sample solution is loaded inside a capillary. Data were collected at 25 °C with an exciting power of IR-laser at 40% and 20% respectively. Mesurements were reproduced on eight independent experiments and thermophoresis signals, collected for all sixteen capillaries, are normalized. Normalized fluorescence is plotted against ligand

concentration that give a sigmoidal curve in which standard error of measurements (SEM for n ≥ 1) are determined for each points.

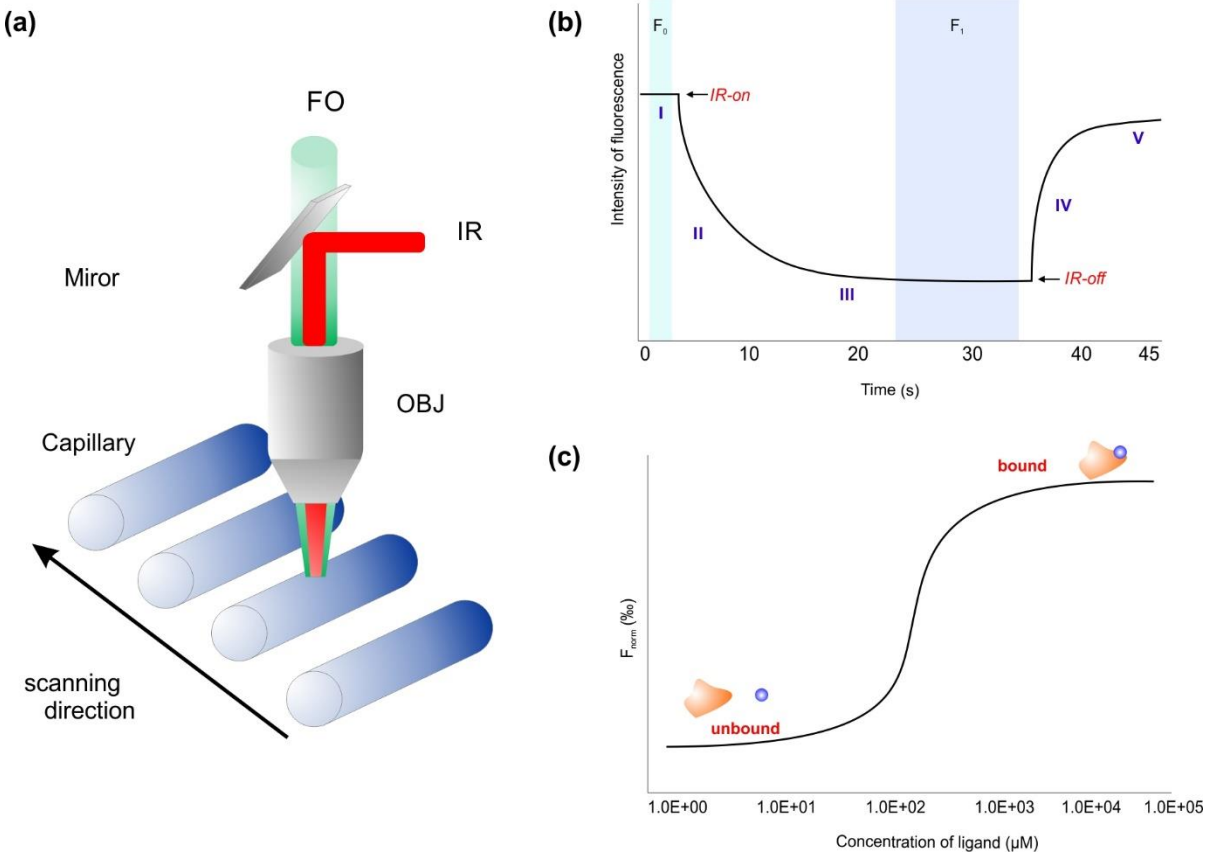


Figure S28: Schematic representation of **(a)** MST set-up, **(b)** acquisition of fluorenscence intensity with a gradient of temperature and **(c)** graph of normalized fluorescence F_{norm} plotted against the concentration of ligand.

d. Results of MST experiments

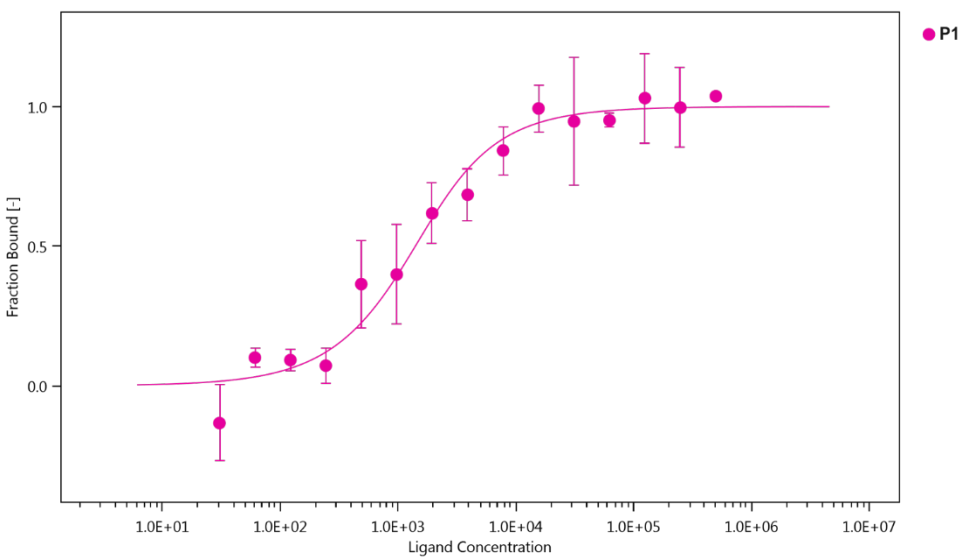


Figure S29 : Determination of binding affinity for peptide **P1** with zVDR determined by MST experiments.

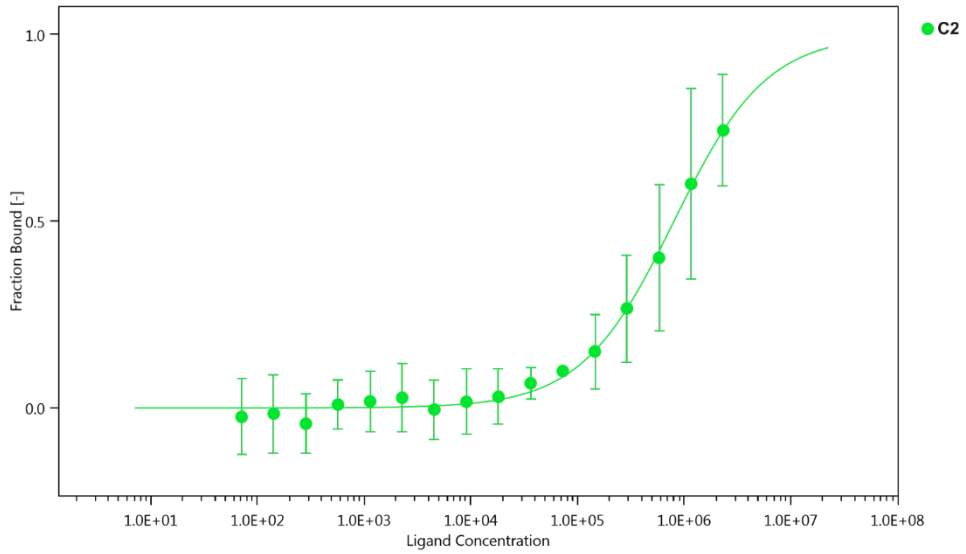


Figure S30: Determination of binding affinity for chimera **C2** with zVDR determined by MST experiments.

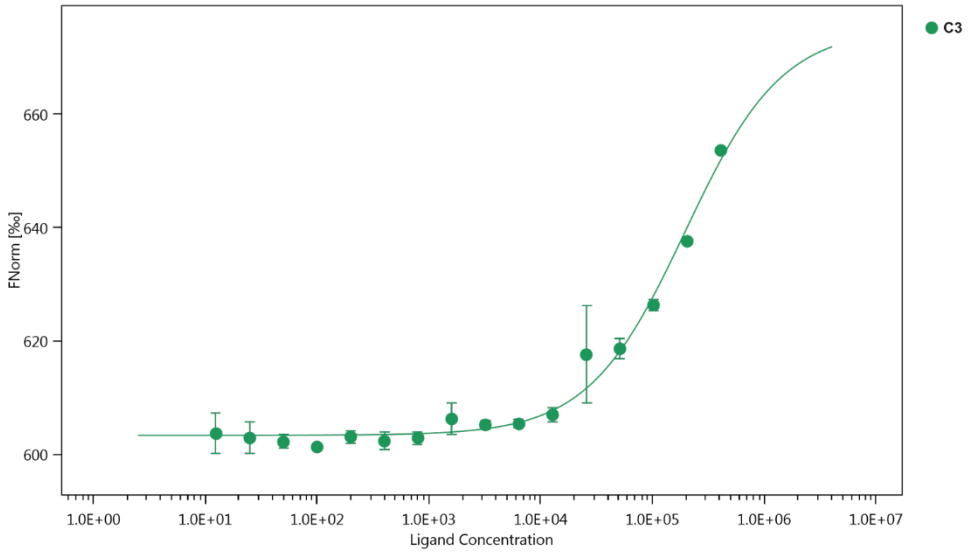


Figure S31: Determination of binding affinity for chimera **C3** with zVDR determined by MST experiments.

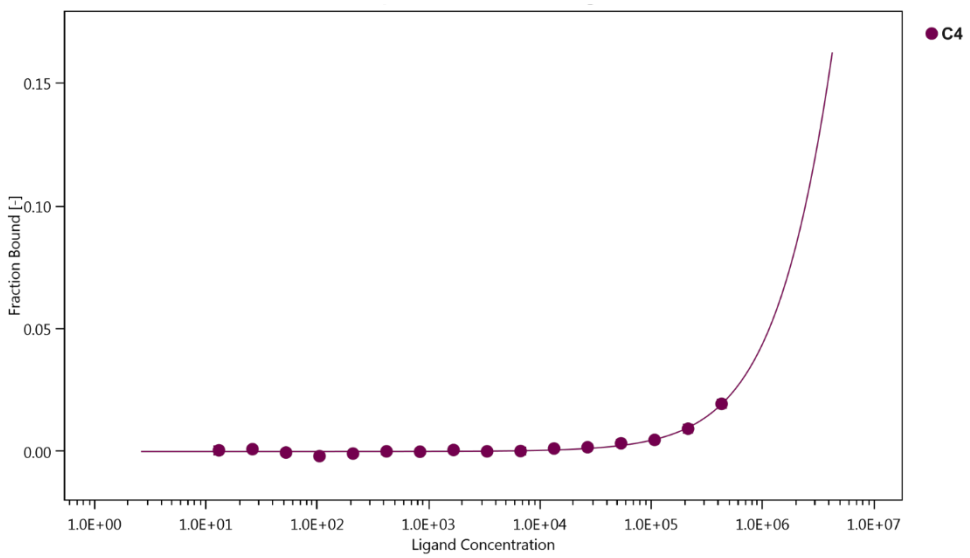


Figure S32: Determination of binding affinity for chimera **C4** with zVDR determined by MST experiments.

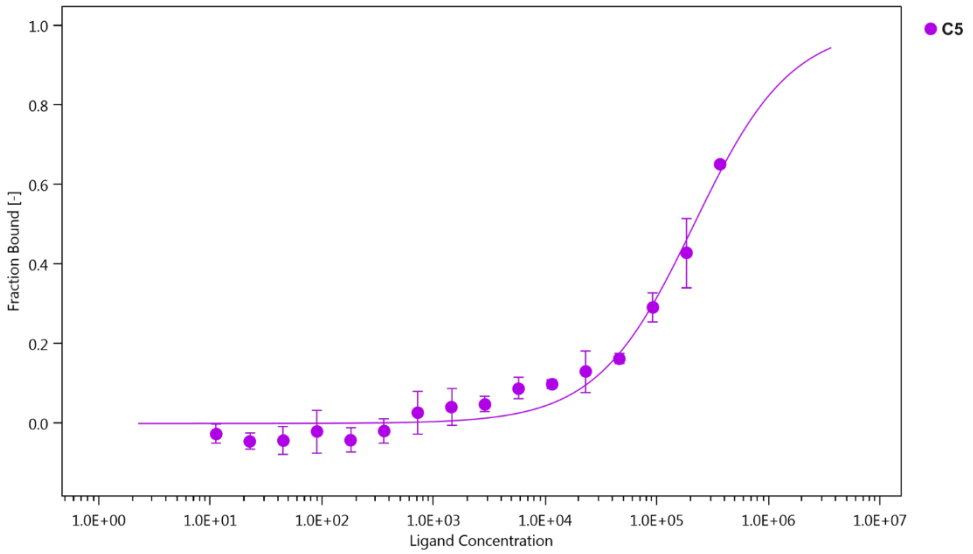


Figure S33: Determination of binding affinity for chimera **C5** with zVDR determined by MST experiments.

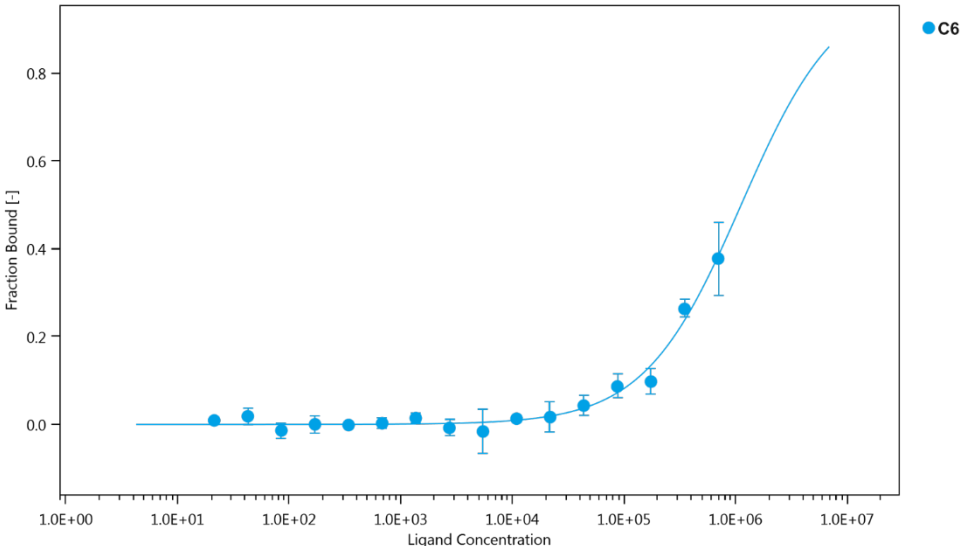


Figure S34: Determination of binding affinity for chimera **C6** with zVDR determined by MST experiments.

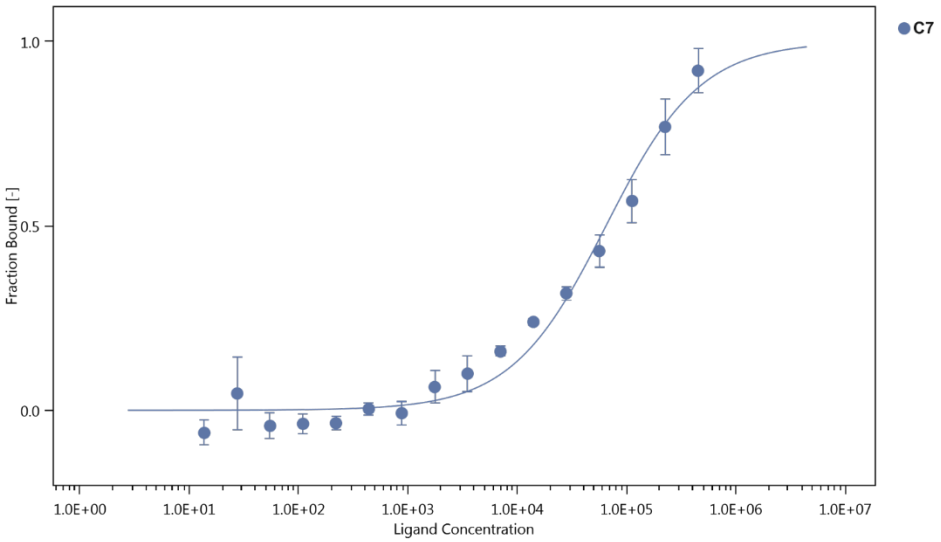


Figure S35: Determination of binding affinity for chimera **C7** with zVDR determined by MST experiments.

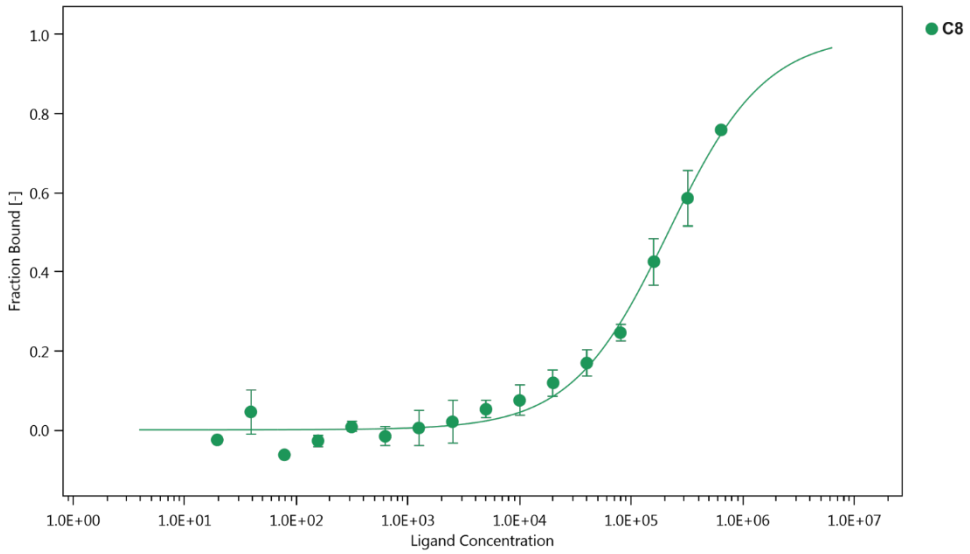


Figure S36: Determination of binding affinity for chimera C8 with zVDR determined by MST experiments.

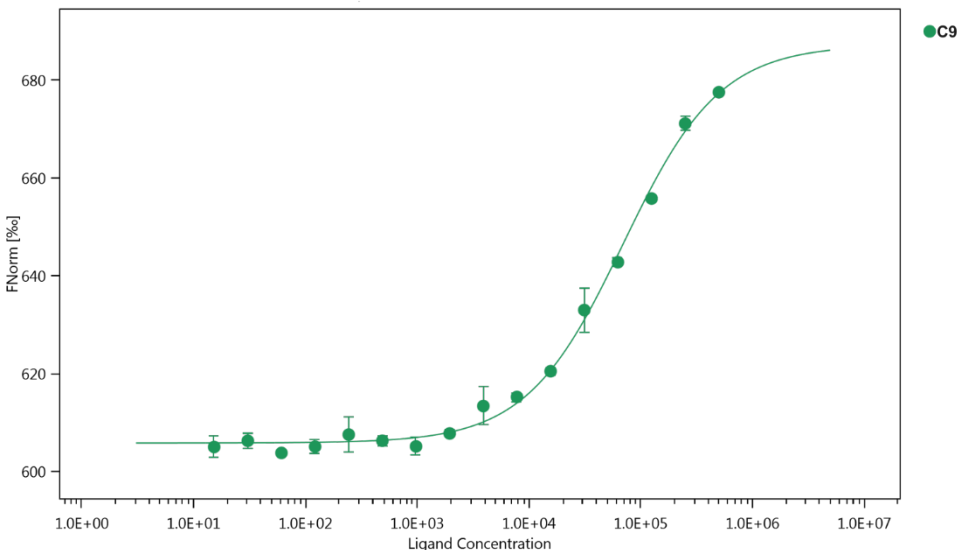


Figure S37: Determination of binding affinity for chimera **C9** with zVDR determined by MST experiments.

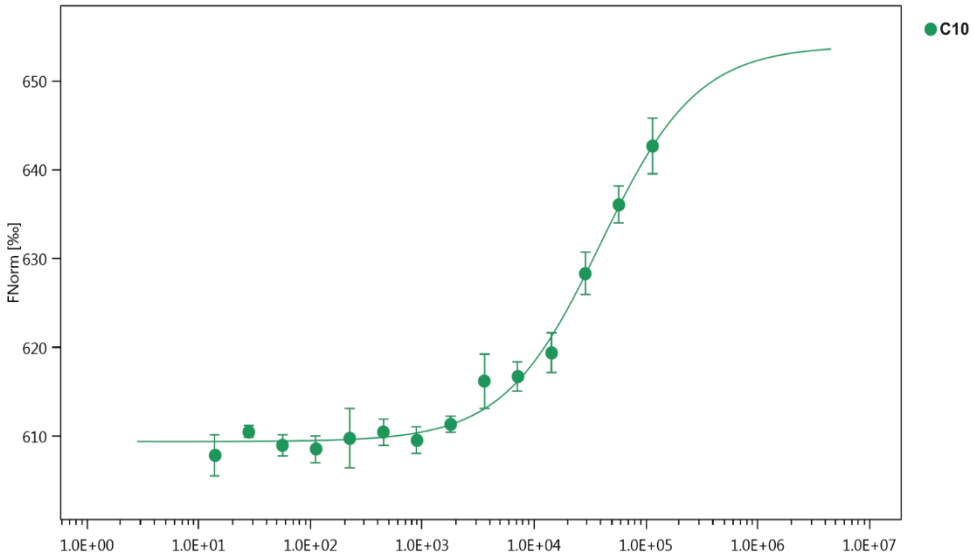


Figure S38: Determination of binding affinity for chimera **C10** with zVDR determined by MST experiments.

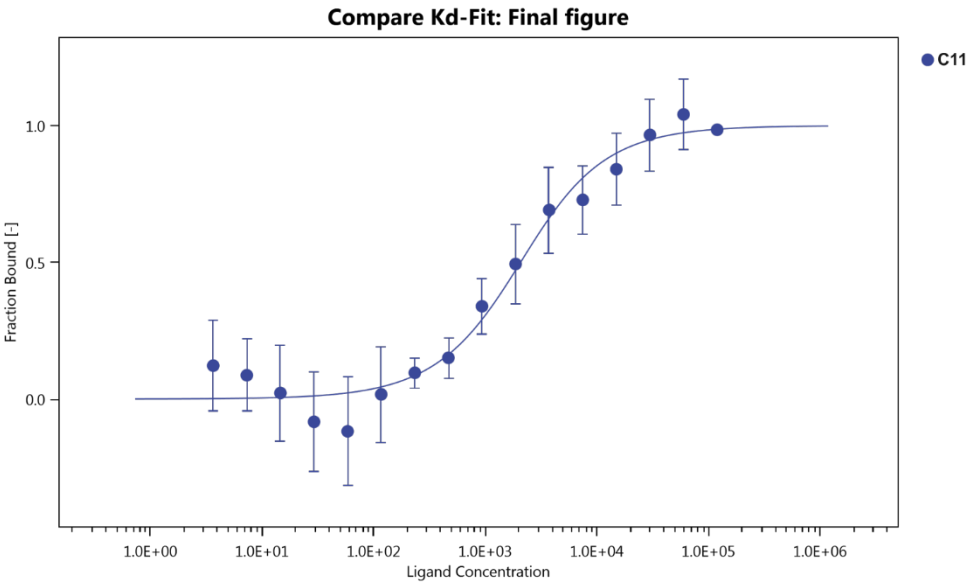


Figure S39: Determination of binding affinity for chimera **C11** with zVDR determined by MST experiments.

e. X-ray crystallography

i. General methods for crystallization

For crystallization trials, lyophilized powders of chimeras s were dissolved with zVDR protein and vitamin D_3 at a total concentration of 10 mg.mL^{-1}. Crystallization trials were performed at 20 °C in standard aqueous sitting drops composed of 0.4 µL of chimera/VDR/VitD_3 solution plus 0.4 µL of crystallization condition by a Mosquito robot (TTP Labtech), using MRC2 or MRC3 crystallization plates (Swissci). MRC plates were conserved in an imaging robot (Rock Imager – Formulatrix) at 20 °C, and crystal growth was monitored regularly. Resolutions of structures were performed by molecular replacement with Phaser software, using the PDB structure of zVDR **2HC4** as a starting model. Refinement and atomic model construction were performed on Phenix and Coot software respectively.

ii. X-ray crystallography results for **C2-C11**

Crystals were obtained in conditions reported in **Table S1**. Diffraction were performed at synchrotron ESRF, beams lines ID30B and ID23 (ESRF) were used for **C2**, **C11** and **C3-C8** respectively.

chimeras	protein and ligand	crystallization condition
C2	zVDR-LBD + $VitD_3$	Sodium acetate 2.5 M, HEPES 0.1 M pH 7.5
C3	zVDR-LBD + $VitD_3$	Lithium sulfate 1.8 M, HEPES 0.1 M pH 7.5
C5	zVDR-LBD + $VitD_3$	Ammonium acetate 3.5 M, TRIS 0.1 M pH 8.0
C6	zVDR-LBD + $VitD_3$	Sodium acetate 3.0 M, MES 0.1 M pH 6.0
C7	zVDR-LBD + $VitD_3$	Sodium acetate 3.0 M, TRIS 0.1 M pH 8.0
C8	zVDR-LBD + $VitD_3$	Sodium acetate 3.0 M, TRIS 0.1 M pH 8.0
C11	zVDR-LBD + $VitD_3$	Lithium sulfate 1.8 M, MES 0.1 M pH 6.0

Table S1: crystallization condition of chimeras **C2** - **C11** in complex with VDR and $VitD_3$.

	C2	C3*	C5	C6	C7	C8	C11
Resolution (Å)	20,4-2,37 (2,55-2,37)	43,0-1,98 (2,05-1,98)	22,0-1,95 (2,02-1,95)	32,96-2,35 (2,53-2,35)	38,0-2,30 (2,47-2,30)	32,9-2,00 (2.09-2.00)	14,9,0-2,00 (2,09-2,00)
Space group	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P6_522$	$P6_522$
Cell dimensions a,b,c (Å) α, β, γ(°)	a = b = 66,09; c = 260,14 α = β = 90 γ = 120	a = b = 66,18; c = 261,37 α = β = 90 γ = 120	a = b = 66,35; c = 262,37 α = β = 90 γ = 120	a = b = 65,91; c = 262,42 α = β = 90 γ = 120	a = b = 66,20; c = 254,22 α = β = 90 γ = 120	a = b = 66,37; c = 253,72 α = β = 90 γ = 120	a = b = 66,11; c = 262,48 α = β = 90 γ = 120
N° reflections	14409	24519	25904	14937	15313	23248	23172
Redundancy	6,0	*nd*	7,3	9,1	13,7	9,4	1,6
R_{syn} (%)[a]	8,4	*nd*	9,3	8,7	7,4	5,9	1,2
Completness (%)	99,0 (98,0)	99,1 (99,0)	99,4 (99,0)	99,6 (99,0)	98,2 (97,0)	99,1 (99,8)	96.2 (80,0)
Mean I/σ (%)	12,3 (1,34)	*nd*	12,3 (1,33)	17,6(1,37)	21,2(1,36)	19,7 (1,36)	16,0 (1,34)
RMS[b] bond lengths (Å) bond angles (°)	0,007 1,10	0,013 1,81	0,008 1,15	0,007 1,13	0,006 1,00	0,006 0,98	0,006 0,87
R_{work} / R_{free} [c] (%)	21,3/24,6 (35,4/41,4)	26,0/29,0 (51,1/55,7)	19,2/21,7 (38,55/42,91)	20,3/22,7 (29,9/34,3)	19,2/23,9 (26,6/34,6)	18,3/21,7 (26,0/28,4)	19,9/22,1 (24,3/26,7)
B factor							
Protein	85,21	*nd*	57,89	70,86	66,04	56,71	49,58
Ligand/ion	71,01	*nd*	52,11	70,27	61,76	58,01	45,96
Water	92,98	*nd*	67,65	77,32	70,37	66,81	55,26

Table S2: Crystallographic parameters for chimeras **C2-C11** with zVDR

iii. X-ray crystallography results for **C14** and **C19**

Crystals were obtained in conditions reported in **Table S3**. Diffraction were performed at synchrotron, beams lines PX2 (SOLEIL) and ID30A (ESRF) were used for **C14** and **C19** respectively.

chimeras	protein and ligand	crystallization condition
C14	zVDR-LBD + $VitD_3$	Ammonium sulfate 1.5 M, BisTrisPropane 0.1 M pH 7.0
C19	zVDR-LBD + $VitD_3$	Lithium sulfate 1.5 M

Table S3: crystallization condition of chimeras **C14** and **C19** in complex with VDR and $VitD_3$.

	C14	C19
Resolution (Å)	48.0-2.20 (2.26-2.20)	19.9-2.30 (2.38-2.30)
Space group	$P6_522$	$P6_522$
Cell dimensions a,b,c (Å) α, β, γ(°)	a = b = 66.47; c = 261.96 α = β = 90 γ = 120	a = b = 66.24; c = 252.52 α = β = 90 γ = 120
N° reflections	18379	15412
Redundancy	2.0	37.1
R_{syn} (%)[a]	1.5	10.9
Completness (%)	99.7 (99.0)	99.3 (99.8)
Mean I/σ (%)	21.6 (1.35)	25.5 (1.35)
RMS[b] bond lengths (Å) bond angles (°)	0.009 1.10	0.006 0.93
R_{work}/R_{free} [c] (%)	18,2/23,1 (27.8/31.4)	19.9/23.9 (29.1/35.0)
B factor		
Protein	65.41	78.85
Ligand/ion	59.21	78.47
Water	66.96	75.59

Table S4: X-ray crystallographic parameters for chimeras **C14** and **C19** with zVDR.

IV. Experimental data of Chapter 3

1. Synthesis of oligoureas **OL1**-**OL5**

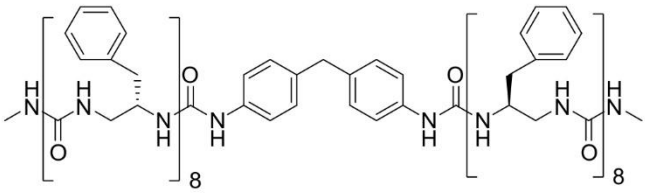
a. General procedure B for oligourea synthesis in solution.

Boc protected-oligourea (0.1 mmol) was treated with TFA (2 mL) at 0 °C for 30 minutes. After completion of the reaction followed by TLC the solvent was evaporated under reduced pressure first and then co-evaporated with cyclohexane. TFA salt was dissolved in ACN (2 mL) at cooled down to 0 °C, DIPEA (0.3 mmol) was added followed by the addition of diisocyanate (0.05 mmol). The reaction mixture was stirred at room temperature overnight. After completion of the reaction, the solvent was evaporated and a saturated $NaHCO_3$ solution was added, a precipitate is formed which was filtered off and washed with various solvents (sat. $NaHCO_3$, 1N $KHSO_4$, water, DCM, brine) and finally dried under high vacuum.

*b. Synthesis and characterization of oligoureas **OL1**-**OL5** and their enantiomers **OLent-1**-**OLent-4.***

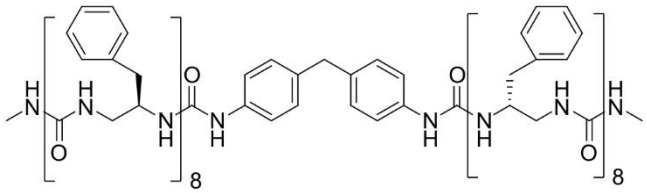
Full characterisation of oligoureas was performed by myself.

Compound **OL1**



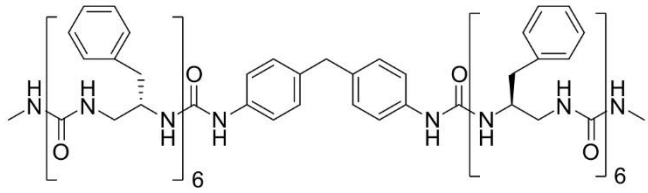
OL1 was synthesized with the general procedure *B* (yield: 77%). **^{1}H NMR** (700 MHz, DMF-d_7) δ (ppm): 8.23 (s, 2H), 7.11-7.43 (m, 88H), 6.69-6.76 (m, 6H), 6.59-6.68 (m, 4H), 6.54-6.59 (m, 2H), 6.53 (d, *J* = 10 Hz), 6.36-6.40 (m, 2H), 6.35 (d, *J* = 9 Hz, 2H), 6.26 (d, *J* = 10 Hz, 2H), 6.21-6.25 (m, 4H), 6.19 (d, *J* = 10 Hz, 2H), 6.12-6.15 (m, 6H), 5.99 (d, *J* = 10 Hz, 2H), 5.96 (d, *J* = 9 Hz, 2H), 4.18-4.29 (m, 7H), 4.00-4.11 (m, 9H), 2.77-2.85 (m, 16H), 2.68 (d, *J* = 4 Hz, 6H), 2.17-2.67 (m, 32H). **^{13}C NMR** (100 MHz, DMF-d_7) δ (ppm): 161.1, 161.0, 161.0, 160.9, 160.9, 160.8, 160.3, 159.7, 157.6, 140.2, 140.0, 139.9, 139.8, 135.9, 130.2, 130.2, 130.1, 130.0, 129.9, 129.2, 129.1, 127.1, 126.9, 119.1, 52.0, 51.9, 51.9, 51.6, 51.5, 51.4, 51.1, 47.1, 46.7, 46.5, 45.4, 44.2, 41.08, 40.8, 40.3, 39.7, 39.5, 39.3, 27.0. MALDI-TOF MS (M_W 3129.7): *m/z* 3152.8 $[M+Na]^+$.

Compound **OL*ent*-1.**



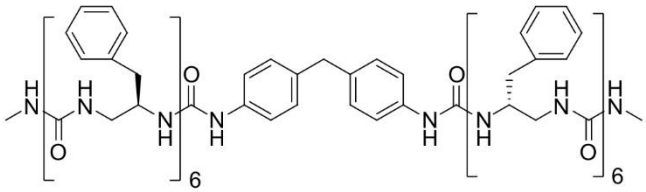
OL*ent*-1 The synthesis was performed with the general procedure *B*, coupling yield of 88%. **^{1}H NMR** (700 MHz, DMF-d_7) δ (ppm): 8.23 (s, 2H), 7.16-7.43 (m, 88H), 6.63-6.79 (m, 8H), 6.27-6.42 (m, 5H), 6.25 (d, *J* = 7 Hz, 2H), 6.23 (d, *J* = 6 Hz, 2H), 6.11-6.20 (m, 5H), 6.00 (d, *J* = 10 Hz, 2H), 5.97 (d, *J* = 9 Hz, 2H), 4.17-4.30 (m, 6H), 4.02-4.10 (m, 9H), 3.92 (s, 2H), 2.78-2.89 (m, 9H), 2.70 (d, *J* = 4 Hz, 6H), 2.19-2.68 (m, 32H). **^{13}C NMR** (100 MHz, DMF-d_7) δ (ppm): 161.1, 161.0, 161.0, 160.9, 160.3, 159.7, 157.6, 130.2, 130.1, 130.1, 130.0, 129.9, 129.3, 129.3, 129.2, 129.1, 127.2, 127.0, 126.9, 119.1, 52.1, 52.0, 51.8, 51.7, 51.6, 51.5, 51.1, 46.8, 46.7, 46.5, 44.3, 41.0, 40.9, 40.4, 39.6, 39.3, 27.1. MALDI-TOF MS (M_W 3129.87): *m/z* 3152.8 $[M+Na]^+$.

Compound **OL2.**



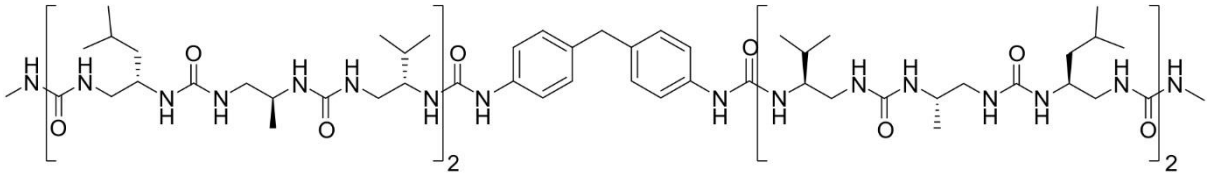
OL2 was synthezised with the general procedure *B*; coupling yield of 71%. **^{1}H NMR** (700 MHz, DMF-d_7) δ (ppm): 7.15-7.31 (m, 68H), 6.59-6.70 (m, 6H), 6.45 (d, *J* = 9 Hz, 2H), 6.44 (d, *J* = 9 Hz, 2H), 6.37 (d, *J* = 9 Hz, 2H), 6.32 (d, *J* = 9 Hz, 2H), 6.23-6.29 (m, 4H), 6.11-6.16 (m, 4H), 6.02 (d, *J* = 9 Hz, 2H), 6.02 (d, *J* = 8 Hz, 2H), 4.17-4.25 (m, 7H), 4.00-4.10 (m, 5H), 3.91 (s, 2H), 2.80-2.90 (m, 12H), 2.70 (d, *J* = 4 Hz, 6 H), 2.30-2.69 (m, 24 H). **^{13}C NMR** (100 MHz, DMF-d_7) δ (ppm): 160.9, 160.9, 160.7, 160.3, 159.7, 157.5, 139.5, 139.3, 139.0, 138.9, 138.7, 136.2, 130.2, 130.1, 130.0, 129.3, 129.2, 129.1, 129.1, 129.1, 129.1, 127.2, 127.2, 127.1, 127.0, 126.9, 119.3, 52.1, 51.9, 51.8, 51.7, 51.5, 51.3, 46.9, 46.5, 46.3, 45.9, 44.2, 41.0, 40.8, 40.3, 39.6, 39.5, 39.3, 27.0 . MALDI-TOF MS(M_W 2425.3): *m/z* 2448.5 $[M+Na]^+$.

Compound **OL*ent*-2.**



OL*ent*-2 was prepared with the general procedure *B*; coupling yield of 94%. **^{1}H NMR** (700 MHz, DMF-d_7) δ (ppm):) 7.16-7.41 (m, 68 H), 6.63-6.70 (m, 6H), 6.44 (d, *J* = 9 Hz, 2H), 6.44 (d, *J* = 10 Hz, 2H), 6.39-6.41 (m, 2H), 6.23-6.31 (m, 8H), 6.10-6.16 (m, 4H), 6.03 (d, *J* = 9 Hz, 2H), 6.01 (d, *J* = 8 Hz, 2H), 4.17-4.30 (m, 6H), 4.01-4.12 (m, 6H), 3.91 (s, 2H), 2.78-2.88 (m, 6H), 2.70 (d, *J* = 4 Hz, 6H), 2.24-2.69 (m, 24H). **^{13}C NMR** (100 MHz, DMF-d_7) δ (ppm): 160.9, 160.7, 160.3, 159.7, 157.6, 136.2, 130.3, 130.1, 130.1, 130.0, 129.3, 129.2, 128.2, 129.1, 129.1, 127.2, 127.2, 127.1, 127.0, 126.9, 119.3, 52.2, 51.9, 51.8, 51.7, 51.5, 51.3, 47.0, 46.5, 45.9, 44.2, 40.8, 40.4, 39.6, 39.5, 39.3, 27.1. MALDI-TOF MS (M_W 2425.3): *m/z* 2448.5 $[M+Na]^+$

Compound **OL3.**

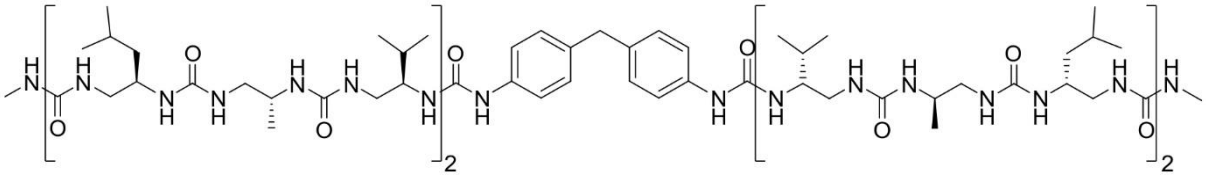


OL3 was synthesised with the general procedure *B*; coupling yield of 91%.**^{1}H NMR** (400 MHz, CD_3OH) δ (ppm): 7.25 (d, *J* = 8 Hz, 4H), 7.07 (d, *J* = 8 Hz, 4H), 6.31-6.52 (m, 10H), 6.20-6.23 (m, 4H), 5.89-6.10 (m, 12H), 5.84 (d, *J* = 9 Hz, 2H), 3.97-4.11 (m, 2H), 3.77-3.92 (m, 8H), 3.50-3.69 (m, 14H), 2.50-2.70 (m, 4H), 2.70 (d, *J* = 3 Hz, 6H), 2.27-2.42 (m, 8H), 1.54-1.78 (m, 8H), 1.14-1.34 (m, 8H), 0.84-1.03 (m, 60H). **^{13}C NMR** (100MHz, CD_3OH) δ (ppm): 162.1, 162.0, 161.8, 161.7, 161.3, 160.9, 158.99, 138.8, 137.1, 130.1, 120.4, 56.1, 47.1, 46.8, 46.3, 46.1, 44.6, 43.8, 42.6, 41.2, 31.9, 31.8, 26.8, 26.2, 25.8, 23.6, 22.7, 22.4, 20.1, 20.0, 18.8, 18.4, 18.3. MALDI-TOF MS (M_W 1793.2): *m/z* 1794.4 $[M+H]^+$.

residues		NH	NH '	$C_\alpha H^a$	$C_\alpha H^b$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	$C_\varepsilon H$	CH_{term}	$\Delta\delta$ ($C_\alpha H$)
CH_{Ar} NHCO			8,23							7,25 7,07 3,85	
V^u	1	5,93	6,03	2,60	3,63	3,82	1,72	0,97			1,03
A^u	2	5,84	6,04	2,31	3,57	3,85	1,01				1,26
L^u	3	6,02	6,50	2,35	3,65	3,91	1,17	1,69	0,89		1,30
V^u	4	6,33	6,47	2,38	3,66	3,58	1,57	0,88			1,28
A^u	5	5,98	6,43	2,34	3,56	4,05	1,01				1,22
L^u	6	6,04	6,39	2,63	3,59	3,89	1,23	1,70	0,91		0,96
	NHMe	6,20								2,70	

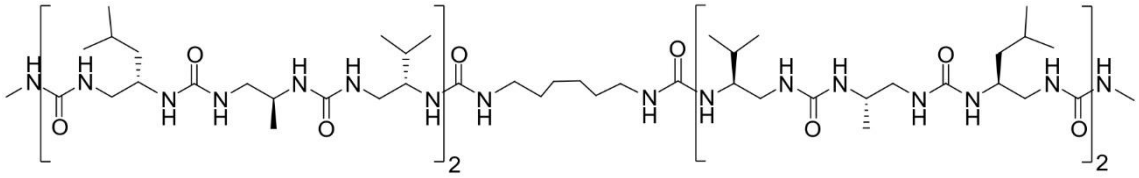
Table S5: ^{1}H NMR Chemical Shifts (ppm) of **OL3** in CD_3OH, 298 K.

Compound **OL*ent*-3**



OL*ent*-3 was prepared with the general procedure *B*; coupling yield of 74%. **^{1}H NMR** (300 MHz, CD_3OH) δ (ppm): 7.25 (d, *J* = 8 Hz, 4H), 7.07 (d, *J* = 8 Hz, 4H), 6.31-6.50 (m, 10H), 6.20-6.23 (m, 4H), 5.87-6.05 (m, 12H), 5.85 (d, *J* = 9 Hz, 2H), 4.04-4.07 (m, 2H), 3.76-3.95 (m, 8H), 3.48-3.73 (m, 14H), 2.50-2.70 (m, 4H), 2.71 (d, *J* = 4 Hz, 6H), 2.27-2.42 (m, 8H), 1.54-1.76 (m, 8H), 1.09-1.28 (m, 8H), 0.84-1.06 (m, 60H). **^{13}C NMR** (75 MHz, CD_3OH) δ (ppm): 162.0, 161.9, 161.8, 161.7, 161.3, 160.8, 158.9, 138.8, 137.0, 130.1, 120.3, 56.1, 47.1, 46.8, 46.3, 46.0, 44.6, 43.7, 42.6, 41.2, 32.0, 31.8, 26.8, 26.2, 25.8, 23.6, 23.5, 22.7, 22.4, 20.1, 20.0, 18.8, 18.5, 18.4, 18.3. MALDI-TOF MS (M_W 1793.2): *m/z* 1816.2 $[M+Na]^+$.

Compound **OL4**



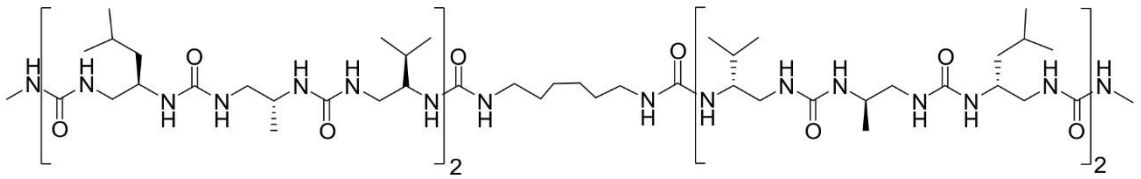
OL4 was prepared with the general procedure *B*; coupling yield of 94%. **^{1}H NMR** (400 MHz, CD_3OH) δ (ppm): 6.45-6.55 (m, 4H), 6.36-6.44 (m, 4H), 6.33 (d, *J* = 10 Hz, 2H), 6.17-6.28 (m,

4H), 6.14 (d, *J* = 10 Hz, 2H), 6.04 (d, *J* = 10 Hz, 2H), 5.92-5.99 (m, 4H), 5.82-5.91 (m, 4H), 5.70 (d, *J* = 10 Hz, 2H), 3.98-4.12 (m, 2H), 3.77-3.98 (m, 6H), 3.49-3.76 (m, 16H), 2.97-3.26 (m, 4H), 2.70 (d, *J* = 4 Hz, 6H), 2.25-2.68 (m, 12H), 1.46-1.79 (m, 8H), 1.34-1.41 (m, 4H), 1.15-1.34 (m, 8H), 1.03 (d, *J* = 6 Hz, 6H), 1.02 (d, *J* = 6 Hz, 6H), 0.87-0.98 (m, 48H). **^{13}C NMR** (100 MHz, CD_3OH) δ (ppm): 162.1, 162.0, 161.9, 161.8, 161.6, 161.4, 160.9, 56.1, 47.2, 46.8, 46.3, 46.0, 44.6, 43.8, 42.8, 41.1, 32.1, 31.8, 31.5, 27.7, 26.8, 26.2, 25.8, 23.6, 23.5, 22.7, 22.4, 20.1, 20.0, 18.8, 18.5, 18.4, 18.2. MALDI-TOF MS (M_W 1711.4): *m/z* 1712.4 $[M+H]^+$.

residues		NH	NH '	$C_\alpha H^a$	$C_\alpha H^b$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	$C_\varepsilon H$	CH_{term}	Δδ ($C_\alpha H$)
Linker NHCO			5,95							3,23-3,07 1,52 1,39	
V^u	**1**	5,70	5,88	2,56	3,61	3,71	1,69	0,93			1,05
A^u	**2**	5,85	6,23	2,36	3,60	3,88	1,04				1,24
L^u	**3**	6,14	6,50	2,33	3,65	3,91	1,20	1,72	0,93		1,32
V^u	**4**	6,33	6,49	2,36	3,63	3,58	1,58	0,88			1,27
A^u	**5**	5,99	6,42	2,35	3,56	4,05	1,04				1,21
L^u	**6**	6,04	6,39	2,64	3,60	3,90	1,26	1,72	0,92		0,96
	NHMe	6,20								2,70	

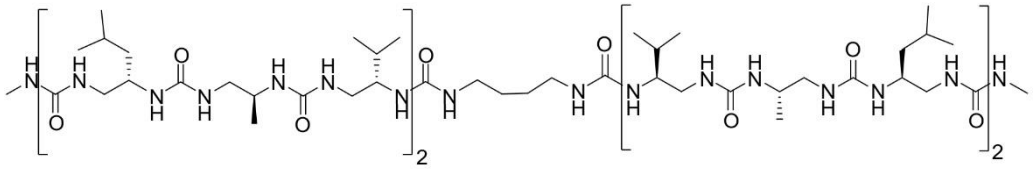
Table S6: ^{1}H NMR Chemical Shifts (ppm) of **OL4** in CD_3OH, 298 K.

Compound **OL*ent*-4.**



OL*ent*-4 was prepared with the general procedure *B*; coupling yield of 96%. **^{1}H NMR** (400 MHz, CD_3OH) δ (ppm): 6.36-6.50 (m, 8H), 6.32 (d, *J* = 10 Hz, 2H), 6.17-6.28 (m, 4H), 6.14 (d, *J* = 10 Hz, 2H), 5.80-6.08 (m, 10H), 5.68 (d, *J* = 9 Hz, 2H), 4.01-4.07 (m, 2H), 3.79-3.98 (m, 6H), 3.52-3.75 (m, 16H), 3.14-3.26 (m, 2H), 2.99-3.07 (m, 4H), 2.71 (d, *J* = 4 Hz, 6H), 2.50-2.67 (m, 4H), 2.21-2.46 (m, 10H), 1.44-1.79 (m, 12H), 1.33-1.43 (m, 4H), 1.16-1.31 (m, 8H), 1.03 (d, *J* = 6 Hz, 6H), 1.02 (d, *J* = 6 Hz, 6H), 0.84-0.96 (m, 48H). **^{13}C NMR** (100 MHz, CD_3OH) δ (ppm): 162.1, 162.0, 161.9, 161.8, 161.6, 161.3, 160.9, 56.2, 56.1, 47.1, 46.8, 46.2, 46.0, 44.6, 43.9, 42.9, 41.1, 32.2, 31.9, 31.6, 27.88, 26.8, 26.3, 25.9, 23.7, 23.7, 22.9, 22.5, 20.2, 20.1, 18.9, 18.6, 18.5, 18.3. MALDI-TOF MS (M_W 1711.4): *m/z* 1734.3$[M+Na]^+$.

Compound **OL5**



OL5 was prepared with general procedure *B* with tetramethylene diisocyanate (TtMDI) as linker. A white solid was obtained and was washed with DCM; yield of coupling of 37%. MALDI-TOF MS (M_W 1683.2): *m/z* 1706.2 $[M+Na]^+$.

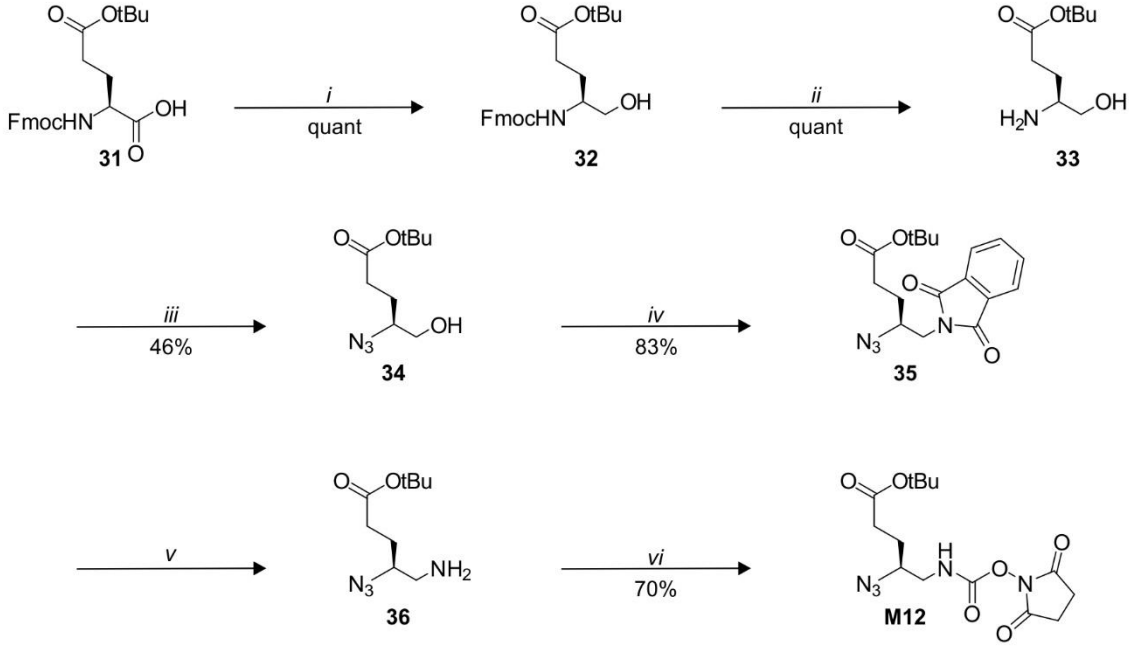
2. X-ray crystallography results of oligourea **OL1**-**OL5**

Crystals were obtained in DMF for oligourea **OL1** and **OL2**, in methanol for **OL4** and in DSMO for **OL3** and **OL5**.

	OL1	**OL2**	**OL3**	**OL4**	**OL5**
Formula	$C_{177}H_{212}N_{36}O_{18}$	$C_{137}H_{164}N_{28}O_{14}$	$C_{85}H_{156}N_{28}O_{14}$	$C_{78}H_{158}N_{28}O_{14}$	$C_{76}H_{154}N_{28}O_{14}$
M	3129,7	2425,30	1793,2	1711,4	1683,2
Crystal system	triclinic	monoclinic	monoclinic	monoclinic	orthorombic
Space group	P-1	$P1\ 2_1/n\ 1$	$P1\ 2_1\ 1$	$P2_1/c$	$P2_1\ 2_1\ 2_1$
Unit cell parameters					
a (Å)	19.94 (4)	13.08 (3)	11.33 (2)	19.32 (4)	13.84 (8)
b (Å)	22.06 (4)	30.62 (6)	33.44 (7)	54.93 (11)	25.56 (15)
c (Å)	22.74 (5)	35.69 (7)	30.58 (6)	19.64 (4)	27.32 (16)
α (°)	92.96 (3)	90	90	90	90
β (°)	94.13 (3)	94.81 (3)	90.00 (3)	91.64 (3)	90
γ (°)	93.47 (3)	90	90	90	90
Volume ($Å^3$)	9943 (3)	14249 (5)	11597 (4)	20848 (2)	9674 (10)
Z	2	4	4	8	8
Temperature (K)	293	293	293	293	373
ρ ($g.cm^{-1}$)	1.120	1.144	1,115	1,107	1.157
unique reflexions[a]	26101	9246	29419	29413	12140
Parameters/Restraints	2290/99	1126/147	2518/238	2285/122	1085/0
R_1[b]	0.0835	0.1843	0.1452	0.1196	0.0795
R_2[c]	0.2317	0.4527	0.3552	0.3180	0.1313
Goodness of fit	1.024	1.713	1.332	1.073	1.023

Table S7: X-ray crystallographic parameters for oligoureas **OL1**-**OL5**.

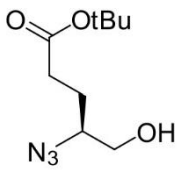
3. Preparation of the azide protected O-succinimidyl carbamate monomer **M12.**



i: 1) IBCF, NMM, THF, -20°C ; 2) $NaBH_4$, H_2O ; *ii*: DBU, THF, rt ; *iii:* $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4.5H_2O$, CH_3CN, H_2O, rt ; *iv*: PPh_3, phtalimide, THF, TA ; *v:* N_2H_4, MeOH, rt ; *vi:* DSC, DCM, rt

Scheme S6: Synthesis of O-succinimidyl carbamates monomers **M12**

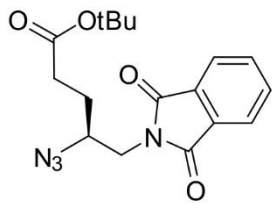
(S)-*tert*-butyl 4-azido-5-hydroxypentanoate 34



Fmoc-(L)-Glu(OtBu)-OH protected α-amino acid (1 equiv., 6 g, 13.53 mmol) was dissolved in anhydrous THF under N_2 atmosphere and cooled to -10 °C. *N*-methylmorpholine (1.2 equiv., 1.70 mL, 16.23 mmol), isobutylchloroformate (1.2 equiv., 2.10 mL, 16.23 mmol) were successively added and the reaction mixture was stirred at -10 °C for 30 min. The resulting white suspension was then filtered off and a solution of $NaBH_4$ (2 equiv., 1.02 g, 27.06 mmol) in water (10 mL) was added dropwise to the filtrate. The reaction was left to react 30 min and then concentrated under reduced pressure. AcOEt was added and the organic layer was washed with 1M $KHSO_4$ solution (x3), saturated solution of $NaHCO_3$ (x3), brine, dried over magnesium sulfate and concentrated under vacuum. N-Fmoc protected amino alcohol **32** was obtained as a white solid (quant.). **32** was dissolved in anhydrous THF under N_2 atmosphere. DBU (1.25 equiv., 2.50 mL, 16.91 mmol) was added dropwise and the reaction was stirred during 30 min at room temperature. THF was removed under reduced pressure and the amine was directly engaged in the next step without further purification. To a solution of the amine in ACN/H_2O mixture (1:1, 40 mL) were successively added K_2CO_3 (n x acid + 0.05 mmol, 2.31 g), $CuSO_4.5H_2O$ (0.01 equiv., 34 mg, 0.135 mmol) and imidazole-1-sulfonyl azide hydrochloride (1.2 equiv., 3.39 g, 16.23 mmol). The reaction mixture was stirred at room temperature for 3 h then concentrated under

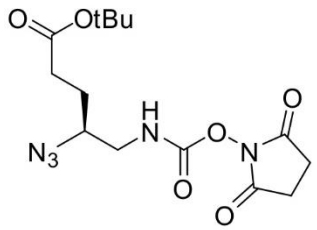
reduced pressure and re-dissolved in AcOEt (50 mL). The organic layer was then washed with $KHSO_4$ three times, twice with brine, dried over magnesium sulfate and concentrated under vacuum to afford, after silica gel flash chromatography (15% AcOEt in cyclohexane), the azide **34** as yellow oil (1.34 g, 46% over three steps). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 3.97-4.07 (m, 1H), 3.84 (dd, *J* = 9 Hz, 14 Hz, 1H), 3.72 (dd, *J* = 5 Hz, 13 Hz, 1 H), 2.93 (dd, *J* = 5 Hz, 13 Hz, 1H), 2.82 (dd, *J* = 9 Hz, 14 Hz, 2H), 1.54-1.59 (m, 2H), 1.32 (s, 9H).

(S)-*tert*-butyl 4-azido-5-(1,3-dioxoisoindolin-2-yl)pentanoate 35



To a solution of triphenylphosphine (1.2 equiv., 1.95 g, 7.46 mmol) and phtalimide (1.2 equiv., 1.10 g, 7.46 mmol) in anhydrous THF (40 mL) at 0 °C and under N_2 atmosphere pressure, was added DIAD (1.2 equiv., 1.47 mL, 7.46 mmol). A solution of **34** in anhydrous THF was added dropwise to the reaction mixture and the solution was stirred overnight at room temperature. The THF was evaporated under reduced pressure and the crude material was purified by silical gel flash chromatography (10% AcOEt in cyclohexane) to afford **35** as a yellow solid with 83% of yield (1.78 g). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 7.86 (dd, *J* = 3 Hz, 5 Hz, 2H), 7.72 (dd, *J* = 3 Hz, 5 Hz, 2 H), 3.69-3.87 (m, 3 H), 2.33-2.48 (m, 2H), 1.71-1.98 (m, 2H), 1.34 (s, 9H).

(S)-*tert*-butyl 4-azido-5-((((2,5-dioxopyrrolidin-1-yl)oxy)carbonyl)amino)pentanoate M12.



To a solution of azido phtalimide derivative **35** (1 equiv., 1.78 g, 5.17 mmol) in MeOH (50 mL) was added hydrazine hydrate 55% (2.5 equiv., 0.73 mL, 23.49 mmol). The reaction mixture was left to react overnight at room temperature and a white precipitate appeared. The reaction mixture was filtered off, washed with AcOEt and the filtrate was concentrated under reduced pressure. The crude material was dissolved in DCM (30 mL). The organic layer was washed twice with a 1M solution of HCl. The combined aqueous layers were neutralized by addition of solid $NaHCO_3$ until pH = 8. The aqueous layer was finally extracted with DCM and the combined layers were washed one time with brine, died over magnesium sulfate and resulting azido amine was directely engaged for activation step without further purification. To a stirred suspension of DSC (1.2 equiv., 1.59 g, 6.20 mmol) in freshly distilled DCM, was added dropwise a solution of azido amine derivative in DCM. The reaction was left to react 2 h and a 1M $KHSO_4$ solution was added. The organic layer was extracted with DCM, washed one time with brine, dried over magnesium sulfate and concentrated under vacuum. The resulting crude was purified by

precipitation in Et_2O to afford activated carbamate **M12** as a white solid with 70% of yield after two steps (1.28 g). **^{1}H NMR ($CDCl_3$, 300 MHz)** δ (ppm): 5.83 (brs, 1H), 3.60-3.68 (m, 1 H), 3.19-3.48n (m, 2H), 2.82 (s, 4H), 2.38 (dt, *J* = 2 Hz, 7 Hz, 2H), 1.76-1.94 (m, 2H), 1.34 (s, 9H). **^{13}C NMR ($CDCl_3$, 75 MHz)** δ (ppm): 172.1, 169.7, 151.7, 81.2, 61.1, 45.0, 31.4, 28.2, 26.9, 25.6. **HRMS (ESI-TOFMS)** m/z cald for $C_{14}H_{21}N_5O_6Na$ $[M+Na]^+$, 378.13895, found 378.13812.

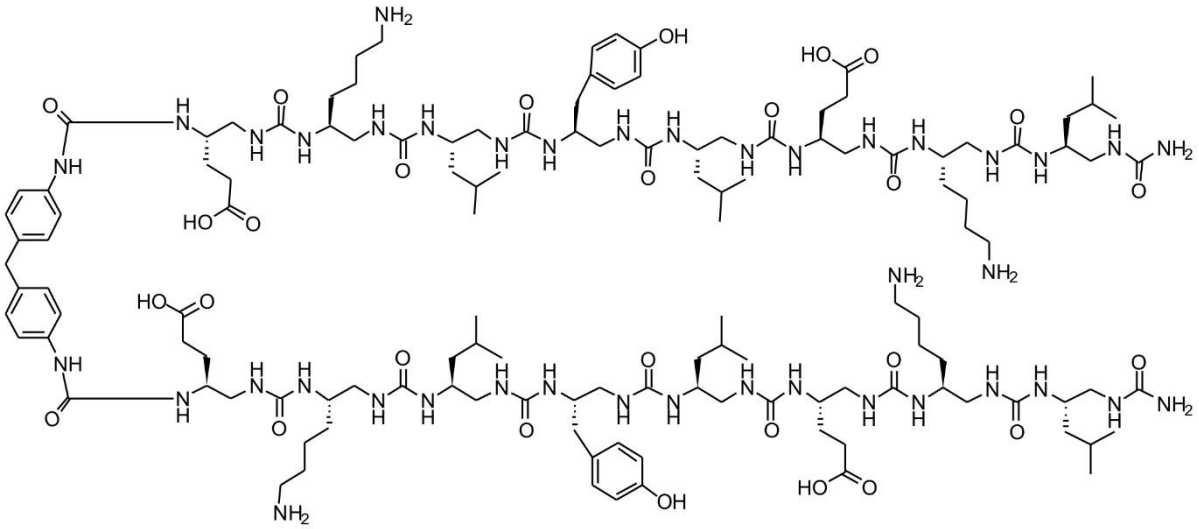
4. Synthesis of oligourea **OL8**-**OL12**

a. General procedure C for on-resin dimerization

Oligourea helices were synthesized on Rink amide MBHA resin (loading 0.87 mmol.g^{-1}) according to general procedure A. Dimerization of helices with diisocyanate linker was performed prior to final cleavage from the resin. After completion of the synthesis of oligourea helices, the resin was swelled in DMF for 30 min. The resin was filtered off and dissolved in DMF (3 mL), DIPEA (6 equiv. relative to reasin loading) and diisocyanate linker (0.3 eq relative to resin loading) were added successively. The reaction mixture was stirred overnight at room temperature. The resin was washed with DMF (6 x 5 mL), and the coupling efficiency was monitored with a chloranil test. Resin beads were then dissolved with a solution of dioxane/H_2O (7:3) for 1 hour at room temperature to quench all diiocyanate groups that didn't react. The resin beads were filtered off and washed with DMF (10 x 5mL), DCM (10 x 5 mL), Et_2O (5 x 5 mL) and dried under vacuum. The cleavage of the oligomer from the resin was performed according to general procedure *A* for 120 minutes.

b. Synthesis of oligoureas

Compound **OL8.**



OL8 was synthesized on 25 µmol scale following the general procedure *C*. The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 30-70%, 20 min, flow = 20 mL.min^{-1}. 20.3 mg was obtained with a global yield of 29%. ESI-MS (M_W 2781.73): m/z 1392.22 $[M+2H]^{2+}$, 928.45 $[M+3H]^{3+}$, 696.57 $[M+4H]^{4+}$, 557.45 $[M+5H]^{5+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 6.05 min. **^{1}H NMR (CD_3OH, 400 MHz)** δ (ppm): 8.23 (s, 2H), 7.24 (d, *J* = 8 Hz, 4H), 7.11 (d, *J* = 8 Hz, 4H), 7.02 (d, *J* = 8 Hz, 4H), 6.67 (d, *J* = 8 Hz, 4H), 6.58-6.64 (m, 2H), 6.49-.55 (m, 4H), 6.37-6.38 (m, 6H), 6.34 (d, *J* = 10 Hz, 2H), 6.20-6.26 (m, 2H), 6.07-6.17 (m, 4H), 5.86-6.07 (m, 18H), 3.7-4.05 (m, 18H), 3.47-3.73 (m, 16H), 2.85-2.97 (m, 12H), 2.65-2.81(m, 4H), 2.50-2.64 (m, 6H), 2.22-2.48 (m, 16H), 1.80-1.92 (m, 2H), 1.56-1.76 (m, 16H), 1.36-1.54 (m, 12H), 1.21-1.34 (m, 12H), 1.21-1.34 (m, 12H), 1.12-1.20 (m, 4H), 0.85-0.97 (m, 36H). **^{13}C NMR (CD_3OH, 100 MHz)** δ (ppm): 179.6, 131.1, 130.1, 120.9, 116.1, 47.1, 47.0, 46.8, 46.7, 40.0, 33.2, 33.1, 31.7, 31.7, 28.9, 28.9, 28.7, 28.7, 27.7, 26.2, 25.8, 24.3, 23.7, 23.6, 26.3, 22.8, 22.4.

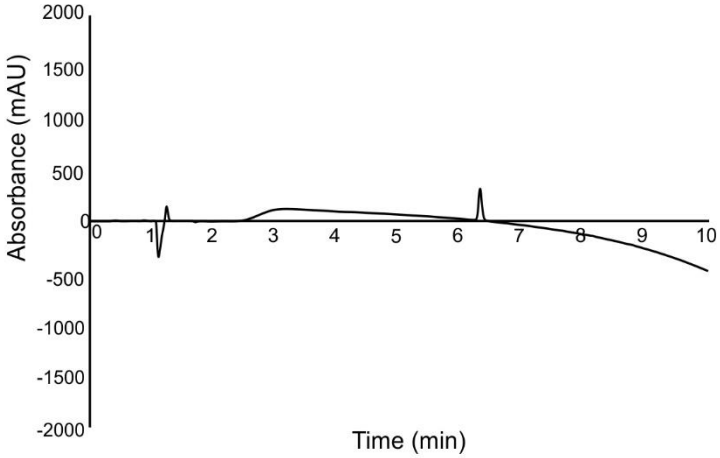
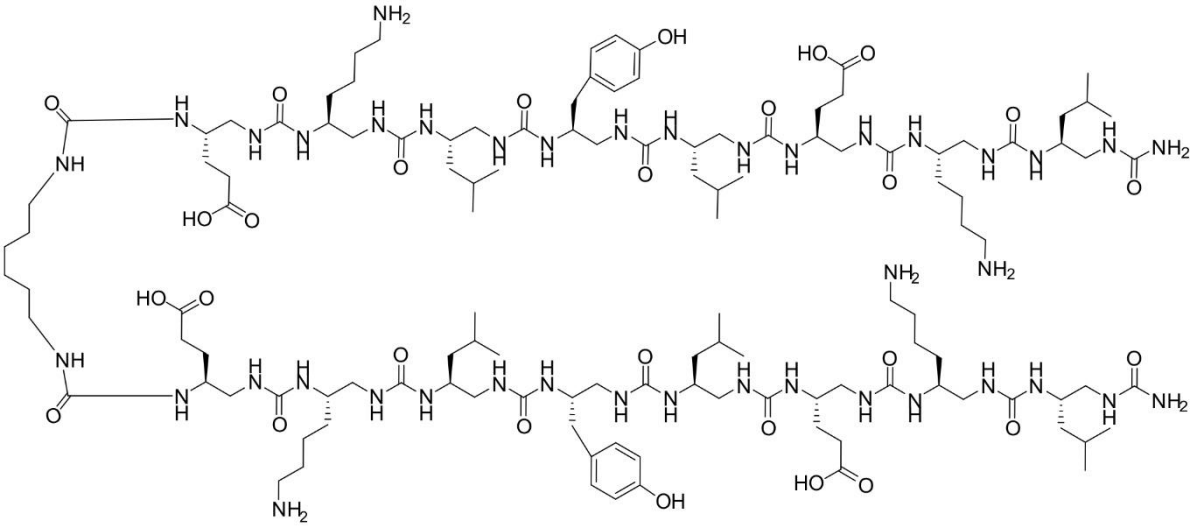


Figure S40: RP-HPLC chromatogram of **OL8** after purification (λ = 200 nm).

residues		NH	NH '	$C_\alpha H^a$	$C_\alpha H^b$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	$C_\varepsilon H$	$C_\zeta H$	CH_{term}	Δδ ($C_\alpha H$)
CH_{Ar} NHCO			8,23								7,24 7,11 3,83	
E^U	**1**	6,03	6,23	3,54	2,39	3,98	1,87-1,59	2,72				1,15
K^U	**2**	5,94	6,01	3,68	2,37	3,89	1,25	1,67	1,42	2,90		1,31
L^U	**3**	6,40	5,92	3,57	2,41	3,84	1,17	1,67	0,88			1,21
Y^U	**4**	6,34	6,53	3,68	2,37	3,94	2,59					1,31
L^U	**5**	5,96	6,60	3,61	2,34	3,87	1,16	1,68	0,89			1,27
E^U	**6**	6,05	6,43	3,57	2,28	3,78	1,40	2,89				1,29
K^U	**7**	5,90	6,44	3,57	2,37	3,91	1,27	1,67	1,42	2,92		1,20
L^U	**8**	6,11	6,51	3,49	2,74	3,90	1,26	1,70	0,92			0,75

Table S8: ^{1}H Chemical shift (in ppm) of **OL8** in CD_3OH (400 MHz) at 5 mM.

Compound **OL9.**



OL9 was synthesized on 25 µmol scale following the general procedure *C*. The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 30-70%, 20 min, flow = 20 mL.min^{-1}. 10.4 mg was obtained with a global yield of 15%. ESI-MS (M_W 2699.75): m/z 1351.23 $[M+2H]^{2+}$, 901.13 $[M+3H]^{3+}$, 676.08 $[M+4H]^{4+}$, 541.06 $[M+5H]^{5+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 7.21 min. **^{1}H NMR (CD_3OH, 400 MHz)** δ (ppm): 7.02 (d, *J* = 8 Hz, 4H), 6.67 (d, *J* = 8 Hz, 4H), 6.39-6.63 (m, 12H), 6.34 (d, *J* = 10 Hz, 2H), 5.94-6.24 (m, 16H), 5.90 (d, *J* = 10 Hz, 2H), 5.64-5.85 (m, 4H), 3.40-4.04 (m, 32H), 3.12-3.24 (m, 2H), 3.00-3.09 (m, 2H), 2.90-2.93 (m, 12H), 2.65-2.79 (m, 4H), 2.54-2.65 (m, 4H), 2.22-2.49 (m, 14H), 1.78-1.90 (m, 2H), 1.59-1.77 (m, 16H), 1.10-1.56 (m, 40H), 0.85-0.98 (m, 38H). **^{13}C NMR (CD_3OH, 100 MHz)** δ (ppm): 179.9, 177.2, 163.1, 162.3, 162.1, 12.0, 161.8, 161.4, 161.2, 157.0, 131.1, 129.9, 116.1, 52.1, 51.3, 50.9, 50.2, 47.1, 46.7, 43.6, 42.7, 42.6, 41.2, 40.7, 39.1, 33.1, 31.8, 31.7, 31.3, 29.0, 28.7, 27.9, 27.7, 26.1, 25.8, 24.8, 23.7, 23.7, 23.6, 23.5, 22.8, 22.8, 22.4.

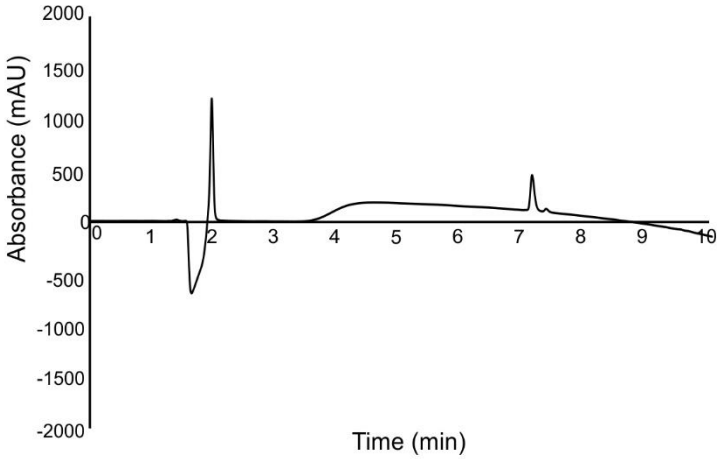
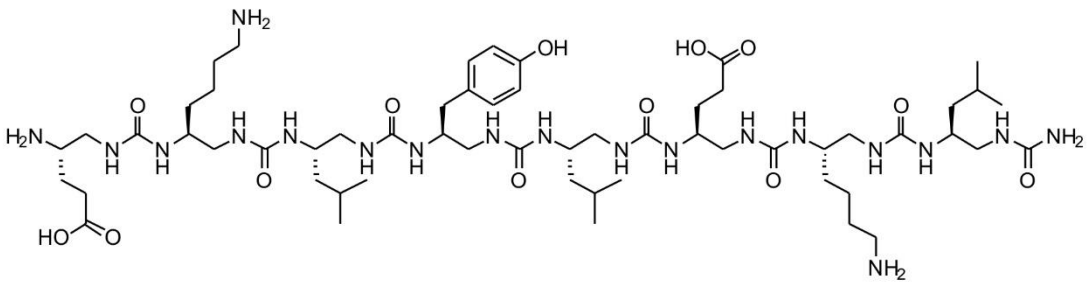


Figure S41: RP-HPLC chromatogram of **OL9** after purification (λ = 200 nm).

residues		NH	NH '	$C_\alpha H^a$	$C_\alpha H^b$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	$C_\varepsilon H$	$C_\zeta H$	CH_{term}	Δδ ($C_\alpha H$)
$(CH_2)_3NHCO$			5,95								3,90-2,36 1,19 0,90	
E^u	**1**	5,79	6,15	3,48	2,35	3,85	1,84-1,54	2,69				1,13
K^u	**2**	5,95	6,21	3,58	2,44	3,71	1,28	1,65	1,48	2,92		1,14
L^u	**3**	6,05	6,42	3,58	2,38	3,86	1,20	1,70	0,90			1,20
Y^u	**4**	6,33	6,54	3,69	2,36	3,95	2,60					1,33
L^u	**5**	5,98	6,59	3,61	2,35	3,88	1,16	1,69	0,90			1,26
E^u	**6**	6,03	6,44	3,58	2,39	3,92	1,69-1,46	2,92				1,19
K^u	**7**	5,90	6,45	3,59	2,38	3,92	1,27	1,67	1,46	2,93		1,21
L^u	**8**	6,11	6,51	3,50	2,75	3,90	1,26	1,70	0,90			0,75

Table S9: ^{1}H Chemical shift (in ppm) of **OL9** in CD_3OH (400 MHz) at 5 mM.

Compound **OL10**



OL10 was synthesized on 75 µmol scale following the general procedure *A*. The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 40.7 mg was obtained with a total yield of 43%. ESI-MS (M_W 1265.83): m/z 1266.59 $[M+H]^+$, 633.75 $[M+2H]^{2+}$, 422.81 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_R = 5.14 min. **^{1}H NMR (CD_3OH, 400 MHz)** δ (ppm): 7.01 (d, *J* = 8 Hz, 2H), 6.67 (d, *J* = 8 Hz, 2H), 6.39-6.57 (m, 4H), 6.06-6.26 (m, 4H), 6.02 (d, *J* = 9 Hz, 1H), 5.94 (d, *J* = 10 Hz, 2H), 5.90 (d, *J* = 10 Hz, 2H), 5.84 (d, *J* = 10 Hz, 1H), 3.42-4.01 (m, 16H), 2.82-2.98 (m, 4H), 2.23-2.80 (m, 10H), 1.93-2.03 (m, 2H), 1.80-1.92 (m, 2H), 1.58-1.75 (m, 9H), 1.12-1.56 (m, 16H), 0.85-0.96 (m, 18H). **^{13}C NMR (CD_3OH, 100 MHz)** δ (ppm): 176.8, 176.1, 162.3, 162.1, 161.8, 161.3, 161.3, 161.2, 161.2, 157.0, 131.0, 130.0, 116.1, 53.3, 52.3, 51.6, 50.9, 46.9, 46.0, 43.5, 42.7, 42.3, 40.7, 40.6, 39.1, 33.2, 33.1, 31.6, 30.8, 28.8, 28.7, 27.9, 26.7, 26.1, 26.1, 25.8, 24.1, 23.7, 23.6, 23.5, 22.7, 22.6, 22.3.

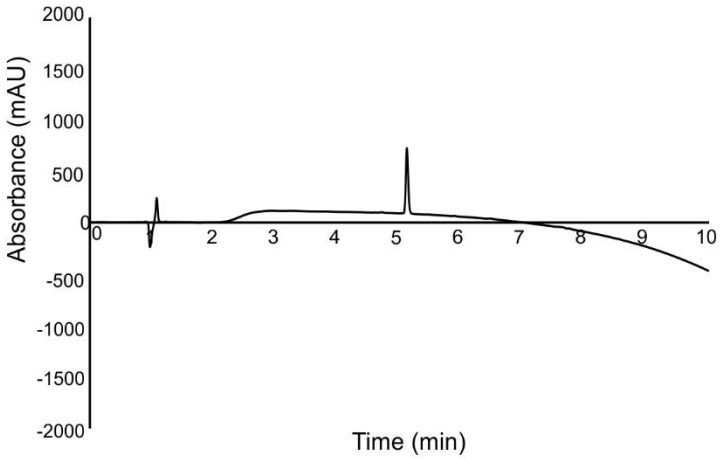
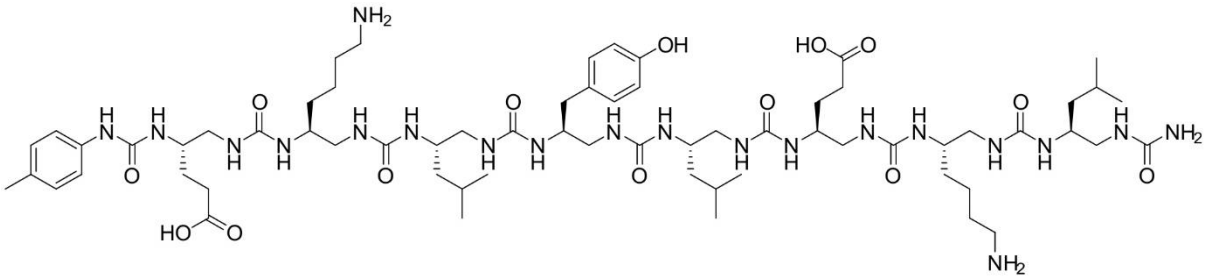


Figure S42: RP-HPLC chromatogram of **OL10** after purification (λ = 200 nm).

residues		NH	NH '	$C_\alpha H^a$	$C_\alpha H^b$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	$C_\varepsilon H$	$C_\zeta H$	Δδ ($C_\alpha H$)
E^u	**1**		6,45	3,54	2,51	3,33	1,84	1,97			1,03
K^u	**2**	6,22	6,22	3,47	2,70	3,72	1,35	1,65	1,47	2,91	0,77
L^u	**3**	5,85	6,18	3,51	2,46	3,86	1,22	1,68	0,90		1,05
Y^u	**4**	6,13	6,50	3,64	2,65	3,96	2,57				0,99
L^u	**5**	5,93	6,51	3,60	2,36	3,87	1,16	1,68	0,90		1,24
E^u	**6**	6,01	6,40	3,57	2,30	3,7	1,40	1,68			1,27
K^u	**7**	5,90	6,42	3,60	2,37	3,89	1,27	1,66	1,44	2,92	1,23
L^u	**8**	6,08	6,49	3,50	2,74	3,89	1,26	1,69	0,92		0,76

Table S10: ^{1}H Chemical shift (in ppm) of **OL10** in CD_3OH (400 MHz) at 5 mM.

Compound **OL11.**



OL11 was synthesized on 25 µmol scale following the general procedure *C* with 1 equiv. of *p*-tolylisocyanate. The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 10.5 mg was obtained with a global yield of 30%. ESI-MS (M_W 1398.90): m/z 1399.74 $[M+H]^+$, 700.34 $[M+2H]^{2+}$, 467.21 $[M+3H]^{3+}$; HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 6.05 min.

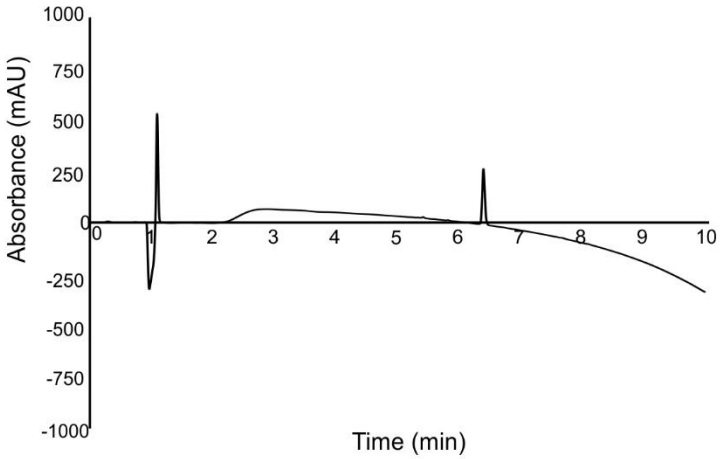
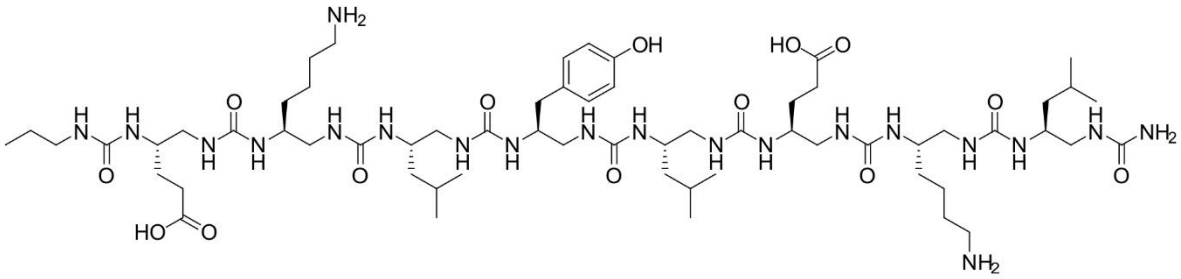


Figure S43: RP-HPLC chromatogram of **OL11** after purification (λ = 200 nm).

Compound **OL12.**



OL12 was synthesized on 25 µmol following the general procedure *C* with 1 equiv. of propylisocyanate. The product was purified by preparative HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 20-60%, 20 min, flow = 20 mL.min^{-1}. 8.9 mg was obtained with a global yield of 26%. ESI-MS (M_W 1350.90): m/z 1351.73 $[M+H]^+$, 676.34 $[M+2H]^{2+}$;HPLC (H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100%, 10 min) t_r = 5.89 min.

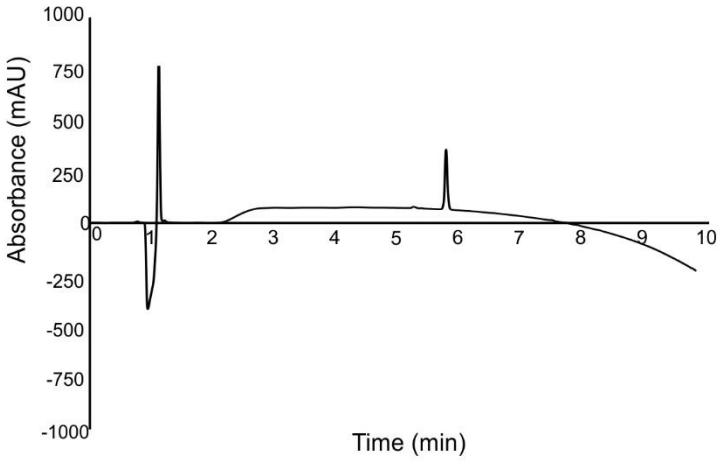


Figure S44: RP-HPLC chromatogram of **OL12** after purification (λ = 200 nm).

5. X-ray crystallography results of watersoluble oligourea **OL12**

For crystallization trials, lyophilized powder of oligourea **OL12** was dissolved to final concentration of 10 mg.mL^{-1} in ddH_2O. Crystallization trial was performed at 20 °C in standard aqueous hanging drop composed of 0.5 µL of oligourea solution plus 0.5 µL of crystallization reagent. Crystals were obtained in 28% PEG400, HEPES-Na 100 mM, 200 mM calcium chloride, pH 7.5.

	OL12
Formula	$C_{61}H_{114}N_{20}O_{14}$
M	1380,88
Crystal system	monoclinic
Space group	C 1 2 1
Unit cell parameters	
a (Å)	36.49 (2)
b (Å)	10.31 (7)
c (Å)	27.24 (3)
α (°)	90
β (°)	125.66 (11)
γ (°)	90
Volume (Å^3)	8332 (17)
Z	4
Temperature (K)	293
ρ (g.cm^{-1})	1.187
unique reflexions[a]	6976
Parameters/Restraints	987/18
R_1[b]	0.0614
R_2[c]	0.1615
Goodness of fit	1.069

Table S11: X-ray crystallographic parameters for oligourea **OL12**.

6. Iodination of oligoureas **OL9** and **OL10**

N-iodosuccinimide was added to a solution of oligourea (1 equiv.) at 4.5 mM dissolved in a mixture of ACN/H_2O (1:1). The reaction was allowed to stir at room temperature, and was directly injected in a semi-preparative RP-HPLC in order to isolate iodinated compound. HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); gradient 10-100 %, 20 min, flow = 4 mL.min^{-1}. ^{1}H NMR (CD_3OH,400 MHz) of compound **OL9-I$_4$** at 2mM confirmed the presence of iodine on the aromatic ring. 2D NMR revealed more signals in density maps than in 2D NMR of **OL9**, which might be in favour of a non-symmetrical conformation in solution, thereof an opened conformation of the dimer in presence of iodine is suggested.

Compound **OL9-I$_4$ 1.**

Reaction was performed on a 0.37 µmol scale (1 mg) and NIS (5 equiv.) 5 min. **OL9-I$_4$** was obtained as a white solide with 25% of yield (0.3 mg). t_r = 10.89 min. ESI-MS (M_W 3204.99): m/z 1603.36 $[M+2H]^{2+}$, 1069.27 $[M+3H]^{3+}$, 802.22 $[M+4H]^{4+}$.

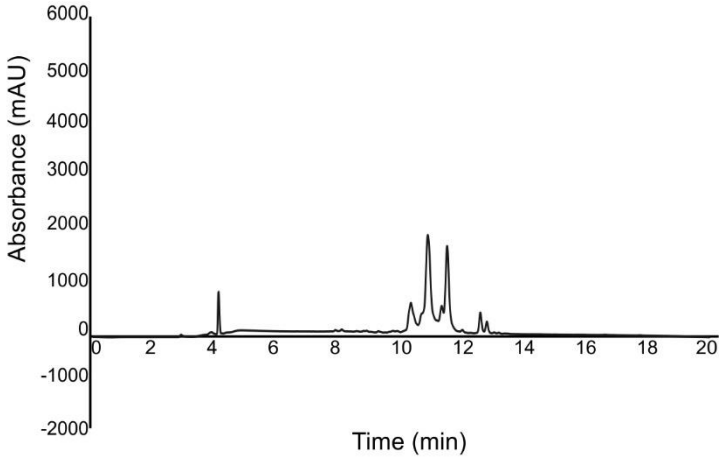


Figure S45: RP-HPLC chromatogram of **OL9-I$_4$** (λ = 200 nm).

Compound **OL9-I$_4$ 2.**

Reaction was performed on a 0.37 µmol scale (1 mg) and NIS (3.5 equiv) 5 min. **OL9-I$_4$** was obtained with 34% of yield (0.4 mg). t_r = 11.15 min. ESI-MS (M_W 3204.99): m/z 1603.27 $[M+2H]^{2+}$, 1069.20 $[M+3H]^{3+}$, 802.16 $[M+4H]^{4+}$.

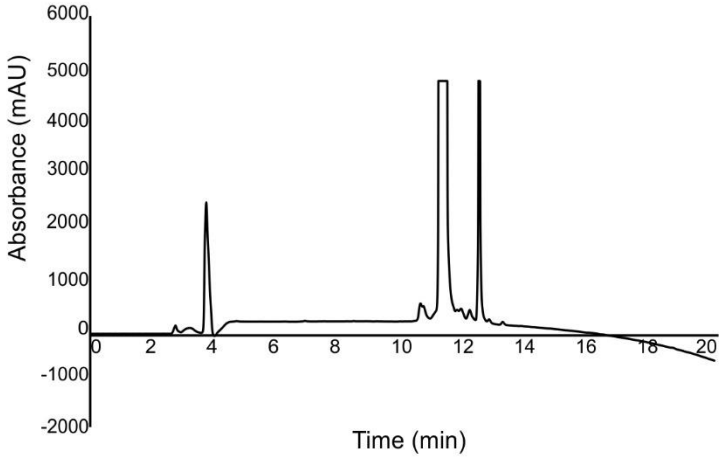


Figure S46: RP-HPLC chromatogram of **OL9-I$_4$** (λ = 200 nm).

Compound **OL9-I$_4$ 3.**

Reaction was performed on a 0.37 µmol scale (1 mg) and NIS (5 equiv.) 2.5 min. **OL9-I$_4$** was obtained with 34% yield (0.4 mg). t_r = 11.22 min. ESI-MS (M_W 3204.99): m/z 1603.26 $[M+2H]^{2+}$, 1069.20 $[M+3H]^{3+}$, 802.16 $[M+4H]^{4+}$.

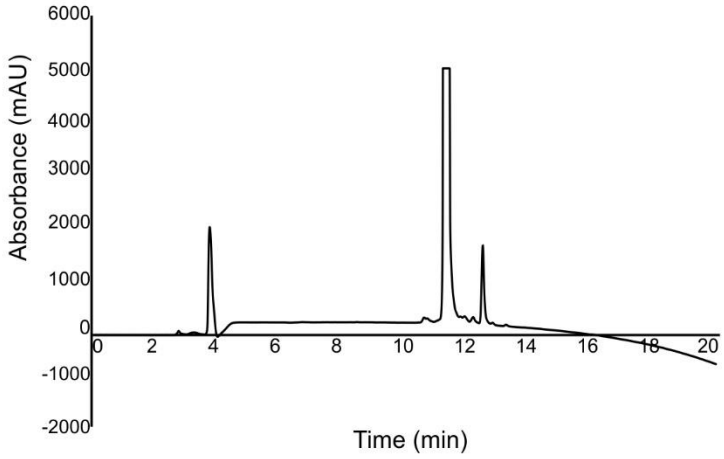


Figure S47: RP-HPLC chromatogram of **OL9-I$_4$** (λ = 200 nm).

Compound **OL9-I$_4$ 4.**

Reaction was performed on a 0.37 µmol scale (1 mg) and NIS (5 equiv.) 7.5 min. **OL9-I$_4$** was obtained as a white solid with 59% yield (0.7 mg). t_r = 11.21 min. ESI-MS (M_W 3204.99): m/z 1603.26 $[M+2H]^{2+}$, 1069.19 $[M+3H]^{3+}$, 802.15 $[M+4H]^{4+}$.

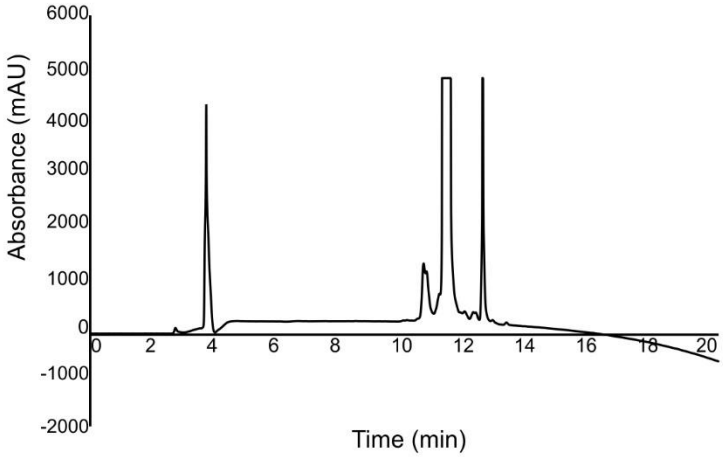


Figure S48: RP-HPLC chromatogram of **OL9-I$_4$** (λ = 200 nm).

Compound **OL10-I$_2$.**

Reaction was performed on a 0.789 µmol scale (1 mg) and NIS (5 equiv.) 10min. **OL10-I$_2$** was obtained as a white solid with 33% of yield (0.4 mg). t_r = 9.90 min. ESI-MS (M_W 1517.62): m/z 1518.80 $[M+H]^{+}$, 759.94 $[M+2H]^{2+}$, 506.98 $[M+3H]^{3+}$.

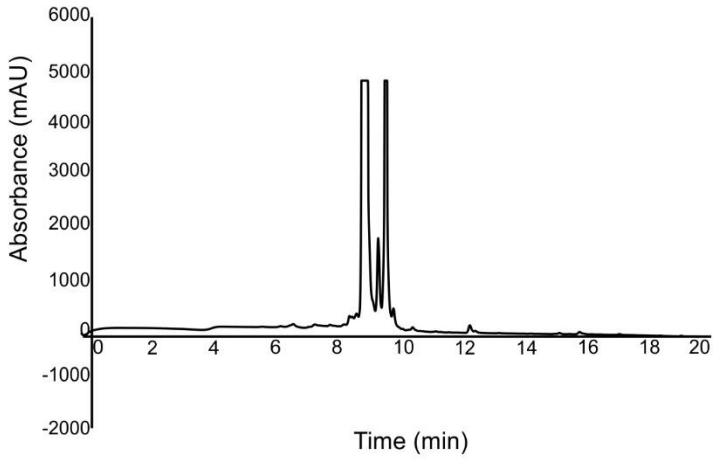


Figure S49: RP-HPLC chromatogram of **OL10-I₂** (λ = 200 nm).

7. <u>Circular dichroism of oligourea **OL8**-**OL12**</u>

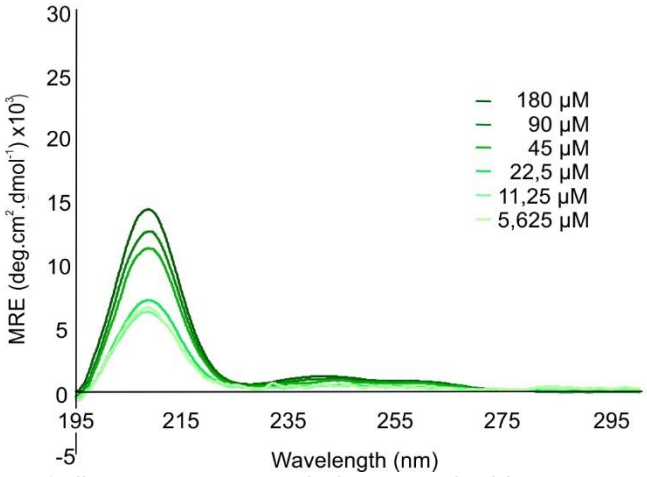


Figure S50: CD spectra of oligourea **OL8** recorded at 20 °C in ddH_2O.

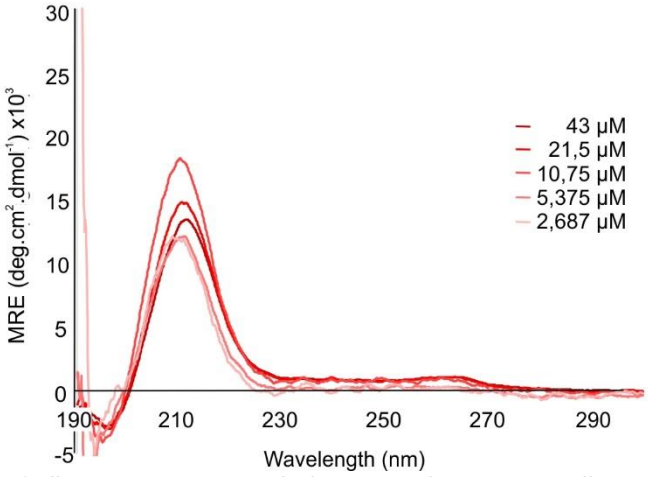


Figure S51: CD spectra of oligourea **OL8** recorded at 20 °C in 20 mM sodium phosphate pH 5.8.

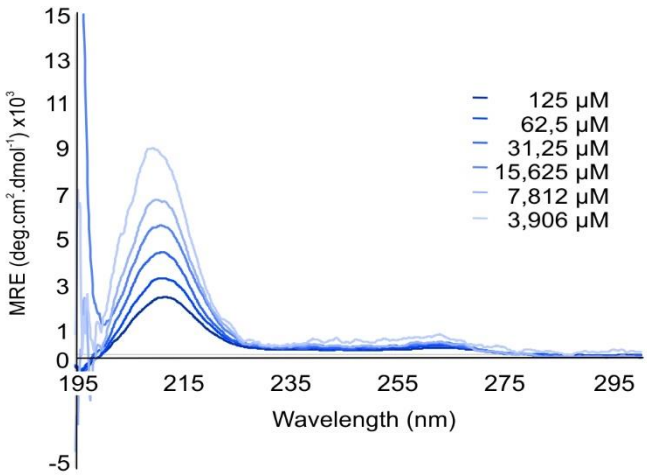


Figure S52: CD spectra of oligourea **OL8** recorded at 20 °C in 20 mM sodium phosphate pH 8.2

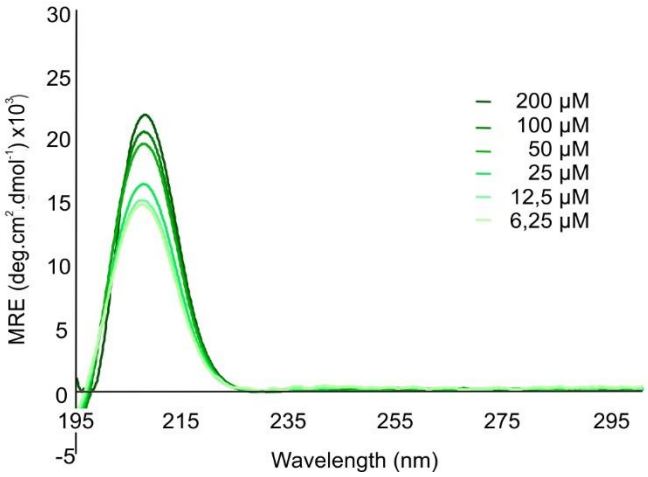


Figure S53: CD spectra of oligourea **OL9** recorded at 20 °C in ddH_2O.

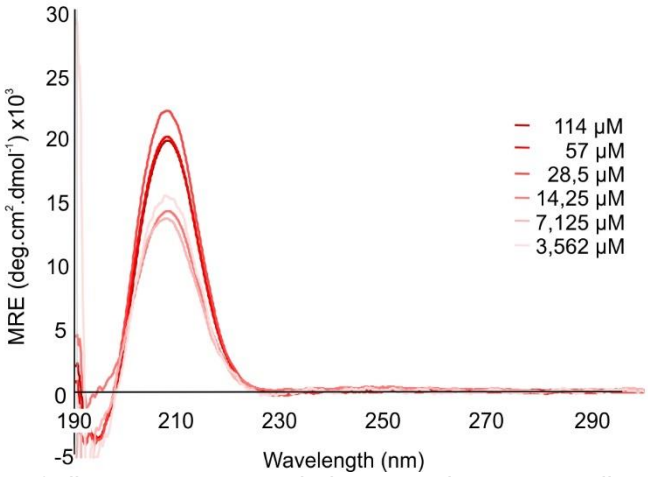


Figure S54: CD spectra of oligourea **OL9** recorded at 20 °C in 20 mM sodium phosphate pH 5.8.

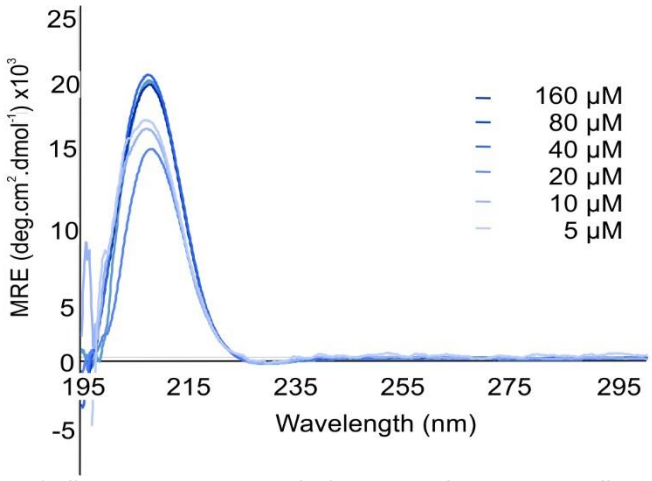


Figure S55: CD spectra of oligourea **OL9** recorded at 20 °C in 20 mM sodium phosphate pH 8.2.

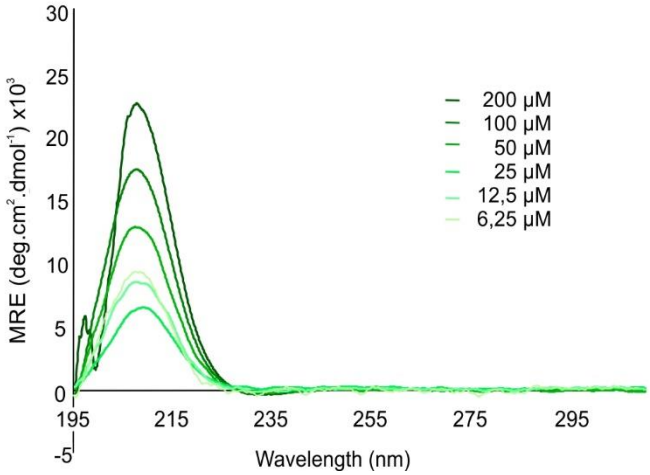


Figure S56: CD spectra of oligourea **OL10** recorded at 20 °C in ddH_2O.

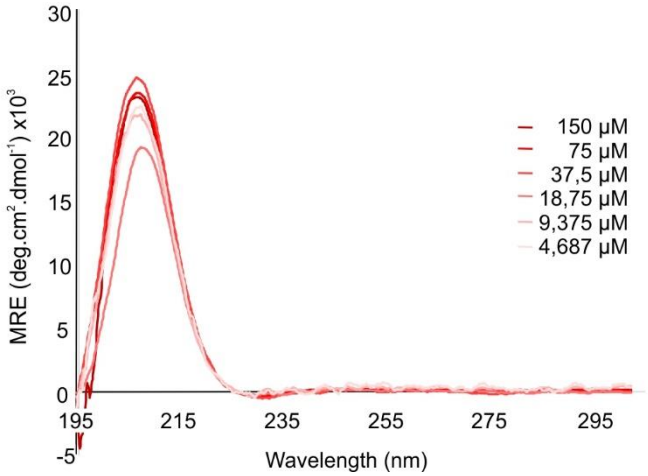


Figure S57: CD spectra of oligourea **OL10** recorded at 20 °C in 20 mM sodium phosphate pH 5.8.

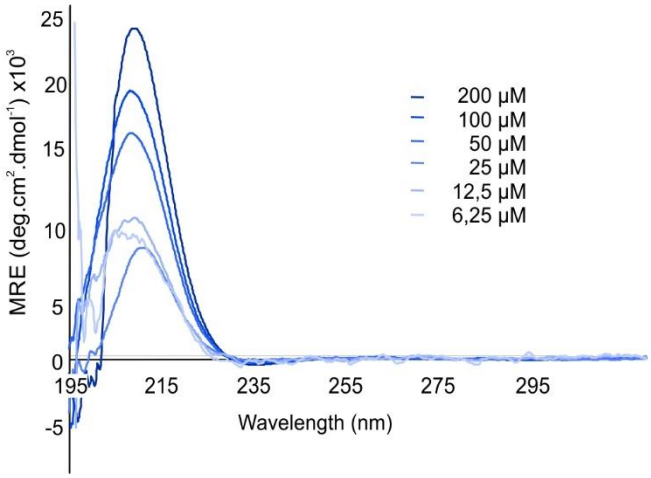


Figure S58: CD spectra of oligourea **OL10** recorded at 20 °C in 20 mM sodium phosphate pH 8.2.

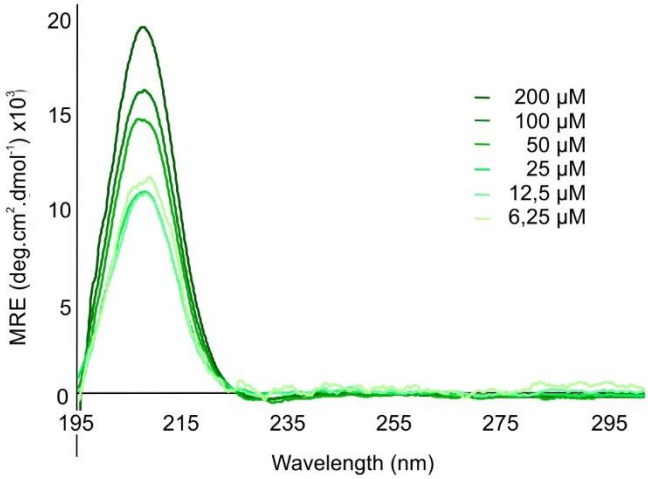


Figure S59: CD spectra of oligourea **OL11** recorded at 20 °C in ddH_2O.

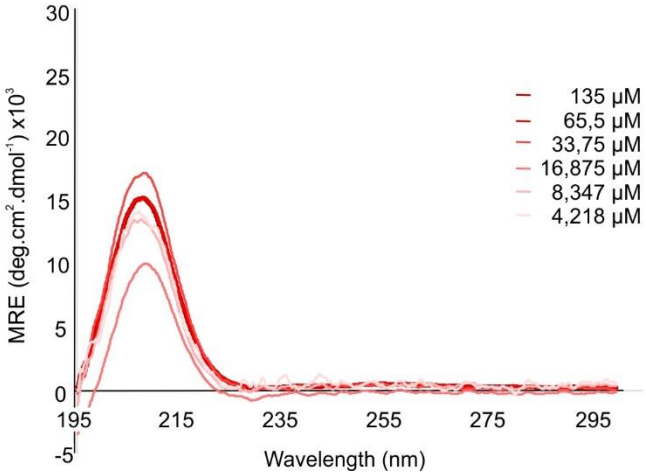


Figure S60: CD spectra of oligourea **OL11** recorded at 20 °C in 20 mM sodium phosphate pH 5.8.

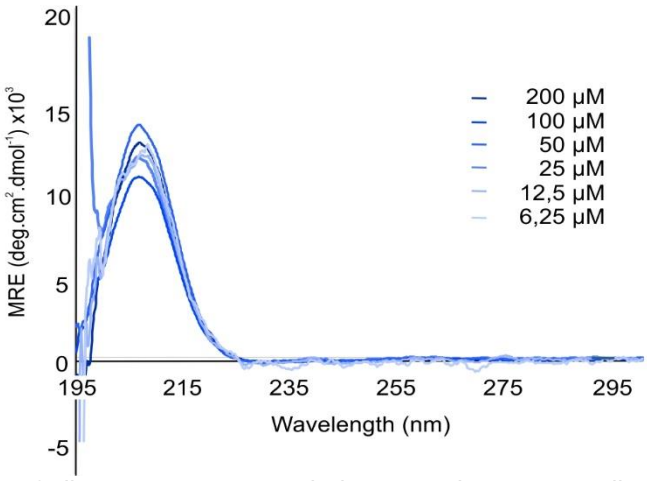


Figure S61: CD spectra of oligourea **OL11** recorded at 20 °C in 20 mM sodium phosphate pH 8.2.

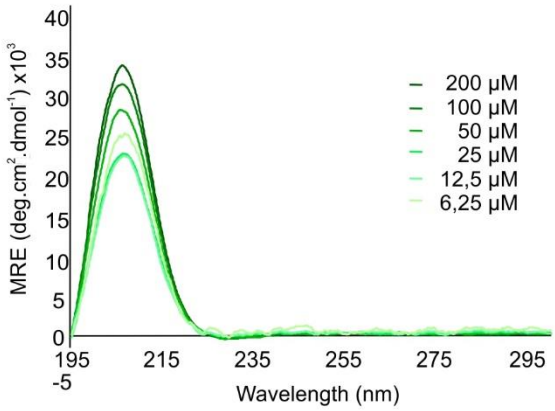


Figure S62: CD spectra of oligourea **OL12** recorded at 20 °C in ddH_2O.

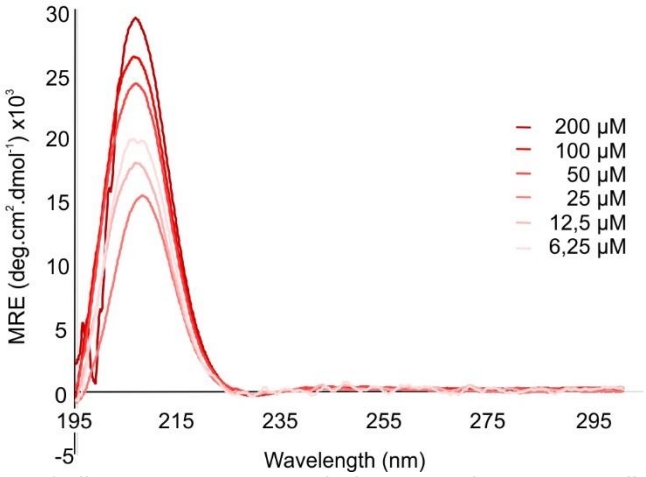


Figure S63: CD spectra of oligourea **OL12** recorded at 20 °C in 20 mM sodium phosphate pH 5.8.

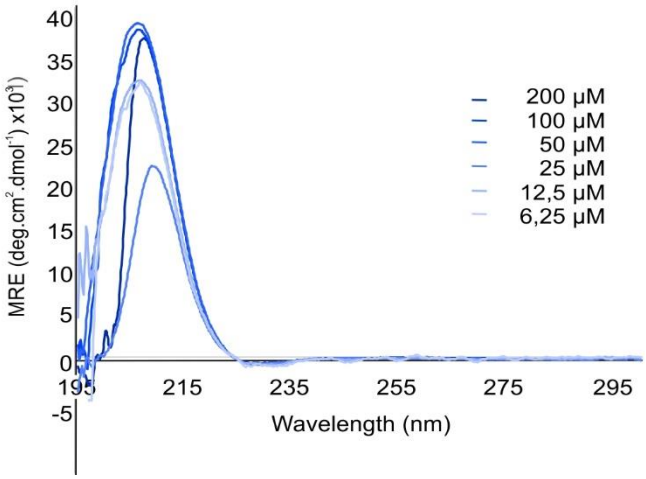


Figure S64: CD spectra of oligourea **OL12** recorded at 20 °C in 20 mM sodium phosphate pH 8.2.

Références bibliographiques

1. T. Vojkovsky *Org. Lett.* **1995**, *8*, 236-237.
2. C. Douat-Casassus, K. Pulka, P. Claudon, G. Guichard *Org. Lett.* **2012**, *14*, 3130-3133.
3. N. Pendem, C. Douat, P. Claudon, M. Laguerre, S. Castano, B. Desbat, D. Cavagnat, E. Ennifar, B. Kauffmann, G. Guichard *J. Am. Chem.* **2013**, *288*, 10914-10922.
4. (a) F. H. Niesen, H. Berglund, M. Vedali *Nat. Protoc.* 2007, *2*, 2212-2221; (b) D. Matulis, J. K. Kranz, F. R. Salemme, M. J. Todd *Biochemistry* 2005, *44*, 5258–5266; (c) G. A. Senisterra *J. Biomol. Screen.* 2006, *11*, 940–948 (2006); (d) M. Vedadi *Proc. Natl. Acad. Sci. USA* 2006, *103*, 15835–15840.
5. D. Ross, M. Gaitan, L.E. Locascio *Anal. Chem.* **2001**, *73*, 4117-4123.
6. (a) S.A.I Seidel, P. M. Dijkman, W. A. Lea, G. van den Bogaart, M. Jerabek-Willemsen, A. Lazic, J. S. Joseph, P. Srinivasan, P. Baaske, A. Watts, D. Braun, S. Durh *Methods.* **2013**, *59*, 301-315; (b) C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun, S. Duhr *Nat. Commun* **2010**, *1*, 100

Inhibition d'interactions protéine-protéine par des foldamères mixtes oligoamide/oligourée

Les interactions protéine–protéine (IPP) jouent un rôle primordial dans les processus physiologiques. L'inhibition de ces interactions ouvre la voie à la conception de nouvelles molécules à visée thérapeutique. Les structures secondaires en hélice α sont fréquemment impliquées dans les interactions entre protéines auxquelles elles peuvent contribuer de manière significative. La conception de molécules, mimant ce motif de reconnaissance et pouvant interagir avec la protéine cible tout en inhibant la reconnaissance avec le partenaire naturel, représente une voie innovante pour trouver de nouveaux candidats médicaments. Les oligomères d'urée aliphatique, une classe de foldamères qui adoptent une structure secondaire en hélice bien définie et proche de l'hélice α, ont été proposés comme mimes d'hélice α pour inhiber les IPP. Au cours de cette thèse, nous nous sommes d'abord intéressés à la conception de foldamères d'oligourée et de chimères oligoamide/oligourée pour cibler des surfaces de protéine. Nous avons sélectionné le récepteur nucléaire de la vitamine D (VDR) comme modèle d'étude en raison de son intérêt thérapeutique, et des connaissances structurales disponibles. Les protéines partenaires de VDR (coactivateurs) interagissant via une courte région structurée en hélice α, nos recherches ont portés sur des mimes d'hélices inspirés des séquences de coactivateurs. Dans une seconde partie, nous nous sommes intéressés à la génération et à l'étude de dimères covalents de foldamères, qui pourraient être utilisés pour couvrir des surfaces d'interaction plus larges. En effet, les interactions protéine-protéine montrent souvent un mode d'interaction plus complexe qu'une simple hélice, faisant intervenir des structures tertiaires et quaternaires de type *coiled coils,* qui peuvent aussi servir de point de départ pour la conception de nouvelles classes d'inhibiteurs.

Mots clés : Foldamères, Oligourée, Interactions protéine-protéine, Structures supersecondaires

Protein-protein interactions inhibition by mixed oligoamide/oligourea foldamers

Protein-protein interactions (PPI) have a key role in physiological processes. The inhibition of these PPI may lead to new therapeutic strategies. Secondary structures in α-helix are frequently involved in protein interactions where they may contribute significantly to binding. Designing molecules which mimic the helical motif for protein surface recognition and inhibition of the natural partner represents an innovative path to discover new drug candidates. Aliphatic urea oligomers, a class of foldamers that adopt a well-defined H-bonded helical secondary structure with good similarity to the α-helix have been proposed as possible α-helix mimics to inhibit protein-protein interactions. The first part of this PhD project was dedicated to the design and synthesis of oligoureas and oligourea/α-peptide chimeras for specific protein surface recognition. We have selected the vitamin D receptor as a potential target, mainly because (i) it is therapeutically relevant; (ii) its protein partner (coactivators) interact through a short region which adopts an α-helical structure upon binding and (iii) structures at atomic resolution were available to enable the design of effective mimetics. In the second part, we investigated methods to generate foldamer covalent dimers that could potentially be used to cover larger interaction surfaces. The rationale is that the binding interface is often more complex than a single helix and may involve tertiary and quaternary structures such as coiled coils which in turns may also serve as a basis for the design of new classes of inhibitors.

Keywords : Foldamers, Oligourea, Protein-protein interactions, Supersecondary structures

Unité de recherche

CBMN- Institut de Chimie et de Biologie des Membranes et des Nano-objets – UMR5248
IECB- Institut Européen de Chimie et Biologie - 2, rue Robert Escarpit – 33607, Pessac France