Protein Surface Recognition with Urea-based foldamers : application to the design of ligands targeting histone chaperone proteins

Johanne Mbianda

THÈSE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR DE

L'UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

SPÉCIALITÉ CHIMIE ORGANIQUE

Par Johanne MBIANDA

PROTEIN SURFACE RECOGNITION WITH UREA-BASED FOLDAMERS: APPLICATION TO THE DESIGN OF LIGANDS TARGETING HISTONE CHAPERONE PROTEINS

Sous la direction de : Dr. Gilles GUICHARD
(co-directeur : Dr. Céline DOUAT)

Soutenance le 08 Octobre 2018

Membres du jury :

M. AITKEN, David	Professeur, Université Paris Sud	Rapporteur
M. ENNIFAR, Eric	Directeur de recherches CNRS, Université de Strasbourg	Rapporteur
Mme BESTEL, Isabelle	Professeur, Université de Bordeaux	Présidente du jury
Mme OCHSENBEIN, Françoise	Chercheur CEA, Université Paris Saclay	Examinateur
Mme DOUAT, Céline	Chargé de recherches, Université de Munich	Co-directeur
M. GUICHARD, Gilles	Directeur de recherches CNRS, Université de Bordeaux	Directeur

Reconnaissance de surfaces de protéines avec des foldamères à base d'urées : Application au design de ligands ciblant une protéine chaperon d'histone

Avec 8,8 millions de décès dénombrés en 2015, le cancer est l'une des plus grandes causes de mortalité dans le monde. De nouvelles stratégies thérapeutiques ont émergé et l'identification de nouvelles cibles biologiques comme notamment la protéine Asf1, un chaperon d'histone H3-H4 surexprimée dans les cellules cancéreuses et en particulier le cancer du sein. Cette protéine possède différentes fonctions dans la cellule et agit à plusieurs endroits par des interactions protéine-protéines. Au cours de cette thèse de doctorat, nous avons développé une stratégie originale de design d'inhibiteurs d'interactions protéine-protéine avec des foldamères peptidomimes à base d'urées. Ces foldamères ont 1) la capacité de se replier en hélice 2,5, rappelant les hélices α des peptides et 2) d'être hautement tolérés et initiateurs d'hélicité lorsqu'ils sont conjugués à des fragments peptidiques. Nous avons développé des oligomères mixtes comprenant une alternance de segment(s) peptidique(s) et multi-urée, appelées chimères, ayant l'avantage de combiner la reconnaissance naturelle de peptides et la forte propension des oligourées à former des hélices stables. Après une étude structurale montrant qu'avec l'insertion d'un court segment à base d'urées dans un peptide hydrosoluble adoptant une conformation en hélice α, la conformation hélicoïdale pour une majorité des chimères est conservée, des composés mimant la partie hélicoïdale C-terminale de l'histone H3 ont été élaborés. Une interaction de l'ordre du micromolaire avec Asf1 a été observée en solution puis validée à l'état solide par cristallographie aux rayons X. En vue d'optimiser la reconnaissance de ces chimères avec la surface d'Asf1 et leur sélectivité, un panel de modifications a été réalisée (i.e. séquence primaire, longueur du segment urée). Nous avons ainsi conçu des chimères α/urée possédant des affinités de liaison pour Asf1 comprises entre le nano- et micromolaire. Le composé le plus prometteur a été internalisé avec succès dans des cellules cancéreuses après conjugaison bioreductible avec un peptide vecteur et pourrait conduire à la mort cellulaire de la lignée tumorale étudiée.

Mots clés : Oligourées; Foldamères ; Inhibition d'Interaction Protéine-Protéine ; Hélices, Cancer

---

Protein Surface Recognition with Urea-based foldamers: Application to the design of ligands targeting histone chaperone proteins

In 2015, 8.8 million of death were due to cancer making it an important cause of death in the world. The necessity to develop new anticancer treatments led to the search and discovery of new biological targets, such as Asf1, a histone chaperone protein of H3-H4 which is overexpressed in cancer cells, in particular in breast cancer. This protein plays a role in different biological processes in cells through protein-protein interactions (PPIs). During this thesis, we developed an original strategy to design inhibitors of PPIs with urea-based peptidomimetics. These foldamers are able to fold into stable 2.5-helix reminiscent to the natural α-helix. Designed urea-based foldamers have been synthesized as hybrid oligomers consisting of α-peptide and oligourea segments. With a combination of the two backbones, these compounds named "chimeras" presents advantages of both species with the natural recognition of α-peptides and the innate helical stability of oligourea. First, a model study was performed to evaluate the impact of the introduction of short urea segments into a long water-soluble peptide. Circular dichroism experiments confirmed that the helical conformation was conserved. New series of compounds that mimic a helical part of H3 were synthesized and their interaction with Asf1 was studied in solution and in solid state using a range of biophysical methods. Several modifications into the sequence were performed (side chain substitution, size of the urea segment or compound) in order to improve the recognition of Asf1 surface as well as their selectivity. We conceived oligourea-peptide chimeras with affinity for Asf1 in the micromolar range. Our best compound linked to a cell penetrating peptide was shown to enter into cells and to induce cell death.

Keywords : Oligourea, foldamers ; protein-protein interaction ; inhibition ; Helix, Cancer

Institut de Chimie et Biologie des Membranes et des Nano-objets-UMR5248
IECB-Institut Européen de Chimie et Biologie
2, rue Robert Escarpit- 33607 Pessac, France

Remerciements

Ce travail de doctorat a été réalisé au sein de l'Institut de Chimie et de Biologie des Membranes et des Nano-objets (CBMN) à l'Institut Européen de Chimie et de Biologie (IECB) et a été financé par une bourse de la Ligue contre le cancer.
Je souhaiterais donc dans un premier temps remercier les organismes financeurs qui ont permis de réaliser mon projet de thèse, mais aussi le Dr. Erick Dufourc ancien directeur et le Dr. Sophie Lecomte, directrice du laboratoire de Chimie et Biologie des Membranes et des Nano-objets (CBMN) ainsi que le Dr. Jean-Louis Mergny, directeur de l'Institut Européen de Chimie et de Biologie (IECB), pour m'avoir accueillie dans leur institut et permis de travailler dans les meilleures conditions.
Je souhaite remercier le Pr. Aitken et le Dr. Ennifar qui ont accepté d'évaluer ce travail en qualité de rapporteurs. Je remercie également le Dr. Ochsenbein et le Dr. Bestel d'avoir accepté d'examiner mon travail et pour l'attention donné à mon manuscrit
Je remercie tout particulièrement le Dr. Gilles Guichard et le Dr. Céline Douat pour m'avoir fait confiance après mon stage de Master 2 afin de déposer une demande de financement à la Ligue contre le cancer et pour avoir été d'excellents directeurs de recherche toujours présents. Je vous remercie pour votre disponibilité, pour votre passion communicative pour la recherche, pout tout ce que vous m'avez appris durant ces quatre années. J'ai eu la chance d'avoir un projet de recherche multidisciplinaire qui m'a passionné, dans un laboratoire doté des meilleurs équipements et de moyens pour les synthèses, tout en bénéficiant d'un encadrement très formateur.
Je tiens également à remercier le Dr. Françoise Ochsenbein de m'avoir accueillie au sein de son équipe au CEA de Saclay pendant trois mois (+ deux semaines) et de m'avoir formé à la biochimie et aux mesures biophysiques. Je remercie également toute l'équipe du CEA pour l'accueil que j'ai eu ainsi que les échanges scientifiques ou non avec tous les membres de l'équipe. Cette collaboration a été très enrichissante autant sur le plan scientifique que personnel. Je voudrais aussi remercier tous les autres collaborateurs de ce travail de thèse : le Dr. May Bakail qui testait les affinités de mes premiers composés et qui a résolu la première structure, Gwenaëlle Moal, le Dr. Raphaël Guerois et le Dr. Seydou Traoré pour les études de modélisation et tous les membres de l'équipe qui ont participé de près ou de loin à ces travaux, notamment lors de nos réunions très enrichissantes au CEA.
Je remercie également le Dr. Jean Dessolin et le Dr. Rémy Bailly du CBMN pour leurs modélisations de chimères, mais aussi tous les autres chercheurs des différentes équipes de
recherche du CBMN ou de différents laboratoires avec qui j'ai eu l'opportunité de travailler et/ou d'échanger durant ces trois années.
Merci à tous ceux qui ont partagé mon quotidien durant ces quatre années et qui l'ont rendu agréable, les membres de mon équipe avec lesquels j'ai pris plaisir à travailler et à partager de bons moments, notamment Christophe (nos week-end « Liberty/Gilson » et tout le reste), Laura

(collée à la paillasse et au bureau), Stéphanie (on a trop géré notre happy hour « Euro »), Camille, Mégane et Jérémie (WES). Je voudrais aussi remercier Arup, Caterina, Diane, Léonie, Gavin, Christel, Morgane, Guillaume, Juliette et Claire. Merci également à la team Garanger, nos anciennes voisines, dont Rosine et surtout ma Laura, mon coup de cœur de ces quatre années (merci pour notre soutien mutuel et tous ces bons moments partagés). Merci aux membres de l'équipe Innis, Huc, Friscourt et Quideau : Caro, Britta, Natasha, Maëlle, Camille, les Antoine, les Simon, Arthur, Pedro et Joan.

Merci également aux membres de l'ADoC où j'ai adoré être présidente et travailler avec un bureau de choc, en particulier Medhi, Amélie, Samy, Caro et Camille. Merci à tous ceux avec qui j'ai passé de supers moments dans les autres laboratoires, notamment au LCPO avec BenJ, Dounia, Boris, Cindy et tous les autres. Je remercie également mes amis pour leur présence de près ou de loin tout au long de ces travaux de thèse, en particulier Barbara et Binta.

Pour finir mais non des moindre, je souhaite infiniment remercier ma famille pour leur soutien sans faille, en particulier ma mère qui a toujours été présente (merci pour tes conseils, ta positivité et tes encouragements), mes frères : Flavien, Tillian et Zoryan (merci d'avoir supporté mes humeurs et pour tous nos fou-rires), mes cousin(e)s dont Nadège et Sandrine, mes tantines, ainsi que mon parrain et ma tante (votre force et votre courage donnent du sens à tous ces travaux que j'ai réalisé pendant quatre ans). Merci à Loïc.

Reconnaissance de surfaces de protéines avec des foldamères à base d'urées : Application au design de ligands ciblant une protéine chaperon d'histone

Résumé détaillé :

La compréhension toujours plus fine des mécanismes d'interactions dans la cellule conduit les scientifiques à innover et proposer de nouveaux traitements thérapeutiques. Il a ainsi été établi que le dysfonctionnement d'interactions protéine-protéine (PPIs) était à l'origine de certaines maladies, dont le cancer. La possibilité d'inhiber ces complexes supramoléculaires offre ainsi de nouvelles perspectives de développement de traitement anticancéreux. Cependant, les interfaces mises en jeux dans ces interactions sont larges et les cibler représente toujours un défi. De nouvelles cibles biologiques ont récemment émergé dont Asf1, une protéine chaperon d'histone surexprimée dans les cellules cancéreuses notamment dans le cancer du sein. L'équipe de Françoise Ochsenbein (CEA Saclay) a identifié les mécanismes d'interactions entre Asf1 et les histones H3-H4. Il a ainsi été mis en évidence que l'interaction se fait en particulier avec la partie C-terminale de l'histone H3 qui adopte une conformation en hélice α. Lors de ces travaux de thèse, nous souhaitions concevoir des oligomères à base d'urée mimant la structure en hélice α qui interagit avec Asf1 en conservant les chaines latérales clés pour l'interaction. Ces peptidomimes sont composés d'unité éthylène diamine pouvant être facilement équipées avec les chaines latérales des acides aminés naturels. Grace à un réseau de liaisons hydrogène à trois centres, les oligourées ont la capacité de se replier en hélice stable comportant 2,5 résidus par tour. Ils appartiennent à la famille des foldamères définit comme l'ensemble des oligomères non naturels capables de s'auto-organiser pour adopter des repliements bien définis, rappelant ceux des biomolécules incluant les peptides. Avec leur propriété de résistance à la protéolyse, les oligourées sont de bons candidats pour développer de nouveaux traitements thérapeutiques. Durant cette thèse, nous avons conçu des composés hybrides peptide/oligourée hydrosolubles (important pour des applications biologiques et analyses biophysiques), appelés chimères. La conception de ces composés avec un squelette mixte pourrait permettre de combiner les avantages des deux types de composés : les propriétés de reconnaissance des peptides naturels et, la stabilité intrinsèque des oligourées aux protéases – et pourrait être d'autant plus modulable pour mimer au mieux les peptides.

Tout d'abord, nous avons réalisé une étude modèle dans le but d'étudier l'impact de l'insertion d'un segment court d'oligourée dans un peptide hélicoïdal long (>30 résidus). Cette insertion de segment avec modification du squelette est assez innovante car dans la littérature, la modification du squelette dans un peptide est généralement ponctuelle limitée à un résidu. Pour cette étude, nous avons sélectionné une séquence mutante de GCN4-p1, un domaine leucine zipper d'un facteur de transcription dont la séquence présente une répétition en heptade *abcdefg* avec les résidus hydrophobes aux positions *a* et *d* caractéristique de peptides formant des faisceaux d'hélices. Ce peptide a été choisi pour son hydrosolubilité et pour sa capacité à

cristalliser. La structure RX de GCN4-p1 indique qu'il dimérise en une double hélice parallèle, appelée coiled-coil(Figure 1).

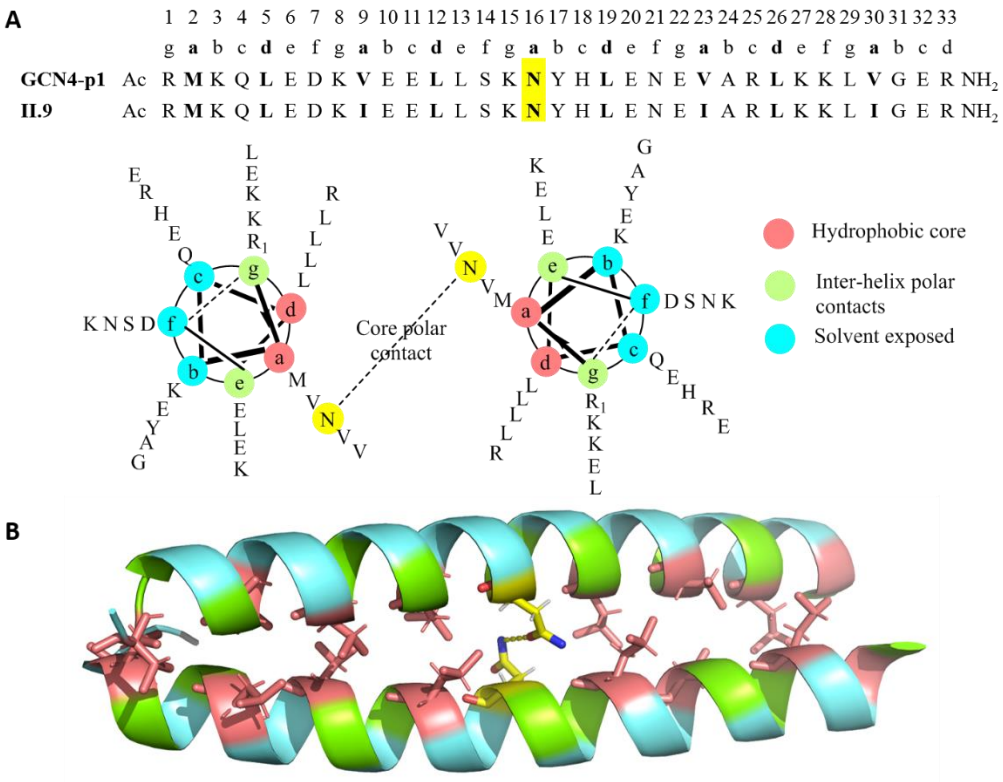


Figure 1: Séquence de GCN4-p1 et du mutant II.9 et structure du dimère GCN4-p1. A Séquence primaire et projection sur roue hélicoïdale du coiled-coil. B Structure cristallographique de GCN4-p1 (pdb=4dmd) à 2.0Å avec les résidus hydrophobes en stick.

Après avoir optimisé la synthèse sur support solide assisté aux micro-ondes du peptide mutant de référence, **II.9**, plusieurs chimères peptide/oligourée ont été synthétisés avec des segments de deux ou trois résidus urée. La partie peptidique était introduite avec un synthétiseur automatique *Liberty Blue*, alors que la partie urée était introduite manuellement avec le système de synthèse micro-ondes *Discover Bio*. Après purification par HPLC en phase inverse et caractérisation par spectrométrie de masse, des expériences de Dichroïsme Circulaire (CD) ont été réalisées. Une première série de chimères consistait à remplacer trois résidus α-aminés par deux résidus urée. Pour cela, après une simple étude de modélisation moléculaire, nous avons inséré deux résidus urée pour remplacer le segment peptidique central Leu13-Ser14-Lys15 afin de ne pas toucher de résidu impliqué dans le cœur hydrophobe. La longueur des chaines latérales mimant celle de la Leu a été variée ainsi que la position de la chaine latérale du résidu Lys sur le C_α ou C_β. Les spectres CD de 200-260 nm ont permis de confirmer que la conformation en hélice était conservée. Cependant, le repliement en hélice est perturbé car le taux d'hélicité diminue en présence du segment urée. Le segment urée a été déplacé en C-terminal de la séquence et des résultats légèrement moins bons ont été observés par CD, indiquant une perturbation plus importante de l'hélice. Les propriétés d'autoassemblage de ces composés ont également été

étudiés par CD avec des expériences de variations de la température de 4 à 100°C et la détermination de la température de transition (Tm). Les Tm obtenus pour les chimères avec deux résidus urée étaient inférieurs à ceux du peptide **II.9** (Tm=71°C) indiquant que l'assemblage du dimère est déstabilisé. Une autre série de composés où quatre acides aminés ont été remplacés par trois résidus urée ont également été synthétisés. Cette substitution impliquant un résidu hydrophobe, les expériences de CD ont montrés que la déstabilisation de l'hélice était encore plus importante. De nombreux efforts ont été réalisés afin de cristalliser ces chimères et malgré l'obtention de cristaux et de données cristallographiques de bonne qualité, des problèmes de phase nous ont empêché de résoudre les structures pour le moment. Ces premiers résultats sont encourageants mais de nouveaux efforts pour des essais cristallographiques seront encore nécessaires afin de comprendre l'impact de l'introduction des urées sur l'assemblage et proposer ainsi de nouvelles séquences.

Dans une seconde partie, nous présentons notre première tentative de design de chimère mimant une hélice α pour l'inhibition d'interactions protéine-protéine (PPIs). Dans un premier temps, nous nous sommes basés sur le peptide **p3** développé par l'équipe de F. Ochsenbein (CEA Saclay) comportant une partie hélicoïdale mimant la partie C-terminal de H3, une partie en brin β mimant la partie C-terminale de H4, les deux parties étant reliées grâce à une s séquence adoptant une conformation étendue. Une structure cristallographique du complexe **p3**-Asf1 a été obtenu par l'équipe de Frannçoise Ochsenbein à 1.8Å (Figure 2B). Nous avons choisi dans un premier temps de nous concentrer sur la partie hélicoïdale de **p3**, que l'on nommera **p3h**, afin de concevoir nos composés hybrides. Par simple superposition de **p3h** avec une structure d'oligourée, nous avons déterminé la nature des chaines latérales d'une première série de chimères en conservant en priorité les hotspots présents dans **p3h** : Leu9, Arg12 et Ile13. La séquence de **p3h** contenant un cap permettant de verrouiller l'hélice, nous avons développé un analogue de **p3h**, le peptide **p5** où le cap a été supprimé afin de déterminer l'effet de l'introduction du segment urée. Une série avec un segment composé de trois urées à différentes positions dans la séquence de **p5** a d'abord été synthétisé et testé par titration isothermale calorimétrique (ITC). Ces composés ont montré une amélioration de la contribution entropique de l'interaction grâce à l'introduction du segment urée ($-T\Delta S$=-2.8 kCal.mol^{-1} pour **ch3** au lieu de -1,4 pour **p5**) compensée par une perte d'enthalpie (ΔH=-4,5 kCal.mol^{-1} pour **ch3** au lieu de -7,6 pour **p5**). La position centrale était la plus favorable avec une chimère **ch3** présentant un K_d=3,6 µM. Des expériences de de RMN hétéronucléaires de type ^{1}H-^{15}N HSQC (Heteronuclear Single-Quantum Correlation) ont permis de mettre en évidence que les résidus d'Asf1 impacté par la présence de **ch3** sont les même que ceux impactés par la présence de **p3h** ou de la partie C-terminale de H3. Ces résultats indiquent donc que la chimère **ch3** agit dans la même région que la partie C-terminale de H3 dans le complexe naturel ce qui est très encourageant.

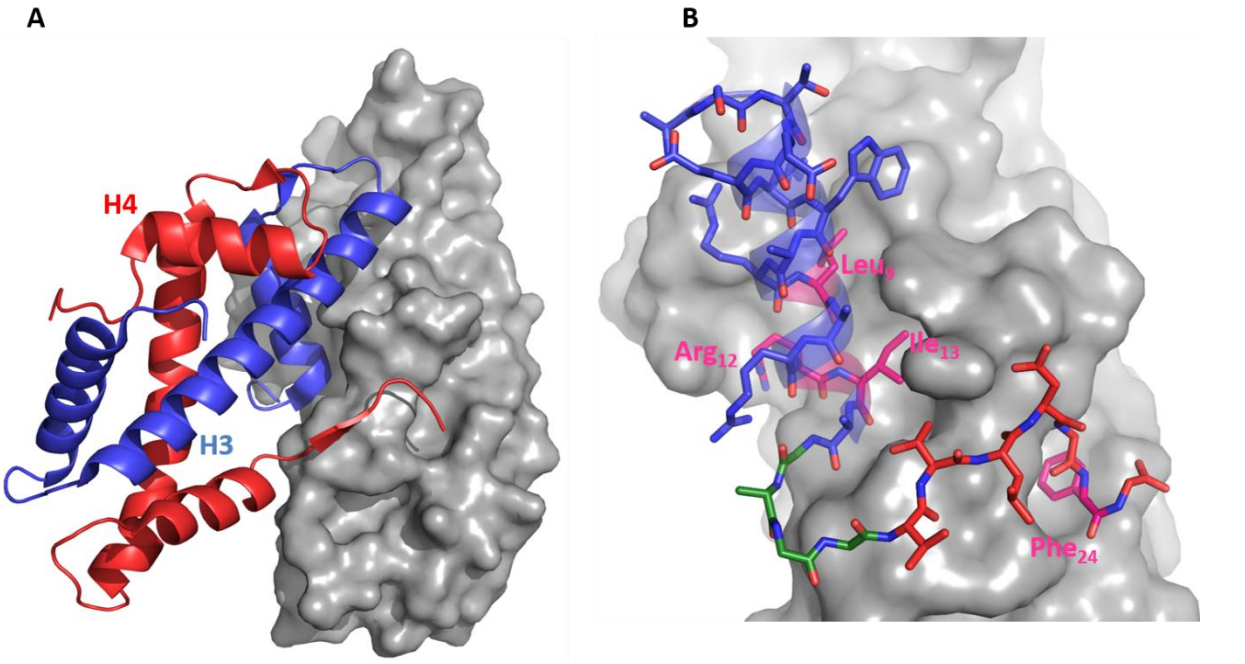


Figure 2: Structure cristallographique du complexe naturel Asf1/H3-H4 (A) et structure de p3 interagissant avec Asf1 (B) à 1.8Å déterminées par l'équipe de Françoise Ochsenbein au CEA Saclay. A la protéine Asf1 est représenté en gris en surface, H3 en bleu et H4 en rouge (pdb 2io5). B La partie en hélice de p3 est représenté en bleu, le linker en vert et la partie mimant le brin β de H4 en rouge. Les hotspots sont représentés en rose.

Dans le but d'augmenter l'affinité de nos composés, une nouvelle série a été synthétisée avec un segment de quatre résidus urée où la nature de la chaine latérale du résidu urée aromatique a été changée. La meilleure affinité a été obtenue pour le composé comprenant un résidu naphtyl (Nal), **ch5**= Ac-EK-Nal^{u}-Arg^{u}-Leu^{u}-Gln^{u}-RIA qui présente un K_d de 2.8 µM. Les expériences RMN HSQC ont confirmé que les résidus affectés par l'addition de **ch5** correspondent à la région d'interaction de la partie C-terminale de H3 que l'on souhaitait mimer. Fait remarquable, la chimère **ch5** a pu être cristallisée en présence de la protéine Asf1. La structure cristallographique du complexe **ch5**-Asf1 a ainsi été obtenue à 2Å, apportant des informations structurales précieuses (Figure 3). Cette structure permet de montrer que **ch5** interagit avec Asf1 par le biais de liaisons hydrogène et d'interactions impliquant des résidus hydrophobes et confirme que les deux résidus hotspots Leu^{u}5 et Ile8 sont bien positionnés dans les poches hydrophobes. Cependant, ces interactions déstabilisent la conformation en hélice de la chimère qui apparait légèrement tordue. De plus, certains résidus de la surface de Asf1, tels que la Tyr112, ont dévié de leur position initiale pour « s'adapter » à la présence de l'oligomère et faire de nouvelles interactions avec le squelette de l'oligomère. Ces données à l'état solide complètent les informations obtenues en solution (RMN et ITC) et seront très utiles pour de futures optimisations visant des ligands avec des affinités de l'ordre du nano molaire. Cette structure est la première concernant un oligomère à base d'urée à la surface d'une protéine et reste à ce jour, l'un des rares exemples de structure montrant un foldamère interagissant avec sa protéine cible.

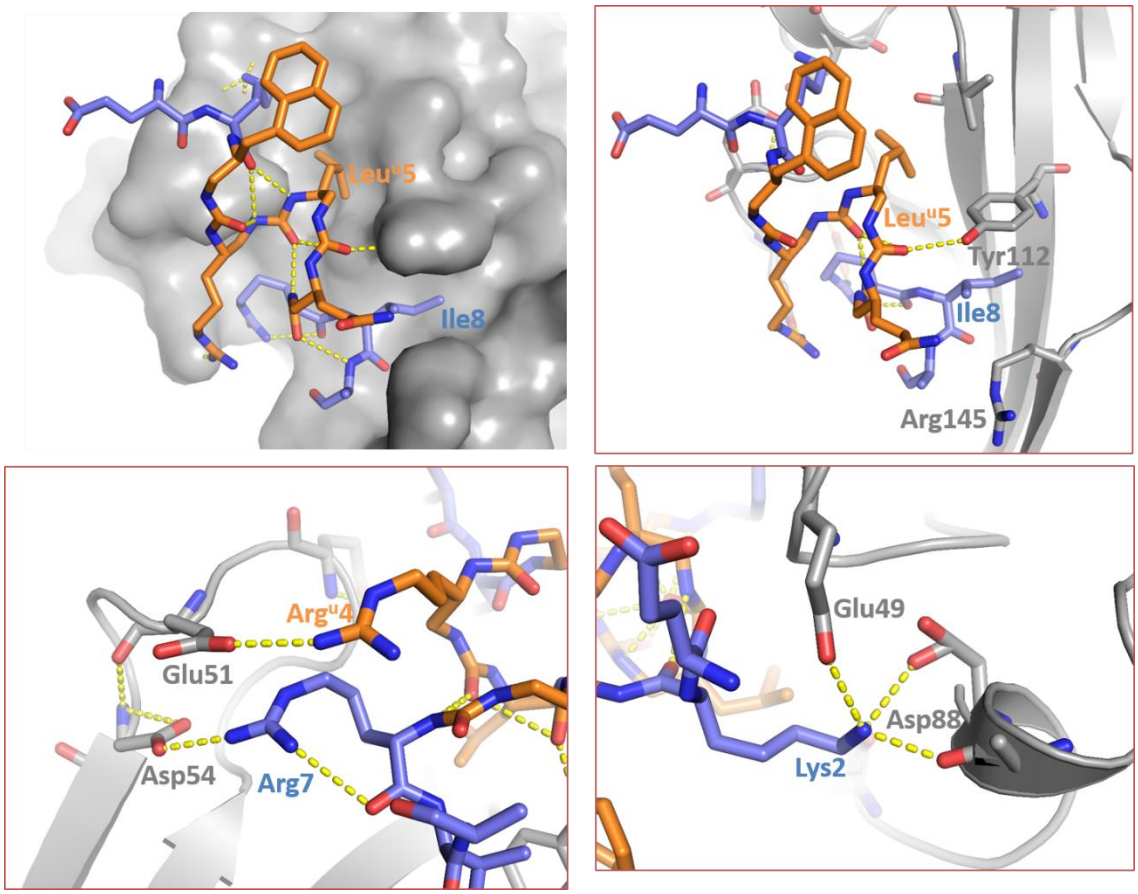


Figure 3: Crystal structure de ch5 en complexe avec la protéine Asf1. Les résidus α-aminés de ch5 sont en bleu et les résidus urée sont en orange. Asf1 est représenté en gris. En haut à gauche, vue globale de la chimère ch5 en interactions avec la surface de Asf1. En encadré, zoom sur les interactions entre résidus.

Cette séquence **ch5** a également été conjuguée avec un peptide vecteur pénétrant les cellules afin d'évaluer l'activité *in cellulo*. Ces études réalisées dans l'équipe de F. Ochsenbein, ont montré que **ch5** couplé au peptide vecteur présente un effet antiprolifératif et conduit à la mort des cellules cancéreuses. Ces résultats très encourageants doivent cependant être approfondis, notamment avec un control négatif, afin de nous assurer que la mort cellulaire est bien due à **ch5**. D'autres séries de composés ont été synthétisés avec des modifications de chaines latérales (modifications de résidus basiques, aromatiques, de chaines alkyles pour remplacer Leu^{u}5 dans **ch5**) mais les meilleurs résultats restaient ceux obtenus pour **ch5**.

Finalement, nous avons cherché à atteindre un second point d'ancrage situé sur une autre face de la protéine Asf1 dans le but d'augmenter la spécificité et l'affinité de nos composés. Dans la structure cristallographique du complexe Asf1/H3-H4, la partie C-terminale de l'histone H4 adopte une conformation en brin β et interagit avec cette autre face de la protéine (Figure 2A). Un hotspot supplémentaire, Phe100 est situé dans une poche hydrophobe de Asf1 et contribue de manière importante à l'interaction. L'équipe de F. Ochsenbein a développé un lien permettant de lier de manière covalente la partie hélicoïdale mimant la partie C-terminal de H3 avec la partie mimant la partie C-terminale de H4, menant au peptide **p3** qui possède une affinité 100 fois

meilleure par rapport à **p3h**. Ce lien est composé de quatre résidus α-aminés (GAGG) qui forment un coude épousant la surface de la protéine et permettent de bien positionner la chaine latérale de la Phe dans la poche hydrophobe d'Asf1 (Figure 2B).

La synthèse de ce peptide **p3** a été optimisée au laboratoire avec l'introduction d'une pseudo-proline permettant d'éviter l'agrégation du peptide. De plus, pour palier à une réaction de formation d'aspartimide observée durant la synthèse et conduisant à une chute importante du rendement, un dipeptide contenant le 2,4-dimethoxybenzyle (Dmb) comme groupement protecteur de l'amide, a été introduit dans la séquence afin d'éviter cette réaction secondaire. La structure à haute résolution de ce peptide **p3** à la surface de Asf1 montre que des liaisons hydrogène ainsi que des ponts salins stabilisent le complexe. Afin d'optimiser les contacts avec la surface et simplifier la boucle formée par le linker, nous avons développé des séquences alternatives dont une avec des résidus aliphatiques mimant les résidus GAGG. Une autre stratégie consistait à raccourcir la taille du composé en réalisant un « pont » plus court entre les deux segments permettant une connexion plus directe à la surface de la protéine. Des peptides analogues de **p3** contenant ces nouveaux linkers ont donc été synthétisés et ont montré une affinité du même ordre que **p3** (K_d=0.01µM).

Fort de ce résultat, nous avons souhaité rallonger la partie C-terminale de **ch5** afin d'atteindre le nouveau point d'ancrage avec le résidu Phe. Tout d'abord, la partie C-terminal de **p3** (contenant le linker GAGG) a été ajoutée à notre meilleure chimère (**ch20** : Ac-EK-Nal^u-Arg^u-Leu^u-Gln^u-RIAGAGGVTLDGFG). Puis, **ch5** a été prolongé en C-terminal avec les linkers « raccourcis » où le nombre de groupes méthylène variait (Figure 4). Les expériences ITC de ces nouvelles séquences ont montré une amélioration du K_d pour **ch20** (1,2 µM) et **ch22** (2,2 µM) alors que **ch21** n'a pas montré d'amélioration (6,2 µM alors que pour **ch5**, K_d=2.8 µM).

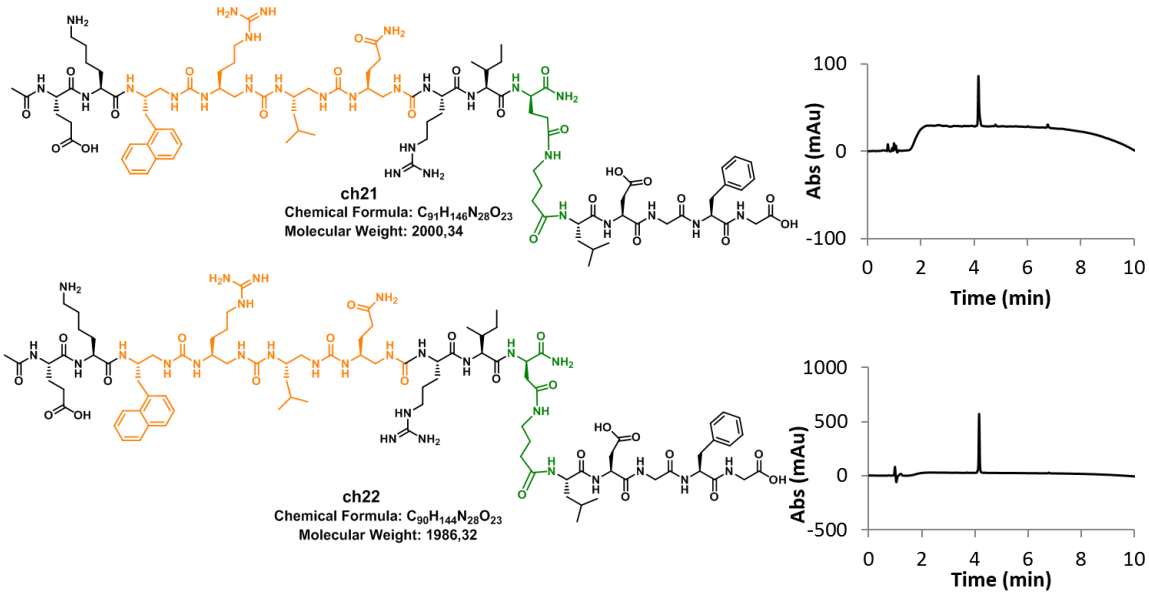


Figure 4: Structure chimique de ch5 allongé en C-terminal aves les linker de type "court-circuit". La partie urée est représentée en orange, le linker en vert et la partie peptidique en noir.

Comparé aux résultats obtenus avec le peptide **p3**, ces résultats restent décevants d'un point de vue du gain d'affinité. Cependant, comparé à **ch5**, on observe un gain considérable d'enthalpie indiquant que de nouveaux contacts ont bien lieu avec la protéine (ΔH varie entre -11 et -13 $kCal \cdot M^{-1}$ pour **ch20**-**ch22** alors qu'il est de -2.2 pour **ch5**) La perte d'entropie dû à l'augmentation de la taille du composé compense ce gain d'enthalpie entrainant une perte d'affinité. Ces résultats semblent indiquer que le linker n'est pas optimal et ne permet pas de positionner au mieux le résidu Phe qui doit être décalé. Afin de comprendre mieux l'effet de ces linkers, des essais de cristallogenèse sont actuellement en cours dans l'équipe de F. Ochsenbein, mais n'ont pas donné de résultats pour le moment. Une structure cristallographique nous permettrait en effet de comprendre la position du linker et de la Phe afin de pouvoir créer de nouveaux designs.

Par ailleurs, nous avons développé un composé comprenant uniquement des résidus urée mimant l'hélice de la partie C-terminale de H3 afin de tester sa capacité à reconnaitre la surface de Asf1. Après des études de modélisation avec Rosetta, une oligourée **ol1** a été synthétisée sur support solide et a montré une affinité très intéressante (K_d= 4.5±0.7 µM). Ces résultats sont très encourageants pour le développement de futures oligourée capables d'inhiber des interactions protéines-protéines.

Cette thèse a donc permis de concevoir des composés à base d'urée capable de mimer la structure en hélice α des protéines. Les chimères hydrosolubles conservent la structure hélicoïdale des peptides et présentent une meilleure stabilité aux protéases. Les composés mimant la partie C-terminale de H3 ont montré qu'ils interagissaient dans la même zone que le peptide natif et après conjugaison avec un peptide pénétrant les cellules, conduisent à la mort cellulaire. Après des études de modélisation, notre première tentative de concevoir un « pur » oligourée capable d'inhiber l'interaction Asf1/H3-H4 semble prometteuse. Finalement, des efforts restent nécessaire pour atteindre de meilleures affinités de l'ordre du nano molaire, notamment en travaillant sur le linker permettant d'atteindre une autre face de la protéine Asf1 en positionnant bien un résidu clé additionnel : la Phe.

Abbreviations and formulas

δ	chemical shift (ppm)
λ	wavelength
%	percent
°	degree
Å	Angström
Abs	Absorbance
ACN	Acetonitrile
Boc	tert-butoxycarbonyl
BOP	(Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate
Bu	Butyl
C	Celsius
Cbz=Z	Benzyloxycarbonyl
CD	Circular dichroism
$CDCl_3$	Deuterochloroform
DCM	Dichloromethane
DIC	Diisopropylcarbodiimide
DIAD	Diisopropylazodicarboxylate
DIEA	N, N'-Diisopropylethylamine
DSC	N,N'-Disuccinimidyl carbonate
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxyde
DNA	Deoxyribonucleic acid
EDT	1,2-Ethanedithiol
Et_2O	Diethylether
eq.	equivalent
ESI-MS	Electrospray ionization mass spectrometry
EtOAc	Ethyl Acetate
Fmoc	Fluorenylmethyloxycarbonyl chloride
g	gram
h	hour
H_2O	water
HBTU	2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexaurophosphate
HCl	Chlorhydric acid
HOBt	Hydoxybenzotriazol
HPLC	High Perform Liquid Chromatography

I_2	Diiode
IBCF	Isobutylchloroformate
*i*Pr	isopropyl
ITC	Isothermal Titration Calorimetry
K_2CO_3	Potassium carbonate
MeOH	Methanol
min	minute
mol	mole
N_2	Nitrogen
$NaBH_4$	Sodium tetrahydroborate
$NaHCO_3$	Sodium bicarbonate
NaI	Sodium iodide
NaN_3	Sodium azide
NH_2-NH_2	Hydrazine
NMM	4-methylmorpholine
NMR	Nuclear Magnetic Resonance
PBS	phosphate buffer saline
Pbf	2,2,4,6,7-Pentamethyldihydrobenzofuran-5-sulfonyl
pH	potential of hydrogen
PMe_3	Trimethylphosphine
PPh_3	Triphenylphosphine
ppm	parts per million
PPI	Protein protein interaction
PS	polystyrene
res	residue
RP-HPLC	reverse phase High Perform Liquid Chromatography
TFA	Trifluoroacetic acid
THF	Tetrahydrofurane
TIS	Triisopropylsilane
t_r	retention time
Trt	tritylphenylmethyl

Nomenclature

To mention α-amino acids, the three letters code will be used such as valine named Val. β-amino acids are preceded by β such as β^3-Alanine named β^3-Ala. γ-amino acids will be preceded by γ sign. For example, γ-Leucine= γ-Leu.

Urea residues will be mentioned by using the three letters code followed by a "u" exponent. For example, the residue derived from the alanine will be named Ala^u. In table, the one letter code follow by a "u" exponent will also be used as well as for amino acids. The carbon atoms of the main chain will be named with greek letters α, β, γ, δ as shown on Figure 1.

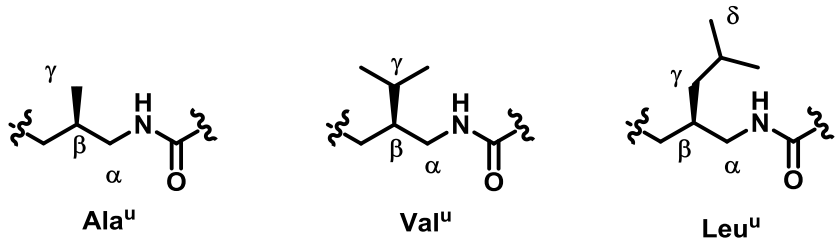
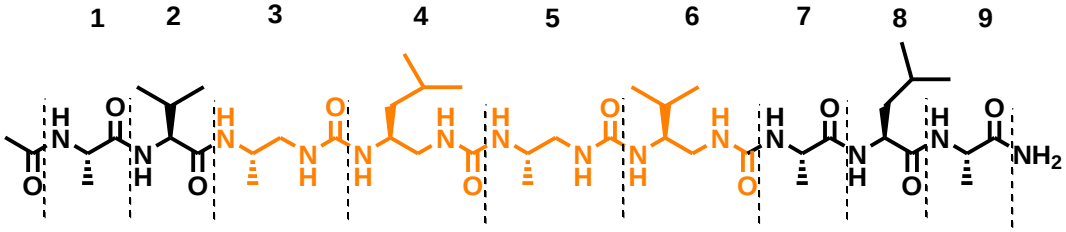


Figure 1: Example of notation for urea residue

In chapter I, the compounds will be mentioned as **I.X** with X corresponding to the number of apparitions of the compound. In chapters II, III and IV, monomers will be named by **MX** (with X their number of apparitions into the manuscript). Synthesized peptides will be named **pX** whereas peptide-oligourea chimera will be named **chX**. Residues will be numbered starting from the N-terminus toward the C-terminus. α-amino acids can be noted with the one letter code whereas urea residues are noted with the three letters code followed by the "u" exponent as shown on Figure 2.



Ac-AV-Ala^u-Leu^u-Ala^u-Val^u-ALA-NH$_2$

Figure 2: Numbering of residues in oligomer

Table of contents

Chapter I: Targeting protein-protein interactions (PPIs) involved in cancer development

Protein-protein interactions (PPIs) are involved in the majority of biological processes such as DNA replication, protein synthesis, apoptosis, transport, etc. PPIs occur between two or more structured proteins that interact together and trigger a biological function. Understanding how protein partners interact with each other is a structural prerequisite for understanding function that is associated. In most cases, the interface of PPIs is a relatively flat and large surface with an average area of 1200-2000 Å^2 whereas the average surface between enzyme and substrate is only 500 Å^2 [1–4]. The physical contact between the two proteins should be specific and the assembly can be stable (macromolecular complexes such as ATP synthase) or more ephemeral (for transient actions)[5,6]. Few key residues, called hot-spots contribute significantly to the binding affinity[7,8]. Wells and coworkers performed mutagenesis experiments with alanine scanning substitution and measured the binding free energy difference (ΔΔG) between mutant and wild type to define hot-spots as residues whose substitution leads to an important drop in free energy (ΔΔG > 2kcal·mol^{-1})[7].

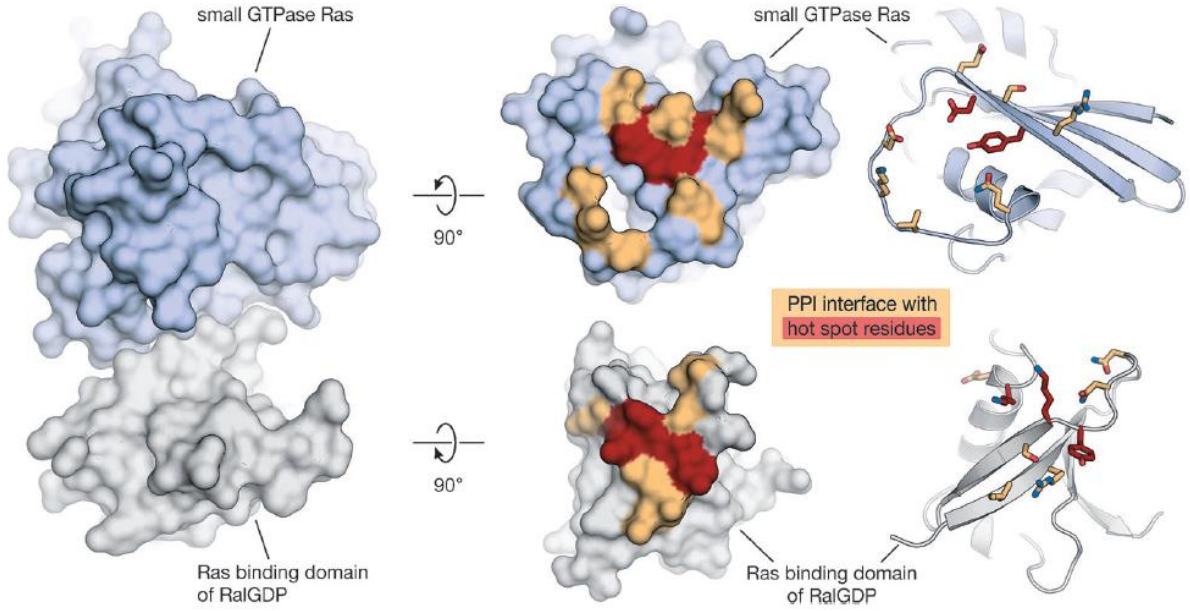


Figure 5: Example of a PPI with its interaction area and hot spots. The crystal structure of the complex between Ras (light blue) and the Ras binding domain of RalGDP (grey) (PDB: 1FLD). Residues involved in the interface are represented in orange and hot spots in red. Interacting residues are shown as sticks. (Adapted from Grossmann et al., 2015)

Over the last two decades, knowledge of PPIs has considerably increased and helped for a better understanding of many biological processes including cellular functions and diseases. PPIs are essential for infection of host by pathogens and thus represent possible targets to combat infectious diseases. Alternatively, deregulation of PPIs may also be associated with the progression of some diseases, including cancer. For these reasons, in recent years, PPIs have become attracting targets for developing new therapeutics. Research on modulators of PPIs has been very active over the last two decades and such modulators are today considered as promising and innovative drug candidates[9,10]. Furthermore, computational tools and chemical technologies (structure-based approach, high throughput screening and diversity-oriented

synthesis (DOS)…)[11] have been developed to increased chemical space and facilitate the design based on structure of inhibitors of PPIs[12]. From a structural point of view, Arora *et al.* examined entries in PDB in 2009 and highlighted that the secondary structure most frequently found at protein-protein interfaces is the α-helix (62%)[13]. Because of the frequency of the α-helix conformation at PPI interfaces, there has been a strong focus on α-helix stabilization and mimicry. In contrast to small molecules that may interact only with a limited number of protein pocket due to their small size, medium sized compounds such as peptides and related mimics by covering a larger surface should increase binding and the specificity of protein recognition[11]. However, the direct use of α-peptides as PPI modulators has major drawbacks: (1) a high sensitivity to enzyme degradation; (2) due to their short sequences, peptides are generally poorly folded in solution and (3) they lack good cell permeability[14,15].

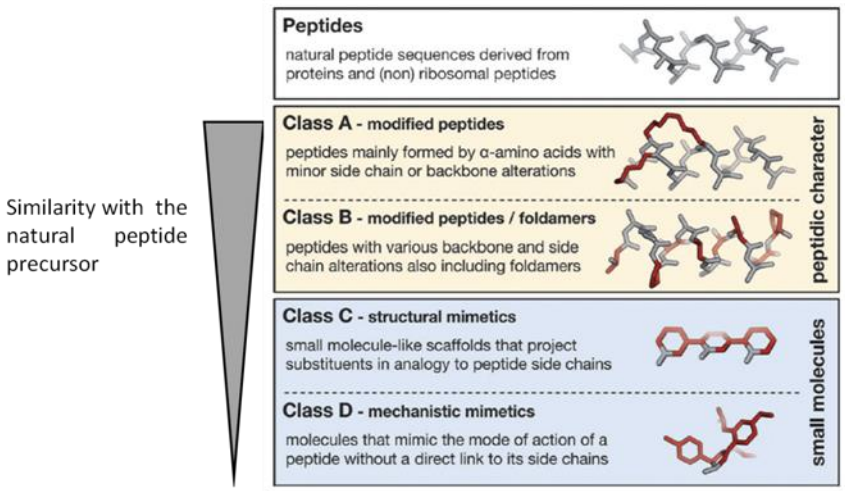


Figure 6: Classification of peptidomimetics. For illustration, a α-helical peptide and corresponding helix mimetics are shown. Modifications are highlighted in red. (Adapted from Grosmann et al.[16])

In this context, being capable to mimic the secondary structure of α-helical peptides has become a major challenge for researchers working in the field of peptide-mimetic chemistry. In the literature, the exact definition of peptidomimetics remains unclear and the classification has been changing with recent advances. In 2015, Grosmann *et al*. defined peptidomimetics as designed molecules that mimic the binding properties of natural peptide precursors[16]. They introduced a new classification depending on the similarity of compounds with the natural peptide precursor (Figure 6). Class A mimetics regroup modified peptides quite similar to the parent ones exhibiting minor modifications (side-chain amino acids or backbones). Class B comprises modified class A mimetics with unnatural amino acids, major backbone alterations and in this class B are also included the field of foldamers research. The Class C belongs to small molecules that do not present anymore features of peptides and where the scaffold projects substituents with the aim to reproduce the spatial arrangement and angular projection of "hot-spot" residue side-chains. Finally, the class D regroups molecules that do not mimic the side chains functionalities but only mimic the mode of action of a bioactive parent peptide. They can be obtained with screening libraries or by affinity optimization of a molecule from class C.

As briefly summarized above, nowadays the field of research on PPIs modulators is extremely large and diverse. However, in this chapter we will focus more on the subject of this PhD. thesis,

namely on the development of foldamers to target PPIs. The term foldamer, has been introduced for the first time in 1998 by Gellman and corresponds to any oligomers "with a strong tendency to adopt a specific compact conformation"[17,18]. These artificial backbones (from polymers to artificial protein sequences) should present well-defined folded conformations and can eventually achieve chemical functions similar to those found in nature[19]. In contrast to peptides, the non-natural backbone of foldamers allows their resistance to proteolysis. These properties make foldamers interesting candidates to develop new drugs targeting PPIs.

In the following paragraphs, we are going to give a non-exhaustive state of the art regarding the design and development of foldamers as PPI modulators in the context of cancer research. Moreover, new strategies to promote foldamer cell penetration as well as new targeted PPIs for innovative anticancer therapies will be detailed in the following sections.

A. Design of foldamers to inhibit PPIs involved in cancer diseases

In recent years, knowledge in cancer biology has been considerably improved and direct links between cellular pathway perturbation and cancer have been established[20]. In cancer, PPIs form signaling hubs and nodes that allow the transmission of signals essential for cell transformation and tumor progression[10]. PPIs play essential roles in linking networks that relay oncogenic signals, enabling the hallmark features of cancers and have also a role in driving and maintaining cancer cell growth.[10] For example, several PPIs such as p53/MDM2 and CDK4/pRB are known to neutralize tumor suppressive functions. Thus, the possibility to disrupt PPIs that are critical for cancer development became a challenging goal for scientists as way to develop innovative anticancer strategies.

The development of PPI inhibitors has considerably progressed and various approaches such as small molecules, recombinant proteins or synthetic molecules have shown some promises[4,10,21]. Several anticancer compounds targeting PPIs have entered clinical trials illustrating the potential of the approach for anticancer therapy[22]. For example, a recently developed p53-derived stapled peptide[23] is currently in Phase I trial (NCT02264613)[24] encouraging the development of new compounds with helical stabilization for clinical lead[25].

However, the design of compounds able to specifically modulate such PPIs remains a huge challenge due to the high diversity of PPIs in term of shapes and sizes. With their medium size, α-peptides could be a good compromise between small molecules and biologics, but as mentioned previously, due to their chemical constitution, peptides present major drawbacks for their use as therapeutics. Concurrently, some foldamers present a high propensity to mimic α-helices and there is growing evidence that they are good candidates to interact with extended protein surfaces, in a high specificity manner. Indeed, several foldamer backbones have been shown to reproduce the main features of protein secondary structures of proteins found at the interface of PPIs while expressing resistance to proteases, suggesting their potential development for future therapies.

1. α-helix mimicry for PPIs inhibition…

Intense research efforts have been conducted to disrupt PPIs that are related to apoptosis or programmed cell death. In particular, restoring apoptosis activation is an attractive strategy that could contribute to tumor regression in cancers where this function has been downregulated [26]. Two distinct signaling pathways are known to control apoptosis activation: the extrinsic pathway with extracellular signals generated by cytotoxic cells and the intrinsic pathway with intracellular signals such as DNA damage (Figure 7). These two pathways are interwoven and both stimulate a family of cysteine proteases called pro-apoptotic caspases.

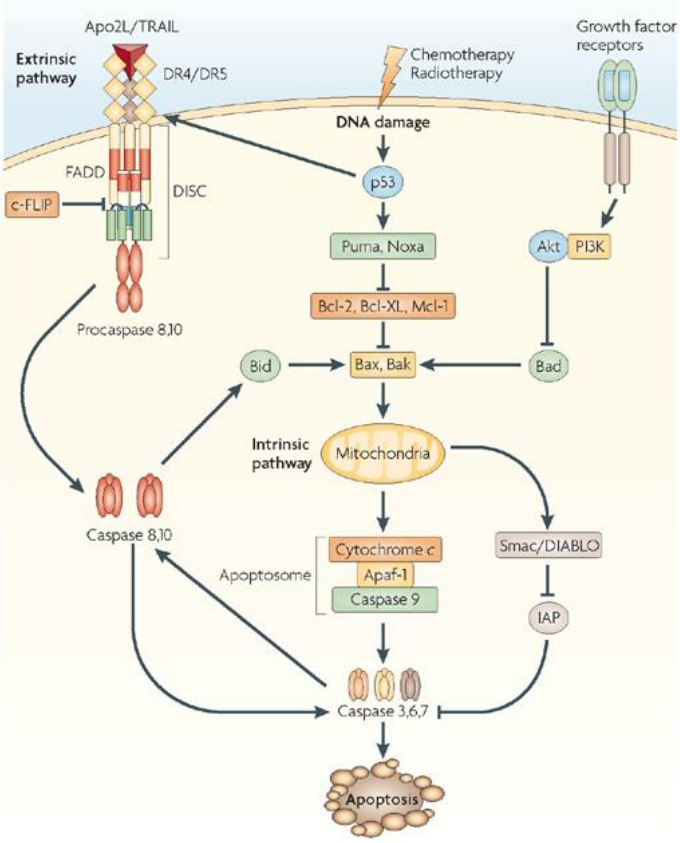


Figure 7: Key steps in apoptotic signaling pathways. Intrinsic pathways: cellular stress activates the p53 tumor-suppressor protein p53 initiates the intrinsic pathway by upregulating Puma and Noxa, which in turn activate Bax and Bak. These two last permeabilize the outer mitochondrial membrane, resulting in efflux of cytochrome c, which binds to Apaf-1 and caspase9 to form apoptosome. Extrinsic pathway: cytotoxic immune cells produce or express pro-apoptotic ligands such as Apo2L/Trail. It binds to the pro-apoptotic receptors DR4 and DR5 on the surface of a target cell. Ligand binding induces receptor clustering and recruitment of the adaptor protein FADD and the initiator caspases 8 and 10 as pro-caspases, forming DISC. This triggers activation of the apical caspases, driving their autocatalytic processing and release into the cytoplasm, where they activate the effector caspases 3, 6 and 7.[26] (adapted from Ashkenazi, Nature 2008)

First, we will focus on the well-studied interaction between p53 and Mdm2.

a) … of the complex P53/MDM2

The protein p53 plays a key role in apoptosis, but also at different levels in the cell process such as DNA repair or angiogenesis. p53 has been discovered in 1979 by the team of Old and has been defined as a transcription factor and a tumor suppressor[27–29]. Its biological role consists in "ringing the bell" when DNA is damaged and p53 can be viewed as the guardian of the genome integrity. Failure of DNA repair mechanisms leads to p53-mediated induction of cell death programs. The loss of p53 tumor-suppressor activity favors the development of cancers.

The human protein double minute 2 (hDM2, MDM2 in mouse) is regulating p53 functions and is itself the product of a gene induced by p53[30,31]. An autoregulated negative feedback loop links the two proteins together in order to maintain a low p53 level in the cell in absence of stress[31]. During cellular stress (DNA damage, oncogene activation or hypoxia for example), p53 is activated to trigger cell death. As MDM2 overexpression blocks p53 pathway for cell arrest or apoptosis and the p53 loss in cancer contributes to the progression of cancer cells, many efforts toward p53-xDM2 interaction have been performed making it an important target for cancer therapies[32].

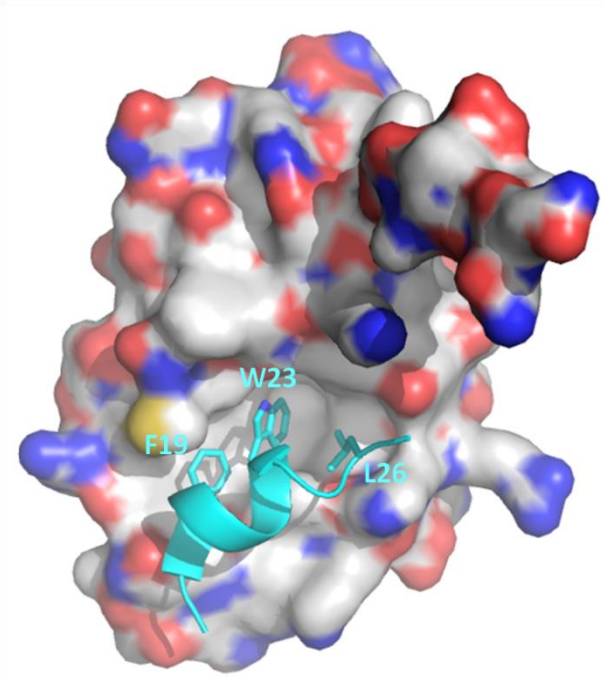


Figure 8: X-ray crystal structure of hDM2 in complex with the p53 N-terminal transactivation domain. hDM2 represented in grey and the p53 helix represented as cartoon in cyan (pdb: 4HFZ). The three key residues are represented in stick illustrating their interaction with the hydrophobic binding cleft of hDM2: F19, W23 and L26

The p53-xDM2 interaction has indeed been extensively studied from both cellular and structural perspectives and can now be viewed as a model system for PPI inhibitor discovery. The first crystal structure of the complex p53/MDM2 was published in 1996 by Pavletich *et al.* at a 2.3 Å resolution (pdb: 1YCQ)[33]. Currently, 100 structures of MDM2 in complex with p53 peptide or analogues are available in the PDB. These structures allowed an understanding at the molecular level of how the binding occurs between the two partners and highlight that upon binding to MDM2, the p53 N-terminal transactivation domain adopts an α-helical conformation where three "hot spot" residues are located in the hydrophobic binding cleft of M(h)DM2: Phe_{19}, Trp_{23} and Leu_{26} (Figure 2A)[18]. The knowledge of the mode of interaction between the two protein partners thanks to these numerous x-ray crystal structures have guided the design of potent M(h)DM2 antagonists that have emerged as potential anticancer therapeutics (from small molecules to peptides and peptidomimetics) to trigger the apoptosis cellular machinery[32].

β-peptides designed to mimic p53 and specifically interact with the surface of hDM2 have been investigated first[34,35]. In contrast to α-peptides, β-peptides consist of residues with an extra α-methylene unit in their backbone and exhibit a high propensity to fold into stable helical secondary structure with as few as four β-amino acids[36]. These aliphatic foldamers adopt a stable 3_{14}-helical conformation whereas α-peptides form a regular 3.6_{13} helix[37] (Figure 9A). The 14-helix presents a little longer rise per residue (1.56 Å) than the α-helix (1.5Å) and the radius is slightly wider (2.7Å and 2.2Å for β- and α-peptides respectively). β-Peptides which are resistant to proteolysis have been used successfully to mimic α-helical peptide epitopes despite significant differences with α-helical peptides (the two helices have opposite polarities),[38]. Schepartz *et al.*

have designed β-decapeptides to mimic the α-helix from the activation domain of p53 ($p53_{AD}$=QETF$_{19}$SDLW$_{23}$KLL$_{26}$PEN) with the hypothesis that the function of p53 peptide would be recapitulated by keeping the side chains of the three hot spot residues, namely Phe19, Trp23 and Leu26[33–35,39]. Although β-peptides adopt a 14-helix conformation with different dimensions (residues per turn and radius) than α-helix, the authors assumed that the β-Phe19, β-Trp23 and β-Leu26 β-residues could position their side-chains in the same *i*/*i*+3 relationship than that observed for the reference peptide (Figure 9B)[39]. Experiments of direct and competition fluorescence polarization have been performed and highlighted one compound, referred here as **I.1,** capable to interact with the surface of hDM2 and to inhibit the interaction with p53 reference peptide (Figure 9B). Structural determination by high-resolution NMR spectroscopy allowed the determination of the structure of **I.1** in solution and revealed a C-terminal part partially unwrapped. This result suggested to the authors to develop a library of analogues of **I.1** with various substitutions on its non-recognition face, in order to obtain more active compounds[35]. More recently, they managed to increase the affinity of their β-peptides to the nanomolar range by introducing non-natural side chains on the binding face[39]. Indeed, they replaced the β-Trp23 by a β-3-trifluoromethylphenylalanine (**I.2,** K_d=28±5 nM) or even β-6-chlorotryptophan (**I.3** K_d=30±5 nM). Moreover, these two compounds are able to bind hDM2 as well as hDMX, a hDM2 related protein identified as another oncologic target.

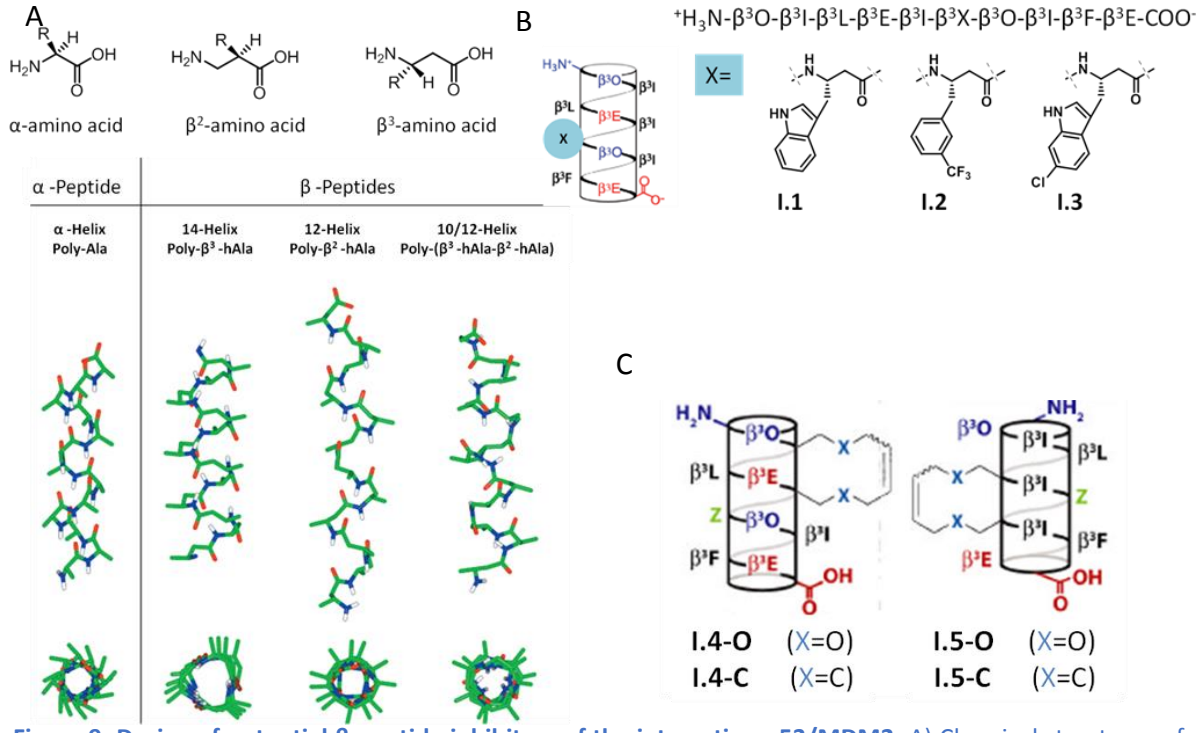
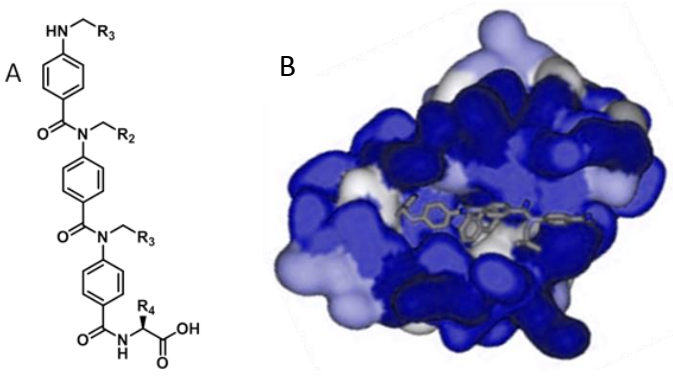


Figure 9: Design of potential β-peptide inhibitors of the interaction p53/MDM2. A) Chemical structures of α- and β-amino acids and structural representation of their corresponding helices: α-helix, 12-helix and 10/12-helix (adapted from Cheng *et al.*, Chemical reviews, 2001 B) Helical representation of β-peptides designed as p53 analogs with the replacement of the aromatic key residue X by unnatural hydrophobic side-chain containing β^3-amino acids. C) Helical representation of β^3-peptides bridged between positions 2 and 5 or 4 and 7, designed as p53 analogs. Z corresponds to 3-(*S*)-3-amino-4-(2-trifluoromethylphenyl)-butyric acid (adapted from Bautista et al., JACS, 2011).

In continuation of this work, a strategy involving peptide structure stabilization and aimed at increasing cell permeability was applied to these β-peptide foldamers[40]. Bridged β-peptides, reminiscent of the stapled peptide strategy, were hence developed by Schepartz *et al.* The bridge was composed of diether or hydrocarbon linkages and was inserted between positions 4 and 7 or 2 and 5 (Figure 9C). For the first position, fluorescence polarization (FP) assays of bridged analogues of **I.2** (**I.5-O** and **I.5C**) have shown a better binding to hDM2 (K_d= 53.9 and 94.1 nM respectively) than the unbridged β-peptide analogues (K_d= 114 and 253 nM)[40]. Regarding the introduction of the bridge at the second position (between 2 and 5), the observed binding of **I.4-O** and **I.4-C** was 4 and 8-fold lower than the unbridged analogues, illustrating the importance of the position of the bridge for protein surface recognition. Concurrently; confocal microscopy confirmed that fluorescently-labeled versions of bridged β-peptides (between positions 4 and 7) were distributed widely in Hela cells (endosomes as well as nuclear and cytosolic compartments) whereas analogues bridged between positions 2 and 5 were not. These results highlight a correlation between the affinity for hDM2 and cell uptake.

Amongst proteomimetic scaffolds that mimic the side chain projections of targeted peptide sequences, aromatic oligoamide foldamers, initially developed by Hamilton et al. with terphenyl

derivatives, are good candidates for the inhibition of PPIs[41–43]. These compounds are of particular interest for several reasons: 1) they exhibit predictable conformations; 2) their side chains may be projected in a similar spatial orientation than key residues of α-helices; 3) they are potentially amenable to library synthesis by resorting to solid-phase strategies[44,45]. A. J. Wilson and coworkers have recently reported the inhibition of p53-hDM2 interaction with *N*-alkylated oligoamides[46], and 3-*O*-alkylated oligoamides[47,48]. They have reported the synthesis of *N*-alkylated oligoamides (with the development of solid phase synthesis based on Fmoc strategy[46]) and their ability to mimic the *i*, *i*+4 and *i*+7 side-chain projections of an α-helix[49,50]. Fluorescence anisotropy competition (FAC) assay (or fluorescence anisotropy competition titration) for inhibition of the p53/hDM2 interaction has been performed on a small library of *N*-alkylated oligoamides and unveiled that a minimal length of three aryl motif was required to ensure a correct inhibition (micromolar range)[46]. The most potent inhibitor of the interaction of $p53_{15-29}$/hDM2 protein in this series, **I.6** (Figure 10), has a half-maximal inhibitory concentrating value (IC_{50}) of 2.8 µM which is comparable with the native $p53_{15-29}$ α-helix (IC_{50}=1.2 µM). Structural studies have also been performed with ^{1}H-^{15}N HSQC experiments and the chemical shifts were mapped on the crystal structure of p53-hDM2 complex (pdb entry: 1YCR) (Figure 10B). The chemical shifts observed during the addition of compound **I.7** illustrated similarities with the addition of the p53 peptide, but also that structural changes occurred upon formation of the hydrophobic cleft of hDM2. Indeed, shifts are observed at both end of the helix binding cleft of hDM2, suggesting that the helix mimetic **I.7** s adopts an extended conformation to bind to the surface of hDM2.



I.6 R_1=iPr; R_2=Nap; R_3=Ph
I.7 R_1=iPr; R_2=Ph; R_3=Ph
I.8 R_1=iPr; R_2=2-Nap; R_3=2-Bu; R_4=2-Bu
I.9 R_1=iPr; R_2=4-Cl-Ph; R_3=2-Bu; R_4=2-Bu

Figure 10: N-alkylated oligoamides for the inhibition of p53-hDM2. A) Chemical structure of N-alkylated oligoamides B)^{1}H-^{15}N HSQC chemical shift mapping with hDM2 compared with the crystal structure of the hDM2-p53 structure (1YCR) is represented in a blue surface and N-alkylated oligoamide in grey stick has been manually docked (adapted from A. J. Wilson *et al.* , Org. Biomol. Chem, 2010).

To complement this structural study in solution, cell experiments have been performed on a library of *N*-alkylated oligoamides to evaluate their cellular activity[51]. A high-content imaging screen was developed and among the 77 members of the oligoamide library, compounds such as **I.8** and **I.9** appeared to be effective at several endpoints (four endpoints were assessed: (1) cell number, (2) the use of an antibody against caspase 3 to identify cells which maybe apoptopic, (3) autophagy and (4) the arrangement of actin filaments)[51]. MTT cell toxicity assays have highlighted

that few compounds including **I.8** and **I.9** present a similar level of toxicity to Nutlin-3 (a small molecule inhibiting the interaction p53-MDM2, currently in preclinical study[52]). FAC assays confirmed that **I.8** and **I.9** inhibit p53-hDM2 in the low micromolar range of IC_{50}. These two compounds were the most effective in dose-response experiments but they lack selectivity as they also showed some activity as inhibitors of another anticancer target (Mcl-1/NOXA-B). As this class of compounds enter the cells, they can reproduce their binding activity in a cellular context. Concurrently, *O*-alkylated oligoamides have been synthesized and also showed potential as α-helix mimetics[53]. Indeed, with their benzamide rigid scaffold, their functional groups allow the mimicry of side chains found at positions *i*, *i*+4 and *i*+7 positions of α-helix similar to *N*-alkylated oligoamides[54]. Wilson *et al.* designed and synthesized a series of compounds as inhibitors of p53-hDM2 and the best compound **I.10,** presented a promising IC_{50} = 1.0±0.11 µM as determined by FAC assays[47]. However, the aqueous solubility is an important parameter for structural and biological studies and **I.10** presents a poor solubility in aqueous environment[55]. The chemical composition of this molecule with aliphatic and aromatic side chains does not help for water solubility. To circumvent this chemical limitation, Wilson et *al.* have integrated a "wet-edge" composed of a hydrophilic ethylene glycol chain that is incorporated along the solvent-exposed surface of **I.10** and obtained the mimetic **I.11** (Figure 11A)[55]. The synthesis route is first based on the preparation of a bifunctional *O*-alkylated aromatic monomer, followed by a succession of iterative coupling with Ghosez's reagent and reduction of nitro-masked groups into an amine group. FAC assays demonstrated that the functionalized compound **I.11** is similar to **I.10** in mimicking the p53 peptide with IC_{50} in micromolar range for both compounds (Figure 11B). These results showed that this orthogonal functionalization has a low impact on the binding affinity but improved solubility in the assay buffer (sodium phosphate buffer pH 7.5). Indeed, qualitative solubility tests were performed on compounds **I.10** and **I.11** in buffer with 10 and 1% of DMSO and **I.11** shows a higher solubility with 1% of DMSO than **I.10**.

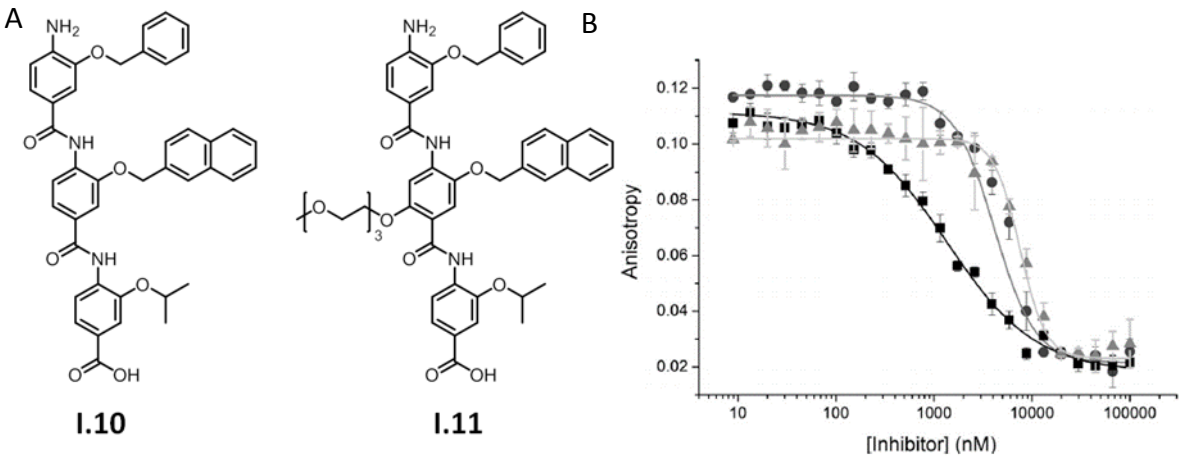


Figure 11: Structure and fluorescence anisotropy assays of *O*-alkylated oligoamides A) Chemical structure of oligoamides **I.10** and its "wet-edged" version **I.11** mimicking the side chain projection of the 3 key hydrophobic residues of p53 transactivation domain. B) Fluorescence anisotropy competition titration data (40mM sodium phosphate buffer pH 7.5). Displacement by native peptide (black squares), trimer **I.10**, (dark grey circles) and PEGylated trimer **I.11** (light-grey triangles) (taken from Prabhakaran, 2013, Euro JOC)

Finally, hybrids α-helix mimetics with aromatic oligoamides (Figure 7 and Figure 12A), such as O-alkoxybenzoic acid and a central α-amino acid have been synthesized and showed their ability to mimic α-helix as they could inhibit p53-hDM2 interaction with IC_{50} varying from 10 to 100 µM (IC_{50} = 0.5 µM for Nutlin-3)[56]. Quantitative structure-activity relationship (QSAR) analysis revealed that the hydrophobicity of the central α-amino acid side chain (the Trp_{23} mimic) is essential for the molecular recognition of hDM2.

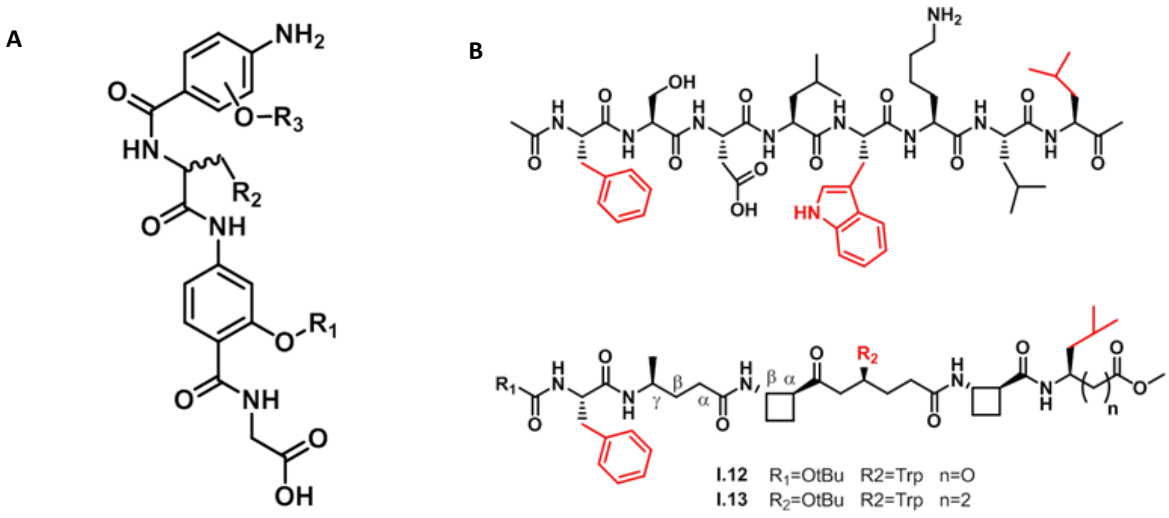


Figure 12: Chemical structure of hybrids α-mimetics of p53. A) Chemical structure of hybrids containing a central amino acid and aromatic oligoamide B) Alignment of key side chains with the p53 segment on the top and the structure of two α/β/γ-peptides helix mimetics.

In collaboration with the group of A. J. Wilson, Aitken et al. have recently reported hybrid α/β/γ-foldamers as inhibitors of the interaction between p53 and hDM2[57]. Previous structural studies of Gellman and Balaram have revealed that the local introduction of β and γ-amino acids within peptide sequence maintain the α-helix[58,59]. Grison et al. have recently showed that the alternation of trans-2-aminocyclobutane carboxylic acid and γ-amino acids was well-tolerated in peptides sequences adopting of 9/8 ribbon or 13-helix structure[60,61]. In this context, α/β/γ-hexapeptides were designed to mimic the projection of the three hot-spot residues of $p53_{19-26}$ peptide (Phe19, Trp23 and Leu26) at *i*, *i*+4 and *i*+7 positions respectively (Figure 12B). NMR, IR and CD experiments data illustrated that compounds fold into 12,13-helicall conformation in solution[57].Furthermore, the proteolytic stability of α/β/γ-peptides was investigated with chymotrypsin and showed that the compounds were still stable after 80 min, whereas $p53_{19-26}$ is fully degraded after 22minutes. Their ability to mimic the N-terminal helical domain of p53 (residues 19-26) was confirmed by FAC assays with the best compounds (**I.12** and **I.13**) showing a half-maximal inhibitory concentration value (IC_{50}) 10-fold higher than the native p53. However, design optimization would be necessary to further improve the affinity of these compounds.
All these foldamers have shown a good potential to inhibit the interaction between p53 and hDM2 but improvement of affinity, solubility and cell penetration remain necessary.

b) … of the Bcl-2 superfamily

As mentioned before, apoptosis is regulated by a complex network of cell-signaling mechanisms involving key regulatory molecules. Apoptosis can be activated by several pathways, including the mitochondrial activation with the release of cytochrome C upon the stimulation by a variety of cell-death signals. Another pathway is the ligation of death receptors leading to the activation of capsase-8 and -3 (Figure 13)[62]. Several PPIs are involved in these pathways and have been defined as being important targets to develop modulators.

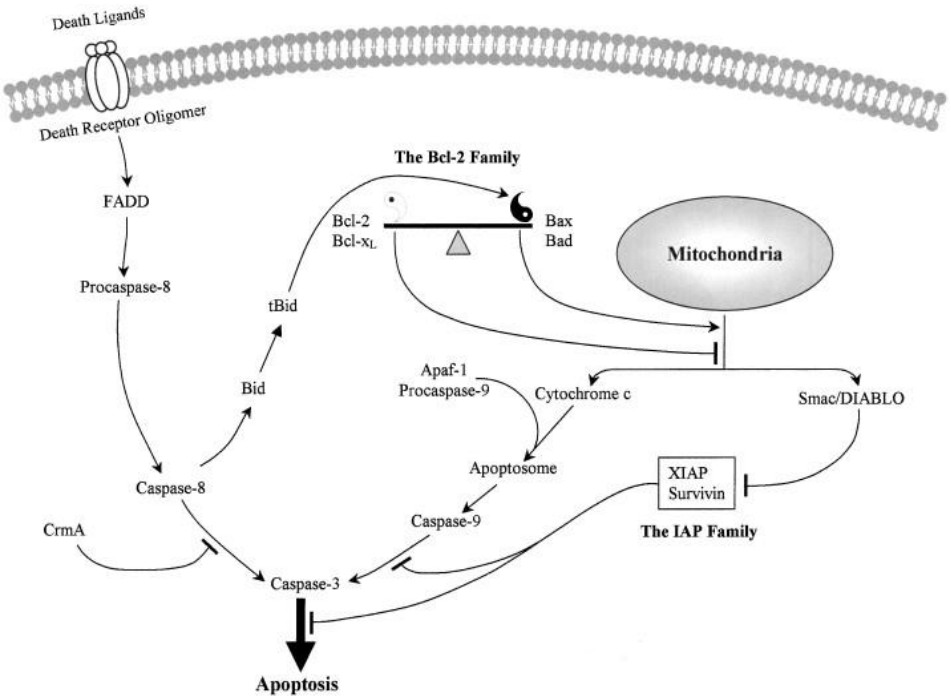


Figure 13: Protein-protein interactions involved in programmed cell death process. A network of PPIs along the two apoptotic pathways involving the mitochondria and death receptor (taken from Huang, *Chemistry and Biology*, Vol. 9, 2002[62]).

The Bcl-2 (B-cell lymphoma 2) family gathers evolutionarily conserved proteins with Bcl-2 homology (BH) domains that control the mitochondrial activation pathway. This family is involved in the regulation of the apoptotic signal and can either promote or inhibit apoptosis[63]. These proteins are structurally different but share four conserved BH domains and they can be defined into three categories: anti-apoptotic members, pro-apoptotic members and BH3-only pro-apoptotic members (Table 1).

Table 1 : The Bcl-2 protein family

	Members	Domains shared
Anti-apoptotic members	Bcl-2, Bcl-x_L, Bcl-W, A1, Mcl-1, BOO	BH1, BH2, BH3 and BH4
Pro-apoptotic members	Bax, Bok, Bcl-x_S, Bak, Bcl-G_L, Bfk,	BH1, BH2 and BH3 BH2 and BH3
BH3-only pro-apoptotic members	Bad, Bik, Bid, Hrk, Bim, Noxa, Puma, Bmf	BH3

In order to understand the pertinence of foldamers as PPI modulators in this context but without going into too much into details, it is important here to remind that the anti-apoptotic members of the Bcl-2 family are preventing mitochondrial outer-membrane permeabilization. They sequester the BH3-only members that normally activate the cell permeabilization, or they inhibit PPIs involving pro-apoptotic members, such as Bax/Bak[63]. Consequently, scientists have made considerable efforts to mimic BH3-only proteins in order to counteract the activity of anti-apoptotic members. The BH3 domain is indeed omnipresent in the Bcl-2 family and adopts a helical conformation which could be mimicked by foldamers. Various designs that led to inhibitors

of anti-apoptotic members have been reported by several groups in the context of new anti-cancer drugs.

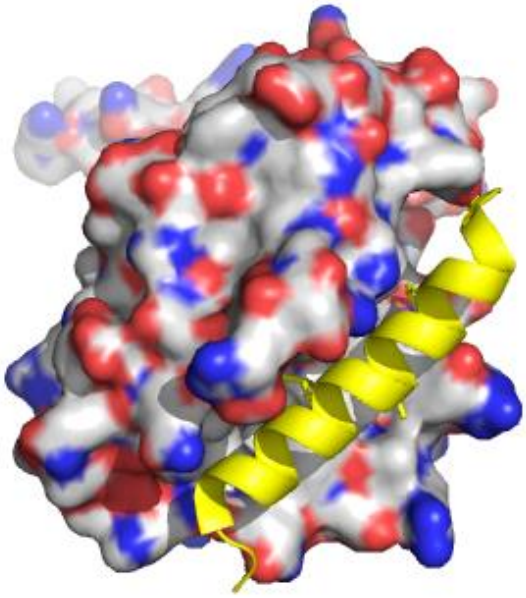
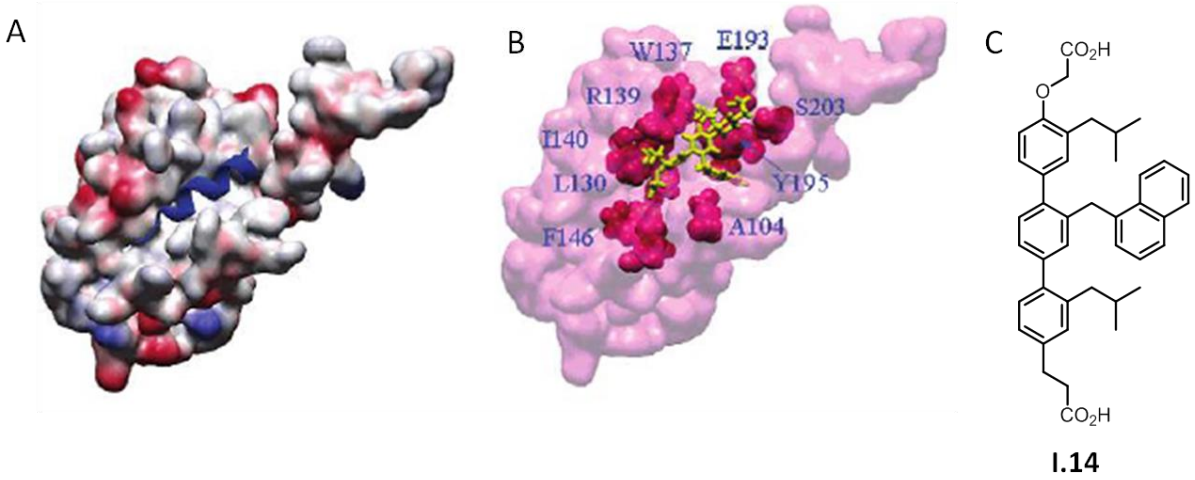


Figure 14: X-ray crystal structure of Bim BH3 peptide (yellow) in complex with Bcl-X_L (pdb entry = 3FDL). Bcl-XL is represented as a surface. The structure shows that the 23-residue α-peptide Bim BH3 adopts a helical conformation.

Solution and X-ray crystal structures indicate that Bcl-2 and its homologs are characterized by a long hydrophobic cleft allowing the binding of the α-helix from the BH3-domain of pro-apoptotic members. The Bim BH3 domain sequence (DMRPEIWIAQELRRIGDEFNAYYARR) which is depicted in Figure 14, adopts an α-helix conformation and presents four key hydrophobic residues in *i*, *i*+4, i+7 and *i*+11 positions[64]. These residues correspond to Ile58 Leu62, Ile65, and Phe69 and are all localized on one face of the helix that is in contact with the protein partner surface. Peptidomimetics, including the foldamers, have been developed to mimic this BH3 domain with the goal to antagonize Bcl-2 family members.

First, terphenyl scaffolds have been developed by Hamilton and coworkers to mimic the BH3 domain of Bak and Bad-proteins that associate with the anti-apoptotic protein Bcl-x_L[42]. A structure-based design was used to determine the best molecular scaffolds that would mainly mimic the side chain projection along one face of the α-helix. Terphenyl scaffolds have the ability to mimic the projection of key recognition motifs on the surface of an α-helix[41,65]. The solution and crystal structures of Bak/Bcl-x_L showed that four hydrophobic residues of Bak peptide are involved in the binding: Val74, Leu78, Ile81 and Ile85[66,67]. They terphenyl-based compounds have substituents appended on the *ortho*-positions to imitate the side-chains of these key hydrophobic residues. The binding affinity of these series of terphenyl α-helix mimics was determined by using a fluorescence polarization assay. The best compound, **I.14**, exhibits a good binding affinity with a K_d = 114 nM[42]. Less hydrophobic compounds showed a lower affinity which illustrates the importance of hydrophobic interactions for the binding to Bcl-x_L. Furthermore, because of the

poor solubility of these aromatic compounds in water, terminal carboxylate groups were added to improve it[65].



I.14

Figure 15: Strategy to inhibit the interaction between Bak BH3-peptide and Bcl-x_L with a terphenyl compound. A) Crystal structure of Bak BH3-peptide/Bcl-x_L complex. B) Docking results and HSQC results with residues shifted. C) Molecular structure of the best terphenyl derivative **I.14**.

Furthermore, docking studies coupled to HSQC NMR experiments with ^{15}N-labelled Bcl-x_L were performed to localize **I.14** on the surface of Bcl-x_L. Shifts were observed for several residues on the surface of Bcl-x_L (Ala104, Leu130, Trp137, Arg139, Ile140, Phe146, Glu193, Tyr195 and Ser203) illustrating that **I.14** interacts in the same binding cleft than the natural helical peptide (Figure 15). The observed affinity of **I.14** in the sub-micromolar range was a real improvement in comparison to small molecule inhibitors of Bcl-x_L reported earlier in the literature with K_D values more in the µM range[68–70].

More than one decade ago now, Gellman and coworkers developed β-, α/β-, (α/β + α)- and (α/β + α)-peptides with the aim to tailor peptidomimetic foldamers that bind tightly to the BH3-recognition cleft of Bcl-x_L[38,71,72]. They initially designed more than 200 β-peptides and α/β-peptides to mimic the BH3 domain of Bak and identified a foldamer with a significant affinity for Bcl-x_L[38,71,72]. The best compound, **I.15** (Figure 16A) is composed of an α/β-segment assembled with an α-peptide tail and exhibited a good affinity for Bcl-x_L (IC_{50} = 0.029 µM). However, this hybrid sequence did not give any binding results with Mcl-1 protein. This absence of interaction appeared to be a significant limitation of this strategy because this foldamer-type ligand could have been more efficient and useful if it could have interacted with multiple anti-apoptotic Bcl-2 protein family[57].

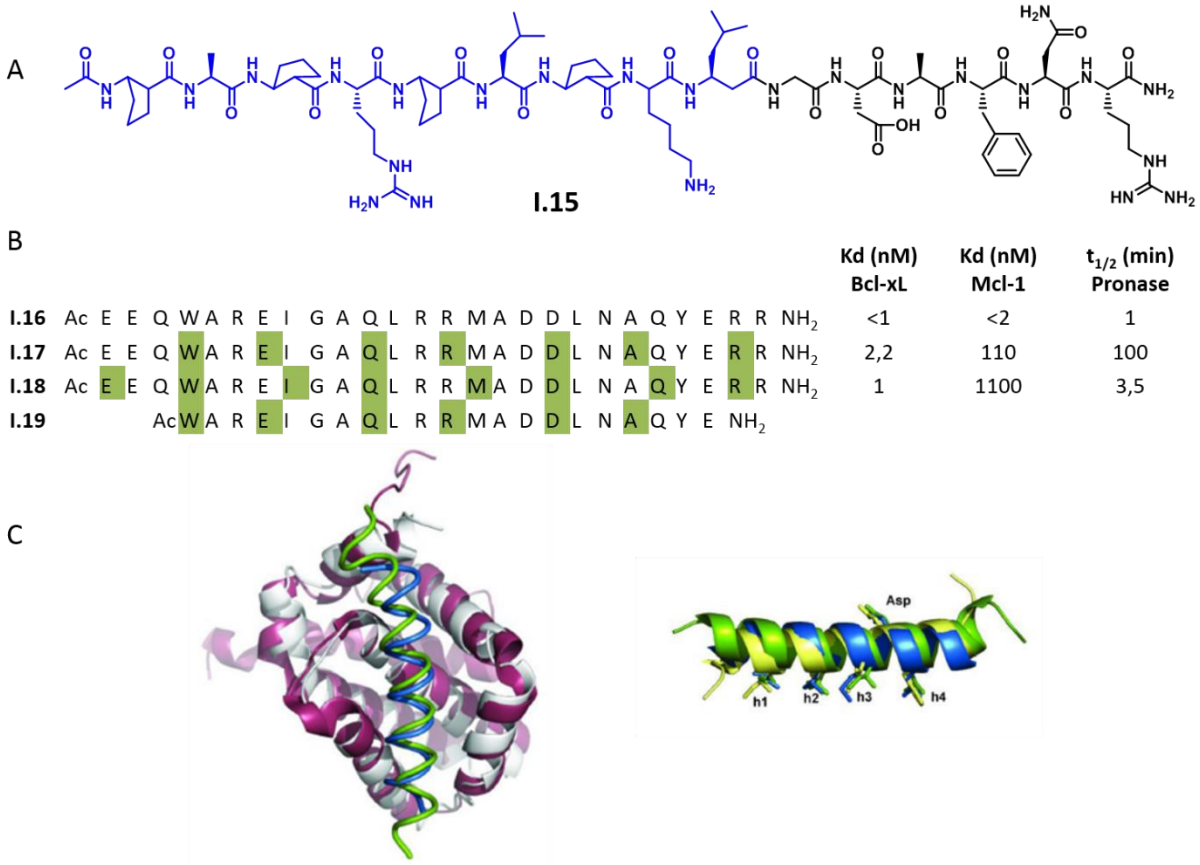


Figure 16: Sequences of α/β-peptides mimicking the BH3 domain of Puma and X-ray crystal structure of the resulting complex at 2.2Å. A) Chemical structure of (β+α / α)-peptide foldamer that mimic well the BH3 domain. B) Sequences of Puma BH3 peptide and α/β-peptides Puma mimetic with dissociation constants determined by direct binding FP and peptide degradation assay to determine half-life at 50µM concentration in the presence of 5 µg/ml pronase. β^3-residues are highlighted in green C) On the left, Puma mimetic **I.19** (navy) bound to Bcl-xL protein (white) (pdb=2YJ1) overlaid with Puma/Mcl-1 complex (PDB = 2ROC) (green/red). On the right, overlay of α/β-peptide **I.19** (navy) with the Puma BH3 domain (green) and the Bim BH3 domain (yellow). Key interacting hydrophobic residue orientations (h1-h4) as well as the conserved Asp residue from the α-peptide BH3 sequence are well-mimicked by **I.19**.[73]

Later on, they described a different approach to design effective α/β-peptide ligands[74]. Starting from a peptide corresponding to the Puma BH3 domain (**I.16**), they prepared 7 analogues containing regularly distributed $\alpha \rightarrow \beta^3$-amino acid residues along the natural protein sequence while keeping the native side-chains. These α/β-peptides possessed an ααβααβ backbone pattern known to adopt a conformation very close to the α-helical conformation[75] and the position of β-residues was moved around the helical periphery (**I.17**-**I.18**) (Figure 16B). Fluorescence polarization assays were performed to determine the inhibition constant (K_i) of the seven α/β-peptides and several chimera-type inhibitors with nanomolar affinities for Bcl-x_L and Mcl-1 were found[72]. The authors showed that a variation in the position of β-residues caused considerable changes in the affinity and selectivity for Bcl-x_L or Mcl-1. Beside the binding affinity, the proteolytic stability was also improved with the insertion of β^3-amino acids. This highlights the fact that the positioning of β-residues is critical for the function of these compounds. Their

results also highlighted that this sequence-based design was more efficient than the structure-based one in engineering promising foldamers as PPI inhibitors. More recently, in collaboration with W. Douglas Fairlie, they reported the synthesis and first structural elucidation of an α/β-peptide mimicking the Puma BH3 domain in complex with Bcl-xL[73]. The α/β-peptide which contains six β^3-amino acids regularly distributed to keep the ααβαααβ backbone adopts an α-helical conformation with β^3-amino acids aligned along one solvent-exposed face of the helix. The crystal structure of **I.19** in complex with Bcl-x_L was obtained at 2.4 Å (pdb= 2YJ1) and confirmed that the BH3-recognition site is involved in the binding with this compound (Figure 16C). The superimposition of helices showed the structural similarities between **I.19**, Puma and Bim BH3 domains but also revealed differences in the orientation of some side-chains such as those of Tyr_{101} and Phe_{105}. Additionally, the radius of curvature indicated that the helix was not straight and can bend to accommodate a binding partner. Furthermore, key binding residues of a native BH3 domain are well-mimicked by this α/β-peptide despite the additional backbone methylene carbon brought by β^3-residues which is quite encouraging for the sequence-based design approach strategy. However, the affinity of **I.19** for Mcl-1 remains low due to the orientation of solvent-exposed residues and may be improved thanks to these structural data.

Concurrently, again in collaboration with the group of Fairlie, Gellman and coworkers published a series of analogues of the Bim BH3 domain that were composed of two different regular α/β patterns throughout the backbone[76]. They studied the local substitution of α-residues in a 18-mer α-peptide **I.20** by β-residues keeping the appropriate side-chain and reinforced the fact that β-residues aligned on one side of the helix can well-mimicry α-helix function. The two ααβαααβ (compound **I.21**) and αααβ (compound **I.22**) patterns have been tested and several α/β-peptide sequences showed effective mimicry of Bim BH3 domain (Figure 17). The ααβαααβ backbone allows an alignment of β^3-residue side chains along one side of the helix. Remarkably, a X-ray crystal structure has been obtained for chimeric α/β-compounds **I.21** and **I.22** in complex with Bcl-x_L.

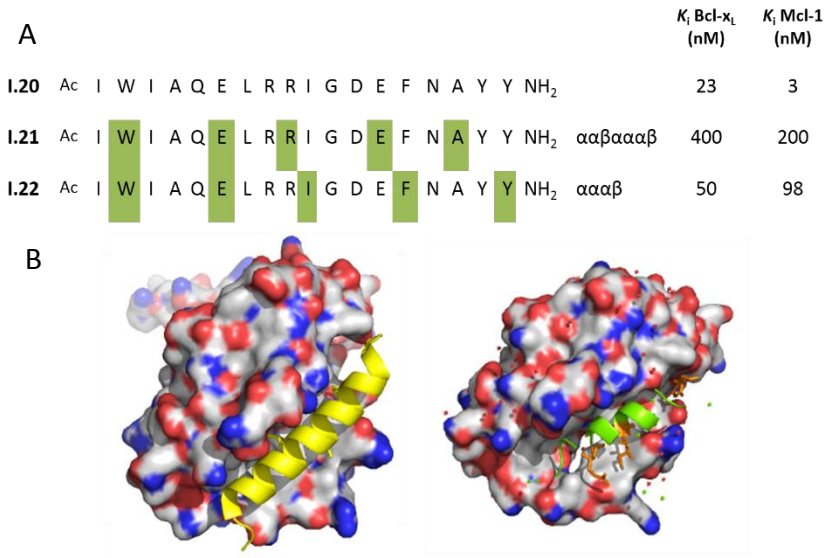


Figure 17: Sequence of natural peptide and two α/β peptide mimicries and crystal structure of one α/β peptide in complex with Bcl-x_L. A) Sequences of natural peptide **I.20** and α/β peptide **I.21** and **I.22**. β^3-Residues are highlighted in green. B) Crystal structure of Bim BH3 peptide (yellow) in complex with Bcl-x_L (pdb entry: 3FDL). The crystal structure of the compound **I.22** (green) in complex with Bcl-x_L reveals the same surface of contact as the one of the natural peptide is involved in the interaction (pdb entry: 4A1W).

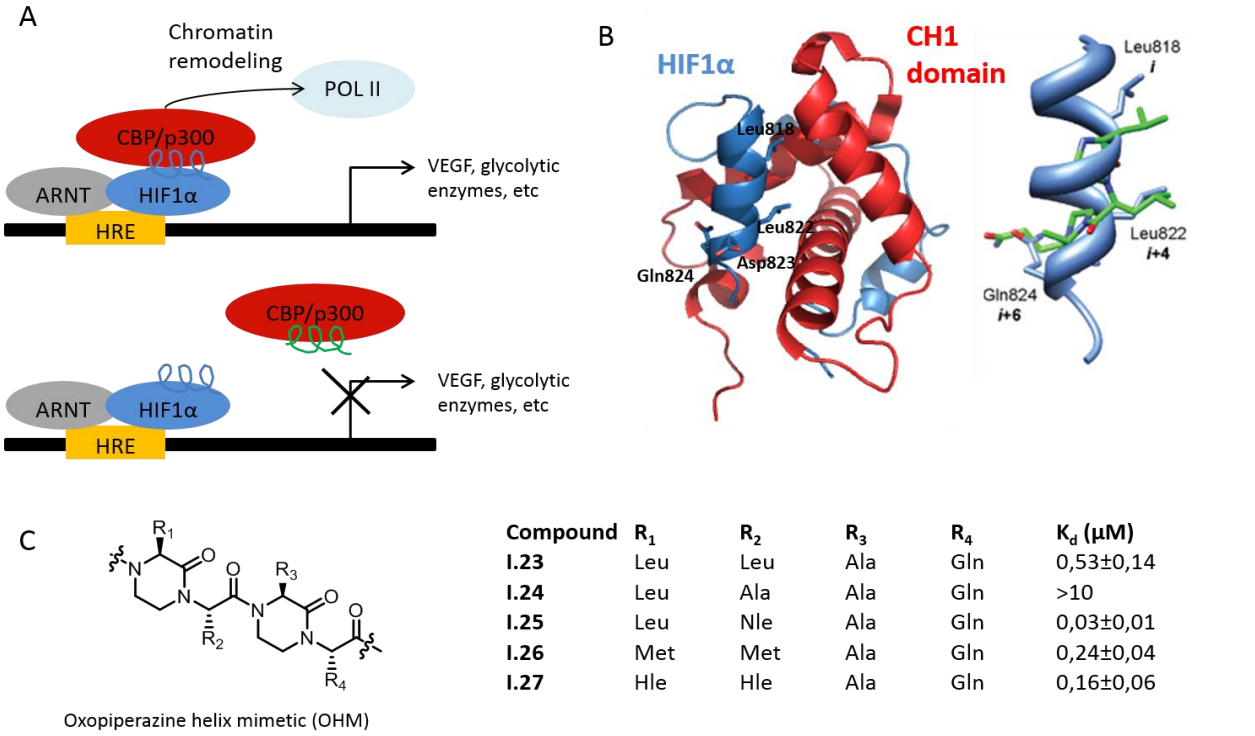
Both crystal structures illustrate that Bim-derived α/β-peptides have huge similarities with the natural complex because they interact on the same area of Bcl-x_L and side chains contacts are conserved. Despite the replacements of five α-amino acids by corresponding β^3-residues, the four key hydrophobic residues defining the BH3 domain remain present in **I.22** with Ile_3, Leu_7, β^3-Ile_{10} and β^3-Phe_{14} (Figure 17). Among the αααβ series, **I.22** exhibited the best affinity for Bcl-x_L (K_i = 50 nM as determined by competition fluorescence polarization) and a similar affinity to Mcl-1. This study allowed the determination of a new substitution pattern for the BH3 domain mimicry in addition to the ααβαααβ pattern, thus expending the possible α/β combinations useful to mimic an α-helix such as here the BH3 domain of Bim.

Other scaffolds such as oligopyridyl[77], oligothienylpyridyl[78] (exhibiting higher solubility), vinylogous amides (enaminones developed by Hamilton and coworkers)[79] or more recently oligoamides (picolinamide and benzamide)[80] led also to good helical mimetics of BH3 domains but several optimizations to increase the affinity and for cell penetration are still needed to make these molecules suitable for a pharmaceutical development. In comparison, The BH3-mimetic ABT-199, a small molecule helix mimetic that interacts with Bcl-2 entered on the market in 2016 as Venetoclax is a promising anti-cancer treatment[81,82].

c) …the complex P300/HIF1α

During this last decade, a new target for cancer therapy research, the hypoxia-inducible factor 1 (HIF1) has emerged. As it accumulates in cancer cells and recruits coactivators CREB-binding protein/p300 (CBP/p300), HIF1 plays a key role in adaptation of tumor cells to hypoxia. This transcription factor is a heterodimer basic helix-loop-helix protein with two subunits: HIF1α and HIF1β[83]. The expression of over 40 genes is activated by HIF1 at transcriptional level[84]. In particular, HIF1α in complex with p300 activates VEGF (vascular endothelial growth factor) and its pro-angiogenic function. HIF1α regulates also gene encoding for glucose transporters leading

cancer cell energy metabolism to be altered[84]. Furthermore, hypoxia inducible genes are involved in the angiogenesis (formation of capillary branches starting from preexisting capillaries) a complex process that involves numerous genes coming from different types of cells. For scientists, targeting the interaction between HIF1α and p300 should allow the control of HIF1α mediated hypoxia signaling. The solution structure of the HIF1α/p300 complex shows that the C-terminal transactivation helix domain (residues 786-826) of HIF1α interacts with the cysteine-histidine rich 1 (CH1) domain of the coactivator p300 (Figure 18A)[85]. Key residues of HIF1α are involved in the interaction with p300: Leu818, Leu822, Asp823 and Gln824 that are located in a hydrophobic pocket of the p300/CBP CH1 domain.



Compound	R_1	R_2	R_3	R_4	K_d (µM)
I.23	Leu	Leu	Ala	Gln	0,53±0,14
I.24	Leu	Ala	Ala	Gln	>10
I.25	Leu	Nle	Ala	Gln	0,03±0,01
I.26	Met	Met	Ala	Gln	0,24±0,04
I.27	Hle	Hle	Ala	Gln	0,16±0,06

Figure 18: Presentation of hypoxia inducible factor in complex with CH1 domain and strategy to inhibit this interaction. A) The resulting complex controls the transcription of hypoxia–inducing genes. B) Solution structure of HIF-1α (blue) in complex with the CH1 domain (red) (pdb entry: 1L8C) and on the right, superimposition of the helix of HIF-1 and an oxopiperazine derivative. C) Chemical structure of oxopiperazine helix mimetics and affinity data[81].

Arora and coworkers have designed oxopiperazine helix mimetics (OHMs, Figure 18) to reproduce the side-chain contacts established by the third helix of HIF1α when binding to p300[86,87]. These small scaffolds are assembled from natural amino acids and ethylene bridge constraints between nitrogen atoms of two neighbor residues. In contrast with the majority of nonpeptidic helix mimetics based on achiral aromatic backbones, OHMs are based on chiral scaffolds and should enhance specificity of binding. Because high resolution structure and computational analysis of the HIF/p300 complex suggested that OHM dimers should mimic the *i*, *i*+4 and *i*+6 or *i*+7 residues of α-helices, three key residue side-chains of the four involved in the natural complex (*i.e.* Leu818, Leu822 and Gln824) were equipped on the R_1, R_2 and R_4 positions respectively[87]. To note, in the

computational design (Rosetta) the R_3 was not predicted to make contacts with the p300 proteins surface and an Ala residue side-chain was installed at this position. The binding affinities of this OMH series (5 compounds synthesized on solid phase) for the p300-CH1 domain were determined by using intrinsic tryptophan fluorescence spectroscopy and a sub-micromolar range affinities were obtained. The best affinity (K_d = 0.53±0.14 µM) was obtained for **I.23** (Figure 18), which contains the native side-chain projections of Leu818, Leu822 and Gln824[87]. *In cellulo* assays have next shown that the down-regulation of several genes implicated in angiogenesis, apoptosis and cell proliferation was observed in cells treated with **I.23**. The authors also performed *in vivo* tests in mice and showed that **I.23** could reduce tumor size without affecting measurable changes in animal body. To optimize HIF mimetics, the authors performed a computational approach with Rosetta to improve the binding affinity for p300-CH1[88]. Predictions were evaluated with the synthesis of analogues **I.24-I.27** that binds p300 with higher affinities than **I.23** such as **I.25** that presents a 13-fold enhancement in binding affinity (K_d=30.2±1.87 nM). This study highlights the efficacy and specificity of OHM derivatives for the inhibition of HIF-p300-CH1. Concurrently, Arora and coworkers also designed and optimized oxopiperazine analogues by relying on a computational approach with Rosetta and applied it for the design of novel p53 activation domain mimics and HIF1α[88].

Another foldamer strategy to inhibit the HIF1α/p300 consisted in designing 3-*O*-alkylated oligobenzamides to mimic the third helix of HIF1α [89]. A. J. Wilson and coworkers conceived a series of these compounds and they even elongated one of their oligobenzamide analogue by conjugating it with the HIF-1α second helix (aa 794-813) with the aim here to increase the binding affinity as well as protein specificity (Figure 19)[90].

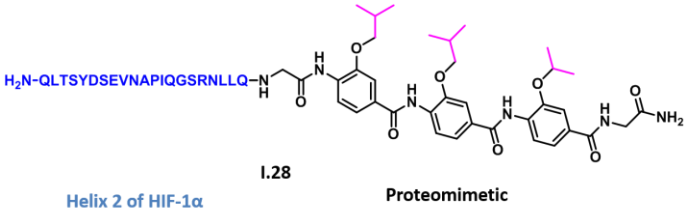


Figure 19: Chemical structure of the hybrid I.28 with one letter code for the amino acids sequence of Helix 2 of HIF-1α in blue.

The resulting molecule **I.28** that referred as "bionic" hybrid by Wilson *et al.*, showed an increased specificity although no real improvement in the binding to p300 (micromolar range) was observed in comparison with the native sequence (HIF-1α$_{794-826}$). This original strategy was performed to increase specificity or to target longer peptide sequences. Perspectives for this target PPI have recently been reported by Wilson and coworkers[91]. In contrast with previous PPIs (hDM2, Bcl-2 family), HIF PPIs are more diverse, complex and larger which renders the development of high-affinity ligands more challenging. Moreover, no crystal structure of the HIF1α/p300 complex are currently available in the PDB while it would help for rational design to develop new inhibitors. Considerable progresses have been realized in the design and synthesis of PPI inhibitors. However, tailoring high affinity ligands with high specificity remains essential to inhibit this action of this promising target in an anticancer therapy context.

d) Other foldamers targeted PPIs involved in cancer

In the literature, other PPIs involved in cancer development have been investigated such as the estrogen receptor (ER) in complex with an estrogen molecule. ER is a transcription factor of the nuclear receptor superfamily that mediates the expression of estrogen-activated genes and is involved in several diseases, in particular breast cancer, making it an important target for new therapeutic research. The estrogen molecule binds ER on its ligand binding domain leading to interaction with DNA-promoters followed by the recruitment of coactivators that are required for the expression of genes regulated by ER. These coactivator proteins present a conserved motif (LXXLL with L for Leucine and X for any amino acid) that is essential for the binding to the ER surface[92]. Regarding the mode of interaction, an α-helix is involved with the key residues localized on two faces of the helix (Figure 20A). In the literature, synthetic scaffolds designed to mimic multiple face contacts have been already reported by several groups. As described previously, Hamilton and coworkers had indeed developed terphenyl molecules capable to mimic *i*, *i*+4 and *i*+7 residues of an α-helix and lead to potent inhibitors of p53/MDM2 and Bcl-x_L/Bak PPIs[41,43,93]. To mimic the LXXLL motif (*i*, *i*+3 and *i*+4 residues), backbone modifications of the terphenyl scaffold were necessary[94]. Hamilton *et al.* had previously developed a biaryl scaffold that closely mimics the projections of *i*, *i*+3 and *i*+4 residues (Figure 20B)[43]. With the expectation to improve the water solubility and bioavailability, they have then designed and synthesized substituted pyridylpyridones. Fluorescence polarization assays indicated micromolar range affinities of compounds for ER, with in particular **I.29** that imitates quite well most of the LXXLL motif, with a K_i of 16 µM or the best compound **I.30** of this series exhibiting a K_i of 4.2 µM underlining here that the introduction of the naphtyl group on the scaffold improved the recognition with the surface of ER. Altogether, these results indicate that bis-heteroaryl scaffold-containing mimicries are able to reproduce the α-helical motif LXXLL of coactivator proteins but there is still room for improving the affinity to ER by for instance performing detailed structure-activity relationship studies.

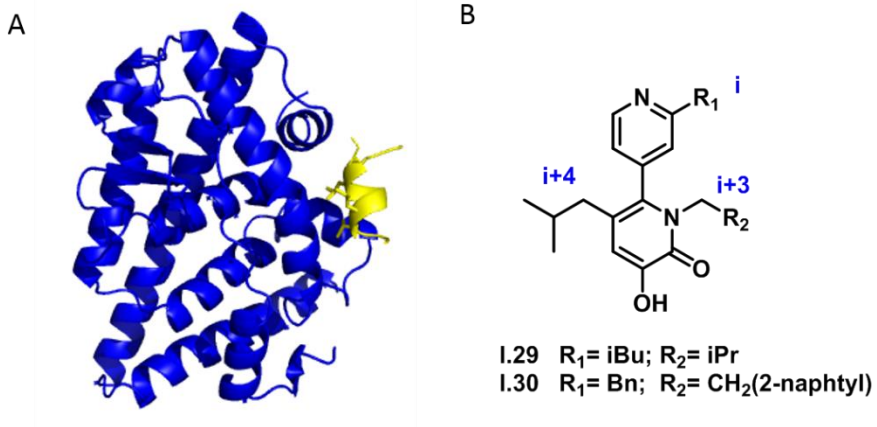
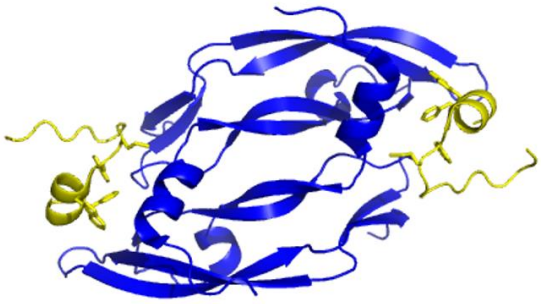


Figure 20: X-ray crystal structure of ER ligand binding domain in complex with peptide GRIP1 coactivator and trisubstituted biaryl scaffold developed by Hamilton and coll. A) X-ray crystal structure of ER ligand binding domain (blue) in complex with peptide GRIP1 coactivator α-helical peptide (yellow) (PDB entry: 3ERD) B) Chemical structure of pyridylpyridone derivatives I.29 and I.30 mimicking *i*, *i*+3 and *i*+4 residues of the α-helix.

Vascular endothelial growth factor (VEGF) plays an important role in the tumor angiogenesis[95]. Its activity is mediated by receptor tyrosine kinases VEGFR-1 and VEGFR-2. It is an important therapeutic target and several antagonists have been developed in anticancer therapies[96,97]. In the field of foldamers, substitutions of α amino acid to β^3-residues have been introduced by Gellman and coworkers in a 19-mer cyclic peptide, binding to $VEGF_{9-108}$ homodimer[98]. This 19-mer disulfide-constrained (C5-C15) peptide I.32 has been selected by phage display for the inhibition of the binding of VEGF with extracellular domains 1-3 of VEGFR-2 and a crystal structure have been obtained with next-best inhibitor **I.31** (Figure 21A)[99]. To realize the local α→β^3 replacement, they actually used a 19-mer cyclic peptide where the Met_{10} was replaced by norleucine to avoid sulfoxide formation (I.33). Hybrids α/β were synthesized and showed significant sub-micromolar affinities to VEGF (in the same range than the native peptide) and this despite a non-classical helix surface recognition. Additionally, a constrained β-amino acid residue Z (aminopyrrolidine carboxyl) was introduced into the sequence as previous studies showed that the introduction of constrained β-residues can favor the helical folding and have a positive effect on the affinity[100]. Furthermore, they demonstrated with the use of proteinase K, the resistance to proteolysis of α/β-peptides compared to the reference cyclic peptide with I.35 presenting a 15-fold higher stability than I.33 (Figure 21B). These results unveiled that foldamer-based design can be employed to mimic a polypeptide even with a discrete irregular conformation which is quite important because most protein recognition surface present irregularities. They next successfully applied this strategy to the Z-domain, a polypeptide composed of three α-helices[101]. These results will be described in the following section.

A



B

																			K_i (μM)	$t_{1/2}$ (min)
					S-S															
I.31	G	C	N	E	C	D	I	A	R	M	W	E	W	E	C	F	E	R	L	NH_2	0,6	
I.32	V	E	P	N	C	D	I	H	V	M	W	E	W	E	C	F	E	R	L	NH_2	0,07	
I.33	V	E	P	N	C	D	I	H	V	^{n}L	W	E	W	E	C	F	E	R	L	NH_2	0,06	1,6
I.34	V	E	P	N	C	D	I	H	V	^{n}L	W	E	W	E	C	F	Z	R	L	NH_2	0,26	
I.35	V	E	P	N	C	D	I	H	V	^{n}L	W	E	W	E	C	F	Z	R	L	NH_2	1,6	24

Z=

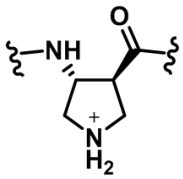


Figure 21: X-ray crystal structure of VEGF in complex with I.31 and sequence of inhibitors of the interaction between VEGF and VEGFR-2. A) VEGF9-108 homodimer (blue) bound to the peptide **I.31** (yellow) (PDB: 1KAT). B) Sequences of α-peptides and α/β-peptides with single-letter code for α-amino acids and ^{n}L for norleucine. β-residues are highlighted in green.

All these foldamers used as α-helical mimetics for PPI inhibition highlighted the potential of foldamer architectures for modern anticancer therapy approaches. Significant challenges remain before to envision possible clinical applications in particular by improving affinity and specificity for the protein targeted as well cell penetration and enzyme stability. After this section of foldamers to efficiently mimic the secondary structure of proteins we will now focus on their ability to mimic their tertiary structure.

2. Helical tertiary structure motifs as modulators of PPIs

As reported in this chapter with the aim to inhibit or modulate PPIs, various classes of modulators designed to mimic the secondary structures of one of the two partners have been identified. However, many biological functions depend on events that occur between highly ordered structures (mainly tertiary or quaternary folds). Therefore, mimicking tertiary structures became particularly challenging for protein interfaces involving more complicated modes of binding. In most of the cases, hot spots are indeed localized on multiple secondary structure motifs that have to be reproduced with synthetic mimics. Small molecules are not the best candidates for these big interfaces because of the large and irregular surfaces they involve during binding. Large polypeptides (such as antibodies) could target these protein surfaces, but their costly production, their proteolytic degradation and their low stability are major disadvantages to their use as therapeutics. Foldamers used as tertiary structure mimicries could be a really good alternative to circumvent these drawbacks. Until now only few studies have focused on the design, synthesis and analysis of foldamers able to mimic the tertiary structure of small proteins[102–106].

To guide the design of PPI inhibitors, scientists have developed computational tools to identify and study critical interfaces mediated by α-helical tertiary structures. Arora and coworkers have sought protein complexes mediated by α-helices in the Protein Data Bank and have developed a database of α-helical dimers involved PPIs (DippDB)[107]. To construct DippDB they applied filters to a previous database originally conceived for single helical interfaces, HippDB[108]. In this study, they focused on the determination of potential PPI targets where a single α-helix mimetic may not be effective and sufficient to promote the inhibition. Coiled coils are helical dimers, where the strands can be associated either in an antiparallel or parallel orientation. They are often encountered in PPIs and mediate several cellular functions[109,110]: Arora and collab. have identified 523 interactions implicated in biological process that require an α-helix dimer mimetic to activate the pathway[107]. Coiled coils are heptad repeats with important hydrophobic residues at positions a and d whereas ionic residues are found at positions e and g. Their interfaces have been classified into three categories: case 1) helical dimer that interacts with a single partner protein through one predominant face; case 2) helical dimer that interacts with two protein partners through two faces and case 3) helical dimer that forms across an interface between partner proteins[107]. The analysis of biophysical properties of these coiled-coil interfaces showed that hot spot residues could be distributed over multiple helices but are localized over a compact area meaning that they could be mimicked by medium-sized molecules or miniproteins. The database thus developed allows a classification of all these interfaces with their PDB identifier codes as well as

the energetic contributions of the hot spot residues found at these interfaces. Over the last years, several examples of tertiary structure mimetics used as PPI inhibitors have emerged and illustrate an approach that can be considered as being highly promising.

Regarding the synthesis, Arora and coworkers described the firsts an approach to modulate PPIs by developing secondary and tertiary protein domain mimics[111]. DippDB was employed to select the best designs and the authors focused on the most common tertiary structure encountered in Nature: the coiled-coils. While α-helices have hot spots with a regular distribution about every seven residues (two helical turns), in case of the α-helix dimers, the distribution of hot spots is approximately 13 residues[107]. In order to mimic this tertiary structure, salt bridge surrogates (SBS) have been applied as new backbones to stabilize original coiled coil sequences by replacing a non-covalent bond by a covalent one with the installation of a bis-triazole linkage, resembling to a salt bridge.

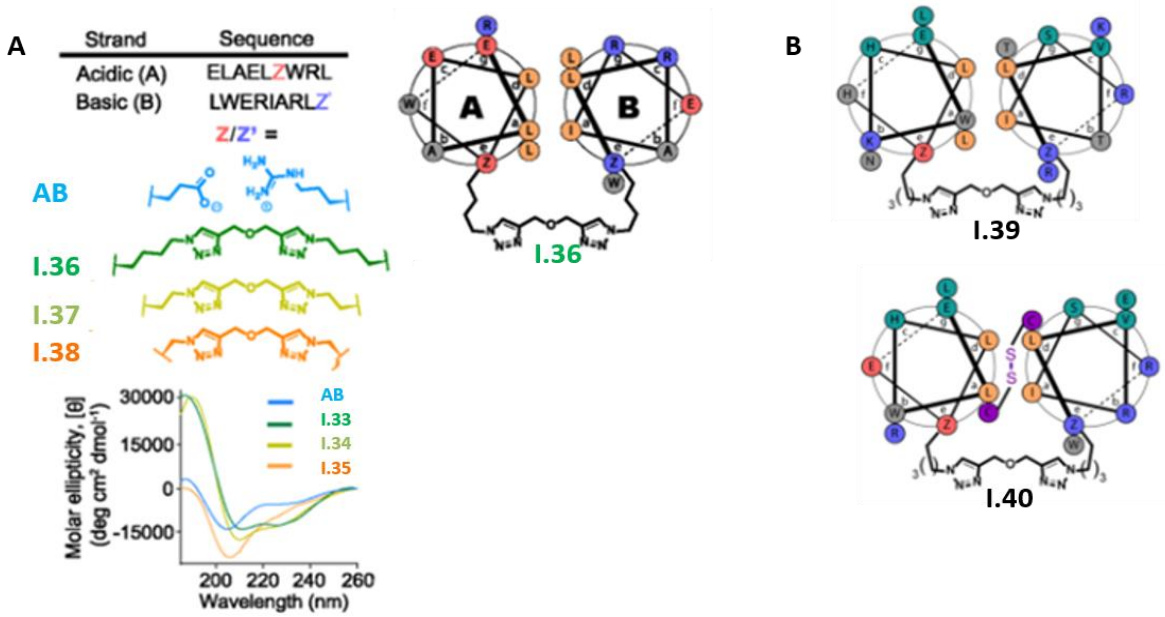


Figure 22: Design of salt bridge surrogate helix dimers as PPI inhibitor. A) Design of cross-linked helix dimers by replacement of an interchain ionic contact with bis-triazole linkers. Bis-triazole linkers of varying lengths resulting from azidolalanine, azidohomoalanine, and azidolysine residues were incorporated at coiled coil positions e/e' to obtain dimers **I.33**, **I.34** and **I.35** respectively. At the bottom, CD spectra of **I.33**-**I.35** in 50 mM aqueous KF, pH 7.4. On the right, helical wheel representation of **I.33**. B) Helical wheel representations depicting sequences of **I.36** and **I.37** to mimic NHR2. (Adapted from Wuo *et al.,* JACS 2015)

They first designed minimal coiled-coil mimics model sequence with a hydrophobic interface and intra and interhelical salt bridges leading to peptides A (Ac-ELAELEWRL-NH_2) and B (Ac-LWERIARLR-NH_2)[112]. As this designed coiled coil did not spontaneously assemble in aqueous solution, they replaced an interhelical ionic bond by a covalent bond to stabilize short helix dimers. Several disulfide designs have been considered as well as bis-triazole bridges of varying lengths inserted at e/e' positions to replace a weak interhelical ionic bond (**I.36**-**I.38**). As the relative helicity of peptides is usually determined with the mean ellipticity at 222 nm, CD analysis

allowed the determination of I.36, derived from azidolysine, as the more helical. This result was confirmed by NMR experiments, in particular total correlation spectroscopy (TOCSY) and nuclear Overhausser effect spectroscopy (NOESY) that indicated a helical tertiary structure.

This approach has then been applied to constrain an antiparallel coiled-coil present in the Nervy homology 2 (NHR2) domain of AMLI-ETO involved in leukemogenesis [112]. This NHR2 domain mediates the oligomerization of AMLI-ETO required for the interaction between NHR2 domain and the NHR2-binding motif (N2B) of E-proteins that occurs through an antiparallel coiled coil[113]. Azidolysine residues have been inserted at e/e' positions of the native sequence of NHR2 leading to compound **I.39**. However, CD experiments showed a nonhelical conformation that could be due to a lack of stabilization of the hydrophobic part. An interhelical disulfide bridge has been added distal from the triazole bridge, leading to I.40. CD analysis showed a significant improvement of the helicity in comparison with I.39. FP assays allowed the determination of affinity for the binding of dimers to N2B and the best value was obtained for I.40 (K_d=53± 20 µM whereas for the NHR2 peptide, K_d=356± 90 µM) which highlighted the importance of the conformation for the molecular recognition.

In another study, Gellman, Horne and coworkers designed α/β-peptides to reproduce the folding of a polypeptide deriving from HIV protein gp41[100]. This protein is localized in the HIV membrane and plays a key role in the viral infection cycle. Indeed, it mediates viral envelope-host cell membrane fusion during HIV cell entry by insertion into the host membrane[114]. The trimeric gp41 undergoes then a structural change with the formation of an antiparallel six-helix bundle[115]. Several α-peptides based on gp41 *N*-terminal heptad repeat (NHR) domain like **I.41** or *C*-heptad repeat (CHR) domain like **I.42**, have been developed as anti-HIV agents and one of them, the drug enfuvirtide (*Fuzeon* by Roche), is currently used clinically[116]. However, in a similar way than for the α-peptides, *in vivo* this drug is quickly degraded by enzymes. In this context, α/β-peptides have been developed for their ability to mimic the structure and function of gp41 subunit and their higher resistance to proteolysis. To perform their design approach, Gellman, Horne and coworkers have first selected a mutated gp41 α-peptide derived from CHR region (**I.43**) having 50% of modification compared to gp41 CHR domain, and by combining a series of Ala substitutions at various positions (*i.e*. Glu3, Glu7, Asn9, Thr12, Ser13, Ser17, Glu21, Ser22, Gln31, Glu32) and a series of *i*, *i*+4 salt bridges (*i.e*. between Glu and Arg at positions 2,6; 10,14 and 16,20; and between Glu and Lys at position 24,28) that enhance the helix component of the resulting α-peptide (Figure 23). Next, they performed local α→β residue substitutions according to the "sequence-based design" method described earlier in this chapter (Figure 23) (see chapter A.1.b). They first performed a replacement by applying their ααβααβ pattern allowing the alignment of β^3-residues on one face of the helix (**I.44**), but FP assays revealed a poor affinity for gp41. Then, they synthesized chimeric α/β-peptides exhibiting a pure α-segment at the N-terminus and an α/β-segment at the C terminus (**I.45** and **I.46**). **I.45** exhibited a very high affinity for gp41 while **I.46** showed a loss of affinity (K_i<0.2 nM and 15 µM respectively), illustrating here that the size and the localization of the inserted α/β segments matter. Cyclic β-residues able to constrain the backbone were incorporated in order to enhance folding propensity improving also the resistance to proteolysis. The high affinity obtained confirmed the utility of backbone stiffening as observed

for **I.47** with K_i=0.3 nM. CD experiments have been performed with 1:1 mixtures of **I.41** and **CHR** analogues. As for the α-peptide **I.43,** α/β-analogues **I.45** and **I.47** showed an induction of helicity, which is consistent with six-helix bundle formation.

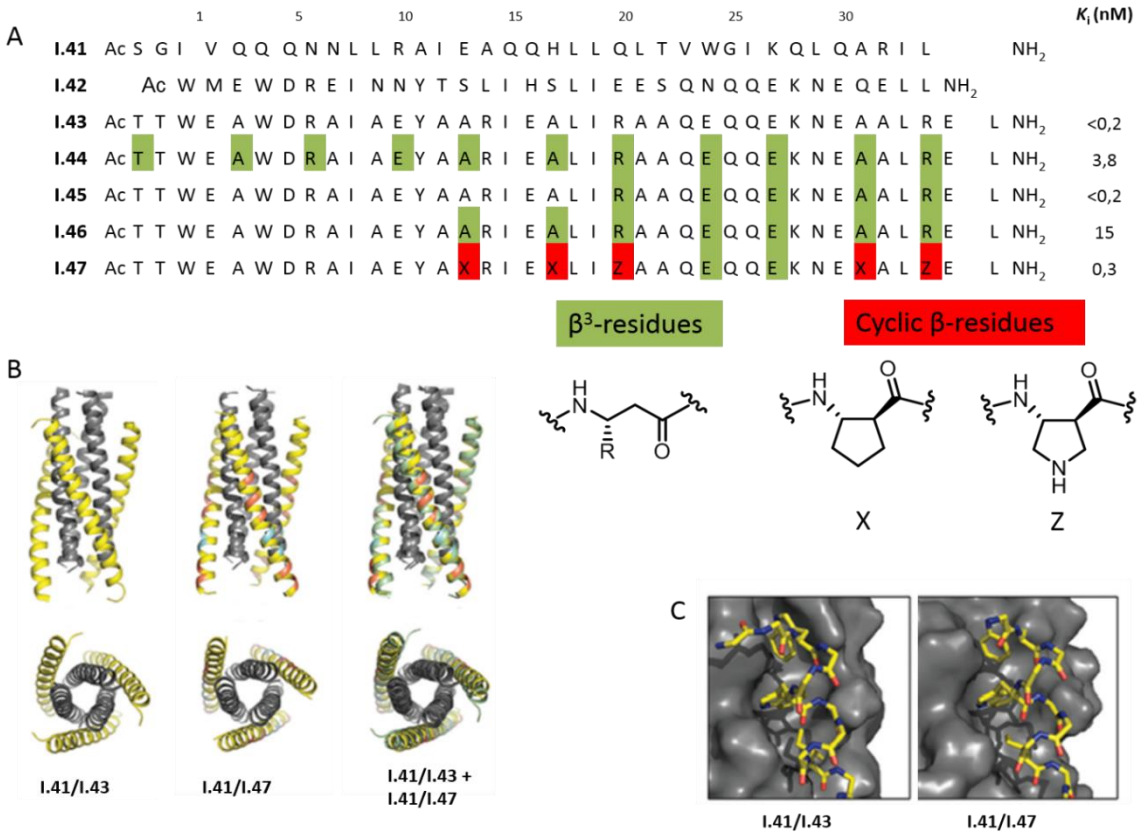


Figure 23: Sequences α-peptide and β³-analogs derived from HIV gp41 with crystal structures A) Structure of the α-peptide and α/β-peptides analogs B) Crystal structures of six-helix bundles formed by NHR α-peptide **I.41** (grey) in complex with α-peptide **I.43** (yellow) and α/β-peptide **I.47** (colored by residue with α in yellow, β³ in blue and cyclic in red) with a view from the side and down the superhelical axis. On the right, superimposition of **I.43** and **I.47** in complex with **I.41** C) **I.43** and **I.47** at the surface of **I.41** (grey). (Taken from Horne et al., PNAS, 2009)

Furthermore, X-ray crystal structures have been obtained for **I.41** in complex with **I.43** and **I.47** and in both cases the crystal structures showed that six-helix bundles obtained are essentially similar to the native CHR sequence. However, at the C-terminus, the two bundles are very close to **I.41** trimer whereas at *N*-terminus they diverge. X-ray crystal structures of both CHR analogs alone revealed a parallel trimeric helix bundle. Overall, this study showed that a long α-helix can be structurally and functionally reproduced by designed foldamers.

Concurrently, following the pioneered work of R. Raines[117–119], Horne and coworkers have developed a general design strategy to mimic the tertiary folding of proteins with unnatural oligomers[120]. They choose the B1 domain of *Streptococcal* protein G (GB1, **I.48**) as a model system because of its well-defined tertiary structure involving all common secondary structures (helix, loop, sheet and turn). They performed backbone substitutions in the native sequence by

introducing several unnatural residues such as β^3-residues, D-α-residue, C_α-methyl and *N*-methyl residues. Various sequences were prepared by individually mutating one secondary structure (helix, loop, sheet or turn) with unnatural backbones (Figure 24). First, the $\alpha \rightarrow \beta^3$ substitutions occurred in the helical part of GB1 on residues not directly involved in hydrophobic core packing (**I.49**). This substitution has also been introduced in the two loops of GB1 by replacing Val_{21} and Asp_{40} by β^3-residues (**I.50**). Another sequence illustrates well the structural consequences of replacing two solvent-exposed residues Glu15 and Trp43 by *N*-Methyl-α-residues in the two sheet parts of GB1 (**I.51**). CD analysis showed a curve similar to that of the native peptide GB1 that confirmed its mimicry, but the midpoint of the thermal transition (T_m) was below the value of the native GB1 (T_m=75.6 °C and 81.4°C respectively). These results illustrate that the destabilization of the tertiary structure occurs even by mutating residues located at solvent-exposed positions. Then, they replaced α-amino acids in the turns of GB1 with a C_α-methyl-α-residue (Aib) in one turn and with D-α-residue in the other turn (**I.52**). Finally, they combined all these modifications within one sequence (**I.53**). They performed CD thermal denaturation experiments to determine thermodynamic data on all these modified sequences. The midpoints of the T_m obtained showed a loss of folding for all these modified compounds, in particular, a huge destabilization is observed for **I.50** with a T_m of only 31°C. The best tolerated mutations were obtained when modification occurred in the loops of GB1 (**I.50**) with a T_m value of 77.6°C underlining here that this localization is the most favorable for modification without a huge impact on the tertiary folding. To note, X-ray crystal structures were obtained for compounds **I.49**, **I.50** and **I.52** and highlighted that the tertiary folding of these three protein analogues is very similar to that of wild-type GB1.

Consequently, Horne *et al.* demonstrated also that the tertiary structure of a protein could be mimicked with unnatural backbones and that helix, loop, sheet and turn can be individually modified at strategic positions without compromising the overall folding[121]. They recently applied this method to design and synthesize foldamer zinc finger domain mimics with the local introduction of a huge variety of unnatural backbones (such as β^3-residue, *N*-Me-α amino acid or Aib for example)[122]. The peptide **I.54** (Figure 24C) is derived from the third finger of specificity factor known for its folding and its metal-binding environment. The incorporation of β^3-residues and *N*-Methyl α amino acid was done for analogues **I.55** and **I.56** with the incorporation of Aib also for the last one. These analogues showed a good mimicry of this new tertiary structure highlighting good design principles. This study may allow a better understanding of the structure and function induced by these backbone modifications[122].

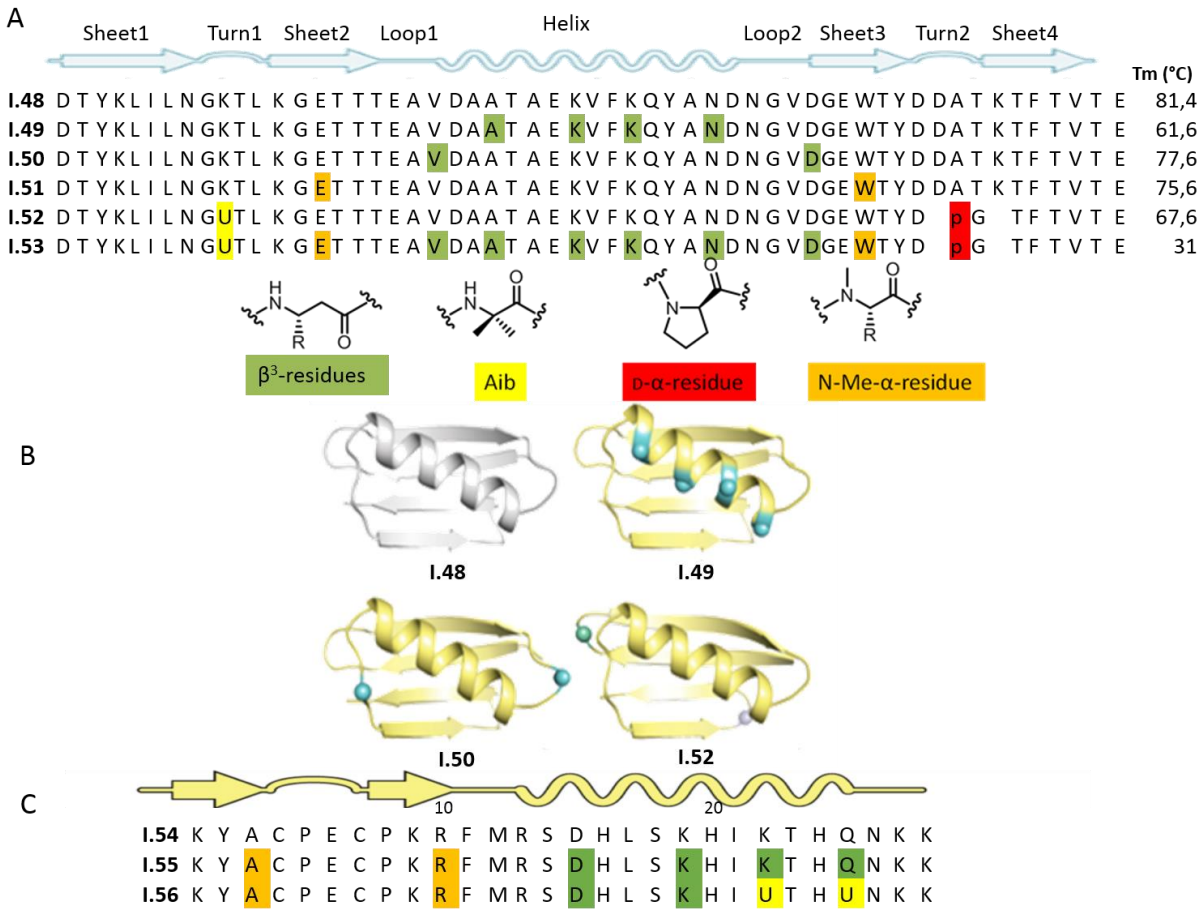


Figure 24 : Sequences and crystal structure of compounds for tertiary structure mimicry A) Sequences and structure map of wild-type GB1 (**I.48**) and backbone-modification analogues (**I.49-I.53**). Chemical structures of modified backbones are also represented. B) Crystal structures of wild-type GB1 (**I.48**) and three analogues (**I.49**, **I.50** and **I.52**) with modified backbones depicted in blue (adapted from Reinert, *J. Am. Chem. Soc.* 2013) C) Primary sequence of peptide **I.54** derived from the third finger of specificity factor 1 and backbone-modification analogues **I.55** and **I.56**

At the end of the previous sub-chapter, we have introduced the work of Gellman and coworkers on α/β-peptides deriving from the three-helix bundle "Z-domain scaffold"[101]. Indeed, Z-domains have been used as protein scaffolds for their capacity to specifically bind to a variety of proteins[123]. The "Z-domain" or "affibody" scaffold is derived from the B-domain in the immunoglobulin binding region of the staphylococcal protein A and has been widely studied[124]. The parent Z-domain, that is composed of 58 residues, interacts with Fc (a portion of IgG, the natural binding partner of protein A) through a large surface of 600 Å^2 formed by helices 1 and 2[125]. Z-domain derivatives have been designed to bind to a variety of protein targets such as VEGF or tumor necrosis factor-α (TNFα)[126]. α/β-peptides were designed to mimic Z-VEGF domain, a peptide composed of 59 amino acids obtained by phage display that binds to the receptor-binding region of VEGF[126]. As residues of Z-VEGF involved in the interaction are localized on helices 1 and 2, a truncation by removing helix 3 was performed to simplify chemical synthesis. Indeed, this truncation had already been

reported by Wells and coworkers with helices 1 and 2 retaining a high affinity for VEGF[127,128]. However, CD experiments suggested that helix 3 is important to stabilize the tertiary structure required for the binding[129]. Gellman and coworkers performed cysteine substitutions of His10 and Pro39 in VEGF sequence in order to create a disulfide bridge and to stabilize the structure (α-VEGF-2, I.57)[101]. The inhibitory constant (K_I) for **I.57**, determined by competition fluorescence polarization assays was in the same range (0.40 μM) than Z-VEGF (0.41 μM). This bridged-peptide was used as starting point to do local α→β replacements in order to bind tightly to VEGF and to improve proteolysis resistance. β-amino acids encompassing a five-membered ring constraint (X and Z in Figure 4) were used to substitute α-residues not localized on the α-helix binding face and several α/β–peptide oligomers showed significant affinity for VEGF. However, it appeared that some regions did not tolerate the introduction of cyclic-β-residues, such as the surrounding area of the central loop, leading to the installation of Aib-type residues in this region. Furthermore, β^3-residues were used to replace α-amino acids in the N-terminal segment and α/β-**I.58** exhibited a better K_I value than Z-VEGF (Figure 4) probably due to the helix stabilization. The impact on biostability of the introduction of unnatural backbones has been evaluated by performing degradation studies in the presence of proteinase K. While Z-VEGF presents a half-life of only 1.6 min and its bridged version **I.57** even faster ($T_{1/2}$= 0.20 min), the two-helix α/β-peptide analogue **I.58** showed a half-life 290-fold higher than **I.57** ($T_{1/2}$= 59 min). These results illustrate that the introduction of unnatural backbones in an α-peptide sequence can considerably improve its overall biostability. This strategy was then applied to Z-domain derivatives binding to other protein targets, leading to biostable α/β-peptide analogues of many Z-domain agents[101]. The authors have indeed extended the introduction of these unnatural backbones into α-peptides to mimic the first two helices of Z-IgG that binds the Fc part of IgG (**I.53**) and the Z-domain that binds the tumor necrosis factor α (TNFα) (**I.54**). Two-helix α/β-peptide analogues were tailored to bind these three targets and the specificity was conserved despite the unnatural backbones. ELISA tests showed that the binding of α/β-peptide analogues was similar to that of the native peptides illustrating the versatility of the Z-domain.

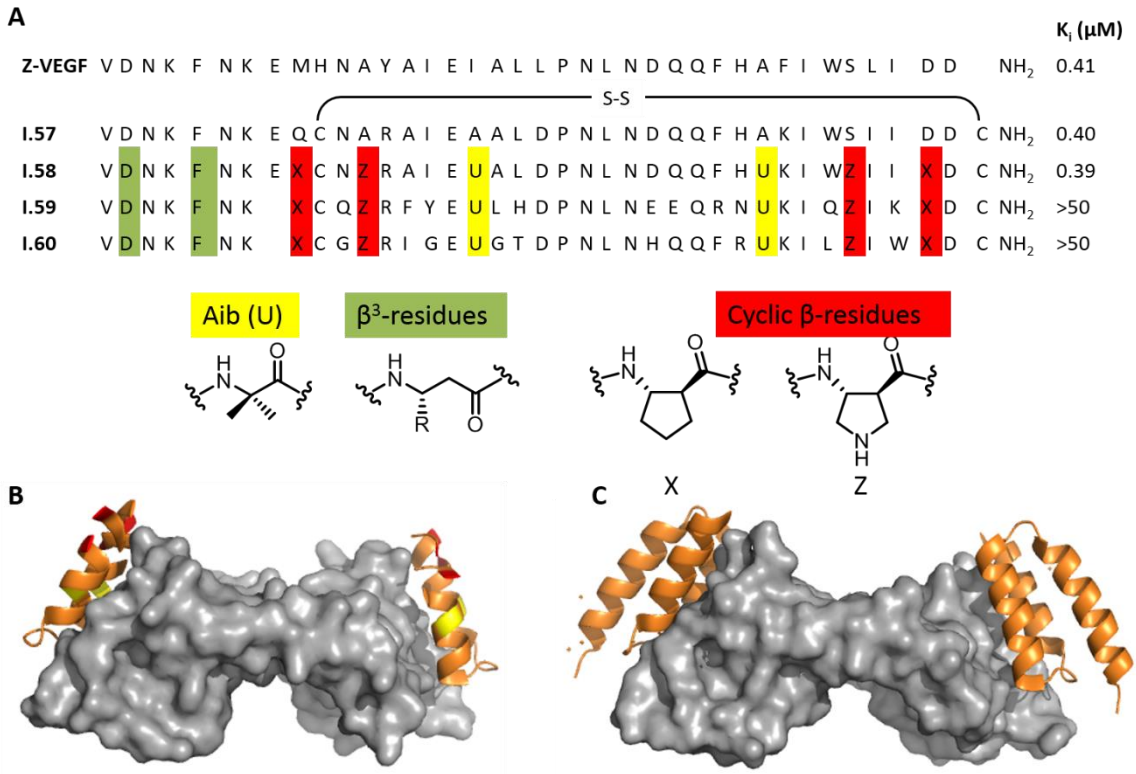


Figure 25: Presentation of Z-VEGF$_{1-38}$ α and α/β-peptides A) Primary sequences of α and α/β-peptides with non-natural peptides in color with their chemical structure. B) Crystal structure of **I.58** (orange) in complex with VEGF (grey) (PDB 4WPB). C) Crystal structure of Z-VEGF (orange) in complex with VEGF (grey) (PDB 3S1K).

In addition, the X-ray crystal structure of **I.52** in complex with VEGF was obtained at 3.1Å (Figure 25B PDB:4WPB) and revealed that I**.52** occupies the same binding site as Z-VEGF with the expected helix-loop-helix conformation. This study highlights the extension of α/β-peptides mimicking single helices to their use to mimic multi helical tertiary structure.

All these studies show the high potential of foldamers as mimics of helical secondary and tertiary structures of peptides and proteins, making them good candidates for the inhibition of PPIs. Their versatility is a huge advantage for protein-like recognition surfaces but several improvements are still necessary to target more complex tertiary structures and to improve the affinity of foldamers for their target proteins. However, their ability to penetrate cell for intracellular target remains particularly challenging.

3. To reach intracellular target with cell-penetrating foldamer

To be considered as potential therapeutic drugs, foldamers targeting an intracellular target must be able to cross the cell membrane. Most often, as for polypeptides and oligonucleotides, their low membrane permeability is a huge limitation for their application as innovative and modern therapeutics. During the last decades, scientists have been focused on seeking solutions to perform intracellular transport of oligomers. The discovery of a positively charged peptide deriving from the homeodomain of Antennapedia with ability of translocation across the plasma

membrane as well as the intracellular transportation of oligomers by conjugation has opened new possibilities[130]. Several cell-penetrating peptides (CPPs) have been developed in order to transport molecule of interest into targeted cells. Several of these CPPs such as HIV-1 TAT[131] (Figure 26A) are rich in arginine residues the guanidinium headgroups facilitating their entry into cell[132]. Usually, CPPs are conjugated covalently or noncovalently to a cargo and an endocytosis pathway is required to promote cell-penetration. However, these complexes have low stability *in vivo* and chemical modifications could improve it. Scientists have handled new approaches with unnatural backbones including foldamers as new classes of molecular transporters. Several non-natural oligomers inspired by control peptides have been designed such as oligocarbamates, guanidine-rich peptoids, D-Arg oligomers[133], hybrids of α-peptide/β-peptoids[134] or L/D-octaarginine[135] and β-peptides[136,137]. These compounds mimic the vehicle function of corresponding Arg-rich α-peptides and are more resistant to proteolytic degradation. However, some compounds do not adopt a regular structure in solution but it does not appear to be required for crossing cell membranes[138].

Concurrently, Huc and coworkers have synthesized water soluble aromatic oligoamide foldamers which have been evaluated for their ability to penetrate into cell[139]. These foldamers were built with 8-amino-2-quinolinecarboxylic acid equipped with aminopropoxy side chains and fold as helices. These helices present a high stability, a smaller helical pitch than α-helices and are wider[140]. They determined that a minimum backbone length of eight residues with cationic side chains were necessary for cell penetration and water solubility. The octamer **I.61** showed a high activity as cell-penetrating foldamer (CPF) on several cell lines and a membrane translocation mechanism was assumed (Figure 26B).

Other foldamers have been directly developed to display cell-penetrating properties such as β-peptides developed by Schepartz and coworkers. They have explored different approaches in order to improve the cell permeability of β-peptides mimicking p53 α-peptide and able to bind HDM2 protein[141]. They inserted β^3-homoarginine residues on the "salt bridge" face of the α-helix. The cationic-patch insertion on this salt bridge face has shown its ability to bind equally to HDM2 and penetrate cells with the same efficiency than the validated CPP $(PRR)_3$ earlier developed by Daniels and Schepartz[142].

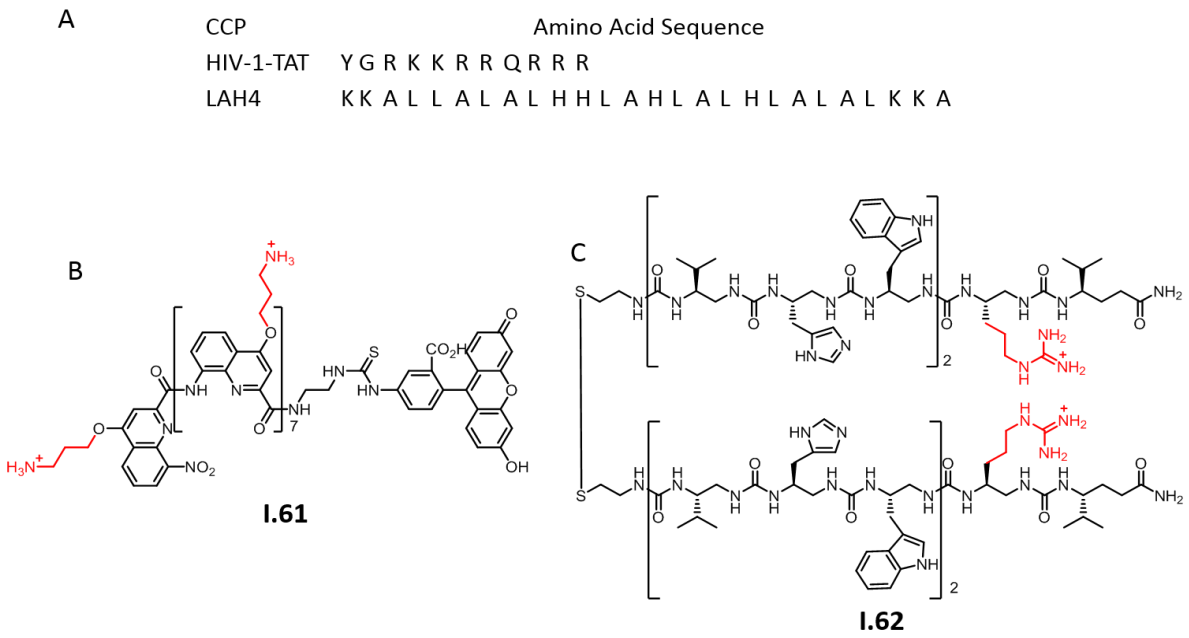


Figure 26: Amino acid sequences of CCPs and chemical structure of CPF. A) Amino acid sequences of two CPPs B) Chemical structure of an oligoamide CPF **I.61** C) Chemical structure of a dimer **I.62** able to do transfection

In the group, Guichard, Douat and coworkers have developed pH-responsive urea-based cell-penetrating foldamers (CPFs) for intracellular delivery of nucleic acids[143]. These CPFs are short aliphatic oligourea peptidomimetic, composed of ethylene diamine units linked through urea bonds, and known to exhibit a high helical propensity reminiscent of the α-helix[144–147]. The sequences of these urea-based CPFs were inspired by (1) a 8-mer antibacterial oligourea mimicking helical host-defense α-peptides[148,149] and (2) the amino acid composition of a cell-penetrating peptide (CPP) LAH4 (KKALLALALHHLAHLALHLALALKKA), a His-rich peptide exhibiting a high transfection ability[150]. Indeed, histidines are pH-responsive residues frequently used in CPPs as they can provide pH-dependent membrane permeation and endosomal buffering referred to as a proton-sponge effect[151]. Several short 8-mer oligourea sequences were designed with the key side chains of LAH4, in particular several His and Arg side chains were localized on the polar surface to increase DNA compaction. To improve the transfection efficacy of the best sequence obtained, the size of the oligomer has been increased in order to improve the assembly properties of the resulting CPF with the plasmid DNA of interest. The best compound, **I.56**, was obtained through thiol-mediated dimerization of this promising sequence and showed a transfection capability without any apparent cytotoxicity in cell (Figure 26).

B. Asf1: a new candidate target protein for anticancer therapy

The Anti-Silencing Function 1 (Asf1) protein is a central histone chaperone with numerous partners and functions. Molecular bases of chromatin structure and function will be presented to

define the biological interest of Asf1 protein, and its implication in human diseases, in particular in cancer, will be discussed in the following part.

1. DNA and chromatin: the support of genetic information must be compacted

The human genome contains 6 billion base pairs of DNA packaged into 23 pairs of chromosomes. Knowing that each base pair measure around 0.34 nanometers, one cell contains around 2 meters of DNA when fully extended. This genetic information must be compacted so that it can be contained in the nucleus of eukaryotic cells (around 10 µM of diameter) while specific segments of the genome need to be accessible to allow the cell to perform its nuclear function and maintain its survival. To organize this compaction and maintain the reading of the genetic information, specialized proteins that bind to and fold the DNA, provide higher levels of organization[152].

In eukaryotes, long linear DNA is associated with these specialized proteins that fold and pack the DNA thread into a more compact structure: the chromosome. The complex between these proteins called histones, and DNA is called chromatin (discovered in 1882 by Flemming). The chromatin can be found in two levels of organization. First, euchromatin is the active form of DNA compaction that make DNA accessible for translation. Then, heterochromatin that is a tightly pack form of DNA and considered as more “passive form”.

At the simplest level, chromatin is found as a double-stranded helical structure of DNA (Figure 27). The DNA is then associated with histone proteins and wraps around them to form complexes called nucleosomes. A linker histone, H1, is associated to the nucleosome to form the chromatosome, allowing the folding of nucleosomes into a 30 nm fiber. Fibers are then compressed and fold to produce the chromatid of chromosome.

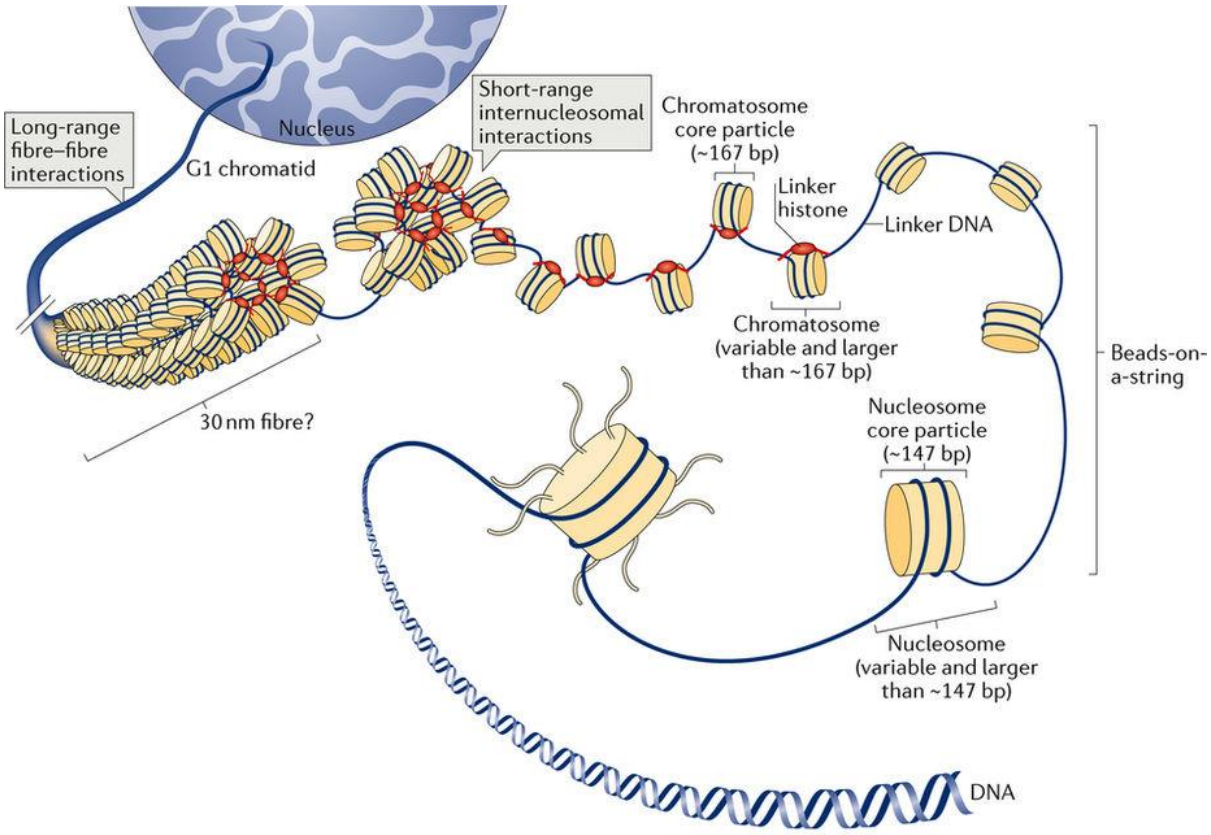


Figure 27: Multiple levels of chromatin folding. DNA compaction within the interphase nucleus occurs through a hierarchy of histone-dependent interactions, including the formation of the nucleosome core particle, strings of nucleosomes (bead-on-a-string arrangement), the chromatosome core particle and 30 nm fibres (the existence of these fibres is debatable in vivo) and the association of individual fibres, which eventually produces tertiary structures (taken from Fyodorov, *Nature reviews*, 2017)[153].

2. The nucleosome is the unit of chromatin

The nucleosome is the chromatin base unit and is repeated all along the chromatin fibers. First electron microscopy characterization of the chromatin structure in the 70's showed that nucleosome are spherical particles of 11 nm diameter that were referred as "beads on a string structure"[154]. The nucleosome unit is indeed connected to the adjacent one through a segment of DNA linker. Nucleosome architecture is composed of about 146 base pairs (bp) of DNA wrapped around an octameric assembly of histone proteins[155]. Histones are small basic, highly positively charged polypeptides with more than 20% of lysines and arginines in the primary sequence. Indeed, their positive charge maximize the interactions with DNA, which is negatively charged. They contain two distinct domains: the histone tail domain and the histone-fold domain. The latter consists of three α-helices (α1, α2 and α3) linked by two loops (L1 and L2) (Figure 28) that participate to the core assembling motif. The histone-fold domain is involved in histone-histone interactions by charge complementarity. Two copies of each histone, namely H2A, H2B, H3 and H4 form the central core histone octamer of nucleosome. H2A and H2B form one heterodimer

whereas H3 and H4 formed the second one. During the formation of the nucleosome, first, two H3-H4 dimers interact together with a head to head arrangement, through H3(α2-α3)-H3(α2-α3) four helix bundle leading to $(H3\text{-}H4)_2$ tetramer called "tetrasome".

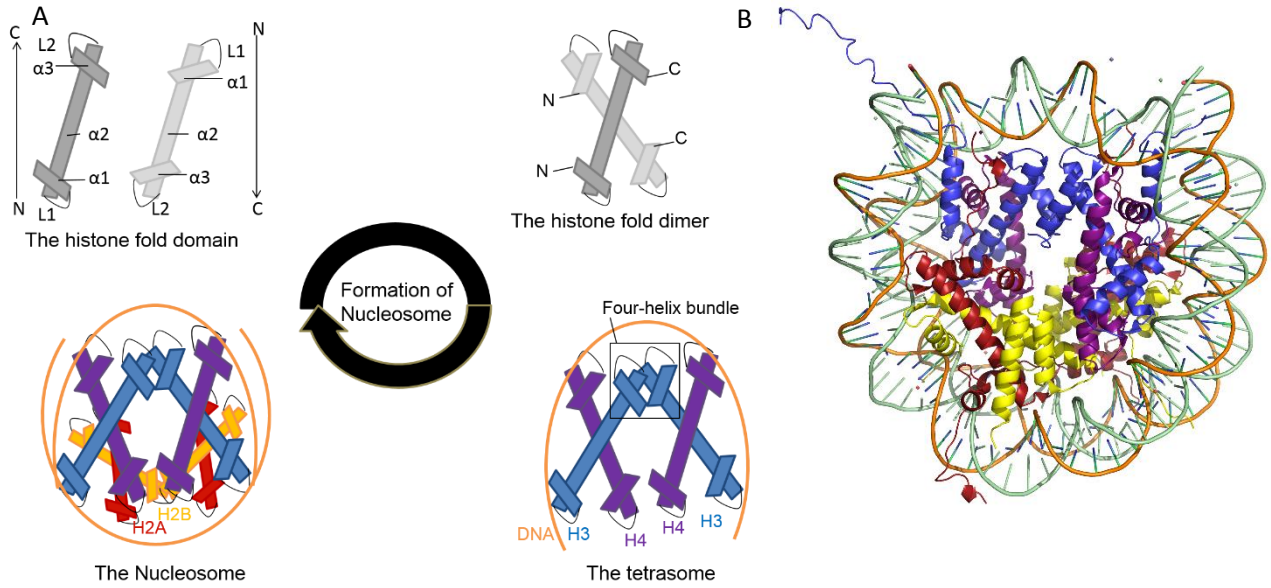


Figure 28: Formation of nucleosome. A) Step of the formation of nucleosome. First, the presentation of the histone fold domain. Second the formation of the dimer. Third, the presentation of the tetrasome with the tetramer $(H3\text{-}H4)_2$ bound to DNA. Finally, the scheme of the nucleosome with the addition of two dimers H2A-H2B. B) Crystal structure of the nucleosome at 2.8 Å with the 146 bp DNA phosphodiester backbones (orange and turquoise), and histone proteins (H3: blue, H4: purple, H2A: red and H2B: yellow) pdb:1AOI

The central position of DNA interacts with the tetrasome and initiates the nucleosome assembly. Finally, each H2A-H2B dimer associates with one side of the tetramer $(H3\text{-}H4)_2$ through a second four-helix bundle between H4(α2-α3) and H2B(α2-α3). The addition of two dimers H2A-H2B completes the assembly of the nucleosome by enrolling the remaining DNA at the entry and the exit of the nucleosome[156].

The histone tail domain is located at the N-terminal part of each core histone proteins and at C-terminal part of H2A histones. These domains are flexible and undefined polypeptides exposed to several regulatory post-translational modifications which play a role in nucleosome dynamics.

The first X-ray crystal structure characterization of the nucleosome was published by Luger et al. at 2.8Å resolution in 1997. It provided molecular insights into the supramolecular assembly between DNA and histones (Figure 28B).[7] Currently, numerous nucleosome core particle structures are found in the protein data bank (98 structures) with core histones from different species and different DNA sequences. The X-ray crystal structure shows that the 147 bp of DNA are wrapped in about 1.7 left-handed superhelix that surround the core histone. Moreover, the histone octamer complex presents several folded domains that are involved in the histone/DNA and histone/histone interactions. In particular, the loops H4 L2 and H3 L1 are interacting together through three hydrogen bonds.

During a DNA-dependent process, nucleosome must be disassembled in order to access free DNA and it should be reassembled when the process is completed to compact again chromatin. Histone proteins are essential for nucleosome assembly and its stabilization. However, histones must be handled upstream and downstream each DNA-dependent event what is achieved by specific proteins called histone chaperones.

3. Histone assembly for the formation of nucleosome requires histone chaperones

Histone chaperones are a group of highly acidic proteins that associate with histones and regulate nucleosome assembly[157]. They can be classified according to their specificity and selectivity for H2A-H2B or H3-H4 which results from specific structural features (Table 2). However, some histone chaperone like the histone Facilitates Chromatin Transcription (FACT) have the ability to bind both H2A-H2B and H3-H4 dimers[158,159]. On a structural point of view, there is low level of conservation among histone chaperones with structural folds used for their function and intrinsically disordered regions. However, they have an overall negative charge which gives them the ability to shield the excess positive charge of histones, which is essential to prevent the aggregation between DNA and histones[160].

Table 2: Selectivity and functions of histone chaperones[161]

Histone selectivity	Histone chaperones	Main functions
H3-H4	Asf1	Transport, Nucleosome assembly
	HIRA	Deposition factor independent of DNA synthesis
	N1/N2	Storage
	Spt6	Transcription initiation and elongation
	Rtt106	Heterochromatic silencing
	CAF1 complex	Deposition factor coupled to DNA synthesis (replication, repair)
	Hif1	Assist HAT
	FACT complex	Transcription elongation
H2A-H2B	Nucleoplasmin	Storage, transport, replication, transcription
	Nap1	Transport, transcription, replication
	Nucleolin	Transcription

Histone chaperones are involved in several functions, such as histone storage, post-translational modifications (PTMs), transport, histone turn over or nucleosome assembly. The chaperones Nucleosome Assembly Protein 1 (Nap1) and Anti Silencing Function 1 (Asf1) are transporting the histones H2A-H2B and H3-H4 dimers respectively, from the cytoplasm (where they are synthesized) to the nucleus with the help of transport factors[162–166]. Otherwise, a soluble pool of soluble histones must be stored to meet the needs of the cell during stress conditions. Nucleoplasmin (NPM2)[167] and Nuclear Autoantigenic Sperm Protein (NASP)[168], are chaperones acting as a reservoir of H2A-H2B and H3-H4 dimers respectively. The recycling of histones also requires histone chaperones, such as FACT and its co-chaperone the Minichromosome Maintenance Complex 2 (MCM2) which recover old $(H3\text{-}H4)_2$ tetramers[168]. Histone chaperones span also cellular processes such as DNA replication, transcription and DNA repair. Indeed, they are directly involved in histone deposition onto DNA during nucleosome assembly. H3-H4 dimers are first deposited involving specific histone chaperones followed by the deposit of H2A-H2B handled by other specific histone chaperones. After nucleosome assembly, a rapid exchange between nucleosomal H2A-H2B dimer and free one is performed by chaperone[169]. Malfunction in histone regulation can alter the barrier for cellular reprogramming, disrupt telomers and centromers functions and challenge DNA replication leading to numerous diseases. These protein chaperones are essential to assist histones with defined functions. One important chaperone, Asf1, is involved in different functions to assist histones H3-H4.

4. Asf1: a central histone H3-H4 protein chaperone involved in cancer

a) Asf1 exists as two paralogs in human

The histone chaperone Asf1 belongs to the replication-coupling assembly factor (RCAF), a protein complex that helps for the assembly of nucleosome onto the newly replicated DNA[170]. This protein chaperone is present in yeast and in numerous eukaryotic organisms with conserved functions such as the replication-coupled (RC) and replication-independent (RI) chromatin assembly pathways[171]. These common properties are due to a conserved N-terminal region of 156-amino acids whereas the C-terminal part could be highly divergent among species. A single isoform of Asf1 is present in yeast whereas mammals, in particular human, possess two isoforms Asf1a and Asf1b with 71% of similarity[172]. Human Asf1a and Asf1b share 61% identity with *Drosophila* Asf1 and 50% with yeast Asf1, illustrating the high conservation of Asf1 among species (Figure 29[172]). The alignment of the primary sequence shows that most of the conservation is localized on the N-terminal parts. Both isoforms Asf1a and Asf1b interact similarly with H3-H4 dimer as well as with the chromatin-assembly factor 1 (CAF1), a complex of three polypeptides (p48, p60 and p150) in the replication-coupled assembly pathway[171] and with MCM2, the helicase protein that unwired DNA ahead of the replication fork. Western-blot analyses have shown that Asf1a and Asf1b are expressed at the same levels during the cell cycle[172] illustrating that there are not restricted to the S-Phase (part of the cell cycle during which DNA is replicated). Even if both isoforms contribute to the replicative function and could maybe substitute each other, there is

nevertheless a distinct physiological distribution and regulation. For example, Asf1a interacts specifically with the histone cell cycle regulator (HIRA) which shows its involvement in replication-independent pathway[173] whereas Asf1b plays an essential role in cell proliferation.

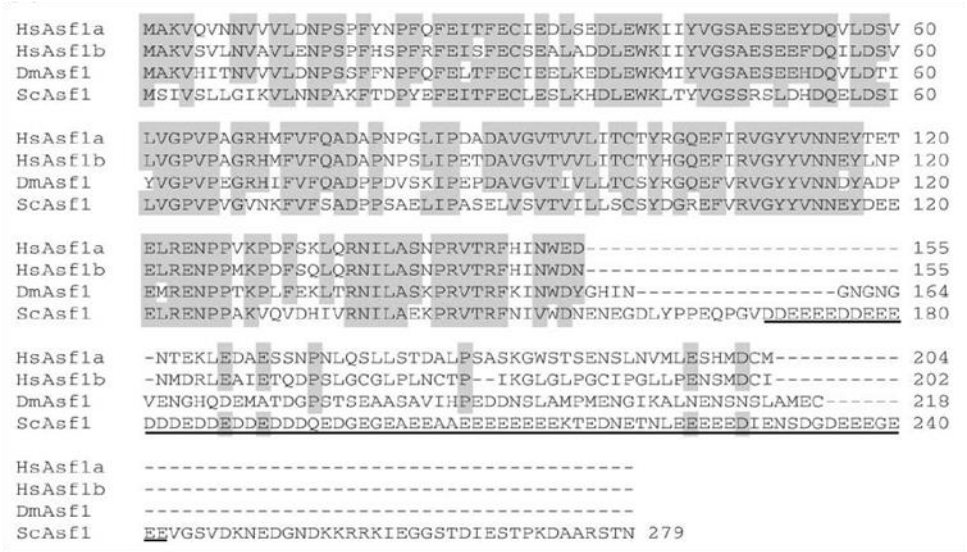


Figure 29: The alignment of Asf1 amino acid sequences. Human Asf1 are noted Hs, *Drosophilia melanogaster* (DmAsf1) and *Saccharomyces cerevisiae* (ScAsf1)[21]

With these two isoforms, Asf1 provide numerous functions and interacts with more than seven partners. Indeed, it has the ability to induce silencing upon overexpression, to bind acetylated histones H3 and H4, to cooperate with CAF1 for the assembly of the nucleosome onto newly replicated DNA[174], and plays a key role for DNA damage response and DNA repair by interacting with the DNA damage checkpoint kinase Rad53[175,176].

b) Asf1 handles histones H3-H4 during their cellular life

Asf1 is handling histone H3-H4 from their synthesis into the cytoplasm to their incorporation into the chromatin. This protects cells from potential cytotoxic effects due to an accumulation of free histones. H3-H4 dimers are first interacting with NASP and diacetylated by the histone actetyltransferase Hat1-RbAp46 complex (pRB-associated proteins p46). Dimers are then transferred to Asf1 in the cytoplasm, followed by the association with Importin-4 which facilitates the penetration into the nucleus[157]. NASP, RbAp46 and Asf1 bind dimers in different regions allowing the different binding to occur simultaneously, facilitating the transfer to Asf1[177]. Inside the nucleus, Asf1 delivers H3-H4 dimers to downstream chaperones such as CAF1 and HIRA, for nucleosome assembly. The protein codanin 1 is also found in cytosolic complex, interacting with Asf1 and negatively regulates histone delivery to CAF1 and HIRA (Figure 30A)[178]. Indeed, the codanin 1 binds to the same domain than these downstream chaperones.

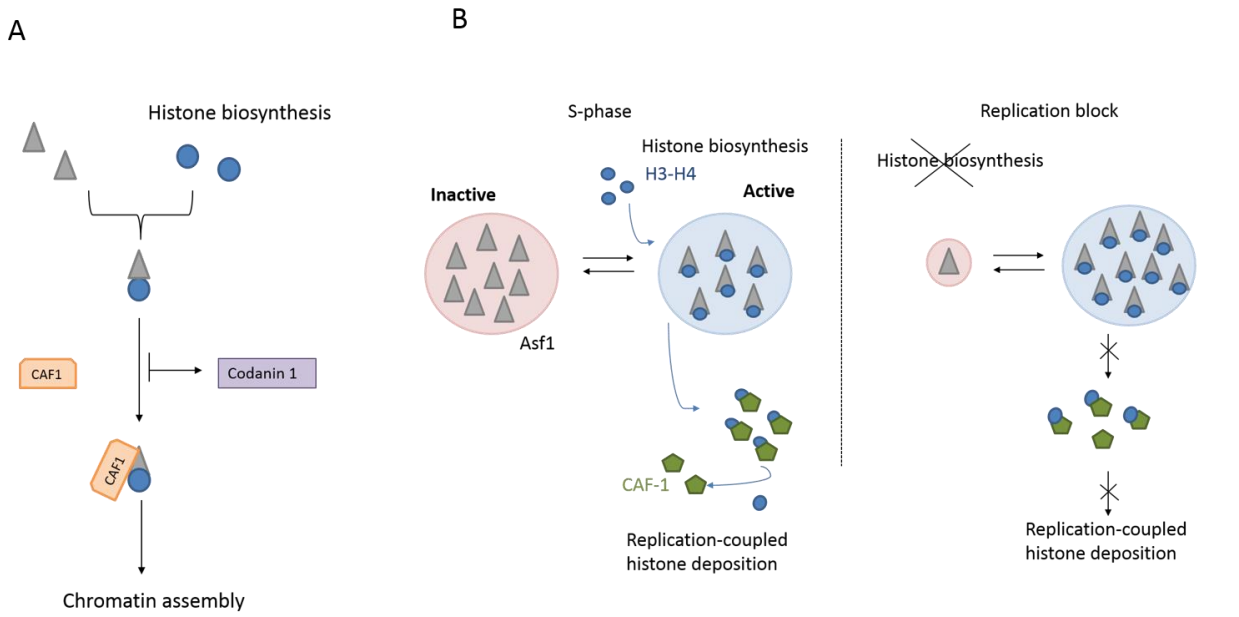


Figure 30: Scheme of Asf1 handling histones H3-H4. A) Model of codanin 1 as negative regulator of Asf1 function. B) Model of the balance between biosynthesis and usage of histones depending on the histone buffer function of Asf1

Moreover, Asf1 is involved in histone buffering and storage. It actually exists as an equilibrium between an active form and a pool of nonactive Asf1 (Figure 30B). When DNA replication is inhibited (as a consequence of replicational stress), the histone deposition is blocked and newly synthesized histones accumulate. Asf1 is then required for a checkpoint to form complexes with H3-H4 dimers[179]. Chaperone NASP which is involved in adjusting the soluble reservoir of H3-H4, may interact simultaneously with Asf1 to bind H3-H4 dimer and protect cells from potential cytotoxic effect. This suggests a balance between biosynthesis and usage of histones illustrating that Asf1 plays a key role to sequester the excess of histones into an active complex. It would allow immediate available histones for their deposition as soon as the DNA synthesis starts again. Then, Asf1 regulates the flow of S phase histones during replication stress.

c) Asf1 plays a central role in nucleosome assembly

During DNA replication two major processes affect chromatin structure: the disruption of pre-existing nucleosome located ahead of replication forks (parental histone segregation) and the new histones deposition (de novo assembly). These parental and newly synthesized histones are incorporated into nucleosomes during S-phase through a process called replication-coupled (RC) nucleosome assembly.

During the disassembly process, Asf1 may act as a histone acceptor to handle the evicted parental H3-H4 histone from nucleosome. These histones are randomly distributed to form new nucleosome downstream of the fork and could be mixed with new H3-H4 dimers. In RC nucleosome assembly, Asf1 synergizes with both MCM2 helicase and CAF1 for the recycling of parental histones H3-H4 and for the deposition of newly synthesized histones onto replicated DNA respectively[179,180]. Asf1 interacts directly with CAF1 through its largest subunit which brings

H3-H4 dimers close enough for their transfer onto DNA. As Asf1 also interacts with MCM2 helicase, the transfer is done at proximity of the replication fork. Indeed, MCM2, that interacts directly with DNA, may capture $(H3\text{-}H4)_2$ tetramer for the dissociation of the nucleosome (tetramer-to-dimers transition) and Asf1 could then separate the tetramer in two dimers. The formation of a complex between MCM2, Asf1 and H3-H4 dimer facilitates the access to replication fork[168,181,182] (Figure 31). The interaction between both chaperones occurs through the dimer H3-H4 as shown in the crystal structure of the complex Asf1-H3-H4-MCM2[183] (pdb: 5C3I).

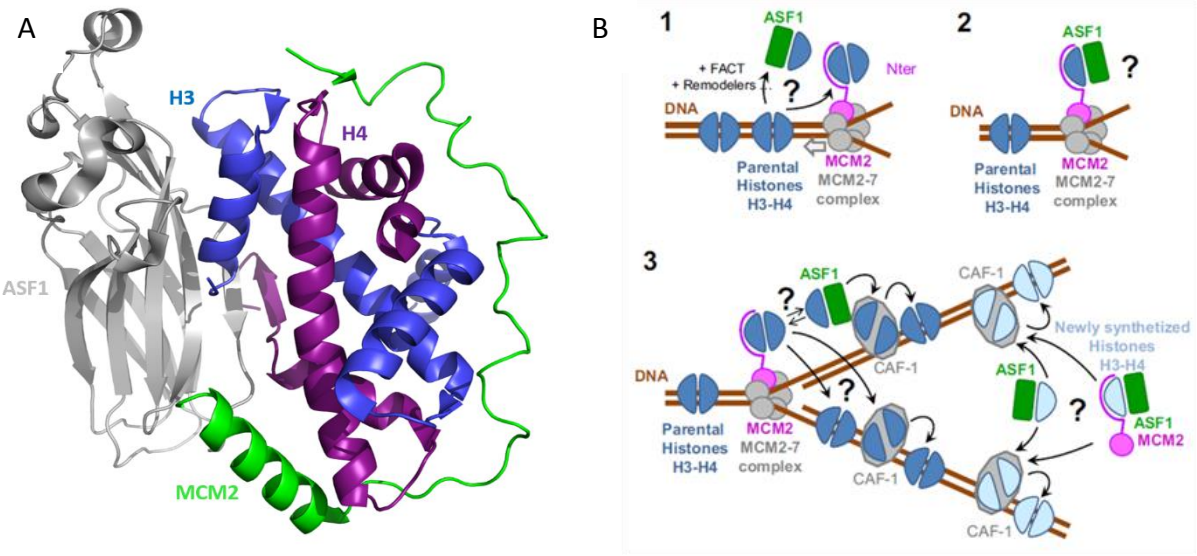


Figure 31: Structure and function of the complex Asf1-H3-H4-MCM2 A) Crystal structure of the complex with Asf1 in grey, histone H3 in blue histone H4 in deep purple and MCM2 in green at 3.5Å (pdb 5C3I) B) Models for the role of the Asf1-H3-H4-MCM2 complex in handling histones during replication. (1) After removal of H2A-H2B by chaperones (not represented), histone chaperones dedicated to H3-H4 could dissociate histones from DNA. The $(H3\text{-}H4)_2$-MCM2 ternary complex could then be transiently formed and protected from its re-association with DNA upstream of the replisome machinery. (2) The Asf1–$(H3\text{-}H4)_2$–MCM2 quaternary complex could constitute the next intermediate step for further protecting and destabilizing the $(H3\text{-}H4)_2$ complex. (3) Parental histone reassembly could proceed by different mechanisms; the tetramer captured by MCM2 the N-terminus of MCM2 could be directly deposited on DNA without splitting or this histone tetramer could be directly transferred to the assembly chaperone CAF-1. Alternatively, histone tetramers split by ASF1 could be reassembled by CAF-1 upon deposition on DNA. Besides, the human pathway dedicated to newly synthesized histones was also shown to involve ASF1 and CAF-1. The major histone fraction associated with MCM2 carries the specific modifications of newly synthesized histones suggesting that the quaternary complex could also transfer histones to CAF-1 for further assembly. All presented pathways remain hypothetical, may not be conserved in all species and may be restricted to some regions of the chromatin. (Adapted from Richet, *Nucl. Ac. Res.*, **2015**)

All these processes are also modulated by histone post-translational modifications (PTMs). Indeed, after binding newly synthesized H3-H4 dimers[184], Asf1 presents dimers for acetylation of Lysines such as Lys_{56} from H3 (by Rtt109-Vps75 complex)[185,186], before to transfer them to CAF1. This acetylation allows a higher binding affinity of CAF1 enhancing de novo nucleosome assembly (Figure 32A) [187].

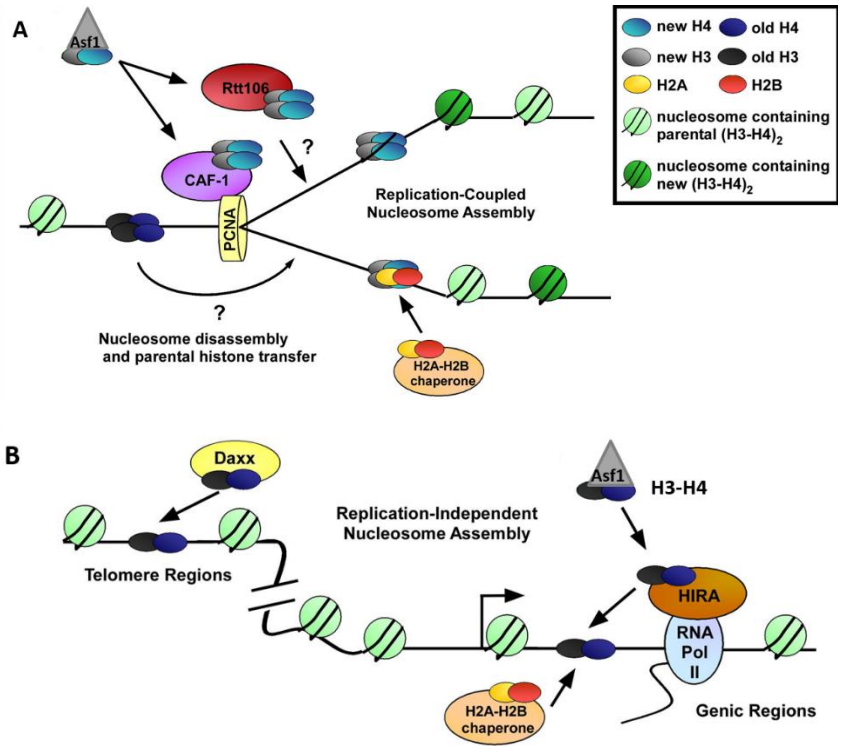


Figure 32: Histone chaperones are key regulators of replication-coupled and replication-independent nucleosome assembly A) Histone chaperones coordinate to regulate DNA replication-coupled nucleosome assembly. Once newly synthesized histone H3–H4 is imported into the nucleus, new H3–H4 of the Asf1-H3–H4 complex is transferred to CAF-1 and Rtt106 for $(H3–H4)_2$ formation and deposition onto newly synthesized DNA. Deposition onto replicated DNA depends, in part, on the interaction between CAF-1 and PCNA. Parental histones are also a source of histones for nucleosome assembly following DNA replication. B) HIRA and Daxx mediate replication-independent nucleosome assembly of H3.3–H4. In human cells, H3.3–H4 of the Asf1a-H3.3–H4 complex is transferred to HIRA for deposition of H3.3–H4 at genic regions, possibly through interactions with RNA polymerase II and dsDNA. Daxx facilitates deposition of H3.3–H4 at telomere regions, although mechanisms by which Daxx-mediated histone deposition is regulated are currently unclear.

Regarding DNA transcription and repair, these processes occur in a replication-independent (RI) manner [157]. Asf1 is the only chaperone involved in both replication-coupled and replication-independent nucleosome assemblies (Figure 32[188])[181]. During RI nucleosome assembly, Asf1 synergizes with HIRA and may promote heterochromatin gene silencing but the mechanism stills unknown. Indeed, in yeast, Asf1, HIRA and H3-H4 dimers form a complex that spreads along silenced domain to enhance transcriptional silencing. The complex between Asf1 and HIRA also interacts with Clr6 complex in order to promote deacetylation of histones which enforces silencing[189]. Finally, the complex affects the occupation and the position of nucleosome in heterochromatin by interacting with another complex called SHREC (HDAC repressor complex/Snf2). In human, only Asf1a interacts with HIRA and creates a facultative heterochromatin called senescence-associated heterochromatin foci (SAHF)[190]. Actually, the complex between Asf1a and HIRA causes condensed chromosome which allows the formation of SAHF and leads to nucleosomal density increase. The complex of Asf1-H3-H4-HIRA has been

solved recently, illustrating another structure of cochaperones targeting an H3-H4 dimer[191].
To sum up, Asf1 assists CAF1 during the assembly of new DNA and is essential for replication-independent chromatin assembly with the HIRA chaperone complex[171]. Asf1 mediates also chromatin disassembly from promoters during transcriptional activation as well as disassembly and reassembly of the chromatin during transcription elongation[192]. All non-DNA bound histones are interacting with Asf1, illustrating its fundamental role in the cell.

After having defined the function and the role played by Asf1 in the cellular process, we will focus on the structural information gathered over the years and on the mode of interaction of Asf1 with the couple H3-H4 histones.

d) Structural insights of the H3-H4 heterodimer in complex with Asf1

The structure of histones H3 and H4 within the nucleosome to understand how these histones interact together with DNA is now well-documented thanks to X-ray and NMR experiments. However, there is still only few structural information regarding the interaction of histones with their chaperones in the absence of DNA. The first X-ray crystal structure of a complex between histones and their chaperone that has been elucidated was that of yeast Asf1 in complex with H3-H4 dimer. It gave precious information regarding how Asf1 handles the soluble H3-H4 histone form and promotes their nucleosomal assembly and disassembly[184]. Later on, the human Asf1 in complex with H3-H4 dimer was successfully crystallized and the 3D structure revealed a similar binding mode between the human and yeast forms of Asf1[193]. The 3D structure shows that Asf1 adopts an immunoglobin-like fold constituted of 9 β-strands connected with three helices in loops and acidic patches that mediate interactions with H3 (Figure 33A).

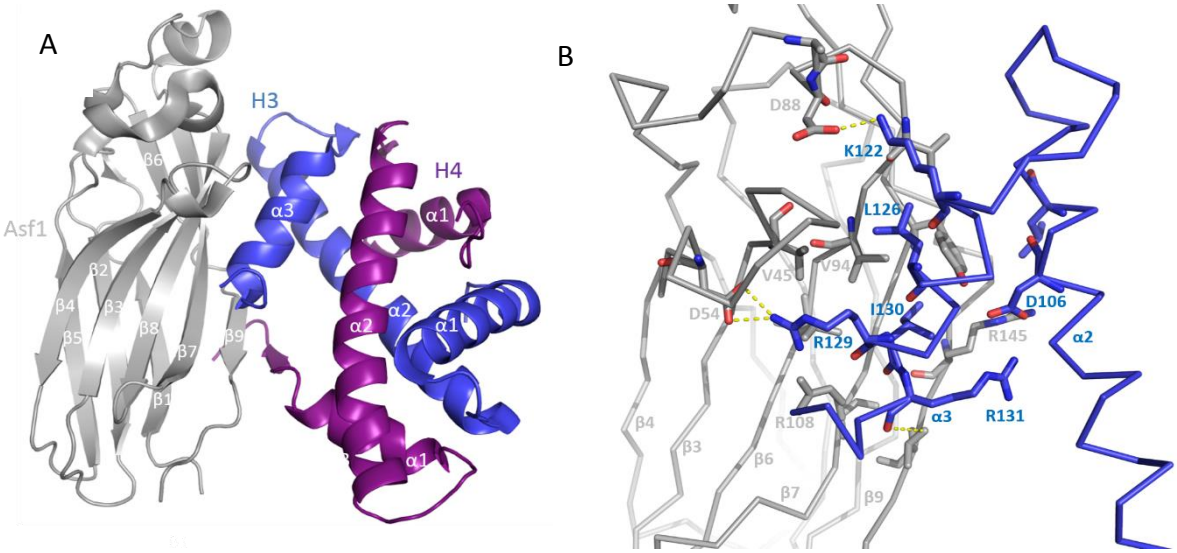


Figure 33: Crystal structure of human Asf1-H3-H4 complex (PDB 2IO5). A: Cartoon representation of Asf1-H3-H4 with Asf1 in grey, H3 in TV Blue and H4 in deep purple. B: Zoom on the interaction between Asf1 (Magenta) and H3 α2 and α3 (deep purple) with the key residues annotated[33].

At the molecular level, Asf1 binds to the fragment 106-135 of H3 located in the C-terminal region of H3 with multiple hydrophobic contacts involving β3, β4 and β6-9 strands of Asf1 (Figure 33B) [184]. Indeed, Asf1 Val45, Val92, Val94, Leu96 and Tyr112 make hydrophobic contacts with Leu109, Leu126 and Ile130 of H3. Furthermore, electrostatic interactions and hydrogen bonds occur between Asp54, Asp88, Arg108, Arg145 and Thr147 of Asf1 and Arg129, Lys122, Asp106 and Arg131 of H3 respectively. This C-terminal helix of H3 is the one involved in the formation of the $(H3\text{-}H4)_2$ tetramer[94] and the interaction of Asf1 with H3 C-terminal part is physically blocking the formation of $(H3\text{-}H4)_2$ tetramer. The main contribution to the binding comes from the H3-α3 helix (residues 122-135) but the H3-α2 helix (residues 106-115) also interacts with Asf1 surface. The C-terminal part of H3 α2 and the loop which connects α2-α3 interacts with residues of Asf1 in the β7-9 strands. The surface contact between the chaperone and H3 is quite extended with a surface about 900 $Å^2$.

Multi-dimensional NMR spectroscopy was performed by Ochsenbein et. al. to get information about the structure in solution of hAsf1 in complex with the histone H3 C-terminal helix[194,195]. Upon titration, the chemical shifts variations defined the binding surface as a concave groove of Asf1[194]. It is mainly hydrophobic with Val94 in the center surrounded by polar residues. The conformation of Asf1 is almost unchanged upon binding to histone H3. Investigations of HSQC experiments revealed that the peptide H3 (122-135) is unfolded in the absence of Asf1 whereas in its presence, the peptide adopts a helical conformation from residues 122 to 131. The authors showed that the interaction Asf1-H3α3 represents 85% of the total surface interaction meaning that it is the main contact zone. Titration experiments determined an affinity of the peptide of 100 ±20 µM. The high-resolution complex of Asf1 with H3-H4 allowed the determination of key residues responsible for the interactions. Asf1 mutants were made with mutations into the binding groove to study their ability to bind H3-H4[194]. A mutation of Val94 (in the center of the binding site) into a positively charged Arg (V94R), was performed and led to an inability to bind H3. As the V94R mutation is located in close contact with Leu126 and Ile130, two key residues of H3, the introduction of positively charged arginine creates steric clashes and electrostatic repulsions leading to a destabilized complex[195]. In the natural complex, residue Asp54 of Asf1a makes salt bridge with Arg129 of H3 whereas Arg108 of Asf1a caps the C-terminus helix. Mutations D54R and R108E bring opposite charges and perturb the formation of salt bridges without leading to steric clashes because of the side chains are able to change their conformation.

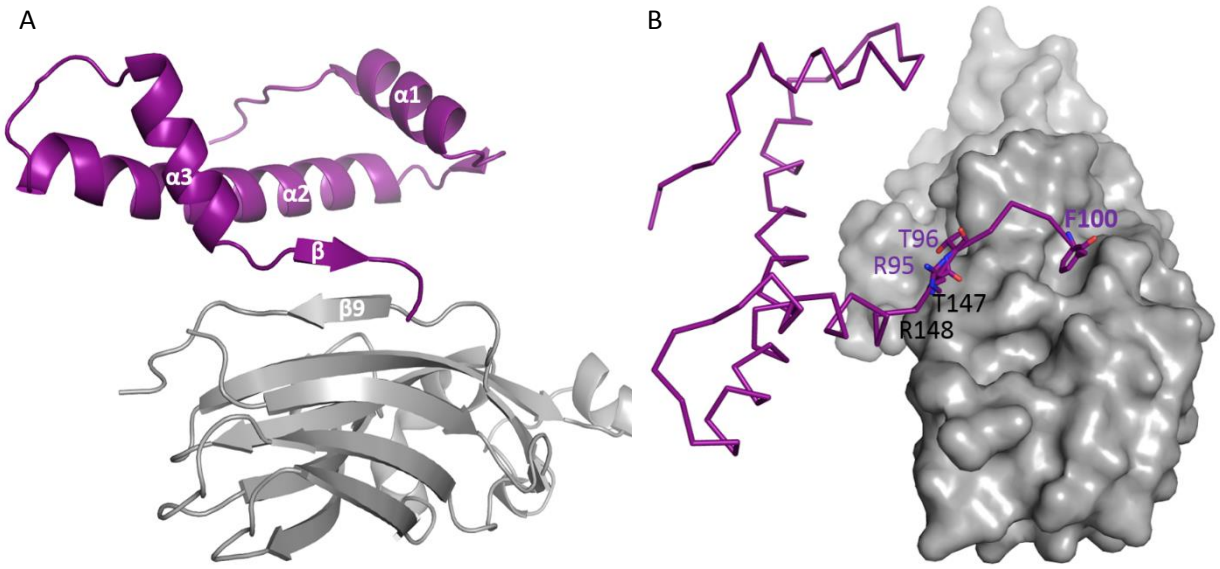


Figure 34: Asf1 interacting with histone H4. A) Representation of H4 in deep purple and Asf1 in grey illustrating the interaction with the C-terminal tail of H4 and β9 of asf1. B) The surface of Asf1 is represented in grey and the histone H4 is represented as ribbon in deep purple with the key residues involved in the interaction. (Pdb 2IO5, 2.7Å)

Of interest, Asf1 is also interacting with the histone H4 through its C-terminal tail whereas the triple helical region of H4 is near Asf1 but does not interact directly with it. The C-terminal tail of H4 undergoes a major conformational change compared to its structure into nucleosome[184]. In the nucleosome, the short C-terminal segment of H4 (residues 96-99) adopts a parallel β-sheet with H2A and folds back over α3 of H4 whereas when interacting with Asf1, this C-terminal part does a rotation of about 180° to form an antiparallel β-sheet with β9 of Asf1 (Figure 34A). This β-strand interacts with residues 144-148 of Asf1 especially through Thr96 that interacts with both H3 and Asf1 through Arg131 and Thr147 respectively. Nevertheless, the highest binding contribution of H4 comes from Phe100 that inserts into a hydrophobic pocket formed by β1 and β9 strands of Asf1 (Figure 34B). Van der Waals interactions are actually found between Phe100 and residues of Asf1 composing the hydrophobic pocket: Leu6, Tyr11, Val109, Pro144 and Val145 and that corresponds to 35% of the total surface interface between Asf1 and H4. The surface of interaction, although smaller than Asf1-H3 surface, involves about 500Å^2. These results concerning the H4 tail suggest that it could act as control for the assembly and disassembly of the $(H3\text{-}H4)_2$ tetramers by Asf1.
In this thesis, we have focused our interest on the interaction between Asf1 and H3-H4 and more specifically on the C-terminal parts of histones H3 and H4.

e) Role of Asf1 in cancer disease

Histone chaperones are involved in the majority of chromatin-dependent processes. Thus, cancer cells might manipulate histones to promote mutagenesis and alter gene expression[20,196]. Studies have shown that cancer cells have abnormal gene expression, aberrant histone PTMs in promoter regions highlighting the changes in chromatin organization during tumorogenesis[196,197]. There is

evidence that Asf1 is involved in several diseases such as Herpes Simplex Virus (HSV)[198],congenital dyserythropoietic anaemia type I (CDAI)[178] and in cancers. In this section, we will only focus on the implication of Asf1 in cancer. It has been previously demonstrated that Asf1 plays an important role in cell proliferation. Proliferation is one of the most relevant physiological context of human cancer and the use of histone chaperones as proliferation marker seems relevant[199]. Studies have shown that the concentration of Asf1 is much higher in several cancer forms than in normal cells. Furthermore, Asf1 has a role in longevity of cells as it facilitates the acetylation of H3 Lys56, which is known to increase lifetime[186,200]. Both Asf1a and Asf1b are involved in the S-phase progression and have an impact in the Alternative Lengthening of Telomeres (ALT) pathway in cancer cells[201,179,180]. In human cancers, this pathway can be diverted to maintain a high proliferation of cancer cells and Asf1 may have a role on it[202].

Corpet and coworkers have highlighted the role of Asf1 in breast cancer[203]. They have investigated the distribution of the two human Asf1 isoforms related to cell proliferation and tumorogenesis. They assigned a key role for Asf1b in proliferation. In order to determine the expression of Asf1 isoforms, they used specific antibodies able to specifically recognize one isoform to the other. Various cell lines were screened during these experiments. In particular, on MCF7 breast carcinoma cells in quiescence (phase of rest) a dramatic decrease of Asf1b was observed whereas in the case of Asf1a a minor decrease of its level of expression was revealed by immunofluorescence in presence of antibodies (Figure 35A). Quantitative RT-PCR was performed to determine the mRNA levels of Asf1a and Asf1b in asynchronous and quiescent human primary fibroblast cells as well as in young, old and senescent IMR90 human diploid primary fibroblasts (Figure 35B). The level of Asf1a mRNA is stable upon quiescence whereas there is a huge slump for Asf1b mRNA level in quiescent cells (about seven-fold). This suggests an important regulation of Asf1b expression. These results were confirmed by inducing quiescence of human primary fibroblast cell line, U2OS osteosarcoma cells and various cell lines. The authors concluded that in contrast to Asf1a, Asf1b is a specific marker of the state of cells (cycling/non-cycling or transiently/permanently arrested).

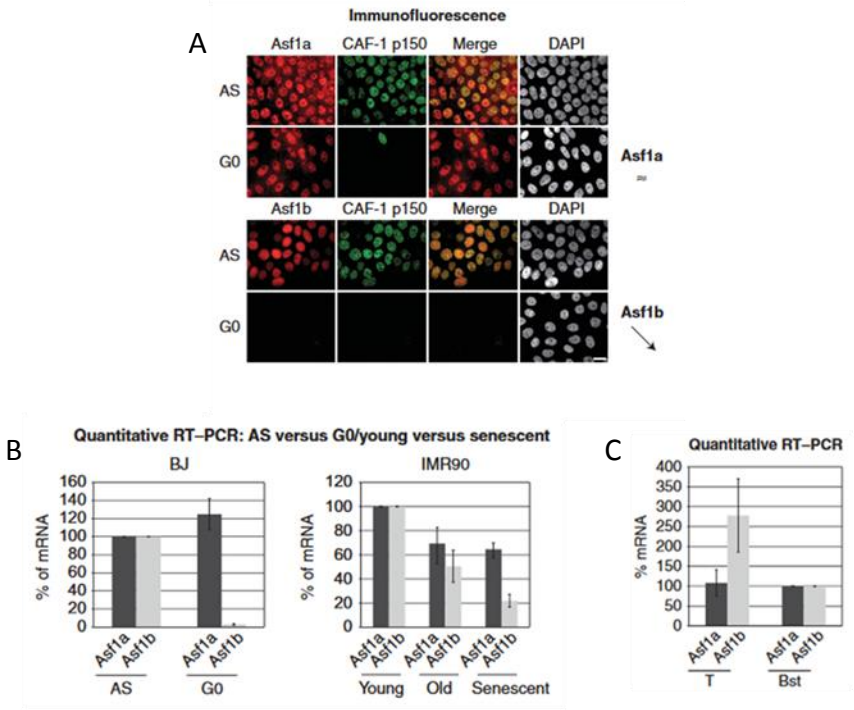


Figure 35: Comparison between the expression of both isoforms Asf1a and Asf1b in function of the cell cycle. (A) Immunofluorescence assay in MCF7 cell line revealing the expression of Asf1a, Asf1b, CAF-1 mRNA in proliferating (AS) and quiescent (G0) state. DAPI was used to stain nuclei and the scale bar is 20µm. (B) Quantitative RT-PCR determination of Asf1a and Asf1b mRNA levels in proliferating AS and G0 primary fibroblasts or in young, old and senescent IMR90 human diploid primary fibroblasts. (C) Quantitative RT-PCR analysis of Asf1a and Asf1b mRNA in tumoral (T) and normal (Bst) mammary cell lines (*Adapted* from Corpet et al. 2011)

Furthermore, Asf1b correlates well with the proliferation status of breast cancer cell lines. Indeed, western-blot and RT-SPR analyses have been performed on tumoral and normal mammary cell lines and revealed an increase in the level of Asf1b mRNA in cancer cells (Figure 35C). In contrast, the level of Asf1a is almost unchanged. To note, the level of Asf1b also correlates with the proliferative status of mammary cells. The depletion of one isomer has distinct effects on cells. Indeed, the depletion of Asf1b leads to an upregulation of Asf1a which suggests a compensating mechanism. Moreover, its depletion alters nuclear morphology and prevents continued-proliferation leading to cell death. In contrast, a depletion of Asf1a as no consequence on the level of Asf1b and on the nuclear morphology.

Consequently, Asf1b can be viewed as a new proliferation marker in breast tumor samples. Indeed, the authors evaluated the correlation between both isoforms Asf1a and Asf1b and clinical parameters to evaluate a potential diagnosis value: the tumor size, the mitotic index and the grade of the tumor (Figure 36). A high correlation with Asf1b was found but none with Asf1a. Finally, Asf1b has a prognostic value in breast cancer. Indeed, patients were divided into two groups with low Asf1b level and high Asf1b level which was correlated to the tumor progression. The level of Asf1b was also associated with the proliferation rate and the aggressiveness of

distinct breast cancers. This overexpression of Asf1b is also present in other types of cancers such as liver, ovarian, lung and skin cancers.

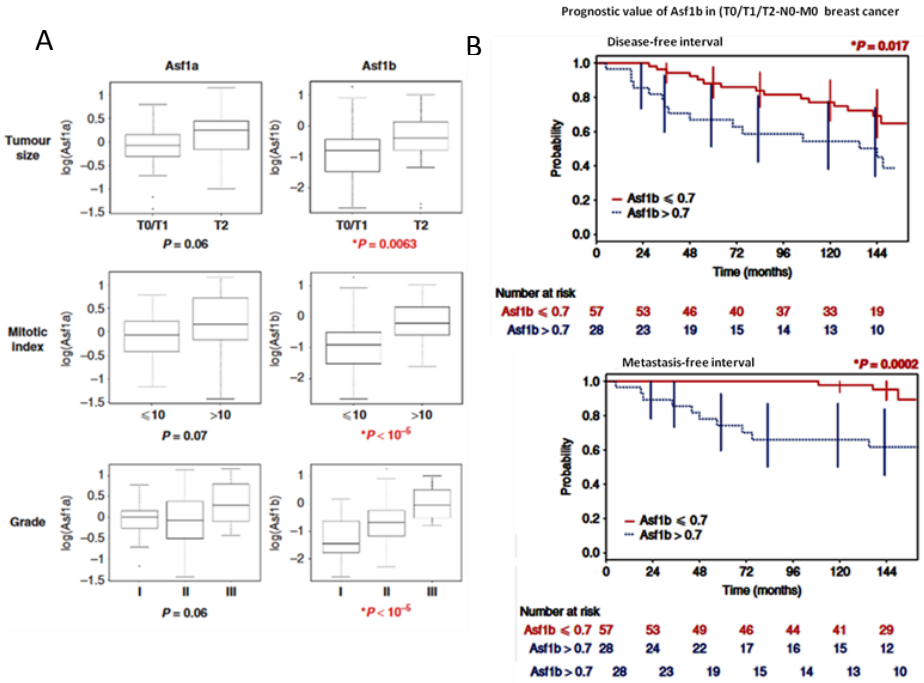


Figure 36: Correlation between Asf1b and prognostic value in breast cancer patients (A) Logarithmic expression levels of Asf1a and Asf1b in breast cancer samples, depending on indicated clinical factors with a >10 years patient follow-up. Boxes represent the 25–75th percentile, brackets: range; black line: median; black dots: outliers. Below each graph the *P*-values determined by a Kruskal–Wallis test are indicated. Red color together with an asterisk * indicates a significant *P*-value (*P*<0.05). (B) Univariate Kaplan–Meier curves of the disease free interval and the occurrence of metastasis in patients expressing low (Asf1b ≤p0.7) or high (Asf1b >0.7) levels of Asf1b (Adapted from Corpet et al. 2011)

Recent studies reported that Asf1a can also be used as prognostic marker in gastrointestinal and colorectal cancers[204]. Compared to non-cancerous cell, Asf1a is overexpressed in GIC tumors. Studies revealed that Asf1a has multiple effects in gastrointestinal cancer (GIC) such as the stimulation of the transcription of β-catenin genes due to tis interaction with Asf1a.

All these studies showing the correlation between the level of expression of Asf1 and cancer highlight that this histone chaperone is a potential target to develop novel anticancer therapies[205]. Chemists just start to show interest in this protein chaperone, but so far, only few studies have been published concerning the inhibition of Asf1. In 2015, Miknis et.al. developed inhibitors of Asf1/H3-H4 by screening a series of small molecules derived from N-acyl hydrazones[206]. After a virtual screening of 139,735 compounds library and a molecular docking, best compounds were classified. Series of compound were prepared and showed IC_{50} varying from 12 to >100 µM. Seol et. al. identified small molecules that inhibit Asf1 and H3 Lys_{56} acetylation only[207]. After screening a library with 260,000 compounds, 49 compounds were selected to be tested in vitro, leading to 6 hits. These compounds decrease the interaction between Asf1 and H3 in the range of 20-50 µM. The activity of these inhibitors is modest and high

concentration is necessary meaning that optimizations as well as the development of other strategies are still needed to generate potent modulators of Asf1.

C. Objectives of the thesis

1. Short historic about oligourea foldamers

Historically, oligourea peptidomimetics were first introduced by Burgess *et al.* in 1995[208]. Later associated to the foldamer family, aliphatic *N,N'*-linked oligoureas are aza-analogs of γ-peptides where amide moieties (CH_2-CO-NH) have been replaced by urea ones (NH-CO-NH). This urea motif presents a carbonyl as H-bond acceptor and two NHs as H-bond donors. Despite similarities between amide and urea bonds such as the flatness, the robustness, the polarity and the ability to make H-bonds, notable differences exist between both chemical bonds. Indeed, the urea dipolar moment is higher than the amide one whereas the rotational energy barrier of urea is lower with 10-12 kcal.mol^{-1} instead of 16-20 kcal.mol^{-1} for amide[209]. The conformational behavior of these oligoureas has been extensively studied by the Guichard group over the last 15 years with detailed structural studies in solution and in the solid state. First, the solution structure of model *N,N'*-linked oligourea sequences (as for example H_2N-Tyr^u-Lys^u-Leu^u-Val^u-Phe^u-Lys^u-Val^u-Tyr^u-Ala^u-NH_2[210]) was determined by NMR spectroscopy and CD experiments and revealed a stable 2.5 helical secondary structure closely related to the (*P*)-2.6$_{14}$ helix of γ-peptides[210,211]. Remarkably, the second NH brought by the urea bond allows the formation of an additional H-bond in comparison with corresponding γ-oligoamide isosteres. With this additional nitrogen, a three-centered H-bond network takes place, leading to the formation of a 12- or 14-membered pseudo cycles (12,14-helix). Next, NMR data recorded in various solvents presenting different polarities (pyridine, MeOH, DMSO, ACN) unveiled that the NH signals were strongly dispersed thus confirming their implication in the H-bond networks[212]. Furthermore, the observation of medium range Nuclear Overhausser effects (NOE) between residues in *i*/*i*+2 and *i*/*i*+3 relationship and repeated all along the sequence, were characteristic of a helical conformation (Figure 37A). Four to five urea residues in the backbone are necessary to promote helix formation in low polarity solvents and it has been demonstrated that helical nature is not strongly affected by the nature and composition of the side chains[212]. Far-UV CD experiments were also useful to study helix formation in solution. CD spectra of helical folded oligoureas recorded in several solvents (MeOH, TFE and ACN) displayed a characteristic helical signature with a minimum negative ellipticity at 188 nm and a maximum of positive ellipticity at 203 nm (Figure 37B)[211]. NMR and CD studies suggest that helical folding is maximized in a low or moderate polarity solvent such as pyridine and acetonitrile but it is worth noting that despite a weaker CD signal the 2.5-helix conformation is still populated in aqueous solution, which is interesting for biological applications[148]. The 2.5-helix nucleation is increased when terminal *N*-caps are introduced in oligourea sequence due to the presence of an additional H-bond acceptor at the N-terminus that avoids repulsive electrostatic interactions between the 2.5-helix macrodipole and the charged N-terminus[211].

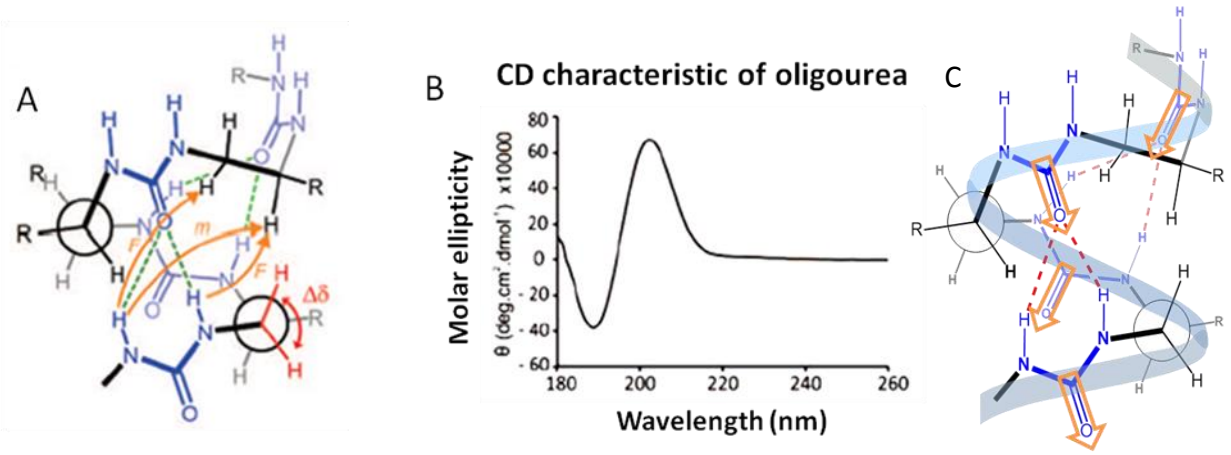


Figure 37: Structural data of oligourea obtained by NMR spectroscopy and CD experiments. A) Structural information obtained by NMR for the helical folding of oligoureas. In orange, characteristic NOE effects of the helix are represented. B) Typical Far UV-CD spectrum of a helical oligourea (model sequence: Boc-Val^{u}-Ala^{u}-Leu^{u}-Val^{u}-Ala^{u}-Leu^{u}-NHMe) recorded in TFE at 0.2 mM. C) Polarity of oligourea

These results were confirmed at the solid state with the X-ray crystal structure characterizations at atomic resolution of several oligourea sequences and the robustness of the 2.5-helix was thereby illustrated[144,146,147]. Like α-helix, the 2.5-helix of *N,N'*-linked oligoureas, whose residues derived from natural L-amino acids, is right-handed with a pitch of 5.1 Å whereas a pitch of 5 Å is observed for γ^4-peptides (Figure 38A). These structural data highlight the analogy between *N,N'*-linked oligoureas and γ-peptides. Interestingly, the helical folding of oligoureas resembles that of α-peptides but some significant differences exist such as the number of residues per turn (2.5 for N,N'-linked oligoureas versus 3.6 for α-peptides). Furthermore, the pitch of the α-helix (5.4 Å) is a little bit higher than for oligoureas whereas the radius is lower (2.3 Å compared to 2.6 Å). These differences and the number of residues per turn in particular, which has a direct impact on the distribution of side chains at the surface of the helical structure, will need to be taken into account when attempting to use oligoureas to design effective mimics of α-helical peptides(Figure 38B).

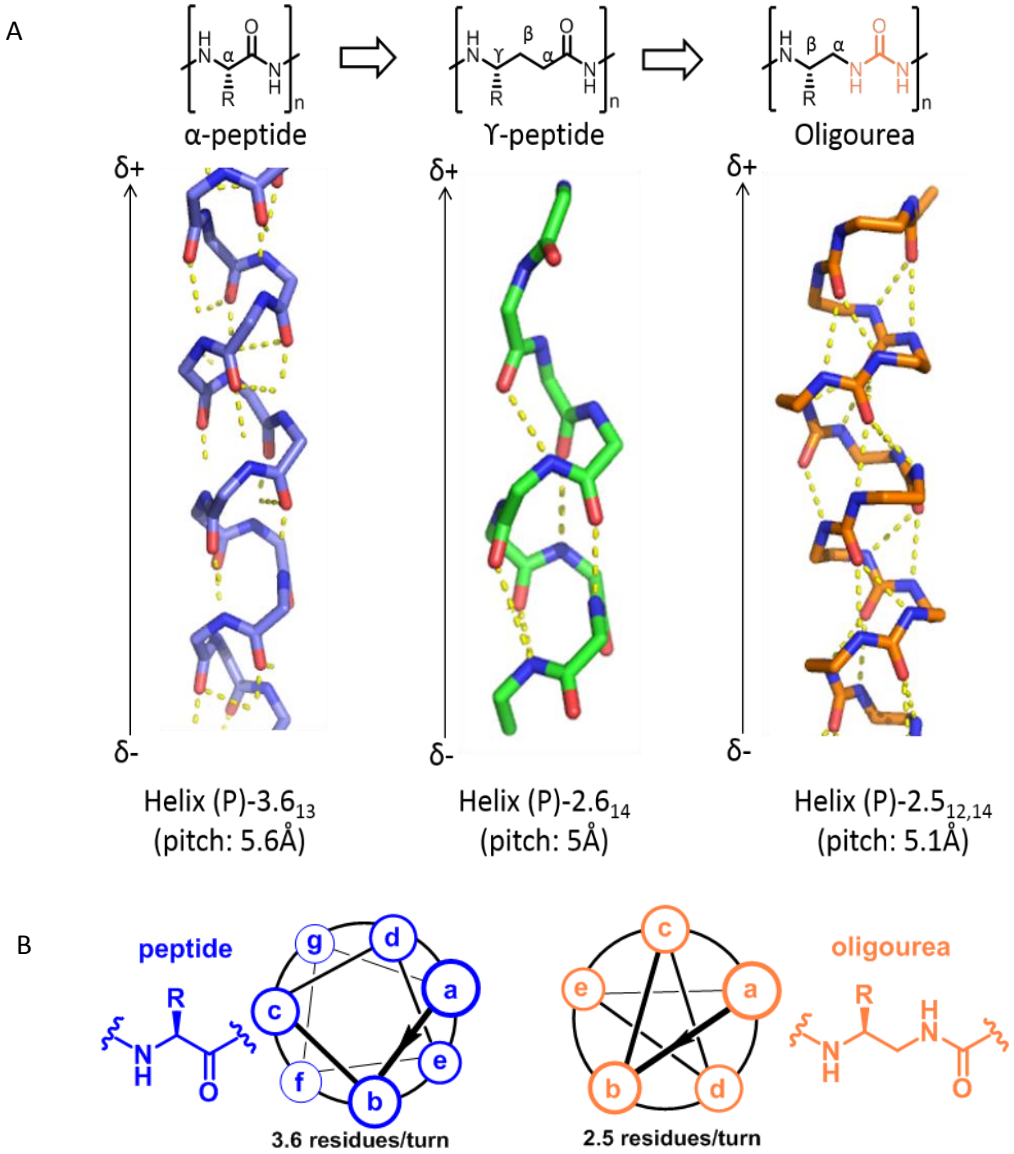


Figure 38: Comparison of foldamer helical structures with the α-helix. A) Structure of α-peptide (in blue), ϒ-peptide (in green) and oligourea (in orange) without side chains B) Diagram wheels of peptide in blue and oligourea in orange.

Moreover, a recent advance in the group was the finding that sequence manipulation can be used to control the self-assembly of short amphiphilic oligoureas in aqueous conditions[213].The design of entirely non-peptidic foldamers able to self-assemble into well-defined three-dimensional protein-like nanostructures have been reported for the first time. These results illustrate an important step concerning the structure-guided design of oligoureas which is essential for specific protein surface recognition and biological applications knowing the close link between structure and function.

Despite the differences between both oligourea and helices, their similarities (polarity, right-handed helix, …) suggested that the two backbones could be combined to develop hybrids

oligomers consisting of both α-peptide and oligourea segments. The combination of the two backbones may bring advantages from both oligoureas and peptides: the natural recognition of α-helix and the innate helical stability of oligoureas. The introduction of oligourea residues into a peptide sequence while keeping the conformational preference may improve its biostability. Recent studies concerning the synthesis and conformational analyses of such hybrid compounds in organic solvents have been reported and confirm the feasibility of this new strategy[214,215]. Indeed, a unique helical conformation is observed in these hybrid systems which is stabilized thanks to a continuum of H-bonds all along the helix. The main topic of this PhD thesis will be to use this technology to develop hybrid compounds as peptide mimics and study their ability to recognize protein surfaces.

2. Oligourea synthesis

The first synthesis of aliphatic oligoureas has been reported by Burgess in 1995 who described a solid phase synthesis approach involving a succession of coupling and deprotection steps of protected isocyanates[208]. Several groups have then reported related procedures for the synthesis of oligoureas such as the use of diversely protected ethylene diamine activated monomers as precursors[216–218].

In the Guichard group, the preferred procedure for activated monomer synthesis consists in the direct conversion of N-protected ethylene diamine units into the corresponding activated carbamate by treatment with *N,N'*-disuccinimidyl carbonate (DSC). This synthetic approach allows the access to activated monomers with a variety of natural or even unnatural side-chains. The synthesis of *N,N'*-linked oligoureas is generally performed either in solution or on solid support depending on the length of the oligomers and the nature of side-chains to be appended. The solid phase synthesis (SPS) of oligoureas has been firstly developed by using a Fmoc strategy[148]. However, it quickly appeared the Fmoc protected group was not compatible with a routine preparation of oligoureas due to a partial instability under SPS conditions that resulted over-insertion of monomeric units with poor overall purities and yields as chain length increases. Furthermore, this strategy was not incompatible with microwave assistance (degradation of N-Fmoc monomer and uncontrolled oligomerization on the resin), increasing the time of coupling and number of equivalents required to perform SPS (up to 4h per coupling step). Concurrently, the Boc strategy has then been developed in solution and showed much higher robustness than the Fmoc strategy on the solid support and under microwave conditions[149]. However, the major drawback of this Boc strategy is that the resin that is generally used, the 4-methylbenzhydrylamine (MBHA), requires a final fluorhydric acid cleavage which is detrimental for routine use. To solve these problems, Douat *et al.* have recently developed a new methodology allowing the use of TFA-labile resins under microwave irradiation[219].They reinvestigated the use of azides as masked amines for SPS of oligoureas. This strategy is based on the coupling under microwave of succinimidyl (2-azido-2substituted ethyl) carbamates (compound **A** Figure 39).

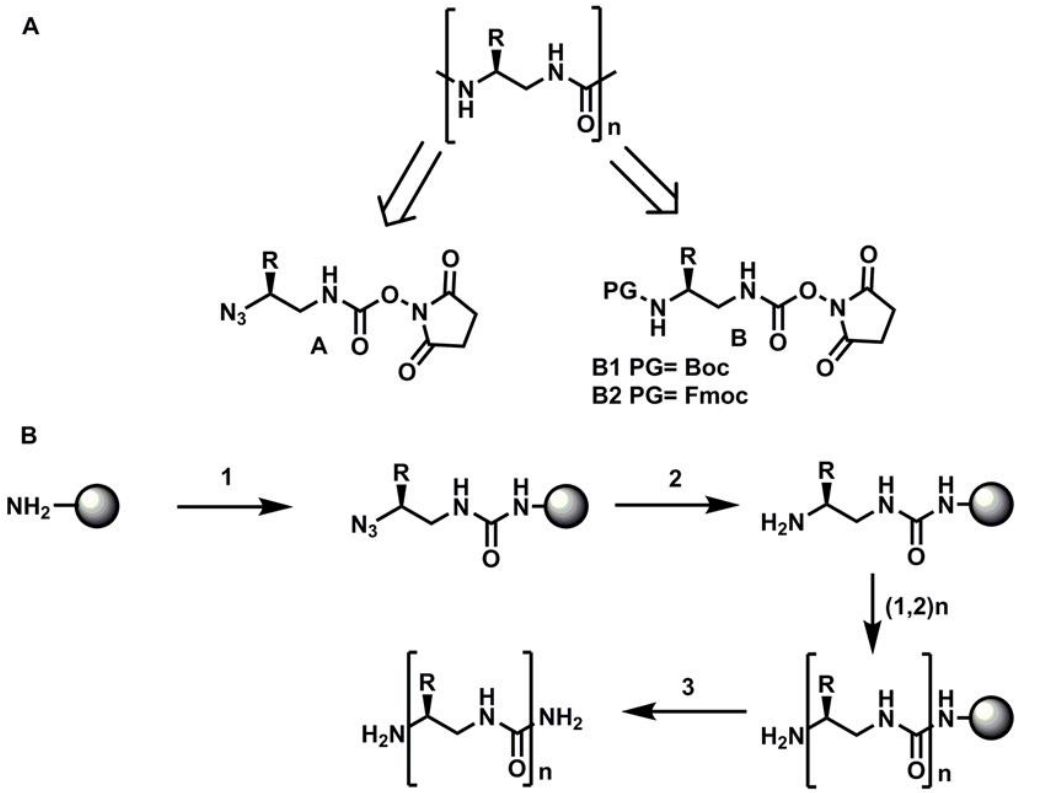


Figure 39: A) Schematic representation of oligourea and formulas of azide monomers A and activated protected monomers B. B) Scheme of the solid phase synthesis approach under microwave assistance with the most recent azide strategy.

The azide reduction is generally performed under Staudinger conditions in the presence of trimethylphosphine (PMe_3) in Dioxane/H_2O (70:30, v/v) solvent mixture. This had an impact on the choice of the resin that should be compatible with polar solvents, notably water used for the Staudinger reaction. For this reason, the NovaPEG Rink amide resin is generally used in the group because its polyethylene glycol matrix is known to present a high swelling in water[220].

This microwave-assisted SPS methodology which allows rapid and reproducible access to N,N'-linked oligoureas exhibits clear benefits such as the enhancement of the rate of synthesis, the reduced amount of activated building blocks required per coupling and the recovery of oligoureas in higher purity and fair to good yields. This azide strategy is now routinely used in the laboratory and has been successfully applied for the synthesis of hybrid sequences in combination with standard peptide synthesis using Fmoc chemistry[214,215].

3. Objectives of the thesis

As mentioned above, water-soluble oligoureas have the ability to form stable 2.5-helix in aqueous solution which is quite relevant for possible biological applications. Their design has been mainly

inspired by the sequences of bioactive peptides and these oligoureas have shown interesting activities such as their capacity to disrupt bacterial membrane and their ability to mimic cell penetrating peptides[143,149]. Oligoureas are resistant to proteolysis and closely mimic the α-helix of peptides which makes them interesting candidate to develop new therapeutics. However, with 2.5 residues per turn of helix, precise mimicry of the projection of all the lateral chains of α-peptides by using an oligourea backbone is challenging. In this context, the combination of oligourea and peptide backbones to form oligourea/peptide chimeras may be a good compromise. Indeed, the mixed backbone may provide additional modularity to generate effective peptide mimics.

In this thesis, we aimed at testing oligoureas and oligourea/peptide chimeras as inhibitors of protein-protein interactions. In particular, we tried to design oligourea/peptide chimeras for the inhibition of a new target involved in cancer: the histone chaperon protein Asf1b. This project was initiated in the context of a collaboration with the team of Françoise Ochsenbein at CEA Saclay who has pioneered the structural characterization of the complex Asf1/H3/H4 combining NMR, crystallography and modeling studies.

Firstly, and as a model study, we strived to study the effect of the insertion short urea segments into long helical peptides on the overall structure by focusing on high resolution structure determination. The number and the position in the sequence of urea residues have been varied leading to the synthesis of several chimeras. Their conformational properties as well as their self-assembling properties were studied in solution and at the solid state.

In the second part of this thesis, we will present our attempts to design effective α-helix mimics for the inhibition of PPIs by using urea based foldamers. We focused our work on oligourea-peptide chimeras designed to mimic the helical C-terminal part of histone H3 and to inhibit the Asf1/H3/H4 interaction which is thought to be relevant in several cancers. The design guided by the structure, the synthesis and affinity of these compounds for the target protein Asf1 as well as our efforts to characterize the protein-foldamer interactions at high resolution will be described in this part.

Finally, we have considered targeting a second binding site located in another face of the protein in order to increase the specificity and the affinity of our compounds for Asf1. In the native Asf1/H3/H4 complex, this second site recognizes the C-terminal tail of histone H4. Strategies to connect the two binding motifs (i.e. the helically folded H3 peptide/foldamer and the C-terminal tail of H4) including the rationale design and fine-tuning of the linking unit will be discussed and the impact of combining two distinct and remote binding motifs in one molecule on the affinity will be reported.

Bibliography

(1) Conte, L. L.; Chothia, C.; Janin, J. The Atomic Structure of Protein-Protein Recognition Sites. *J. Mol. Biol.* **1999**, *285* (5), 2177–2198.
(2) Bakail, M.; Ochsenbein, F. Targeting Protein–Protein Interactions, a Wide Open Field for Drug Design. *Comptes Rendus Chim.* **2016**, *19* (1–2), 19–27.
(3) Azzarito, V.; Long, K.; Murphy, N. S.; Wilson, A. J. Inhibition of α-Helix-Mediated Protein–Protein Interactions Using Designed Molecules. *Nat. Chem.* **2013**, *5* (3), 161–173.
(4) J. Wilson, A. Inhibition of Protein–Protein Interactions Using Designed Molecules. *Chem. Soc. Rev.* **2009**, *38* (12), 3289.
(5) De Las Rivas, J.; Fontanillo, C. Protein–Protein Interactions Essentials: Key Concepts to Building and Analyzing Interactome Networks. *PLoS Comput. Biol.* **2010**, *6* (6), e1000807.
(6) Keskin, O.; Gursoy, A.; Ma, B.; Nussinov, R. Principles of Protein−Protein Interactions: What Are the Preferred Ways For Proteins To Interact? *Chem. Rev.* **2008**, *108* (4), 1225–1244.
(7) Clackson, T.; Wells, J. A. A Hot Spot of Binding Energy in a Hormone-Receptor Interface. *Science* **1995**, *267* (5196), 383–386.
(8) DeLano, W. L. Unraveling Hot Spots in Binding Interfaces: Progress and Challenges. *Curr. Opin. Struct. Biol.* **2002**, *12* (1), 14–20.
(9) Zinzalla, G.; Thurston, D. E. Targeting Protein–Protein Interactions for Therapeutic Intervention: A Challenge for the Future. *Future Med. Chem.* **2009**, *1* (1), 65–93.
(10) Ivanov, A. A.; Khuri, F. R.; Fu, H. Targeting Protein–Protein Interactions as an Anticancer Strategy. *Trends Pharmacol. Sci.* **2013**, *34* (7), 393–400.
(11) Milroy, L.-G.; Grossmann, T. N.; Hennig, S.; Brunsveld, L.; Ottmann, C. Modulators of Protein–Protein Interactions. *Chem. Rev.* **2014**, *114* (9), 4695–4748.
(12) Johnson, D. K.; Karanicolas, J. Ultra-High-Throughput Structure-Based Virtual Screening for Small-Molecule Inhibitors of Protein–Protein Interactions. *J. Chem. Inf. Model.* **2016**, *56* (2), 399–411.
(13) Bullock, B. N.; Jochim, A. L.; Arora, P. S. Assessing Helical Protein Interfaces for Inhibitor Design. *J. Am. Chem. Soc.* **2011**, *133* (36), 14220–14223.
(14) Mason, J. M. Design and Development of Peptides and Peptide Mimetics as Antagonists for Therapeutic Intervention. *Future Med. Chem.* **2010**, *2* (12), 1813–1822.
(15) Marshall, S. A.; Lazar, G. A.; Chirino, A. J.; Desjarlais, J. R. Rational Design and Engineering of Therapeutic Proteins. *Drug Discov. Today* **2003**, *8* (5), 212–221.
(16) Pelay-Gimeno, M.; Glas, A.; Koch, O.; Grossmann, T. N. Structure-Based Design of Inhibitors of Protein-Protein Interactions: Mimicking Peptide Binding Epitopes. *Angew. Chem. Int. Ed.* **2015**, n/a-n/a.
(17) Appella, D. H.; Christianson, L. A.; Klein, D. A.; Powell, D. R.; Huang, X.; Barchi Jr, J. J.; Gellman, S. H. Residue-Based Control of Helix Shape in β-Peptide Oligomers. *Nature* **1997**, *387*, 381.
(18) Gellman, S. H. Foldamers: A Manifesto. *Acc. Chem. Res.* **1998**, *31* (4), 173–180.
(19) Guichard, G.; Huc, I. Synthetic Foldamers. *Chem. Commun.* **2011**, *47* (21), 5933.
(20) Hanahan, D.; Weinberg, R. A. Hallmarks of Cancer: The Next Generation. *Cell* **2011**, *144* (5), 646–674.
(21) Mullard, A. Protein–Protein Interaction Inhibitors Get into the Groove. *Nat. Rev. Drug Discov.* **2012**, *11* (3), 173–175.

(22) Tovar, C.; Graves, B.; Packman, K.; Filipovic, Z.; Xia, B. H. M.; Tardell, C.; Garrido, R.; Lee, E.; Kolinsky, K.; To, K.-H.; et al. MDM2 Small-Molecule Antagonist RG7112 Activates P53 Signaling and Regresses Human Tumors in Preclinical Cancer Models. *Cancer Res.* **2013**, *73* (8), 2587–2597.

(23) Chang, Y. S.; Graves, B.; Guerlavais, V.; Tovar, C.; Packman, K.; To, K.-H.; Olson, K. A.; Kesavan, K.; Gangurde, P.; Mukherjee, A.; et al. Stapled -Helical Peptide Drug Development: A Potent Dual Inhibitor of MDM2 and MDMX for P53-Dependent Cancer Therapy. *Proc. Natl. Acad. Sci.* **2013**, *110* (36), E3445–E3454.

(24) Meric-Bernstam, F.; Saleh, M. N.; Infante, J. R.; Goel, S.; Falchook, G. S.; Shapiro, G.; Chung, K. Y.; Conry, R. M.; Hong, D. S.; Wang, J. S.-Z.; et al. Phase I Trial of a Novel Stapled Peptide ALRN-6924 Disrupting MDMX- and MDM2-Mediated Inhibition of WT P53 in Patients with Solid Tumors and Lymphomas. *J. Clin. Oncol.* **2017**, *35* (15_suppl), 2505–2505.

(25) Wachter, F.; Morgan, A. M.; Godes, M.; Mourtada, R.; Bird, G. H.; Walensky, L. D. Mechanistic Validation of a Clinical Lead Stapled Peptide That Reactivates P53 by Dual HDM2 and HDMX Targeting. *Oncogene* **2017**, *36* (15), 2184–2190.

(26) Ashkenazi, A. Directing Cancer Cells to Self-Destruct with pro-Apoptotic Receptor Agonists. *Nat. Rev. Drug Discov.* **2008**, *7* (12), 1001–1012.

(27) DeLeo, A. B.; Jay, G.; Appella, E.; Dubois, G. C.; Law, L. W.; Old, L. J. Detection of a Transformation-Related Antigen in Chemically Induced Sarcomas and Other Transformed Cells of the Mouse. *Proc. Natl. Acad. Sci. U. S. A.* **1979**, *76* (5), 2420–2424.

(28) Soussi, T. The History of P53. *EMBO Rep.* **2010**, *11* (11), 822–826.

(29) Crawford, R. E. Understanding the Total Meaning. *Foreign Lang. Ann.* **1983**, *16* (3), 209–211.

(30) Barak, Y.; Juven, T.; Haffner, R.; Oren, M. Mdm2 Expression Is Induced by Wild Type P53 Activity. *EMBO J.* **1993**, *12* (2), 461–468.

(31) Moll, U. M.; Petrenko, O. The MDM2-P53 Interaction. *Mol. Cancer Res.* **2003**, *1* (14), 1001–1008.

(32) Chène, P. Inhibiting the P53–MDM2 Interaction: An Important Target for Cancer Therapy. *Nat. Rev. Cancer* **2003**, *3* (2), 102–109.

(33) Kussie, P. H.; Gorina, S.; Marechal, V.; Elenbaas, B.; Moreau, J.; Levine, A. J.; Pavletich, N. P. Structure of the MDM2 Oncoprotein Bound to the P53 Tumor Suppressor Transactivation Domain. *Science* **1996**, *274* (5289), 948–953.

(34) Kritzer, J. A.; Lear, J. D.; Hodsdon, M. E.; Schepartz, A. Helical β-Peptide Inhibitors of the P53-HDM2 Interaction. *J. Am. Chem. Soc.* **2004**, *126* (31), 9468–9469.

(35) Kritzer, J. A.; Luedtke, N. W.; Harker, E. A.; Schepartz, A. A Rapid Library Screen for Tailoring β-Peptide Structure and Function. *J. Am. Chem. Soc.* **2005**, *127* (42), 14584–14585.

(36) Schreiber, J. V.; Frackenpohl, J.; Moser, F.; Fleischmann, T.; Kohler, H.-P. E.; Seebach, D. On the Biodegradation of β-Peptides. *ChemBioChem* **2002**, *3* (5), 424–432.

(37) Seebach, D.; Overhand, M.; Kühnle, F. N. M.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H. β-Peptides: Synthesis by Arndt-Eistert Homologation with Concomitant Peptide Coupling. Structure Determination by NMR and CD Spectroscopy and by X-Ray Crystallography. Helical Secondary Structure of a β-Hexapeptide in Solution and Its Stability towards Pepsin. *Helv. Chim. Acta* **1996**, *79* (4), 913–941.

(38) Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. β-Peptides: From Structure to Function. *Chem. Rev.* **2001**, *101* (10), 3219–3232.

(39) Harker, E. A.; Daniels, D. S.; Guarracino, D. A.; Schepartz, A. β-Peptides with Improved Affinity for HDM2 and HDMX. *Bioorg. Med. Chem.* **2009**, *17* (5), 2038–2046.

(40) Bautista, A. D.; Appelbaum, J. S.; Craig, C. J.; Michel, J.; Schepartz, A. Bridged ?3-Peptide Inhibitors of P53-HDM2 Complexation-Correlation Between Affinity and Cell Permeability. *J. Am. Chem. Soc.* **2010**, *132* (9), 2904–2906.

(41) Orner, B. P.; Ernst, J. T.; Hamilton, A. D. Toward Proteomimetics: Terphenyl Derivatives as Structural and Functional Mimics of Extended Regions of an α-Helix. *J. Am. Chem. Soc.* **2001**, *123* (22), 5382–5383.

(42) Kutzki, O.; Park, H. S.; Ernst, J. T.; Orner, B. P.; Yin, H.; Hamilton, A. D. Development of a Potent Bcl-XL Antagonist Based on α-Helix Mimicry. *J. Am. Chem. Soc.* **2002**, *124* (40), 11838–11839.

(43) Yin, H.; Hamilton, A. D. Strategies for Targeting Protein-Protein Interactions With Synthetic Agents. *Angew. Chem. Int. Ed.* **2005**, *44* (27), 4130–4163.

(44) Saraogi, I.; Hebda, J. A.; Becerril, J.; Estroff, L. A.; Miranker, A. D.; Hamilton, A. D. Synthetic α-Helix Mimetics as Agonists and Antagonists of IAPP Amyloid Formation. *Angew. Chem. Int. Ed Engl.* **2010**, *49* (4), 736–739.

(45) Shaginian, A.; Whitby, L. R.; Hong, S.; Hwang, I.; Farooqi, B.; Searcey, M.; Chen, J.; Vogt, P. K.; Boger, D. L. Design, Synthesis, and Evaluation of an α-Helix Mimetic Library Targeting Protein-Protein Interactions. *J. Am. Chem. Soc.* **2009**, *131* (15), 5564–5572.

(46) Campbell, F.; Plante, J. P.; Edwards, T. A.; Warriner, S. L.; Wilson, A. J. N-Alkylated Oligoamide α-Helical Proteomimetics. *Org. Biomol. Chem.* **2010**, *8* (10), 2344.

(47) Plante, J. P.; Burnley, T.; Malkova, B.; Webb, M. E.; Warriner, S. L.; Edwards, T. A.; Wilson, A. J. Oligobenzamide Proteomimetic Inhibitors of the P53–HDM2 Protein–Protein Interaction. *Chem. Commun.* **2009**, No. 34, 5091.

(48) Azzarito, V.; Prabhakaran, P.; Bartlett, A. I.; Murphy, N. S.; Hardie, M. J.; Kilner, C. A.; Edwards, T. A.; Warriner, S. L.; Wilson, A. J. 2-O-Alkylated Para-Benzamide α-Helix Mimetics: The Role of Scaffold Curvature. *Org. Biomol. Chem.* **2012**, *10* (32), 6469.

(49) Campbell, F.; Plante, J.; Carruthers, C.; Hardie, M. J.; Prior, T. J.; Wilson, A. J. Macrocyclic Scaffolds Derived from P-Aminobenzoic Acid. *Chem. Commun.* **2007**, *0* (22), 2240–2242.

(50) Campbell, F.; Wilson, A. J. An 'Impossible' Macrocyclisation Using Conformation Directing Protecting Groups. *Tetrahedron Lett.* **2009**, *50* (19), 2236–2238.

(51) Barnard, A.; Long, K.; Martin, H. L.; Miles, J. A.; Edwards, T. A.; Tomlinson, D. C.; Macdonald, A.; Wilson, A. J. Selective and Potent Proteomimetic Inhibitors of Intracellular Protein–Protein Interactions. *Angew. Chem. Int. Ed.* **2015**, *54* (10), 2960–2965.

(52) Burgess, A.; Chia, K. M.; Haupt, S.; Thomas, D.; Haupt, Y.; Lim, E. Clinical Overview of MDM2/X-Targeted Therapies. *Front. Oncol.* **2016**, *6*.

(53) Plante, J.; Campbell, F.; Malkova, B.; Kilner, C.; Warriner, S. L.; Wilson, A. J. Synthesis of Functionalised Aromatic Oligamide Rods. *Org. Biomol. Chem.* **2007**, *6* (1), 138–146.

(54) Ahn, J.-M.; Han, S.-Y. Facile Synthesis of Benzamides to Mimic an α-Helix. *Tetrahedron Lett.* **2007**, *48* (20), 3543–3547.

(55) Prabhakaran, P.; Barnard, A.; Murphy, N. S.; Kilner, C. A.; Edwards, T. A.; Wilson, A. J. Aromatic Oligoamide Foldamers with a "Wet Edge" as Inhibitors of the α-Helix-Mediated P53- *h* DM2 Protein-Protein Interaction: Inhibitors of Protein-Protein Interactions. *Eur. J. Org. Chem.* **2013**, *2013* (17), 3504–3512.

(56) Azzarito, V.; Rowell, P.; Barnard, A.; Edwards, T. A.; Macdonald, A.; Warriner, S. L.; Wilson, A. J. Probing Protein Surfaces: QSAR Analysis with Helix Mimetics. *ChemBioChem* **2016**, *17* (8), 768–773.

(57) Grison, C. M.; Miles, J. A.; Robin, S.; Wilson, A. J.; Aitken, D. J. An α-Helix-Mimicking 12,13-Helix: Designed α/β/γ-Foldamers as Selective Inhibitors of Protein-Protein Interactions. *Angew. Chem. Int. Ed.* **2016**, *55* (37), 11096–11100.

(58) Shin, Y.-H.; Mortenson, D. E.; Satyshur, K. A.; Forest, K. T.; Gellman, S. H. Differential Impact of β and γ Residue Preorganization on α/β/γ-Peptide Helix Stability in Water. *J. Am. Chem. Soc.* **2013**, *135* (22), 8149–8152.

(59) Karle, I. L.; Pramanik, A.; Banerjee, A.; Bhattacharjya, S.; Balaram, P. ω-Amino Acids in Peptide Design. Crystal Structures and Solution Conformations of Peptide Helices Containing a β-Alanyl-γ-Aminobutyryl Segment. *J. Am. Chem. Soc.* **1997**, *119* (39), 9087–9095.

(60) M. Grison, C.; Robin, S.; J. Aitken, D. The Discovery of 9/8-Ribbons, β/γ-Peptides with Curved Shapes Governed by a Combined Configuration-Conformation Code. *Chem. Commun.* **2015**, *51* (90), 16233–16236.

(61) M. Grison, C.; Robin, S.; J. Aitken, D. 13-Helix Folding of a β/γ-Peptide Manifold Designed from a "Minimal-Constraint" Blueprint. *Chem. Commun.* **2016**, *52* (50), 7802–7805.

(62) Huang, Z. The Chemical Biology of Apoptosis: Exploring Protein-Protein Interactions and the Life and Death of Cells with Small Molecules. *Chem. Biol.* **2002**, *9* (10), 1059–1072.

(63) Delbridge, A. R.; Grabow, S.; Strasser, A.; Vaux, D. L. Thirty Years of BCL-2: Translating Cell Death Discoveries into Novel Cancer Therapies. *Nat Rev Cancer* **2016**, *16* (2), 99–109.

(64) Lee, E. F.; Sadowsky, J. D.; Smith, B. J.; Czabotar, P. E.; Peterson-Kaufman, K. J.; Colman, P. M.; Gellman, S. H.; Fairlie, W. D. High-Resolution Structural Characterization of a Helical α/β-Peptide Foldamer Bound to the Anti-Apoptotic Protein Bcl-XL. *Angew. Chem. Int. Ed.* **2009**, *48* (24), 4318–4322.

(65) Ernst, J. T.; Kutzki, O.; Debnath, A. K.; Jiang, S.; Lu, H.; Hamilton, A. D. Design of a Protein Surface Antagonist Based on α-Helix Mimicry: Inhibition of Gp41 Assembly and Viral Fusion. *Angew. Chem. Int. Ed.* **2002**, *41* (2), 278–281.

(66) Muchmore, S. W.; Sattler, M.; Liang, H.; Meadows, R. P.; Harlan, J. E.; Yoon, H. S.; Nettesheim, D.; Chang, B. S.; Thompson, C. B.; Wong, S.-L.; et al. X-Ray and NMR Structure of Human Bcl-XL, an Inhibitor of Programmed Cell Death. *Nature* **1996**, *381* (6580), 335.

(67) Sattler, M.; Liang, H.; Nettesheim, D.; Meadows, R. P.; Harlan, J. E.; Eberstadt, M.; Yoon, H. S.; Shuker, S. B.; Chang, B. S.; Minn, A. J.; et al. Structure of Bcl-XL-Bak Peptide Complex: Recognition between Regulators of Apoptosis. *Science* **1997**, *275* (5302), 983–986.

(68) Yu, N.; Aramini, J. M.; Germann, M. W.; Huang, Z. Reactions of Salicylaldehydes with Alkyl Cyanoacetates on the Surface of Solid Catalysts: Syntheses of 4H-Chromene Derivatives. *Tetrahedron Lett.* **2000**, *41* (36), 6993–6996.

(69) Degterev, A.; Lugovskoy, A.; Cardone, M.; Mulley, B.; Wagner, G.; Mitchison, T.; Yuan, J. Identification of Small-Molecule Inhibitors of Interaction between the BH3 Domain and Bcl-XL. *Nat. Cell Biol.* **2001**, *3* (2), 173–182.

(70) Enyedy, I. J.; Ling, Y.; Nacro, K.; Tomita, Y.; Wu, X.; Cao, Y.; Guo, R.; Li, B.; Zhu, X.; Huang, Y.; et al. Discovery of Small-Molecule Inhibitors of Bcl-2 through Structure-Based Computer Screening. *J. Med. Chem.* **2001**, *44* (25), 4313–4324.

(71) Sadowsky, J. D.; Schmitt, M. A.; Lee, H.-S.; Umezawa, N.; Wang, S.; Tomita, Y.; Gellman, S. H. Chimeric (α/β + α)-Peptide Ligands for the BH3-Recognition Cleft of Bcl-x_L: Critical Role of the Molecular Scaffold in Protein Surface Recognition. *J. Am. Chem. Soc.* **2005**, *127* (34), 11966–11968.

(72) Sadowsky, J. D.; Fairlie, W. D.; Hadley, E. B.; Lee, H.-S.; Umezawa, N.; Nikolovska-Coleska, Z.; Wang, S.; Huang, D. C. S.; Tomita, Y.; Gellman, S. H. (α/B+α)-Peptide Antagonists of BH3 Domain/Bcl-XL Recognition: Toward General Strategies for Foldamer-Based Inhibition of Protein−Protein Interactions. *J. Am. Chem. Soc.* **2007**, *129* (1), 139–154.

(73) Lee, E. F.; Smith, B. J.; Horne, W. S.; Mayer, K. N.; Evangelista, M.; Colman, P. M.; Gellman, S. H.; Fairlie, W. D. STRUCTURAL BASIS OF Bcl-XL RECOGNITION BY A BH3-MIMETIC α/β-PEPTIDE GENERATED VIA SEQUENCE-BASED DESIGN. *Chembiochem* **2011**, *12* (13), 2025–2032.

(74) Horne, W. S.; Boersma, M. D.; Windsor, M. A.; Gellman, S. H. Sequence-Based Design of α/β-Peptide Foldamers That Mimic BH3 Domains. *Angew. Chem. Int. Ed.* **2008**, *47* (15), 2853–2856.

(75) Horne, W. S.; Price, J. L.; Keck, J. L.; Gellman, S. H. Helix Bundle Quaternary Structure from α/β-Peptide Foldamers. *J. Am. Chem. Soc.* **2007**, *129* (14), 4178–4180.

(76) Boersma, M. D.; Haase, H. S.; Peterson-Kaufman, K. J.; Lee, E. F.; Clarke, O. B.; Colman, P. M.; Smith, B. J.; Horne, W. S.; Fairlie, W. D.; Gellman, S. H. Evaluation of Diverse α/β-Backbone Patterns for Functional α-Helix Mimicry: Analogues of the Bim BH3 Domain. *J. Am. Chem. Soc.* **2012**, *134* (1), 315–323.

(77) Voisin-Chiret, A. S.; Muraglia, M.; Burzicki, G.; Perato, S.; Corbo, F.; Sopková-de Oliveira Santos, J.; Franchini, C.; Rault, S. Synthesis of New Phenylpyridyl Scaffolds Using the Garlanding Approach. *Tetrahedron* **2010**, *66* (40), 8000–8005.

(78) De Giorgi, M.; Voisin-Chiret, A. S.; Sopková-de Oliveira Santos, J.; Corbo, F.; Franchini, C.; Rault, S. Design and Synthesis of Thienylpyridyl Garlands as Non-Peptidic Alpha Helix Mimetics and Potential Protein–Protein Interactions Disruptors. *Tetrahedron* **2011**, *67* (34), 6145–6154.

(79) Adler, M. J.; Hamilton, A. D. Oligophenylenaminones as Scaffolds for α-Helix Mimicry. *J. Org. Chem.* **2011**, *76* (17), 7040–7047.

(80) Yap, J. L.; Cao, X.; Vanommeslaeghe, K.; Jung, K.-Y.; Peddaboina, C.; Wilder, P. T.; Nan, A.; MacKerell, A. D.; Smythe, W. R.; Fletcher, S. Relaxation of the Rigid Backbone of an Oligoamide-Foldamer-Based α-Helix Mimetic: Identification of Potent Bcl-XL Inhibitors. *Org. Biomol. Chem.* **2012**, *10* (15), 2928.

(81) Souers, A. J.; Leverson, J. D.; Boghaert, E. R.; Ackler, S. L.; Catron, N. D.; Chen, J.; Dayton, B. D.; Ding, H.; Enschede, S. H.; Fairbrother, W. J.; et al. ABT-199, a Potent and Selective BCL-2 Inhibitor, Achieves Antitumor Activity While Sparing Platelets. *Nat. Med.* **2013**, *19* (2), 202–208.

(82) Itchaki, G.; Brown, J. R. The Potential of Venetoclax (ABT-199) in Chronic Lymphocytic Leukemia. *Ther. Adv. Hematol.* **2016**, *7* (5), 270–287.

(83) Wang, G. L.; Semenza, G. L. Purification and Characterization of Hypoxia-Inducible Factor 1. *J. Biol. Chem.* **1995**, *270* (3), 1230–1237.

(84) Hirota, K.; Semenza, G. L. Regulation of Angiogenesis by Hypoxia-Inducible Factor 1. *Crit. Rev. Oncol. Hematol.* **2006**, *59* (1), 15–26.

(85) Freedman, S. J.; Sun, Z.-Y. J.; Poy, F.; Kung, A. L.; Livingston, D. M.; Wagner, G.; Eck, M. J. Structural Basis for Recruitment of CBP/P300 by Hypoxia-Inducible Factor-1 Alpha. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99* (8), 5367–5372.

(86) Kushal, S.; Lao, B. B.; Henchey, L. K.; Dubey, R.; Mesallati, H.; Traaseth, N. J.; Olenyuk, B. Z.; Arora, P. S. Protein Domain Mimetics as in Vivo Modulators of Hypoxia-Inducible Factor Signaling. *Proc. Natl. Acad. Sci.* **2013**, *110* (39), 15602–15607.

(87) Lao, B. B.; Grishagin, I.; Mesallati, H.; Brewer, T. F.; Olenyuk, B. Z.; Arora, P. S. In Vivo Modulation of Hypoxia-Inducible Signaling by Topographical Helix Mimetics. *Proc. Natl. Acad. Sci.* **2014**, *111* (21), 7531–7536.

(88) Lao, B. B.; Drew, K.; Guarracino, D. A.; Brewer, T. F.; Heindel, D. W.; Bonneau, R.; Arora, P. S. Rational Design of Topographical Helix Mimics as Potent Inhibitors of Protein–Protein Interactions. *J. Am. Chem. Soc.* **2014**, *136* (22), 7877–7888.

(89) Burslem, G. M.; Kyle, H. F.; Breeze, A. L.; Edwards, T. A.; Nelson, A.; Warriner, S. L.; Wilson, A. J. Small-Molecule Proteomimetic Inhibitors of the HIF-1α-P300 Protein-Protein Interaction. *ChemBioChem* **2014**, *15* (8), 1083–1087.

(90) Burslem, G. M.; Kyle, H. F.; Breeze, A. L.; Edwards, T. A.; Nelson, A.; Warriner, S. L.; Wilson, A. J. Towards "Bionic" Proteins: Replacement of Continuous Sequences from HIF-1α with Proteomimetics to Create Functional P300 Binding HIF-1α Mimics †Electronic Supplementary Information (ESI) Available: Detailed Experimental Procedures and Characterisation; Additional Biophysical Data. See DOI: 10.1039/C6cc01812b Click Here for Additional Data File. *Chem. Commun. Camb. Engl.* **2016**, *52* (31), 5421–5424.

(91) Burslem, G. M.; Kyle, H. F.; Nelson, A.; Edwards, T. A.; Wilson, A. J. Hypoxia Inducible Factor (HIF) as a Model for Studying Inhibition of Protein–Protein Interactions. *Chem Sci* **2017**, *8* (6), 4188–4202.

(92) Chang, C.; Norris, J. D.; Grøn, H.; Paige, L. A.; Hamilton, P. T.; Kenan, D. J.; Fowlkes, D.; McDonnell, D. P. Dissection of the LXXLL Nuclear Receptor-Coactivator Interaction Motif Using Combinatorial Peptide Libraries: Discovery of Peptide Antagonists of Estrogen Receptors α and β. *Mol. Cell. Biol.* **1999**, *19* (12), 8226–8239.

(93) Ernst, J. T.; Becerril, J.; Park, H. S.; Yin, H.; Hamilton, A. D. Design and Application of an α-Helix-Mimetic Scaffold Based on an Oligoamide-Foldamer Strategy: Antagonism of the Bak BH3/Bcl-XL Complex. *Angew. Chem.* **2003**, *115* (5), 553–557.

(94) Becerril, J.; Hamilton, A. D. Helix Mimetics as Inhibitors of the Interaction of the Estrogen Receptor with Coactivator Peptides. *Angew. Chem. Int. Ed.* **2007**, *46* (24), 4471–4473.

(95) Ferrara, N.; Kerbel, R. S. Angiogenesis as a Therapeutic Target. *Nature* **2005**, *438* (7070), 967–974.

(96) Ferrara, N.; Hillan, K. J.; Gerber, H.-P.; Novotny, W. Case History: Discovery and Development of Bevacizumab, an Anti-VEGF Antibody for Treating Cancer. *Nat. Rev. Drug Discov.* **2004**, *3* (5), 391–400.

(97) Tarhini, A. A.; Frankel, P.; Margolin, K. A.; Christensen, S.; Ruel, C.; Shipe-Spotloe, J.; Gandara, D. R.; Chen, A.; Kirkwood, J. M. Aflibercept (VEGF Trap) in Inoperable Stage III or Stage IV Melanoma of Cutaneous or Uveal Origin. *Clin. Cancer Res.* **2011**, *17* (20), 6574–6581.

(98) Haase, H. S.; Peterson-Kaufman, K. J.; Lan Levengood, S. K.; Checco, J. W.; Murphy, W. L.; Gellman, S. H. EXTENDING FOLDAMER DESIGN BEYOND ?-HELIX MIMICRY: ?/?-PEPTIDE INHIBITORS OF VASCULAR ENDOTHELIAL GROWTH FACTOR SIGNALING. *J. Am. Chem. Soc.* **2012**, *134* (18), 7652–7655.

(99) Fairbrother, W. J.; Christinger, H. W.; Cochran, A. G.; Fuh, G.; Keenan, C. J.; Quan, C.; Shriver, S. K.; Tom, J. Y. K.; Wells, J. A.; Cunningham, B. C. Novel Peptides Selected to Bind Vascular Endothelial Growth Factor Target the Receptor-Binding Site. *Biochemistry* **1998**, *37* (51), 17754–17764.

(100) Horne, W. S.; Johnson, L. M.; Ketas, T. J.; Klasse, P. J.; Lu, M.; Moore, J. P.; Gellman, S. H. Structural and Biological Mimicry of Protein Surface Recognition by α/β-Peptide Foldamers. *Proc. Natl. Acad. Sci.* **2009**, *106* (35), 14751–14756.

(101) Checco, J. W.; Kreitler, D. F.; Thomas, N. C.; Belair, D. G.; Rettko, N. J.; Murphy, W. L.; Forest, K. T.; Gellman, S. H. Targeting Diverse Protein–Protein Interaction Interfaces with α/β-Peptides Derived from the Z-Domain Scaffold. *Proc. Natl. Acad. Sci.* **2015**, *112* (15), 4552–4557.

(102) Delsuc, N.; Léger, J.-M.; Massip, S.; Huc, I. Proteomorphous Objects from Abiotic Backbones. *Angew. Chem. Int. Ed.* **2007**, *46* (1–2), 214–217.

(103) Lee, B.-C.; Zuckermann, R. N.; Dill, K. A. Folding a Nonbiological Polymer into a Compact Multihelical Structure. *J. Am. Chem. Soc.* **2005**, *127* (31), 10999–11009.

(104) Price, J. L.; Steinkruger, J. D.; Gellman, S. H. Detection and Analysis of Chimeric Tertiary Structures by Backbone Thioester Exchange: Packing of an α Helix against an α/β-Peptide Helix. *Angew. Chem. Int. Ed Engl.* **2010**, *49* (2), 368–371.

(105) Qiu, J. X.; Petersson, E. J.; Matthews, E. E.; Schepartz, A. Toward β-Amino Acid Proteins: A Cooperatively Folded β-Peptide Quaternary Structure. *J. Am. Chem. Soc.* **2006**, *128* (35), 11338–11339.

(106) Petersson, E. J.; Schepartz, A. Toward β-Amino Acid Proteins: Design, Synthesis, and Characterization of a Fifteen Kilodalton β-Peptide Tetramer. *J. Am. Chem. Soc.* **2008**, *130* (3), 821–823.

(107) Watkins, A. M.; Wuo, M. G.; Arora, P. S. Protein–Protein Interactions Mediated by Helical Tertiary Structure Motifs. *J. Am. Chem. Soc.* **2015**, *137* (36), 11622–11630.

(108) Bergey, C. M.; Watkins, A. M.; Arora, P. S. HippDB: A Database of Readily Targeted Helical Protein–Protein Interactions. *Bioinformatics* **2013**, *29* (21), 2806–2807.
(109) Rose, A.; Meier, I. Scaffolds, Levers, Rods and Springs: Diverse Cellular Functions of Long Coiled-Coil Proteins. *Cell. Mol. Life Sci. CMLS* **2004**, *61* (16), 1996–2009.
(110) Woolfson, D. N. The Design of Coiled-Coil Structures and Assemblies. *Adv. Protein Chem.* **2005**, *70*, 79–112.
(111) Sawyer, N.; Watkins, A. M.; Arora, P. S. Protein Domain Mimics as Modulators of Protein–Protein Interactions. *Acc. Chem. Res.* **2017**, *50* (6), 1313–1322.
(112) Wuo, M. G.; Mahon, A. B.; Arora, P. S. An Effective Strategy for Stabilizing Minimal Coiled Coil Mimetics. *J. Am. Chem. Soc.* **2015**, *137* (36), 11618–11621.
(113) Sun, X.-J.; Wang, Z.; Wang, L.; Jiang, Y.; Kost, N.; Soong, T. D.; Chen, W.-Y.; Tang, Z.; Nakadai, T.; Elemento, O.; et al. A Stable Transcription Factor Complex Nucleated by Oligomeric AML1–ETO Controls Leukaemogenesis. *Nature* **2013**, *500* (7460), 93–97.
(114) Gallo, S. A.; Finnegan, C. M.; Viard, M.; Raviv, Y.; Dimitrov, A.; Rawat, S. S.; Puri, A.; Durell, S.; Blumenthal, R. The HIV Env-Mediated Fusion Reaction. *Biochim. Biophys. Acta BBA - Biomembr.* **2003**, *1614* (1), 36–50.
(115) Tan, K.; Liu, J.; Wang, J.; Shen, S.; Lu, M. Atomic Structure of a Thermostable Subdomain of HIV-1 Gp41. *Proc. Natl. Acad. Sci. U. S. A.* **1997**, *94* (23), 12303–12308.
(116) Matthews, T.; Salgo, M.; Greenberg, M.; Chung, J.; DeMasi, R.; Bolognesi, D. Case History: Enfuvirtide: The First Therapy to Inhibit the Entry of HIV-1 into Host CD4 Lymphocytes. *Nat. Rev. Drug Discov.* **2004**, *3* (3), 215–225.
(117) Arnold, U.; Hinderaker, M. P.; Nilsson, B. L.; Huck, B. R.; Gellman, S. H.; Raines, R. T. Protein Prosthesis: A Semisynthetic Enzyme with a β-Peptide Reverse Turn. *J. Am. Chem. Soc.* **2002**, *124* (29), 8522–8523.
(118) Nilsson, B. L.; Soellner, M. B.; Raines, R. T. Chemical Synthesis of Proteins. *Annu. Rev. Biophys. Biomol. Struct.* **2005**, *34*, 91–118.
(119) Holmgren, S. K.; Bretscher, L. E.; Taylor, K. M.; Raines, R. T. A Hyperstable Collagen Mimic. *Chem. Biol.* **1999**, *6* (2), 63–70.
(120) Reinert, Z. E.; Lengyel, G. A.; Horne, W. S. Protein-like Tertiary Folding Behavior from Heterogeneous Backbones. *J. Am. Chem. Soc.* **2013**, *135* (34), 12528–12531.
(121) Reinert, Z. E.; Horne, W. S. Protein Backbone Engineering as a Strategy to Advance Foldamers toward the Frontier of Protein-like Tertiary Structure. *Org Biomol Chem* **2014**, *12* (44), 8796–8802.
(122) George, K. L.; Horne, W. S. Heterogeneous-Backbone Foldamer Mimics of Zinc Finger Tertiary Structure. *J. Am. Chem. Soc.* **2017**, *139* (23), 7931–7938.
(123) Nord, K.; Nilsson, J.; Nilsson, B.; Uhlén, M.; Nygren, P. A. A Combinatorial Library of an Alpha-Helical Bacterial Receptor Domain. *Protein Eng.* **1995**, *8* (6), 601–608.
(124) Löfblom, J.; Feldwisch, J.; Tolmachev, V.; Carlsson, J.; Ståhl, S.; Frejd, F. Y. Affibody Molecules: Engineered Proteins for Therapeutic, Diagnostic and Biotechnological Applications. *FEBS Lett.* **2010**, *584* (12), 2670–2680.
(125) Nilsson, B.; Moks, T.; Jansson, B.; Abrahmsén, L.; Elmblad, A.; Holmgren, E.; Henrichson, C.; Jones, T. A.; Uhlén, M. A Synthetic IgG-Binding Domain Based on Staphylococcal Protein A. *Protein Eng. Des. Sel.* **1987**, *1* (2), 107–113.
(126) Fedorova, A.; Zobel, K.; Gill, H. S.; Ogasawara, A.; Flores, J. E.; Tinianow, J. N.; Vanderbilt, A. N.; Wu, P.; Meng, Y. G.; Williams, S.-P.; et al. The Development of Peptide-Based Tools for the Analysis of Angiogenesis. *Chem. Biol.* **2011**, *18* (7), 839–845.
(127) Braisted, A. C.; Wells, J. A. Minimizing a Binding Domain from Protein A. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93* (12), 5688–5692.

(128) Starovasnik, M. A.; Braisted, A. C.; Wells, J. A. Structural Mimicry of a Native Protein by a Minimized Binding Domain. *Proc. Natl. Acad. Sci.* **1997**, *94* (19), 10080–10085.
(129) Webster, J. M.; Zhang, R.; Gambhir, S. S.; Cheng, Z.; Syud, F. A. Engineered Two-Helix Small Proteins for Molecular Recognition. *ChemBioChem* **2009**, *10* (8), 1293–1296.
(130) Lindgren, M.; Hällbrink, M.; Prochiantz, A.; Langel, Ü. Cell-Penetrating Peptides. *Trends Pharmacol. Sci.* **2000**, *21* (3), 99–103.
(131) Schwarze, S. R.; Hruska, K. A.; Dowdy, S. F. Protein Transduction: Unrestricted Delivery into All Cells? *Trends Cell Biol.* **2000**, *10* (7), 290–295.
(132) Copolovici, D. M.; Langel, K.; Eriste, E.; Langel, Ü. Cell-Penetrating Peptides: Design, Synthesis, and Applications. *ACS Nano* **2014**, *8* (3), 1972–1994.
(133) Wender, P. A.; Mitchell, D. J.; Pattabiraman, K.; Pelkey, E. T.; Steinman, L.; Rothbard, J. B. The Design, Synthesis, and Evaluation of Molecules That Enable or Enhance Cellular Uptake: Peptoid Molecular Transporters. *Proc. Natl. Acad. Sci.* **2000**, *97* (24), 13003–13008.
(134) Foged, C.; Franzyk, H.; Bahrami, S.; Frokjaer, S.; Jaroszewski, J. W.; Nielsen, H. M.; Olsen, C. A. Cellular Uptake and Membrane-Destabilising Properties of α-Peptide/β-Peptoid Chimeras: Lessons for the Design of New Cell-Penetrating Peptides. *Biochim. Biophys. Acta BBA - Biomembr.* **2008**, *1778* (11), 2487–2495.
(135) Purkayastha, N.; Eyer, K.; Robinson, T.; Dittrich, P. S.; Beck, A. K.; Seebach, D.; Kolesinska, B.; Cadalbert, R. Enantiomeric and Diastereoisomeric (Mixed) L/ D-Octaarginine Derivatives – A Simple Way of Modulating the Properties of Cell-Penetrating Peptides. *Chem. Biodivers.* **2013**, *10* (7), 1165–1184.
(136) Seebach, D.; Namoto, K.; Mahajan, Y. R.; Bindschädler, P.; Sustmann, R.; Kirsch, M.; Ryder, N. S.; Weiss, M.; Sauer, M.; Roth, C.; et al. Chemical and Biological Investigations of β-Oligoarginines. *Chem. Biodivers.* **2004**, *1* (1), 65–97.
(137) Umezawa, N.; Gelman, M. A.; Haigis, M. C.; Raines, R. T.; Gellman, S. H. Translocation of a β-Peptide Across Cell Membranes. *J. Am. Chem. Soc.* **2002**, *124* (3), 368–369.
(138) Pasco, M.; Dolain, C.; Guichard, G. Foldamers in Medicinal Chemistry. In *Comprehensive Supramolecular Chemistry II*; Elsevier, 2017; pp 89–125.
(139) Iriondo-Alberdi, J.; Laxmi-Reddy, K.; Bouguerne, B.; Staedel, C.; Huc, I. Cellular Internalization of Water-Soluble Helical Aromatic Amide Foldamers. *ChemBioChem* **2010**, *11* (12), 1679–1685.
(140) Delsuc, N.; Kawanami, T.; Lefeuvre, J.; Shundo, A.; Ihara, H.; Takafuji, M.; Huc, I. Kinetics of Helix-Handedness Inversion: Folding and Unfolding in Aromatic Amide Oligomers. *ChemPhysChem* **2008**, *9* (13), 1882–1890.
(141) Harker, E. A.; Schepartz, A. Cell-Permeable β-Peptide Inhibitors of P53/HDM2 Complexation. *ChemBioChem* **2009**, *10* (6), 990–993.
(142) Daniels, D. S.; Schepartz, A. Intrinsically Cell-Permeable Miniature Proteins Based on a Minimal Cationic PPII Motif. *J. Am. Chem. Soc.* **2007**, *129* (47), 14578–14579.
(143) Douat, C.; Aisenbrey, C.; Antunes, S.; Decossas, M.; Lambert, O.; Bechinger, B.; Kichler, A.; Guichard, G. A Cell-Penetrating Foldamer with a Bioreducible Linkage for Intracellular Delivery of DNA. *Angew. Chem. Int. Ed.* **2015**, *54* (38), 11133–11137.
(144) Fischer, L.; Claudon, P.; Pendem, N.; Miclet, E.; Didierjean, C.; Ennifar, E.; Guichard, G. The Canonical Helix of Urea Oligomers at Atomic Resolution: Insights Into Folding-Induced Axial Organization. *Angew. Chem. Int. Ed.* **2009**, *49* (6), 1067–1070.
(145) Guichard, G.; Violette, A.; Chassaing, G.; Miclet, E. Solution Structure Determination of Oligoureas Using Methylene Spin State Selective NMR at 13C Natural Abundance. *Magn. Reson. Chem.* **2008**, *46* (10), 918–924.

(146) Fremaux, J.; Fischer, L.; Arbogast, T.; Kauffmann, B.; Guichard, G. Condensation Approach to Aliphatic Oligourea Foldamers: Helices with N-(Pyrrolidin-2-Ylmethyl)Ureido Junctions. *Angew. Chem.* **2011**, *123* (48), 11584–11587.

(147) Pendem, N.; Douat, C.; Claudon, P.; Laguerre, M.; Castano, S.; Desbat, B.; Cavagnat, D.; Ennifar, E.; Kauffmann, B.; Guichard, G. Helix-Forming Propensity of Aliphatic Urea Oligomers Incorporating Noncanonical Residue Substitution Patterns. *J. Am. Chem. Soc.* **2013**, *135* (12), 4884–4892.

(148) Violette, A.; Fournel, S.; Lamour, K.; Chaloin, O.; Frisch, B.; Briand, J.-P.; Monteil, H.; Guichard, G. Mimicking Helical Antibacterial Peptides with Nonpeptidic Folding Oligomers. *Chem. Biol.* **2006**, *13* (5), 531–538.

(149) Claudon, P.; Violette, A.; Lamour, K.; Decossas, M.; Fournel, S.; Heurtault, B.; Godet, J.; Mély, Y.; Jamart-Grégoire, B.; Averlant-Petit, M.-C.; et al. Consequences of Isostructural Main-Chain Modifications for the Design of Antimicrobial Foldamers: Helical Mimics of Host-Defense Peptides Based on a Heterogeneous Amide/Urea Backbone. *Angew. Chem. Int. Ed.* **2010**, *49* (2), 333–336.

(150) Kichler, A.; Leborgne, C.; März, J.; Danos, O.; Bechinger, B. Histidine-Rich Amphipathic Peptide Antibiotics Promote Efficient Delivery of DNA into Mammalian Cells. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (4), 1564–1568.

(151) Pichon, C.; Roufaï, M. B.; Monsigny, M.; Midoux, P. Histidylated Oligolysines Increase the Transmembrane Passage and the Biological Activity of Antisense Oligonucleotides. *Nucleic Acids Res.* **2000**, *28* (2), 504–512.

(152) Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter, P. Chromosomal DNA and Its Packaging in the Chromatin Fiber. **2002**.

(153) Fyodorov, D. V.; Zhou, B.-R.; Skoultchi, A. I.; Bai, Y. Emerging Roles of Linker Histones in Regulating Chromatin Structure and Function. *Nat. Rev. Mol. Cell Biol.* **2017**.

(154) Woodcock, C. L.; Frado, L. L.; Hatch, C. L.; Ricciardiello, L. Fine Structure of Active Ribosomal Genes. *Chromosoma* **1976**, *58* (1), 33–39.

(155) van Holde, K. E. *Chromatin*; Rich, A., Series Ed.; Springer Series in Molecular Biology; Springer New York: New York, NY, 1989.

(156) Hammond, C. M.; Strømme, C. B.; Huang, H.; Patel, D. J.; Groth, A. Histone Chaperone Networks Shaping Chromatin Function. *Nat. Rev. Mol. Cell Biol.* **2017**, *18* (3), 141–158.

(157) Ransom, M.; Dennehey, B. K.; Tyler, J. K. Chaperoning Histones during DNA Replication and Repair. *Cell* **2010**, *140* (2), 183–195.

(158) Belotserkovskaya, R.; Oh, S.; Bondarenko, V. A.; Orphanides, G.; Studitsky, V. M.; Reinberg, D. FACT Facilitates Transcription-Dependent Nucleosome Alteration. *Science* **2003**, *301* (5636), 1090–1093.

(159) Winkler, D. D.; Muthurajan, U. M.; Hieb, A. R.; Luger, K. Histone Chaperone FACT Coordinates Nucleosome Interaction through Multiple Synergistic Binding Events. *J. Biol. Chem.* **2011**, *286* (48), 41883–41892.

(160) Rocha, W.; Verreault, A. Clothing up DNA for All Seasons: Histone Chaperones and Nucleosome Assembly Pathways. *FEBS Lett.* **2008**, *582* (14), 1938–1949.

(161) De Koning, L.; Corpet, A.; Haber, J. E.; Almouzni, G. Histone Chaperones: An Escort Network Regulating Histone Traffic. *Nat. Struct. Mol. Biol.* **2007**, *14* (11), 997–1007.

(162) Keck, K. M.; Pemberton, L. F. Histone Chaperones Link Histone Nuclear Import and Chromatin Assembly. *Biochim. Biophys. Acta* **2012**, *1819* (3–4), 277–289.

(163) Dennehey, B. K.; Tyler, J. Histone Chaperones in the Assembly and Disassembly of Chromatin. In *Fundamentals of Chromatin*; Springer, New York, NY, 2014; pp 29–67.

(164) Chook, Y. M.; Süel, K. E. Nuclear Import by Karyopherin-Bs: Recognition and Inhibition. *Biochim. Biophys. Acta BBA - Mol. Cell Res.* **2011**, *1813* (9), 1593–1606.

(165) Mosammaparast, N.; Ewart, C. S.; Pemberton, L. F. A Role for Nucleosome Assembly Protein 1 in the Nuclear Transport of Histones H2A and H2B. *EMBO J.* **2002**, *21* (23), 6527–6538.
(166) Alvarez, F.; Muñoz, F.; Schilcher, P.; Imhof, A.; Almouzni, G.; Loyola, A. Sequential Establishment of Marks on Soluble Histones H3 and H4. *J. Biol. Chem.* **2011**, *286* (20), 17714–17721.
(167) Crystal Structure and Function of Human Nucleoplasmin (Npm2): A Histone Chaperone in Oocytes and Embryos http://pubs.acs.org/doi/pdf/10.1021/bi2006652 (accessed Oct 12, 2017).
(168) Gurard-Levin, Z. A.; Quivy, J.-P.; Almouzni, G. Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics. *Annu. Rev. Biochem.* **2014**, *83* (1), 487–517.
(169) Xu, M.; Long, C.; Chen, X.; Huang, C.; Chen, S.; Zhu, B. Partitioning of Histone H3-H4 Tetramers during DNA Replication-Dependent Chromatin Assembly. *Science* **2010**, *328* (5974), 94–98.
(170) Tyler, J. K.; Adams, C. R.; Chen, S.-R.; Kobayashi, R.; Kamakaka, R. T.; Kadonaga, J. T. The RCAF Complex Mediates Chromatin Assembly during DNA Replication and Repair. *Nature* **1999**, *402* (6761), 555–560.
(171) Tagami, H.; Ray-Gallet, D.; Almouzni, G.; Nakatani, Y. Histone H3.1 and H3.3 Complexes Mediate Nucleosome Assembly Pathways Dependent or Independent of DNA Synthesis. *Cell* **2004**, *116* (1), 51–61.
(172) Silljé, H. H. W.; Nigg, E. A. Identification of Human Asf1 Chromatin Assembly Factors as Substrates of Tousled-like Kinases. *Curr. Biol.* **2001**, *11* (13), 1068–1073.
(173) Daganzo, S. M.; Erzberger, J. P.; Lam, W. M.; Skordalakes, E.; Zhang, R.; Franco, A. A.; Brill, S. J.; Adams, P. D.; Berger, J. M.; Kaufman, P. D. Structure and Function of the Conserved Core of Histone Deposition Protein Asf1. *Curr. Biol.* **2003**, *13* (24), 2148–2158.
(174) Sharp, J. A.; Fouts, E. T.; Krawitz, D. C.; Kaufman, P. D. Yeast Histone Deposition Protein Asf1p Requires Hir Proteins and PCNA for Heterochromatic Silencing. *Curr. Biol.* **2001**, *11* (7), 463–473.
(175) Emili, A.; Schieltz, D. M.; Yates, J. R.; Hartwell, L. H. Dynamic Interaction of DNA Damage Checkpoint Protein Rad53 with Chromatin Assembly Factor Asf1. *Mol. Cell* **2001**, *7* (1), 13–20.
(176) Hu, F.; Alcasabas, A. A.; Elledge, S. J. Asf1 Links Rad53 to Control of Chromatin Assembly. *Genes Dev.* **2001**, *15* (9), 1061–1066.
(177) Murzina, N. V.; Pei, X.-Y.; Zhang, W.; Sparkes, M.; Vicente-Garcia, J.; Pratap, J. V.; McLaughlin, S. H.; Ben-Shahar, T. R.; Verreault, A.; Luisi, B. F.; et al. Structural Basis for the Recognition of Histone H4 by the Histone-Chaperone RbAp46. *Structure* **2008**, *16* (7), 1077–1085.
(178) Ask, K.; Jasencakova, Z.; Menard, P.; Feng, Y.; Almouzni, G.; Groth, A. Codanin-1, Mutated in the Anaemic Disease CDAI, Regulates Asf1 Function in S-Phase Histone Supply. *EMBO J.* **2012**, *31* (8), 2013–2023.
(179) Groth, A.; Ray-Gallet, D.; Quivy, J.-P.; Lukas, J.; Bartek, J.; Almouzni, G. Human Asf1 Regulates the Flow of S Phase Histones during Replicational Stress. *Mol. Cell* **2005**, *17* (2), 301–311.
(180) Groth, A.; Rocha, W.; Verreault, A.; Almouzni, G. Chromatin Challenges during DNA Replication and Repair. *Cell* **2007**, *128* (4), 721–733.
(181) Clément, C.; Almouzni, G. MCM2 Binding to Histones H3-H4 and ASF1 Supports a Tetramer-to-Dimer Model for Histone Inheritance at the Replication Fork. *Nat. Struct. Mol. Biol.* **2015**, *22* (8), 587–589.
(182) Richet, N.; Liu, D.; Legrand, P.; Velours, C.; Corpet, A.; Gaubert, A.; Bakail, M.; Moal-Raisin, G.; Guerois, R.; Compper, C.; et al. Structural Insight into How the Human Helicase Subunit MCM2 May Act as a Histone Chaperone Together with ASF1 at the Replication Fork. *Nucleic Acids Res.* **2015**, *43* (3), 1905–1917.
(183) Wang, H.; Wang, M.; Yang, N.; Xu, R.-M. Structure of the Quaternary Complex of Histone H3-H4 Heterodimer with Chaperone ASF1 and the Replicative Helicase Subunit MCM2. *Protein Cell* **2015**, *6* (9), 693–697.

(184) English, C. M.; Adkins, M. W.; Carson, J. J.; Churchill, M. E. A.; Tyler, J. K. Structural Basis for the Histone Chaperone Activity of Asf1. *Cell* **2006**, *127* (3), 495–508.
(185) Han, J.; Zhou, H.; Li, Z.; Xu, R.-M.; Zhang, Z. Acetylation of Lysine 56 of Histone H3 Catalyzed by RTT109 and Regulated by ASF1 Is Required for Replisome Integrity. *J. Biol. Chem.* **2007**, *282* (39), 28587–28596.
(186) Fillingham, J.; Greenblatt, J. F. A Histone Code for Chromatin Assembly. *Cell* **2008**, *134* (2), 206–208.
(187) Li, Q.; Zhou, H.; Wurtele, H.; Davies, B.; Horazdovsky, B.; Verreault, A.; Zhang, Z. Acetylation of Histone H3 Lysine 56 Regulates Replication-Coupled Nucleosome Assembly. *Cell* **2008**, *134* (2), 244–255.
(188) Burgess, R. J.; Zhang, Z. Histone Chaperones in Nucleosome Assembly and Human Disease. *Nat. Struct. Mol. Biol.* **2013**, *20* (1), 14–22.
(189) Yamane, K.; Mizuguchi, T.; Cui, B.; Zofall, M.; Noma, K.; Grewal, S. I. S. Asf1/HIRA Facilitate Global Histone Deacetylation and Associate with HP1 to Promote Nucleosome Occupancy at Heterochromatic Loci. *Mol. Cell* **2011**, *41* (1), 56–66.
(190) Zhang, R.; Poustovoitov, M. V.; Ye, X.; Santos, H. A.; Chen, W.; Daganzo, S. M.; Erzberger, J. P.; Serebriiskii, I. G.; Canutescu, A. A.; Dunbrack, R. L.; et al. Formation of MacroH2A-Containing Senescence-Associated Heterochromatin Foci and Senescence Driven by ASF1a and HIRA. *Dev. Cell* **2005**, *8* (1), 19–30.
(191) Daniel Ricketts, M.; Frederick, B.; Hoff, H.; Tang, Y.; Schultz, D. C.; Singh Rai, T.; Grazia Vizioli, M.; Adams, P. D.; Marmorstein, R. Ubinuclein-1 Confers Histone H3.3-Specific-Binding by the HIRA Histone Chaperone Complex. *Nat. Commun.* **2015**, *6* (1).
(192) Schwabish, M. A.; Struhl, K. Asf1 Mediates Histone Eviction and Deposition during Elongation by RNA Polymerase II. *Mol. Cell* **2006**, *22* (3), 415–422.
(193) Natsume, R.; Eitoku, M.; Akai, Y.; Sano, N.; Horikoshi, M.; Senda, T. Structure and Function of the Histone Chaperone CIA/ASF1 Complexed with Histones H3 and H4. *Nature* **2007**, *446* (7133), 338–341.
(194) Mousson, F.; Lautrette, A.; Thuret, J.-Y.; Agez, M.; Courbeyrette, R.; Amigues, B.; Becker, E.; Neumann, J.-M.; Guerois, R.; Mann, C.; et al. Structural Basis for the Interaction of Asf1 with Histone H3 and Its Functional Implications. *Proc. Natl. Acad. Sci.* **2005**, *102* (17), 5975–5980.
(195) Agez, M.; Chen, J.; Guerois, R.; van Heijenoort, C.; Thuret, J.-Y.; Mann, C.; Ochsenbein, F. Structure of the Histone Chaperone Asf1 Bound to the Histone H3 C-Terminal Helix and Functional Insights. *Structure* **2007**, *15* (2), 191–199.
(196) Esteller, M. Cancer Epigenomics: DNA Methylomes and Histone-Modification Maps. *Nat. Rev. Genet.* **2007**, *8* (4), 286–298.
(197) Shortt, J.; Ott, C. J.; Johnstone, R. W.; Bradner, J. E. A Chemical Probe Toolbox for Dissecting the Cancer Epigenome. *Nat. Rev. Cancer* **2017**, *17* (3), 160–183.
(198) Oh, J.; Ruskoski, N.; Fraser, N. W. Chromatin Assembly on Herpes Simplex Virus 1 DNA Early during a Lytic Infection Is Asf1a Dependent. *J. Virol.* **2012**, *86* (22), 12313–12321.
(199) Polo, S. E.; Theocharis, S. E.; Klijanienko, J.; Savignoni, A.; Asselain, B.; Vielh, P.; Almouzni, G. Chromatin Assembly Factor-1, a Marker of Clinical Value to Distinguish Quiescent from Proliferating Cells. *Cancer Res.* **2004**, *64* (7), 2371–2381.
(200) Feser, J.; Truong, D.; Das, C.; Carson, J. J.; Kieft, J.; Harkness, T.; Tyler, J. K. Elevated Histone Expression Promotes Life Span Extension. *Mol. Cell* **2010**, *39* (5), 724–735.
(201) O'Sullivan, R. J.; Karlseder, J. Telomeres: Protecting Chromosomes against Genome Instability. *Nat. Rev. Mol. Cell Biol.* **2010**, *11* (3), 171–181.

(202) Heaphy, C. M.; Wilde, R. F. de; Jiao, Y.; Klein, A. P.; Edil, B. H.; Shi, C.; Bettegowda, C.; Rodriguez, F. J.; Eberhart, C. G.; Hebbar, S.; et al. Altered Telomeres in Tumors with ATRX and DAXX Mutations. *Science* **2011**, *333* (6041), 425–425.

(203) Corpet, A.; De Koning, L.; Toedling, J.; Savignoni, A.; Berger, F.; Lemaître, C.; O'Sullivan, R. J.; Karlseder, J.; Barillot, E.; Asselain, B. Asf1b, the Necessary Asf1 Isoform for Proliferation, Is Predictive of Outcome in Breast Cancer. *EMBO J.* **2011**, *30* (3), 480–493.

(204) Liang, X.; Yuan, X.; Yu, J.; Wu, Y.; Li, K.; Sun, C.; Li, S.; Shen, L.; Kong, F.; Jia, J.; et al. Histone Chaperone ASF1A Predicts Poor Outcomes for Patients With Gastrointestinal Cancer and Drives Cancer Progression by Stimulating Transcription of β-Catenin Target Genes. *EBioMedicine* **2017**, *21*, 104–116.

(205) Im, J.-S.; Keaton, M.; Lee, K. Y.; Kumar, P.; Park, J.; Dutta, A. ATR Checkpoint Kinase and CRL1βTRCP Collaborate to Degrade ASF1a and Thus Repress Genes Overlapping with Clusters of Stalled Replication Forks. *Genes Dev.* **2014**, *28* (8), 875–887.

(206) Miknis, G. F.; Stevens, S. J.; Smith, L. E.; Ostrov, D. A.; Churchill, M. E. A. Development of Novel Asf1–H3/H4 Inhibitors. *Bioorg. Med. Chem. Lett.* **2015**, *25* (4), 963–968.

(207) Seol, J.-H.; Song, T.-Y.; Oh, S. E.; Jo, C.; Choi, A.; Kim, B.; Park, J.; Hong, S.; Song, I.; Jung, K. Y.; et al. Identification of Small Molecules That Inhibit the Histone Chaperone Asf1 and Its Chromatin Function. *BMB Rep.* **2015**, *48* (12), 685–690.

(208) Burgess, K.; Shin, H.; Linthicum, D. S. Solid-Phase Syntheses of Unnatural Biopolymers Containing Repeating Urea Units. *Angew. Chem. Int. Ed. Engl.* **1995**, *34* (8), 907–909.

(209) Kumler, W. D.; Fohlen, G. M. The Dipole Moment and Structure of Urea and Thiourea1. *J. Am. Chem. Soc.* **1942**, *64* (8), 1944–1948.

(210) Hemmerlin, C.; Marraud, M.; Rognan, D.; Graff, R.; Semetey, V.; Briand, J.-P.; Guichard, G. Helix-Forming Oligoureas: Temperature-Dependent NMR, Structure Determination, and Circular Dichroism of a Nonamer with Functionalized Side Chains. *Helv. Chim. Acta* **2002**, *85* (11), 3692–3711.

(211) Violette, A.; Averlant-Petit, M. C.; Semetey, V.; Hemmerlin, C.; Casimir, R.; Graff, R.; Marraud, M.; Briand, J.-P.; Rognan, D.; Guichard, G. N,N'-Linked Oligoureas as Foldamers: Chain Length Requirements for Helix Formation in Protic Solvent Investigated by Circular Dichroism, NMR Spectroscopy, and Molecular Dynamics. *J. Am. Chem. Soc.* **2005**, *127* (7), 2156–2164.

(212) Fischer, L.; Guichard, G. Folding and Self-Assembly of Aromatic and Aliphatic Urea Oligomers: Towards Connecting Structure and Function. *Org. Biomol. Chem.* **2010**, *8* (14), 3101.

(213) Collie, G. W.; Pulka-Ziach, K.; Lombardo, C. M.; Fremaux, J.; Rosu, F.; Decossas, M.; Mauran, L.; Lambert, O.; Gabelica, V.; Mackereth, C. D.; et al. Shaping Quaternary Assemblies of Water-Soluble Non-Peptide Helical Foldamers by Sequence Manipulation. *Nat. Chem.* **2015**, *7* (11), 871–878.

(214) Fremaux, J.; Mauran, L.; Pulka-Ziach, K.; Kauffmann, B.; Odaert, B.; Guichard, G. α-Peptide-Oligourea Chimeras: Stabilization of Short α-Helices by Non-Peptide Helical Foldamers. *Angew. Chem.* **2015**, *127* (34), 9954–9958.

(215) Mauran, L.; Kauffmann, B.; Odaert, B.; Guichard, G. Stabilization of an α-Helix by Short Adjacent Accessory Foldamers. *Comptes Rendus Chim.* **2016**.

(216) Kim, J.-M.; Bi, Y.; Paikoff, S. J.; Schultz, P. G. The Solid Phase Synthesis of Oligoureas. *Tetrahedron Lett.* **1996**, *37* (30), 5305–5308.

(217) Guichard, G.; Semetey, V.; Rodriguez, M.; Briand, J.-P. Solid Phase Synthesis of Oligoureas Using O-Succinimidyl-(9H-Fluoren-9-Ylmethoxycarbonylamino) Ethylcarbamate Derivatives as Activated Monomers. *Tetrahedron Lett.* **2000**, *41* (10), 1553–1557.

(218) Boeijen, A.; van Ameijde, J.; Liskamp, R. M. J. Solid-Phase Synthesis of Oligourea Peptidomimetics Employing the Fmoc Protection Strategy. *J. Org. Chem.* **2001**, *66* (25), 8454–8462.

(219) Douat-Casassus, C.; Pulka, K.; Claudon, P.; Guichard, G. Microwave-Enhanced Solid-Phase Synthesis of N , N′ -Linked Aliphatic Oligoureas and Related Hybrids. *Org. Lett.* **2012**, *14* (12), 3130–3133.

(220) García-Martín, F.; Quintanar-Audelo, M.; García-Ramos, Y.; Cruz, L. J.; Gravel, C.; Furic, R.; Côté, S.; Tulla-Puche, J.; Albericio, F. ChemMatrix, a Poly(Ethylene Glycol)-Based Support for the Solid-Phase Synthesis of Complex Peptides. *J. Comb. Chem.* **2006**, *8* (2), 213–220.

.

Chapter II: A model system to investigate α-helix mimicry with peptide/oligourea hybrids

During my PhD work, we sought to evaluate the structural consequences of inserting a short urea segment in a peptide sequence known to adopt a stable α-helical conformation. Recent investigations in the group have indeed revealed that peptide and oligourea can be assembled together to form one unique helix in organic solvents[1] and aqueous environment as well (data not published) (Chapter 1 C.3.). To extend this work and study the extent to which the α-helix can be mimicked by an oligourea insert, we thought that it would be interesting to apply this concept to peptides that have the capacity to self-assemble into defined tertiary structures like leucine-zipper domains. To do so, we have selected a model peptide that has been extensively studied in the literature: the yeast transcription factor GCN4 that would allow us to study the impact of a short urea segment insertion (two to three units) within a quite long α-peptide (> 30 residues) on the overall helicity as well as on the stability of the oligomerization states.

A. Peptides derived from the Leucine zipper domain of the transcription factor GCN4

1. A parallel two-stranded coiled-coil

a) The structure of GCN4 is based on heptad repeat

The yeast transcription factor GCN4 regulates gene expression and can activate more than 30 genes required for the biosynthesis of α-amino acids. The 33-residue long GCN4 leucine zipper (GCN4-p1) is an α-helical coiled-coil presenting an *abcdefg* heptad repeat. Positions *a* and *d* are occupied by hydrophobic residues and polar residues generally elsewhere (Figure 40, XOOXOOO where X is hydrophobic and O is polar). An amphiphilic α-helix assembly is then obtained with hydrophobic stripes along one face of the helix.

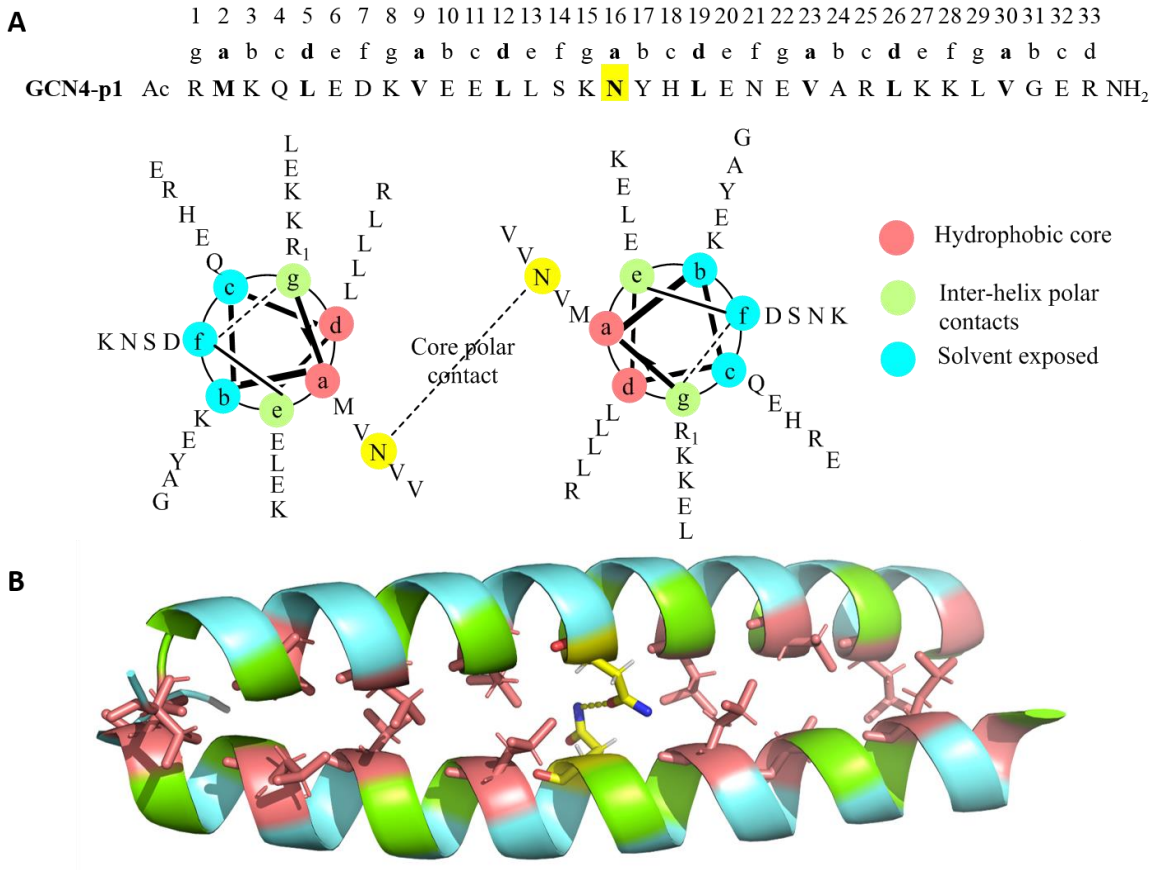


Figure 40: Sequence and structure of GCN4-p1 dimer. A. Primary sequence of GCN4-p1 and helical wheel projection of the coiled-coil. **B**. X-ray crystal structure of GCN4-p1 dimer (pdb: 4dmd) at 2.0 Å resolution, in cartoon representation. Hydrophobic residues are depicted in stick and the polar contact between the two Asn residues is highlighted in yellow

The α-residue composition and shape of these hydrophobic interfaces are crucial for the helical packing and for coiled coil folding. In GCN4-p1, *a* and *d* heptad positions are mainly occupied by Val and Leu residues respectively. Interestingly, an Asn residue is found at the central *a* position (position 16) in GCN4-p1 sequence and this residue plays a key role in the folding process by triggering coiled coil assembling through the formation of an interchain polar contact in the hydrophobic core of the dimer (see Figure 40A, helical wheel projection representation). GCN4 was one of the first coiled-coil proteins to be crystallized and for which the X-ray crystal structure has been solved at 1.8Å[2]. This x-ray crystal structure shows that the peptide forms a parallel, two-stranded coiled coil of α-helices that are packed as in the "knobs-into-holes" model proposed in 1953 by Crick (Figure 40B[3]). Today, 157 X-ray crystal structures associated to GCN4 sequences are referenced in pubmed website. The wild type sequence GCN4-p1 generally form a right-handed homodimer with a parallel orientation[4].

b) Fine tuning of the stoichiometry by sequence modulation

Mutations in the sequence of GCN4 are known to change its oligomerization state[4]. In this leucine zipper coiled coil, Leu residues are present at about 80% of all *d* positions. Substitutions of these leucine with other hydrophobic residues may influence the overall structure of the resulting coiled–coils. In a seminal work, Alber, Kim and coworkers have

designed GCN4 mutants by altering the hydrophobic core and characterized their structures in solution and in the solid state[5]. Mutations at *a*/*d* heptad positions have been simultaneously performed at four *a* positions and four *d* positions along GCN4-p1 sequence (Met2 remained unchanged). Circular dichroism experiments showed that each mutant remained α-helical and present a midpoint of the thermal transition (T_m) higher than that of the GCN4-p1 native sequence. These peptide mutants were found to exist as dimeric, trimeric and tetrameric states in the crystal (Table 3).

Table 3 : Sequences of mutant peptides with T_m and folding associated

Compound	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	Tm (°C)	No of helices
	g	**a**	b	c	**d**	e	f	g	**a**	b	c	**d**	e	f	g	**a**	b	c	**d**	e	f	g	**a**	b	c	**d**	e	f	g	**a**	b	c	d		
GCN4-p1	Ac	R	**M**	K	Q	**L**	E	D	K	**V**	E	E	**L**	L	S	K	**N**	Y	H	**L**	E	N	E	**V**	A	R	**L**	K	K	L	**V**	G	E	R	NH_2	53	2
II.1	Ac	R	**M**	K	Q	**L**	E	D	K	**I**	E	E	**L**	L	S	K	**I**	Y	H	**L**	E	N	E	**I**	A	R	**L**	K	K	L	**I**	G	E	R	NH_2	>100	2
II.2	Ac	R	**M**	K	Q	**I**	E	D	K	**I**	E	E	**I**	L	S	K	**I**	Y	H	**I**	E	N	E	**I**	A	R	**I**	K	K	L	**I**	G	E	R	NH_2	>100	3
II.3	Ac	R	**M**	K	Q	**I**	E	D	K	**L**	E	E	**I**	L	S	K	**L**	Y	H	**I**	E	N	E	**L**	A	R	**I**	K	K	L	**L**	G	E	R	NH_2	>100	4
II.4	Ac	R	**M**	K	Q	**V**	E	D	K	**L**	E	E	**V**	L	S	K	**L**	Y	H	**V**	E	N	E	**L**	A	R	**V**	K	K	L	**L**	G	E	R	NH_2	81	3
II.5	Ac	R	**M**	K	Q	**L**	E	D	K	**V**	E	E	**L**	L	S	K	**V**	Y	H	**L**	E	N	E	**V**	A	R	**L**	K	K	L	**V**	G	E	R	NH_2	95	2,3
II.6	Ac	R	**M**	K	Q	**L**	E	D	K	**L**	E	E	**L**	L	S	K	**L**	Y	H	**L**	E	N	E	**L**	A	R	**L**	K	K	L	**L**	G	E	R	NH_2	>100	3

Interestingly, the occurrence of Ile residues at *a* positions conserved a dimer topology (**II.1**) whereas when located at *d* positions this residue disfavors dimeric state and leads to tetramers (**II.3**). At the contrary, the presence of Ile residues at both *a* and *d* positions induced trimeric assembly. This study illustrates that the distribution of β-branched residues on *a* and *d* positions governs the coiled-coil oligomerization process. Concurrently, the substitution of Asn_{16} with Val results in a mixed population of dimers and trimers even though the T_m exceeded the one recorded for GCN4-p1. This central polar residue appeared to be critical for driving the dimeric structure by forming buried hydrogen bonds.

More recently, Horne and Oshaben have designed a series of dimeric coiled coil peptides derived from GCN4-p1 and exhibiting an identical quaternary folded structures[6]. The mutants of GCN4-p1 were modulated to improve the stability of the dimeric assembly (Figure 41). As mentioned above, the composition of the hydrophobic core appeared to be critical for the quaternary structure stability. Val residues at *a* position in the GCN4-p1 heptad sequence were mutated by Ile residues in order to increase the stability of the dimeric folds. Each sequence contains one Val to Ile mutation (**II.7**), two (**II.8**) or three mutations (**II.9**). The authors maintained the presence of Asn_{16} all along the series of mutants despite its polar nature as it helps for the specific obtaining of a dimer oligomerization.

A

Compound	Primary Sequence	Tm (°C)
	1 5 10 15 20 25 30	
	g **a** b c **d** e f g **a** b c **d** e f g **a** b c **d** e f g **a** b c **d** e f g **a** b c d	
GCN4-p1	Ac R **M** K Q **L** E D K **V** E E **L** L S K **N** Y H **L** E N E **V** A R **L** K K L **V** G E R NH_2	62
II.7	Ac R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **V** A R **L** K K L **V** G E R NH_2	69
II.8	Ac R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **I** A R **L** K K L **V** G E R NH_2	73
II.9	Ac R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	78

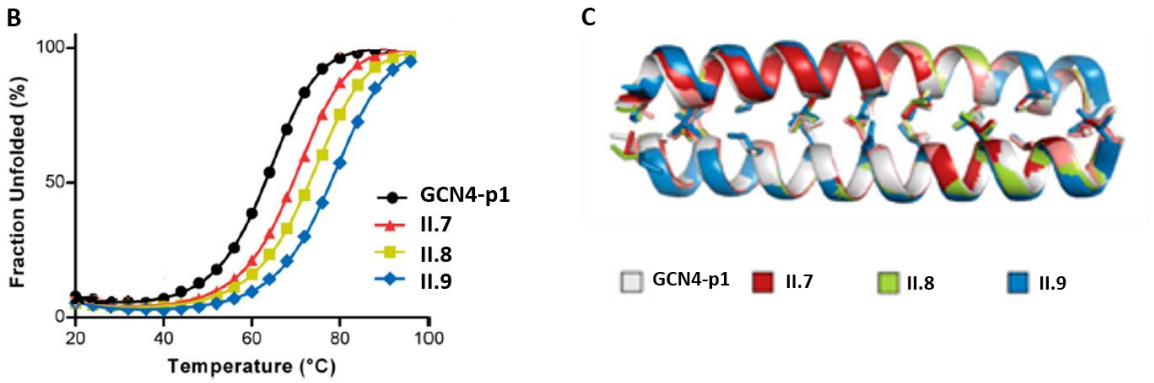


Figure 41: Sequence and CD curves of mutant peptides of GCN4-p1. A. The primary sequence of GCN4-p1 and hydrophobic core mutants (**II.7**-**II.9**) with their corresponding $T_m$values. **B.** CD thermal melts of GCN4-p1 and mutants at 100µM in 10mM pH 7 phosphate buffer[6]. **C.** Comparison of the X-ray crystal structures of peptides **II.7**-**II-9**. Overlay of residues 2-30 with calculated backbone Root-Mean Square Deviations to peptide GCN4-p1 of 0.32Å for **II.7** and **II.8** and 0.77Å for **II.9**.

CD experiments were carried out on the different sequences and the resulting CD curves confirmed that all the mutants retained an α-helical fold similar to that of the native GCN4-p1 sequence and CD thermal melts curves showed that each peptide mutant has a similar cooperative unfolding transition corresponding to transition of a folded to unfolded state. All the peptide or protein folding process results from a cooperative effect[7]. Indeed, amino acids cooperate and create intramolecular interactions to drive the peptide to fold. For all peptide mutants, the shape of the CD thermal melt curves is identical, with a sigmoid profile indicating that when the peptide starts to unfold, all the structuration process is fast. It allows the peptide to pass quickly to a folded state to an unfolded one. The T_m value was determined as the temperature corresponding to 50% of fraction unfolded. The best T_m value (78 °C) was obtained for the sequence **II.9** (encompassing three Val → Ile substitutions) that is higher than the T_m of GCN4-p1 obtained with the same experimental conditions (T_m= 62 °C). The T_m increased as the number of Val→Ile mutation increased. These results highlight that mutations of residues involved in the hydrophobic core do not really disrupt the overall helicity of the resulting coiled-coils but influence the overall stability of the helix assemblies and consequently of the quaternary structure. The authors obtained X-ray crystal structures of mutants that highlight the parallel chain topology. Furthermore, this structural analysis on GCN4-p1 analogues confirms that overall the packing of the hydrophobic core remained conserved for all mutants.

Besides modulation of the amino acid sequence that influences the folding behavior of GCN4 coiled coil, previous study by Horne and coworkers showed that the environment has also an impact on the folding[3]. Biophysical studies suggested a dimeric and a trimeric form of GCN4-p1 depending on the environment, hypothesis that was further confirmed by the high-resolution crystal structure obtained[3]. The importance of the environmental conditions highlights the complexity of the folding behavior GCN4 (that is still considered as a simplest coiled-coil).

These results could be potentially applied to design new oligomers deriving from native GCN4-p1 sequence and predict their folding. They are important to take into account for future studies that will focus on one oligomerization state. Concurrently, side chains modifications to design GCN4 mutant peptides have underlined the possibility to tune the quaternary structure but the modification of the peptide backbone in these systems still remains challenging.

2. α/β peptides to mimic GCN4 coiled-coil

Horne, Gellman and coworkers studied the structural consequences of β^3-amino acid insertions on the self-assembly process of GCN4 peptide sequences[8]. They have developed a "bottom-up" approach by performing a systematic $\alpha \rightarrow \beta^3$ replacements at *b* and *f* positions in the native GCN4-p1sequence that induced substantial changes in the overall quaternary structure of the resulting GCN4-α/β-hybrid foldamer (**II.10**)[8]. Despite the fact that the incorporated β^3-residues were located at solvent exposed positions, CD analysis at 100 μM concentration of the α/β-peptide sequence **II.10** suggested that only little helix formation was still present. Nonetheless, the authors managed to grow crystals and the X-ray crystal structure revealed that α/β-peptide **II.10** formed a parallel helix bundle in the solid state with each bundle containing three α/β-peptide molecules. Interestingly, although the presence of β^3-residues alters the overall stoichiometry of the self-association, each subunit retained an α-helix-like conformation regardless the additional CH_2 per helical turn. To recover a stable and defined quaternary structure, the authors next selected **II.3** as a model where Leu and Ile residues were installed at most *a* position at *d* positions respectively (Table 1). The α- to β^3-amino acid replacement was again performed at *b* and *f* positions in the sequence of **II.3**[8]. In this case, the X-ray crystal structure solved at a resolution of 2.0 Å revealed that the resulting hybrid peptide **II.11**, self-assemble in a four-helix bundle like the parent peptide as shown the. Interestingly, a difference in the self-association behavior was again observed between the solution and the study in the solid state: while the data in solution were in favor of a trimeric bundle, the X-ray crystal structure revealed a stoichiometry of four molecules in parallel orientation mode. These results illustrate that altering the backbone while keeping the native side chains does not necessarily allow a complete mimicry of the α-peptide structure stoichiometry.

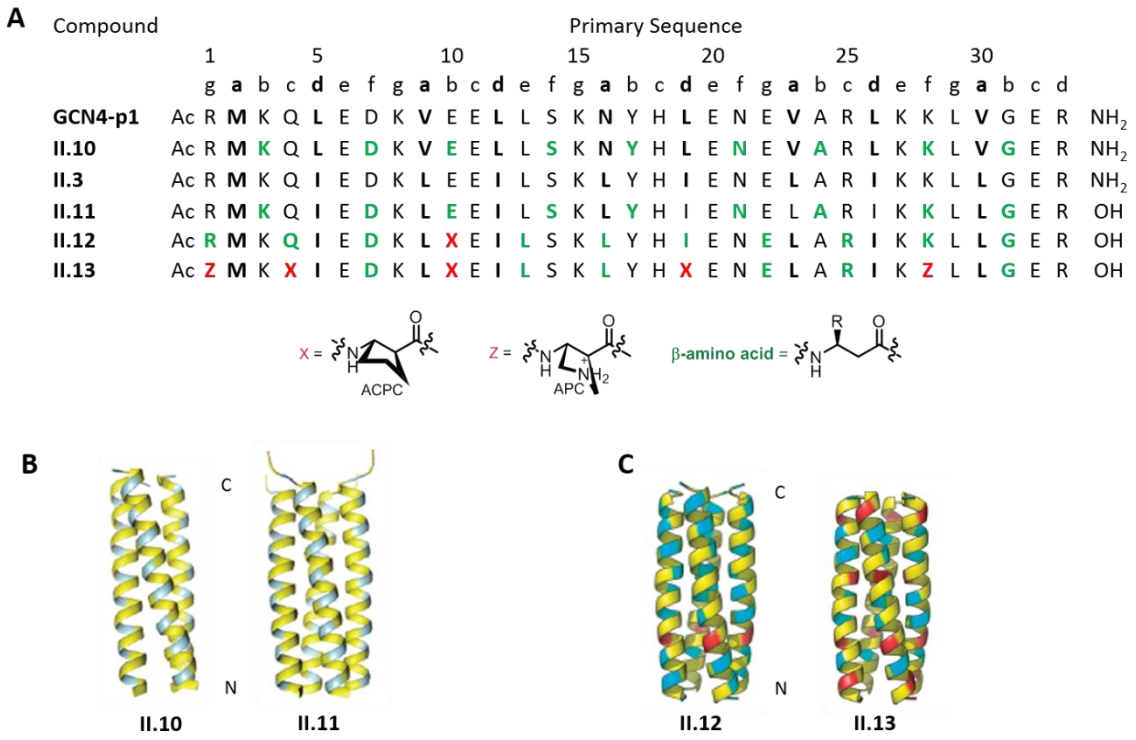


Figure 42: Sequence of peptide foldamers with their crystal structure[7,8]**.** A. Amino acid sequences of α/β-peptides (**II.10-II.13**) and α-peptides GCN4-p1 and **II.3**. Hydrophobic *a* and *d* position residues are highlighted with boldface type. α-amino acids are abbreviated with the one letter code and β^3-amino acids are highlighted in green. Cyclic β-amino acids are in red with the following abbreviation: X=ACPC, Z=APC. B. Crystal structure of α/β-peptide **II.10** (PDB: 2OXJ) with β^3-amino acids in cyan. C. Crystal structures of α/β-peptides **II.12** (PDB: 3C3G), and **II.13** (PDB: HEY).

Later on, they reported new α/β-peptide foldamer (**II.12**-**II.13**) analogues of **II.3** containing regular replacement of α→β-residues with an ααβ repeat pattern (Figure 42)[9]. Because each α→β^3 replacement brings an additional flexible bond to the backbone, cyclic β-amino acids residues were introduced at selected positions to constrain the backbone and enhance the helix stability. Pyrrolidine-based residue (APC) were used to replace basic β-residues whereas ACPC were used to replace β^3-residues with nonpolar chains. **II.12** contains one cyclic residue and in **II.13**, five cyclic residues out of the eleven β-positions. The high-resolution crystal structures solved for the two compounds confirmed that they both display a tetrameric quaternary structure as the parent α-peptide **II.3**. CD spectra show that the α/β-peptide **II.13** displays high levels of folding and assembly. For **II.12,** the introduction of only one cyclic residue destabilizes the structure and causes changes in the self-association stoichiometry as **II.12** is found as monomer, tetramer and pentamer in aqueous solution[9]. More recently, the same group evaluated the effect of introducing γ-residues in α/β-peptide helices and showed that the γ-residue preorganization is less important than the one of β-residue which is essential for helix stability[10]. Indeed, the helical propensity of γ-residue was not optimally suited to the α-helix mimetic conformation.

To conclude, local α-residue to β-residue replacements have been used extensively by the Gellman group to explore the effect of residue homologation on α-helix bundle formation. Although regularly spaced β-residue are α-helix compatible, this work has also revealed the subtle impact of such backbone modifications on the physical behavior of the corresponding

α/β-peptide hybrids. In my PhD work, we wanted to evaluate the possibility of inserting a short urea segment in place of an α-peptide one while maintaining the overall folding properties of the resulting α/urea-hybrid oligomer compared to the cognate peptide. To do so, we considered replacing several consecutive α- positions with a urea segment, herein referred as the "block foldamer" approach in the GCN4-p1 model sequence. We hoped that this strategy would allow us to determine the consequences at different structural levels of peptide→oligourea replacements and delineate the requirements for systematic oligourea replacement within α-peptides and α-helix mimicry.

B. Design of α/urea chimeras to mimic the quaternary structure of GCN4 – Molecular modeling studies

1. Combine amino acid and urea residues in one unique helix

Combining the non-natural backbone of foldamers and α-peptides in a single chain is promising in order to modulate the tertiary structure of proteins[11,12]. Because of physical and structural similarities between the α-helix and the oligourea 2.5-helix such as polarity (carbonyl orientation) and pitch[13,14], the insertion of urea segment(s) into a peptide thus seems feasible. As mentioned in the previous chapter, recent studies in our group showed that the introduction of aliphatic oligourea at the C- or N-terminus of a short α-peptide can propagate the helical conformation[1]. The high-resolution structural data reported show a continuous intramolecular H-bond network allowing the combination of the α-helix in the peptide segment and a canonical 2.5-helix in the oligourea segment (Figure 43[1]).

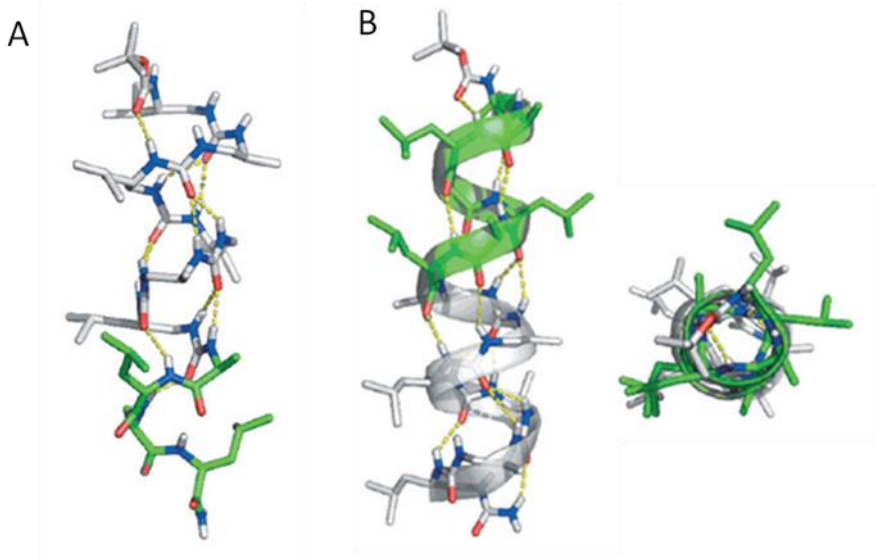


Figure 43: X-ray crystal structure of two oligourea/peptide chimeras. A) X-ray crystal structure of a chimera with the peptide part (green) and the introduction of a urea cap at the *N*-terminus side (grey) B) X-ray crystal structure of a chimera with the peptide part in green and the introduction of a urea cap at the *C*-terminus with the lateral view on the left and corresponding top view on the right.

The α-helix and the 2.5-helix of oligourea are connected together through a unique hydrogen network (C=O…H-N; *i*, *i*+3 and *i*, *i*+4). At the junction, hydrogen bonds occur between carbonyl and NHs of the two types of residues. However, the insertion of urea segment may cause

changes in the peptide part of the chimera. Parameters have indeed been calculated with the Helanal program[15] using $^{\alpha}$C carbon of amino acid residues only (Table 4). Results reveal few differences between the two peptide parts of the helix with first the urea segment located at the N-terminus and second at the C-terminus (**A** and **B** in Figure 43 respectively). The peptide segment of the helix **A** displays almost the same parameters than a regular α-helix whereas the helix **B** shares value of Ψ angle which is similar to that of the 3_{10}-helix as well as the H-bonding pattern (*i*, *i*+3 C=O…H-N).

Table 4: Average of the peptide parameters and torsion angles obtained from X-ray crystal structure of helices A and B calculated with the Helenal program using $^{\alpha}$C carbons of amino acid residues only[1]. A comparison between the α- and 3_{10}-helices is presented.

	Helix A	Helix B	α-helix	3_{10}-helix
(Φ, Ψ) [°]	(-66, -40)	(-69, -29)	(-63, -42)	(-57, -30)
Residues per turn	3.6	3.6	3.63	3.24
Radius [Å]	2.3	2.2	2.3	
Unit height [Å]	1.5	1.6	1.56	1.94
Rise per turn [Å]	5.4	5.8	5.4	6.29

This study illustrates the compatibility of α-peptides with oligoureas and the feasibility of the formation of hybrid helix peptide/oligourea. Two distinct segments were connected together and the resulting hybrid chimeras retain properties of both species. Moreover, the introduction of short urea caps could nucleate and stabilize the formation of an α-helical conformation[1]. However, the original study has been carried out in organic solvent and aqueous solution remains a more challenging environment. To have foldamers exhibiting water-solubilizing side-chains on their surface is indeed mandatory for biological applications. In order to develop a general approach to mimicry peptide folding and self-association, the insertion of an oligourea segment has been studied in model peptide GCN4-p1.

We first reasoned on the possibility to substitute one α-helix turn. As a first step, we thought to replace three consecutive amino acids by two consecutive urea units (diad) that actually corresponds to a little less than a full turn (3.6 residues per turn in α-helix and 2.5 residues per turn in oligourea helix). In a second step, we wanted to replace four consecutive α-residues by three urea units (triad) corresponding this time to a little bit more than a full turn of α-helix. The superimposition of four amino acids by three urea residues illustrates that for one full α-helical turn, the size is almost similar (Figure 44). In contrast to our previous studies mentioned above, we decided to insert urea diads or triads in the middle of the sequence, thus generating three distinct segments (αααX^{u}X^{u}ααα). Due to the importance of the hydrophobic effect, involving residues at positions *a* and *d* for the self-assembly of coiled-coil, we have decided to substitute residues at positions *e*, *f* and *g* to start. These positions have indeed the advantage to present three consecutive amino acids that are

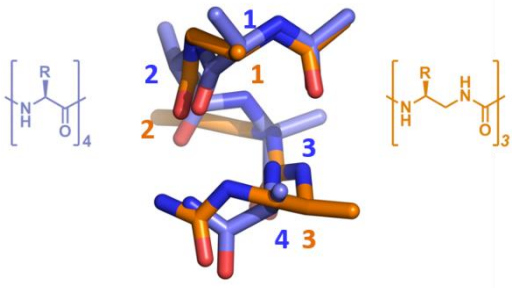


Figure 44: Overlay of 4 amino acids in α-helical conformation in blue and 3 urea residues in orange

oriented toward solvent. These amino acids could be replaced by a urea diad and may have less impact on the overall self-assembly. As starting point, we decided to substitute the α-amino acid residues at the central position, Leu13Ser14Lys15, by two urea units.

2. Substituting 2 urea units (diad) for 3 consecutive α amino acid residues

Starting from the crystal structure of GCN4-p1, a molecular modeling study has been performed by Jean Dessolin and Remy Baily (@ CBMN-Modelling of biomolecules and numerical imaging team) with the insertion of a short oligourea segment replacing three consecutive α-amino acid residues. A replacement of three α-amino acids (Leu13Ser14Lys15) at *e*, *f* and **g** central position by two urea residues encompassing a methyl side chain (noted A^u) was modeled with a classic dynamic study under Amber force fields. The superimposition of GCN4-p1 and its modified version GCN4-p1-2A^u (**II.14:** RMKQLEDKVEELA^{u}A^{u}NYHLENEVARLKKLVGER) was done (Figure 45).

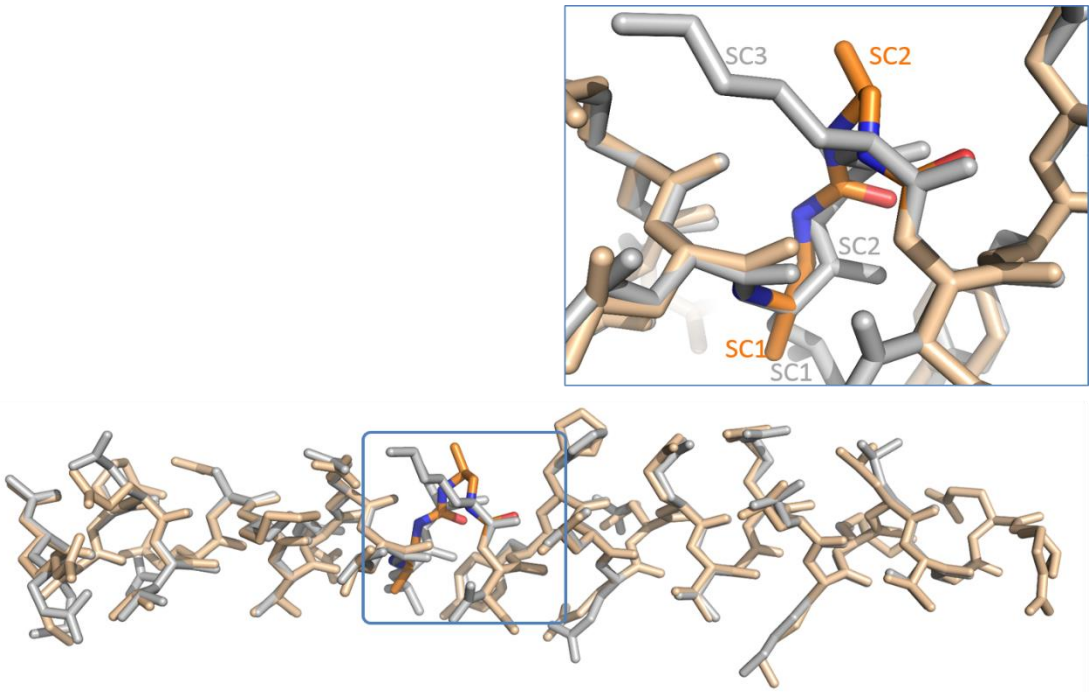


Figure 45: Superimposition of GCN4-p1 (grey) and GCN4-p1-2A^u (model) where the LSK segment was replaced by two urea residues (orange).

Swapping three consecutive residues of an α-turn with a urea diad appears to be well-tolerated by the α-helix. With two urea residues replacing three amino acids (LSK), only two lateral chains can be mimicked. The position and orientation of the methyl side chains of the two urea residues seem to match reasonably well with those of Leu13 and Lys15 side-chains respectively. However, it could be interesting to see if mimicry can be improved by shifting the side-chain corresponding to the Lys residue from the β-carbon to the α-position of the urea backbone. However, to maintain a proper screw sense of the helix in the urea segment, the configuration at this α-carbon has to be inverted to a (*R*) configuration (all *L*-amino-acids are (*S*) except L-cysteine which is (R))[13].

The replacement of these residues at *e*, *f* and *g* positions might not have a huge impact on the hydrophobic core and self-association process because the selected residues have their side-

chains exposed to the solvent. Moreover, we hypothesized that a modification in the central position would not disturb too much the nucleation of the helix because less than one helical turn has been replaced.

3. Substituting 3 urea units (triad) for 4 consecutive α amino acid residues

Another molecular modeling study has been performed, this time with the replacement of four consecutive α-amino-acids by 3 urea residues. The superimposition of GCN4-p1 and GCN4-p1-3A^{u} (with three urea residues, **II.15:** RMKQLEDKVEEA^{u}A^{u}A^{u}NYHLENEVARLKKLVGER) was made for direct comparison of side chain orientation (Figure 46). In comparison with the first in silico study, here the objective was to replace the entire turn of an α-helix by a full turn of 2.5 oligourea helix. In the same manner as presented above, the urea segment was introduced in the middle of the α-peptide sequence. The fit between GCN4-p1 and **II.15** was more challenging than for the previous model, illustrating a more drastic change due to the introduction of the triad.

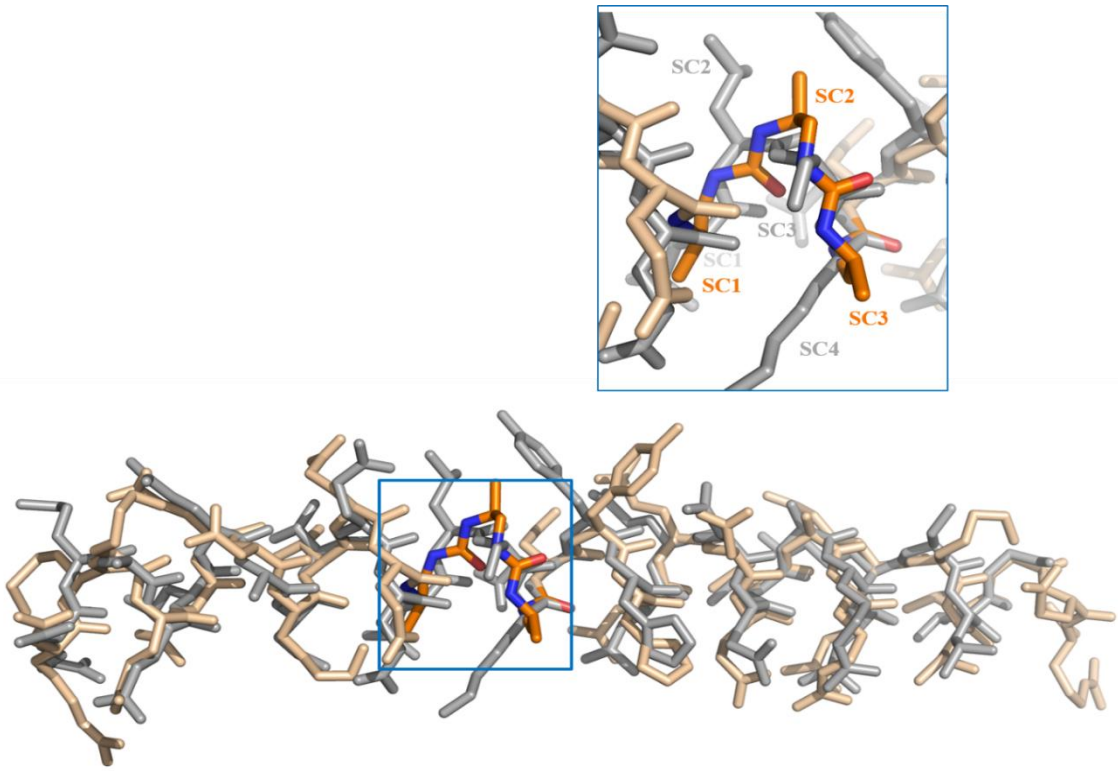


Figure 46: Superimposition of GCN4-p1 in grey and a model of GCN4-p1-4A^{u}. The dimer of GCN4-p1 is represented in grey with one helix in surface and the other one superimposed with the GCN4-p1 modified in light pink and urea residues in orange. A zoom on the urea segment superimposition is shown and the side chains 1, 2 and 3 (noted SC1, SC2 and SC3) are localized.

The Leu12Leu13Ser14Lys15 segment was replaced by three Ala^{u} residues and after minimization, the superimposition of the two sequences gave several clues regarding the positioning and orientation of the side chains of urea-type units. In contrast with the previous modelling (three α to two Ala^{u} mutations), this time a hydrophobic α-residue, the Leu12, is swapped by a urea residue. The Leu12 side-chain in GCN4-p1 is indeed involved in a hydrophobic contact with its counterpart located on the other oligomer. It is therefore important to reproduce this key interaction with the first urea side chain by keeping the isobutyl side-chain on the CHβ. Next, based on the result of the modelling, the side chain SC2

in orange in Figure 46 of the second urea unit appear to mimic reasonably well the side chain orientation of the third α-residue chain (Ser14) and the third side chain SC3 of the oligourea segment (in orange) appears to be close enough to the side chain of the fourth α–Lys15 to expect a similar projection of the linear butylamine chain. Consequently, the triad of urea units Leu^{u}-Ser^{u}-Lys^{u} could potentially replace the Leu-Leu-Ser-Lys α-peptide segment in an attempt to mimic the 3.6-helical turn.

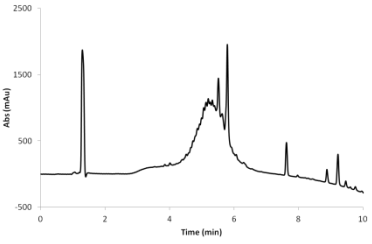
C. Local replacement of consecutive α amino acids residues by ureido units

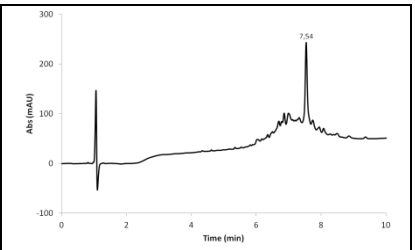
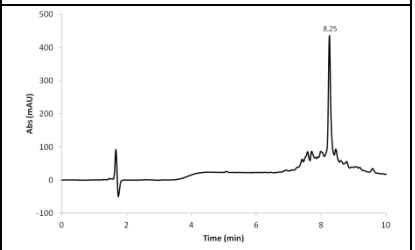
After bibliographic studies, we decided to work with the sequence **II.9**, the mutant peptide of GCN4-p1 that was reported to better stabilize the dimeric assembly[6]. The sequence of **II.9** was used as reference for direct comparison with sequences containing urea replacements. The first step was then to synthesize the peptide and to perform CD thermal melting experiments to compare our results with the literature.

1. Synthesis and optimization of the reference peptide II.9

With the aim to unveil potential synthetic difficulties, we first decided to reproduce the solid phase synthesis (SPS) of the reference GCN4-VI peptide[6]. The synthesis of this reference peptide **II.9** was thus performed by using classical *N*-Fmoc chemistry on commercially available low loading TFA-labile resins in order to reduce as much as possible aggregation issues. Moreover, all SPS were performed under microwave assistance to increase coupling and deprotection reaction rates and get higher purity of the crudes. Besides, during my PhD project, the group has acquired an automated microwave peptide synthesizer, the *Liberty-Blue* from CEM that allows a high reproducibility of the coupling/deprotection cycles and pilots the dispensing of the different reagents within the reaction vessel all along the synthesis.
In a first attempt, the synthesis of **II.9** was carried out on Rink amide NOVAPEG resin (loading 0.46 mmol.g^{-1}) in a 0.050 mmol scale. Iterative couplings of *N*-Fmoc-protected amino acids were done twice in presence of (Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium (BOP) and *N,N'*-diisopropylethylamine (DIEA) as coupling reagents (Table 5). After final capping by acetylation of the main chain, the peptide was released from the support by performing a TFA cleavage using a *"cocktail mixture"* composed of TFA/TIS/H_2O/EDT: (92.5:2.5:2.5:2.5, v/v/v/v).

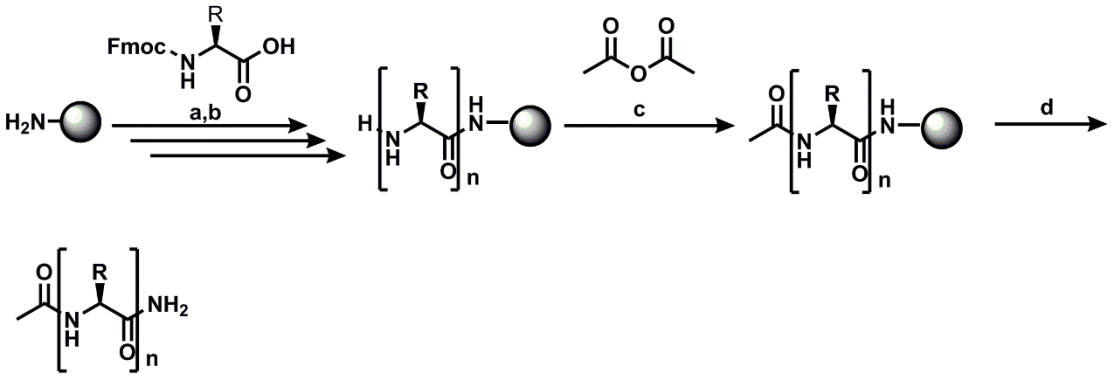
Table 5: Comparison of three syntheses of II.9 with different resins and different coupling reagents

Synthesis	Resin	Coupling agents	Coupling conditions	Profile
SII.1	NOVAPEG (0.46 mmol.g^{-1})	BOP/DIEA	2 x 75°C/30W/5min	

SII.2	Polystyrene LL (0.37 mmol.g^{-1})	BOP/DIEA	2 x 75°C/30W/5min	
SII.3	Polystyrene LL (0.37 mmol.g^{-1})	HBTU/HOBt /DIEA	2 x 75°C/30W/5min	

The peptide was then precipitated with cold Et_2O and the crude material was directly analyzed by analytical RP-HPLC. The HPLC analysis revealed a crowded profile with the presence of a main peak that was further analyzed by mass spectrometry (ESI positive mode) and corresponded to **II.9** with a relatively low purity of 28% (see Table 3 entry 1). The presence of so many impurities surrounding the main peak rendered its purification by semi-preparative RP-HPLC extremely tricky with limited recovery of the expected reference peptide **II.9** in sufficient quantity for structural study. Importantly, the main objective of this study being to introduce oligourea segment in the model sequence, the synthesis of the peptide segments must be as good as possible because the introduction of the urea units within the sequence could bring anticipated synthesis issues in particular partial azide reduction of the first introduced urea unit (**SII.1**).

We next repeated the SPS of **II.9** by replacing the NOVAPEG resin by the more conventional Rink amide MBHA resin. We have indeed already observed that this polymer matrix gave much better results for oligourea synthesis than the NOVAPEG one and this even if this matrix encompassing a PEG shell was reported to give better peptide syntheses in comparison to polystyrene matrix[16,17]. This second synthesis of **II.9** (**SII.2**) was carried out following the same coupling/deprotection cycle conditions as those described in Scheme 1. The RP-HPLC chromatogram after TFA cleavage and Et_2O precipitation showed this time the presence of a major peak with a purity of 42% and much less impurities surrounding it (Table 5). As expected, ESI-MS analysis confirmed that the main peak corresponded to the peptide **II.9**. Changing the resin matrix clearly improved the overall synthesis of **II.9**, underlining here a positive impact of a non-polar polystyrene resin compared to a PEG one.



a: Fmoc -Xaa-OH (5 eq), HBTU (6 eq) HOBt (6 eq), DIEA (12 eq), 75°C (except for Fmoc-His(Trt)-OH at 50°C), 30 W, 5 min; b: 20% piperidine, DMF (90°C/35W/50s) c: Acetic anhydride/ DCM 1:1 (15 min, r.t.); d: TFA/TIS/H_2O/EDT (92.5:2.5:2.5:2.5, v,v,v,v)

Scheme 1: General scheme for the solid phase synthesis of peptides including **II.9**

In a last synthesis optimization, we decided to evaluate the use of another coupling reagent than BOP which tends to disappear from peptide chemical suppliers. Instead of using *in situ* phosphonium-type coupling reagent, we applied N,N,N',N'-Tetramethyl-O-(1H-benzotriazol-1-yl) uronium hexafluorophosphate (HBTU), an uronium-based coupling reagent in combination with 1-Hydroxybenzotriazole (HOBt), commonly used to reduce the risk of epimerization during the condensation cycle by enhancing the coupling reaction rates. The DIEA was kept as base during the course of this third SPS (**SII.3**). After TFA cleavage and peptide precipitation, the HPLC chromatogram showed the presence of a major peak corresponding to **II.9** with a good purity of 70%, significantly higher than the previous synthesis (Figure 47). The full peptide **II.9** was purified by semi-preparative HPLC and obtained with an overall yield of 10%.

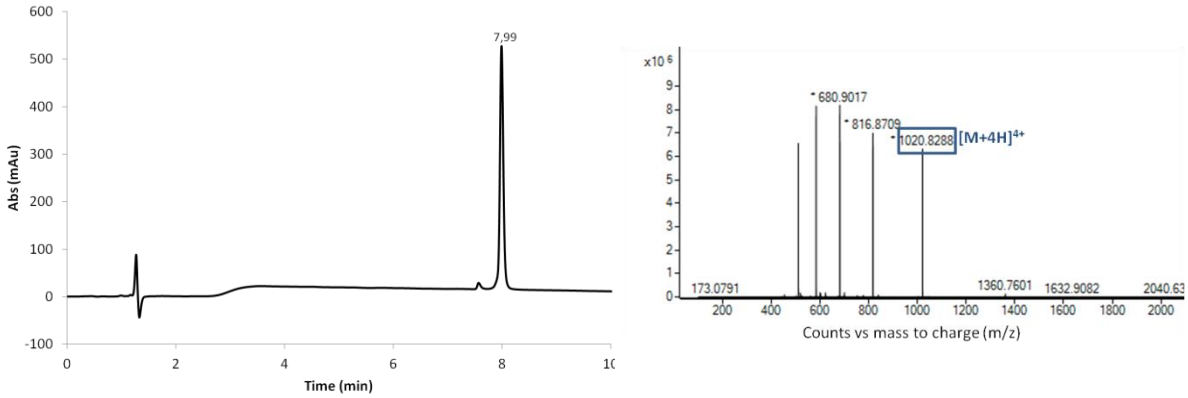


Figure 47: RP-HPLC of II.9 pure with a gradient of 10-100% ACN in water and mass spectrometry of the peptide.

The synthesis of the peptide **II.9** has been optimized and the peptide was obtained in a good yield paving the way to a possible introduction of urea units within the backbone of **II.9**. It is important to note that in the literature (including supporting information), it is difficult to find detailed experimental information about the synthesis, purification and yields of GCN4-p1 and

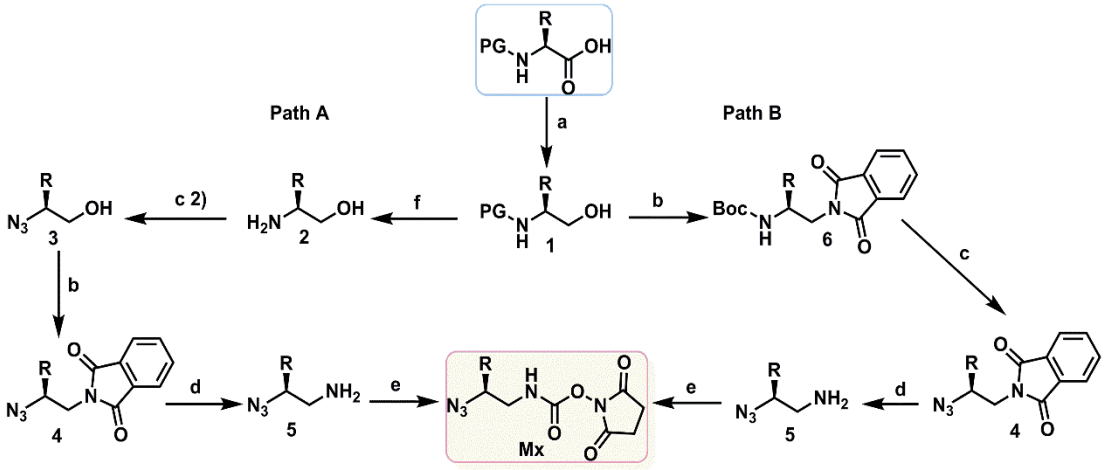
related peptide analogues. Then, it is difficult to compare our yields with what has been already published in the literature.

2. Synthesis of urea type monomers and solid phase chimera synthesis

a) Substituting 2 urea units (diad) for 3 consecutive α-residues

i) Solution synthesis of azide monomers activated as succinimidyl carbamates

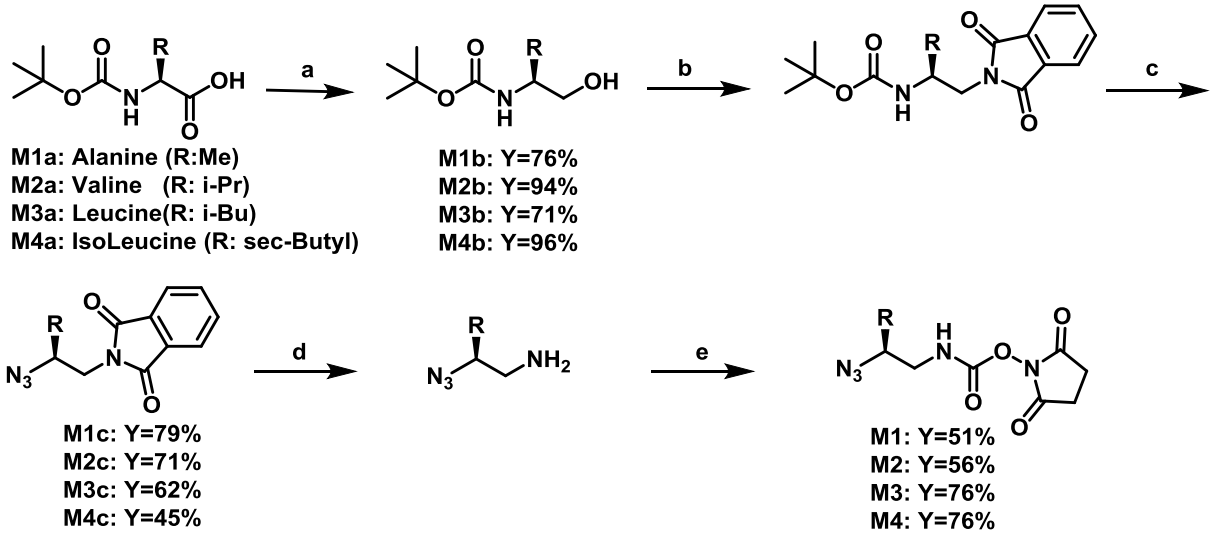
Few years ago, the group has developed a new solid phase synthesis methodology of oligoureas and related hybrids compatible with the use of TFA-labile resins[16]. The main objective of this strategy was to preclude the use of the Boc strategy that requires a final HF cleavage which is not compatible with a daily use of a solid phase approach and raise safety issues. To do so, a new set of azide-type building blocks were developed where the azide group was used as a masked amine allowing the use of orthogonal side-chain protections that can be removed during the TFA cleavage. This approach is now routinely used in the laboratory and most natural but also non-natural amino acid side-chains can be appended on azide-type monomers. Depending on the nature of the side-chain, two optimized synthesis routes were developed to give access to the different building blocks[16] (Scheme 2). In path A, commercially available protected α-amino acids (Boc, Fmoc or Z) are first reduced into alcohols **1**, deprotected before conversion into the corresponding azide **3** by using a recently reported "shelf-stable" diazo transfer reagent: imidazole-1-sulfonyl azide hydrochloride[18–20]. The alcohol is then converted into a phthalimide intermediate **4** through a Mitsunobu reaction and subsequently reduced into the corresponding amine **5** in presence of hydrazine hydrate. The final step consists in the activation of the resulting primary amine with *N'N'*-disuccinimidyl carbonate (DSC). During the preparation of the first series of azide building blocks it has been observed that this route was not compatible with short aliphatic side-chains (*i.e*. R= CH_3, $CH(CH_3)_2$, CH_2-$CH(CH_3)_2$, -$(CH_2)_3$-CH_3) because the azido alcohol intermediates are too volatile to guarantee correct overall yield. Consequently, the route B was envisioned and mainly consists in performing the Mistunobu reaction before the conversion of the amine into azide.



a: 1) NMM, IBCF, THF, -20°C; 2) $NaBH_4$, H_2O; b: PPh_3, Phtalimide, DIAD, THF, 25°C; c: 1) TFA; 2) $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4{:}5H_2O$, $CH_3CN{:}H_2O$(1:1), 25°C; d: $N_2H_4.H_2O$, MeOH, 70°C,e: DSC, CH_2Cl_2, 25°C; f: TFA or DBU, 25°C.

Scheme 2: General procedure for the preparation of succinimidyl (2-azido-2-substituted ethyl) carbamate (azide monomers) **Mx** with two routes depending on the nature of the side chain.

With the aim to substitute several consecutive α-residues by urea ones in the sequence of **II.9**, we needed to have in hands several activated azide-type monomers. Firstly, activated carbamates deriving from Ala, Val, Leu and Ile have been synthesized through the route B (Scheme 3) starting from Boc-*L*-Ala-OH, Boc-*L*-Val-OH, Boc-*L*-Leu-OH and Boc-*L*-Ile-OH respectively. Activated monomers N_3-Ala^U-OSu (**M1)**, N_3-Val^U-OSu, (**M2)**, N_3-Leu^U-OSu (**M3**) and N_3-Ile^U-OSu (**M4)** were obtained with 28%, 37%, 33% and 32% overall yield respectively.

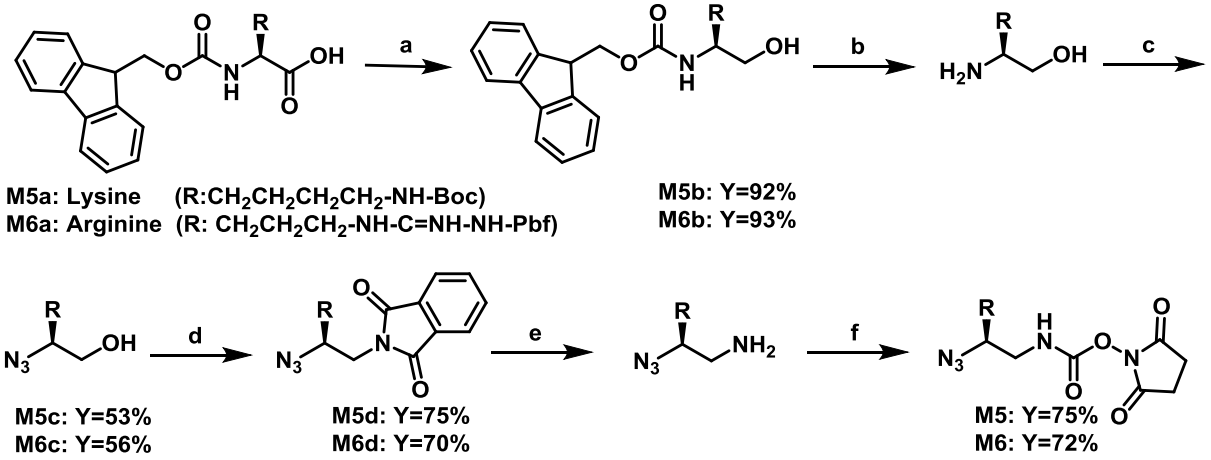


a: 1) NMM, IBCF, THF, -20°C; 2) $NaBH_4$, H_2O; b: PPH_3, Phthalimide, DIAD, THF, r.t;
c: 1) TFA; 2) $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4{:}5H_2O$, $CH_3CN{:}H_2O$(1:1), r.t; d: $N_2H_4.H_2O$, MeOH, 70°C,e: DSC, CH_2Cl_2,r.t.

Scheme 3 : Synthesis of succinimidyl (2-azido-2-substituted ethyl) carbamate M1-M4 starting from *N*-Boc-protected amino acids.

Secondly, succinimidyl (2-azido-2-substituted ethyl) carbamate with positively charged side chains have been synthesized starting from corresponding *N*-Fmoc-*L*-Lys(Boc)-OH and *N*-Fmoc-*L*-Arg(Pbf)-OH following route A (Scheme 4). To note removal of Fmoc protecting group

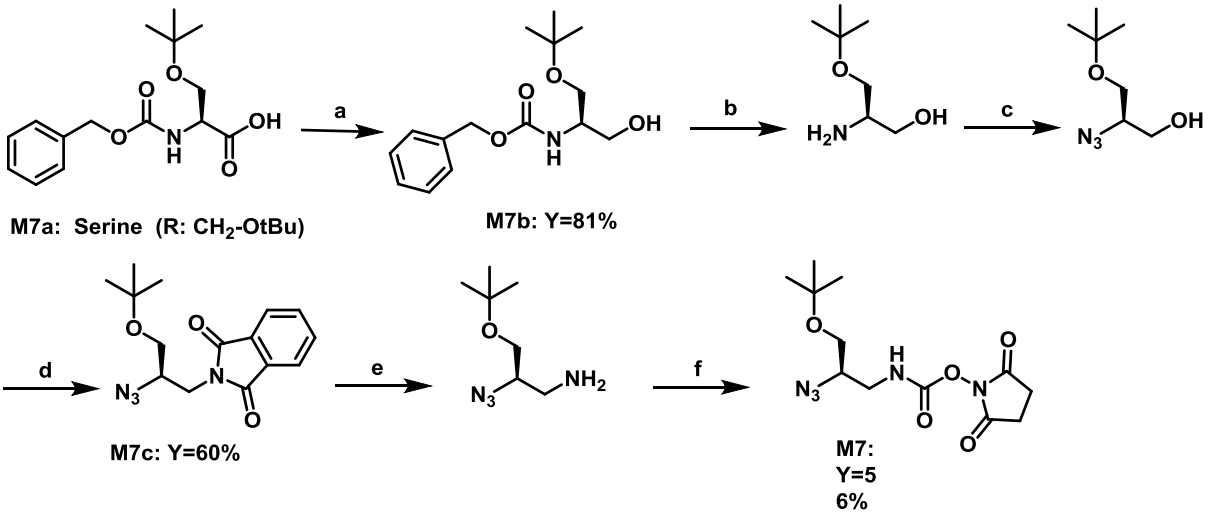
in solution was carried out in presence of DBU as base to prevent the formation of the adduct between dibenzofulvene and piperidine which can be problematic to eliminate in solution from the newly formed primary amine. Azide activated monomers N_3-Lys(Boc)^U-OSu (**M5**) and N_3-Arg(Pbf)^U-OSu (**M6**) were obtained in 27% and 26% yield respectively.



a: 1) NMM, IBCF, THF, -20°C; 2) $NaBH_4$, H_2O; b: DBU, EtOAc, r.t.; c: $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4:5H_2O$, $CH_3CN:H_2O$(1:1), r.t.; d: PPH_3, Phthalimide, DIAD, THF, r.t.; e:$N_2H_4.H_2O$, MeOH, 70°C; f: DSC, CH_2Cl_2, r.t.

Scheme 4 : Synthesis of azide monomers activated as succinimidyl carbamates oM5 and M6 starting from Fmoc-Lys(Boc)-OH and Fmoc-Arg(Pbf)-OH.

Route A was applied to perform the synthesis of N_3-Ser(OtBu)^U-OSu (**M7**) which was recovered in overall 27% yield.



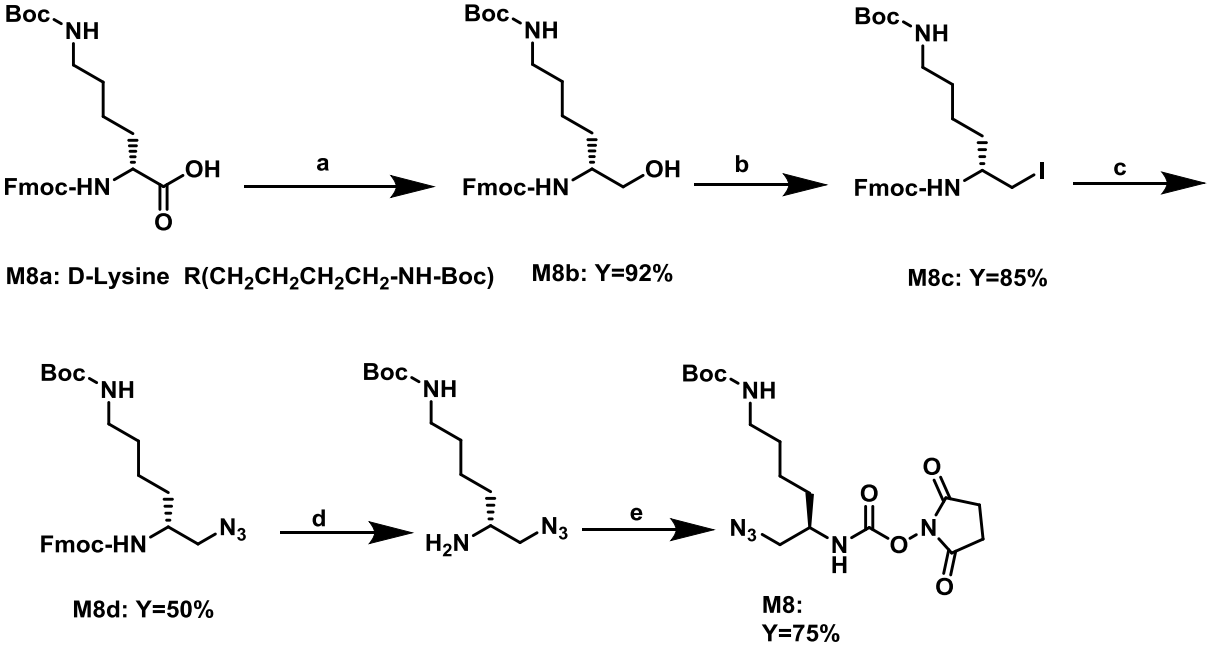
a: 1) NMM, IBCF, THF, -20°C; 2) $NaBH_4$, H_2O; b: H_2, Pd/C, MeOH, r.t.; c: $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4:5H_2O$, $CH_3CN:H_2O$(1:1), r.t.; d: PPH_3, Phthalimide, DIAD, THF, r.t.; e:$N_2H_4.H_2O$, MeOH, 70°C; f: DSC, CH_2Cl_2, r.t.

Scheme 5: Synthesis of succinimidyl (2-azido-2-substituted ethyl) carbamate M7 starting from Z-protected amino serine derivative.

ii) Solution synthesis of a monomer with inversed configuration

As evoked in the section related to the design of the hybrid sequences, we thought to prepare some building blocks with the side-chain shifted on the second carbon, to potentially evaluate

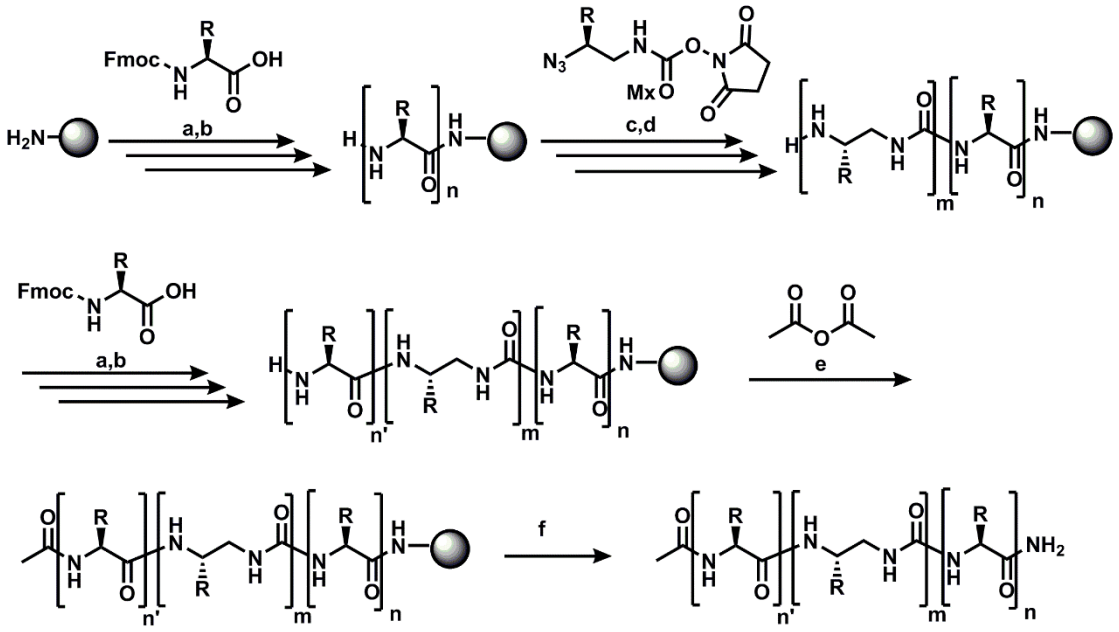
and improve the level of mimicry that can be achieved when replacing the LSK segment (residues13-15) in **II.9** (Chapter II, B. 2.). In particular, we focused on the Lys side chain that could be potentially better mimicked by shifting the side chain to the C_α carbon along with an inversion of the absolute configuration on this carbon ((*S*)→ (*R*)). N_3-Lys(Boc)$_{inv}^{U}$-OH **M8** was therefore prepared starting from corresponding *D*-amino acid and following the synthetic route detailed below.



Scheme 6: The synthesis of succinimidyl (2-azido-1-substituted ethyl) carbamate M8, bearing a shifted lysine side chain

iii) Solid phase synthesis of chimeras

During this PhD work, I focused my attention on the solid phase synthesis of chimera oligomers due to its robustness, its flexibility and its rapidity. Syntheses were performed on a rink amide Polystyrene resin (loading: 0.42 mmol.g^{-1}) on a 50 µmole scale. Taking into account the high number of α-amino acids for the peptide part (30 residues), the two peptide parts were synthesized with the automatic synthesizer, *Liberty Blue*. Concerning the addition of the two urea residues, they are added manually under microwave assistance (CEM Discover), using azide monomers activated as carbamate of succinimide. The coupling occurs with 1.5 eq of monomers **Mx** in presence of a base and after the Staudinger reaction under microwave irradiation allows the regeneration of the free amine for the next coupling (Scheme 7). the second peptide segment is next introduced by stepwise coupling of the N-Fmoc amino acids on the Liberty Blue in presence of HBTU, HOBt and DIEA. After final acetylation, compounds were cleaved from the resin with TFA/Tris/H_2O/EDT (92.5:2.5:2.5:2.5).



a: Fmoc -Xaa-OH (5 eq), HOBt (6 eq), HBTU (6 eq), DIEA (12 eq), 75°C (except for Fmoc-His-OH at 50°C), 30 W, 5 min; b: 20% piperidine, DMF (90°C/35W/50s); c:1.5 eq A, 3 eq DIEA 50°C 50W 2 x 15 min; d:10 eq. PMe_3 in THF (1M), Dioxane, H_2O (7:3 v/v), 50°C, 50W, 2 x 15min; e: Acetic anhydrid/ DCM 1:1 (15 min, r.t.); f: TFA/TIS/H_2O/EDT (92.5:2.5:2.5:2.5, v,v,v,v)

Scheme 7: General scheme of solid phase synthesis of α/urea chimeras

After precipitation in Et_2O and lyophilization, compounds were purified on semi-preparative HPLC. They were then characterized by RP-HPLC and mass spectrometry.

Table 6 : Sequences of chimeras oligomers with the LSK segment replaced by two urea residues

Compound	Sequence	t_R (min)	M_W (g.mol^{-1})	Yield (%)
	1 5 10 15 20 25 30			
	g **a** b c **d** e f g **a** b c **d** e f g **a** b c **d** e f g **a** b c **d** e f g **a** b c d			
II.9	Ac R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	7,98	4077	10
II.16	Ac R **M** K Q **L** E D K **I** E E **L** L^u K^u **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	7,56	4048	8
II.17	Ac R **M** K Q **L** E D K **I** E E **L** L^u $K_{inv}{}^u$ **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	7,5	4048	8
II.18	Ac R **M** K Q **L** E D K **I** E E **L** I^u K^u **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	7	4048	9
II.19	Ac R **M** K Q **L** E D K **I** E E **L** V^u K^u **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	6,68	4034	8
II.20	Ac R **M** K Q **L** E D K **I** E E **L** A^u K^u **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	6,49	4006	11

Five chimeric sequences **II.16** to **II.20** were prepared by replacing the Leu-Ser-Lys tripeptide by various diurea sequences. Oligomer **II.17** is analogous to **II.16** but contains the Lys^u analogue with the shifted side chain. Oligomers **II.18-II.20** were prepared to evaluate the best way to mimic the hydrophobic Leu13 side chain.

The chromatogram RP-HPLC of the crude oligomer chimera **II.17** shows that the synthesis can be performed without any special difficulty. The major peak which corresponds to the desired product as confirmed by mass spectrometry was isolated with a purity of 96% (Figure 48).

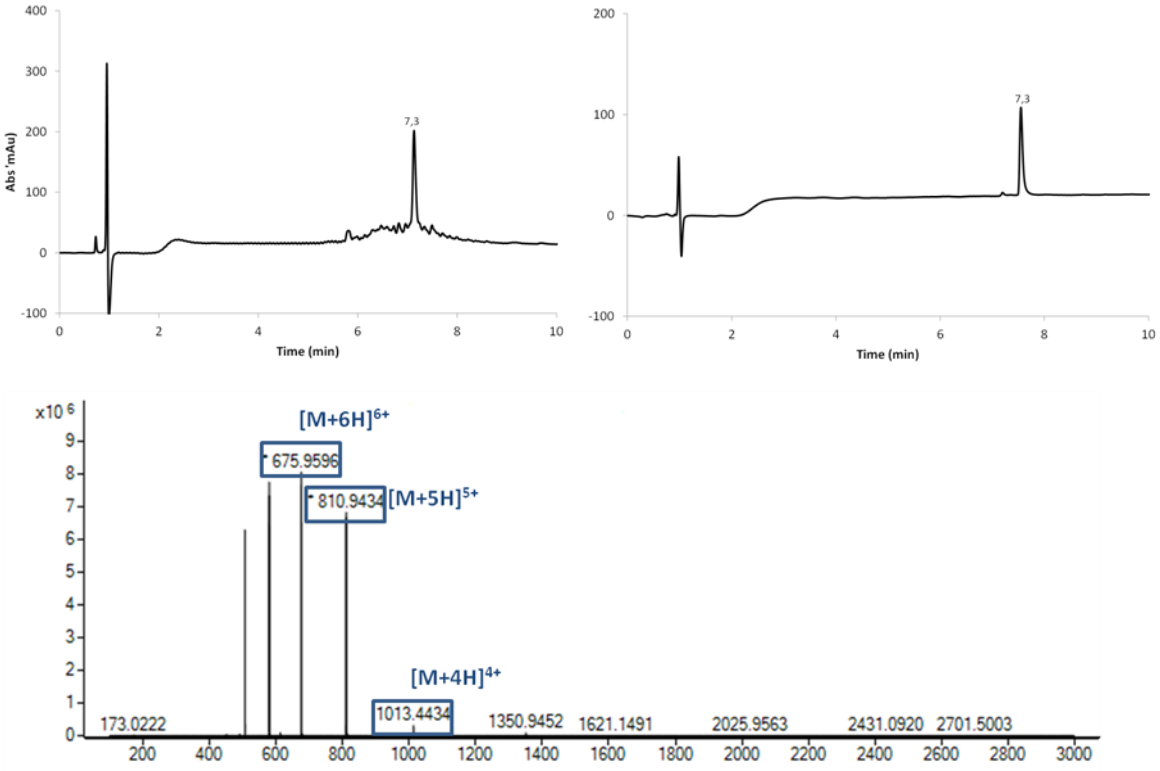


Figure 48:RP-HPLC Chromatograms of the chimera II.17 crude (left) and after purification (right) and Mass spectrometry analysis

We next wanted to study the impact of replacing one residue involved in the hydrophobic core by a urea derivative. We thus focused on modifications of tripeptide Leu12-Ser14 and Ile23-Arg25. In compound, **II.21**, Leu12 was mimicked by the introduction of N_3-Leu^U-OSu, Ser14 in *i*+2 position by Ser^U (**II.21**). Finally, we wanted to replace a residue involved in the hydrophobic core in another region of the peptide. The Ile23 was then mimicked by Ile^U and Ala^U whereas the residue i+2, Arg25 was replaced by Arg^U (**II.22** and **II.23**).

Table 7: Sequences of chimera oligomers where the urea segment is introduced in the hydrophobic core

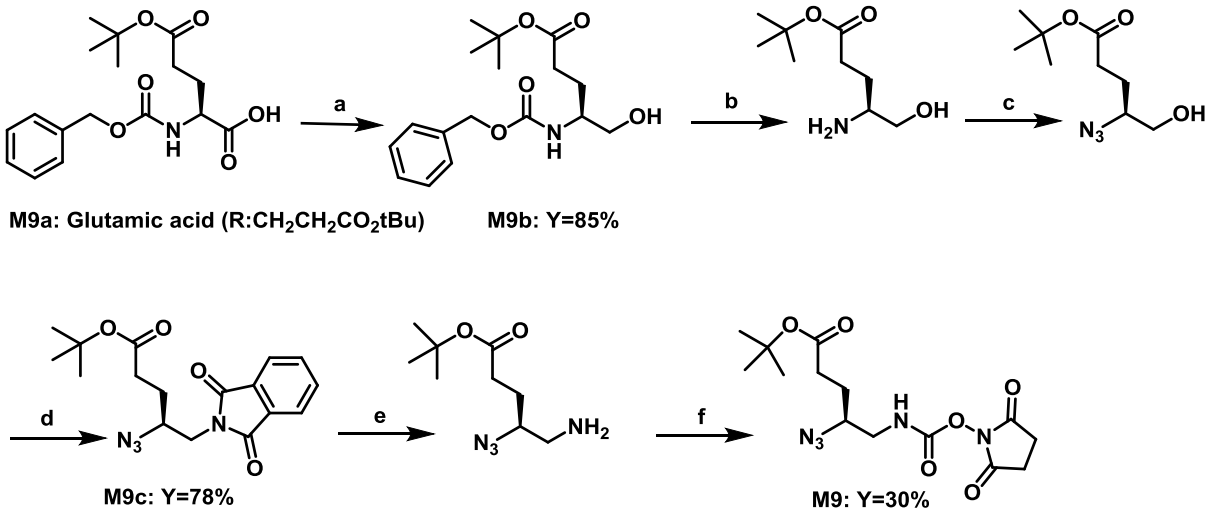
Compound	Sequence	t_R (min)	M_W (g.mol^{-1})	Yield (%)
	1 5 10 15 20 25 30			
	g **a** b c **d** e f g **a** b c **d** e f g **a** b c **d** e f g **a** b c **d**			
II.9	Ac R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	7,98	4077	10
II.21	Ac R **M** K Q **L** E D K **I** E E **L^u** **S^u** K **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	5,9	4022	11
II.22	Ac R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **I^u** **R^u** **L** K K L **I** G E R NH_2	7,23	4078	16
II.23	Ac R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **A^u** **R^u** **L** K K L **I** G E R NH_2	6,29	4038	11

All these sequences after synthesis and purification were then used for structural analysis.

b) Substituting 3 urea units (triad) for 4 consecutive α amino acid residues

i) Solution synthesis of azide monomers activated as succinimidyl carbamate

Azide monomer **M9** with a carboxylate side chain and activated as a succinimyl carbamate was prepared from commercially available Z-L-Glu(O*t*Bu)-OH via the route A (see scheme 8)C.2)a)i) In these conditions, monomers N_3-Glu(O*t*Bu)^U-OSu (**M9**) was obtained with an overall yield of 20%. The monomers **M2**, **M3**, and **M5** required for the synthesis of this series of oligomers were described in Part C.2) a) i)).



a: 1) NMM, IBCF, THF, -20°C; 2) $NaBH_4$, H_2O; b: H_2, Pd/C, MeOH, r.t.; c: $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4:5H_2O$, $CH_3CN:H_2O$(1:1), r.t.; d: PPH_3, Phthalimide, DIAD, THF, r.t.; e:$N_2H_4.H_2O$, MeOH, 70°C; f: DSC, CH_2Cl_2, r.t.

Scheme 8: Multi-step synthesis of activated monomer N_3-Glu(OtBu)^U-OSu M10.

ii) Solid phase synthesis of chimeras

Syntheses of chimeras were performed on a Rink amide polystyrene resin (loading: 0.42 mmol.g^{-1}) on a 50 µmole scale as described in paragraph C.2.a)iii). The peptide segments were assembled on the *Liberty Blue,* while the three consecutive urea residues were introduced manually under microwave conditions on a *Liberty Bio* apparatus. Two sequences have been designed to replace four consecutive α-residues by three urea ones (Chapter II B.3.). We focused on the same region as before, meaning in the middle of the GCN4-IV sequence and selected two peptide segments to replace containing Leu12 either as the first or the last residue.

In the case of sequence **II.25**, whose frame of oligourea replacement is shifted compared to **II.24**, it is worth noting that two α-residues involved in the hydrophobic core (*a* and *d* positions) are replaced. The Leu mutation in position 12 was kept as a reference mutation point and α-residues from 9 to 11 were substituted by a Val^U and Glu^U to mimic the side chains in positions 9 and 11 respectively. The choice of replacing the Ile side-chain by a shorter one like the Val^U was driven by the fact that we made the assumption that the shortening of this side-chain could compensate increased diameter of the oligourea helix, thus providing a better orientation/positioning of the branched propyl side-chain.

Table 8: Different reference peptides and chimeras II.24-II.25 sequences synthesized that encompassed 4 amino acids to 3 urea substitutions (in blue). In right columns physical chemistry data of the 3 synthesized chimeras.

	Sequence	t_R (min)	M_W (g.mol^{-1})	Yield (%)
	1 5 10 15 20 25 30			
	g **a** b c **d** e f g **a** b c d e f g **a** b c **d** e f g **a** b c **d**			
II.9	Ac R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	7,98	4077	10
II.24	Ac R **M** K Q **L** E D K **I** E E L^u A^u K^u **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	6,38	4035	15
II.25	Ac R **M** K Q **L** E D K V^u E^u L^u L S K **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	6,27	4006	20

After purification by semi-preparative RP-HPLC and lyophilization, all α/urea chimeras were used to perform structural investigations in solution and in the solid state.

c) Mutation of Asn16: switching from dimeric to trimeric assembly

As described in chapter A.1.a, GCN4-p1 peptide forms a homodimer where two Asn16 are forming inter-helical hydrogen bonds (Figure 49)[4,21]. This key residue is known to be determinant for the oligomerization state because it has been reported that mutating this residue lead to other oligomerization states. Indeed, the peptide sequence where Val residues occupied most of the *a* positions including the Asn16 position (only the Met2 at the N-termini was conserved) and Leucine residues at *d* positions leads to a mixed bundle population comprising dimeric and trimeric conformations[5]. In other mutated sequences, Asn16 was replaced by isoleucine or leucine as well as residues at *a* position leading to dimeric, trimeric or even tetrameric forms[5]. In the GCN4-p1 sequence, the residue Asn16 has also been replaced by valine or other polar residues: threonine, serine, and glutamine[22]. These mutations induced a change in the bundle stoichiometry with the formation of three-helix bundles that have been further characterized in the solid state by X-ray crystallography. Altogether these studies highlighted the pivotal role played by this polar residue in the stoichiometry of the bundle assembly through hydrogen bond contacts between the peptide chains.

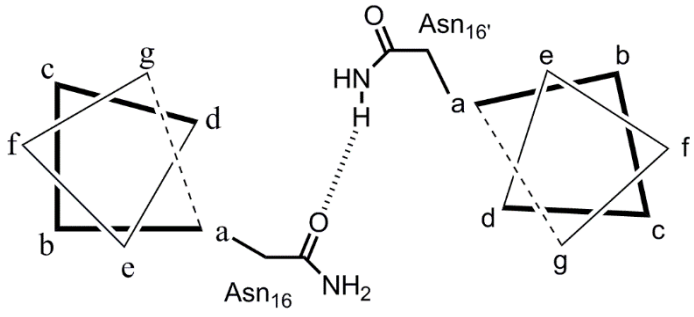


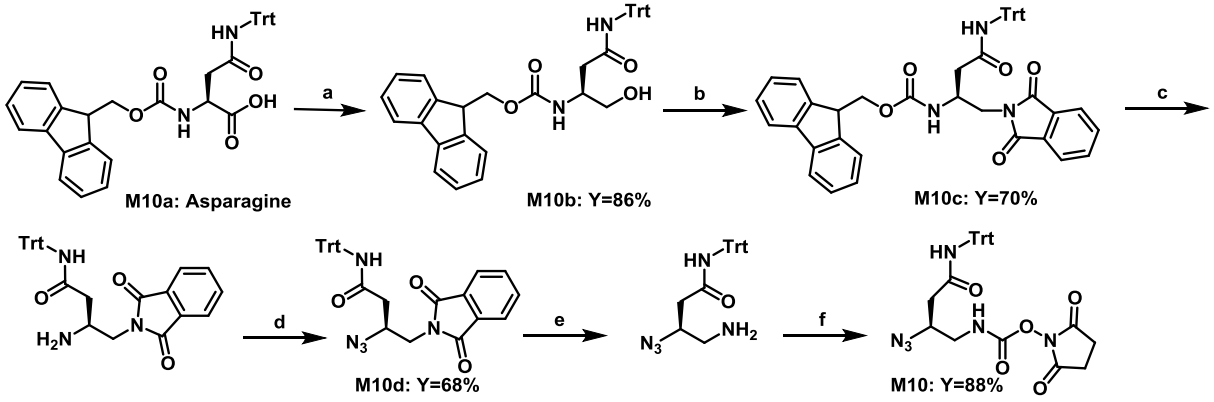
Figure 49: Heptad arrangement of GCN4-p1 showing the interhelical hydrogen bond between the carboxamide groups of Asparagines

A new strategy with the aim to replace the α-Asn only by its urea counterpart was developed. This would allow to study the influence of a change in the environment of this residue only

and the consequences of an additional C in the backbone due to the replacement of 1 α-residue by 1 urea.

i) Solution synthesis of activated azide monomer with Asn side chain

A new urea monomer (**M10**) derived from commercially available Fmoc-*L*-Asn(Trt)-OH, was synthesized. The preparation of this new building block was performed following the route B described in the general procedure and the monomer **M10** was obtained with a global yield of 34% (Scheme 9).



a: 1) NMM, IBCF, THF, -20°C; 2) $NaBH_4$, H_2O; b: PPH_3, Phthalimide, DIAD, THF, r.t.; c: DBU, EtOAc, r.t.; d: $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4:5H_2O$, $CH_3CN:H_2O(1:1)$, r.t.; e:$N_2H_4.H_2O$, MeOH, 70°C; f: DSC, CH_2Cl_2, r.t.

Scheme 9: Synthetic scheme for the preparation of succinimidyl (2-azido-2-substituted ethyl) carbamate M10 (N_3-Asn(Trt)^U-OSu)

ii) Asn16→Leu16 peptide and chimeras synthesis

In the previous designed chimeras, most of the short oligourea segments have been introduced in the center of the sequence of **II.9**, near the Asn16. We thought that if the lateral chain of Asn16 was shifted due to the introduction of the oligourea segment, it could influence the oligomerization state and disturb the dimer conformation. We decided to replace Asn16 with an α-Leu residue in the sequence of **II.17**, in order to investigate if this substitution could have a positive impact on the oligomerization state as well as bundle stabilization (*i.e.* **II.26** in Table 9). Concurrently, the corresponding peptide sequence **II.27** incorporating Leu residues in all *a* position (at the exception of Met_2) and Ile residues in all *d* position has been synthesized as well as its chimera version incorporating the Leu^U Lys^U diad in place of the Leu13Ser14Lys15 segment (**II.28**)[8]. As described above, the SPS of these three sequences was performed on a low loading Rink amide-MBHA resin (loading: 0.42 $mmol.g^{-1}$) on a 50 µmole scale. In addition to these modifications we envisioned to prepare the chimera **II.29** where the Asn16 was replaced by its urea counterpart (*i.e.* Asn^U, N^U in Table 9). The aim of this single α→urea mutation was to study the structural consequences of having installed one extra carbon C_α plus a urea bond in the backbone of **II.26** on the folding and oligomerization state.

Table 9: Sequences synthesized to study the influence of Asn16 in the oligomerization state

		Sequence	t_R (min)	M_W (g.mol^{-1})	Yield (%)
		1 5 10 15 20 25 30			
		g **a** b c **d** e f g **a** b c **d** e f g **a** b c **d** e f g **a** b c **d** e f g **a** b c d			
II.9	Ac	R **M** K Q **L** E D K **I** E E **L** L S K **N** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	7,98	4077	10
II.26	Ac	R **M** K Q **L** E D K **I** E E **L** **L^u** **K$_{inv}^u$** **L** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	7.83	4047	8
II.27	Ac	R **M** K Q **I** E D K **L** E E **I** L S K **L** Y H **I** E N E **L** A R **I** K K L **L** G E R NH_2	7,95	4076	11
II.28	Ac	R **M** K Q **I** E D K **L** E E **I** **L^u** **K$_{inv}^u$** **L** Y H **I** E N E **L** A R **I** K K L **L** G E R NH_2	7,76	4047	10
II.29	H_2N	R **M** K Q **L** E D K **I** E E **L** L S K **N^u** Y H **L** E N E **I** A R **L** K K L **I** G E R NH_2	6.99	4064	12

3. Circular Dichroism monitored thermal melting data

Circular dichroism (CD) is a spectrometric method that is often used to assess information on the secondary and tertiary protein structures. With their chromophores, proteins present characteristic CD spectra with peptide bond absorption in far UV region (240-180 nm) giving information concerning the secondary structure[23]. The environment of aromatic amino acid side chains is observed in the near UV region (320-260 nm) and allows some additional knowledge on the tertiary structure of the protein investigated. CD experiments also allow the determination of the melting point, the reversibility of structural changes or the percentage of secondary structure elements. The main advantages of CD technique are the convenience and the quick access to structural data. Furthermore, only small amount of material is required to obtain good quality spectra and samples can eventually be recovered as this technique is non-destructive.

To evaluate the impact of the insertion of urea block in the peptide sequence of **II.9**, CD seemed to be a spectrometric method of choice. CD spectra of GCN4-p1 indicate that an α-helical structure is adopted by the peptide in aqueous solution[4]. The characteristic data of α-helices are found with a maximum near 190 nm and two minima at 208 and 222 nm. The CD intensity at 222 nm indicates the helical folding of GCN4-p1 because this wavelength corresponds to the negative maximum of the molar ellipticity [θ] for α-peptides known to adopt an α-helix conformation. CD spectra of oligoureas are different and present a characteristic signature with a positive maximum of ellipticity around 203 nm and a negative maximum of ellipticity around 188 nm[24]. Then, in our case, the two wavelengths near 190 and 203 nm are not judicious points to focus because signal of peptide and oligourea overlaps. Because the CD spectrum of helical aliphatic oligoureas is generally silent in this region, the wavelength at 222 nm can be used to monitor α-helix formation. The change of helicity will be followed to evaluate the impact of the introduction of a urea segment in the peptide sequence.

a) CD for chimeras with urea diad

The chimera oligomers with the insertion of a diad and the reference peptide **II.9** have been first studied in solution with CD analysis. All experiments have been performed in a phosphate buffer (10 mM, pH 7) at 100 µM concentration[6]. By applying the same experimental conditions

as those already reported by Horne and coworkers we had a mean to compare the data obtained with the chimera series to that of the native peptide **II.9** and evaluate the reliability of CD spectroscopy in the context of this structural investigation.

All the CD spectra were recorded from 260 to 200 nm at 20°C and we focused our attention to the shape of the curve around 222 nm. Moreover, here it is important to underline that that we assume that at this wavelength there is not a direct contribution of the urea chromophore in the observed cotton effect of the α/urea chimera and the values at 222 nm were considered to only reflected the contribution of the peptide segments to the overall folding. Thereby, the lowest this value would be, the more stable the helical folding/assembling of the molecule should be. As a comparison, in the literature GCN4-p1 has a negative maximum of the mean residue ellipticity (MRE) $[\theta]_{MRE}$ at 222 nm of -33 deg.cm^2.dmol^{-1}res^{-1} x 10^3 [5].

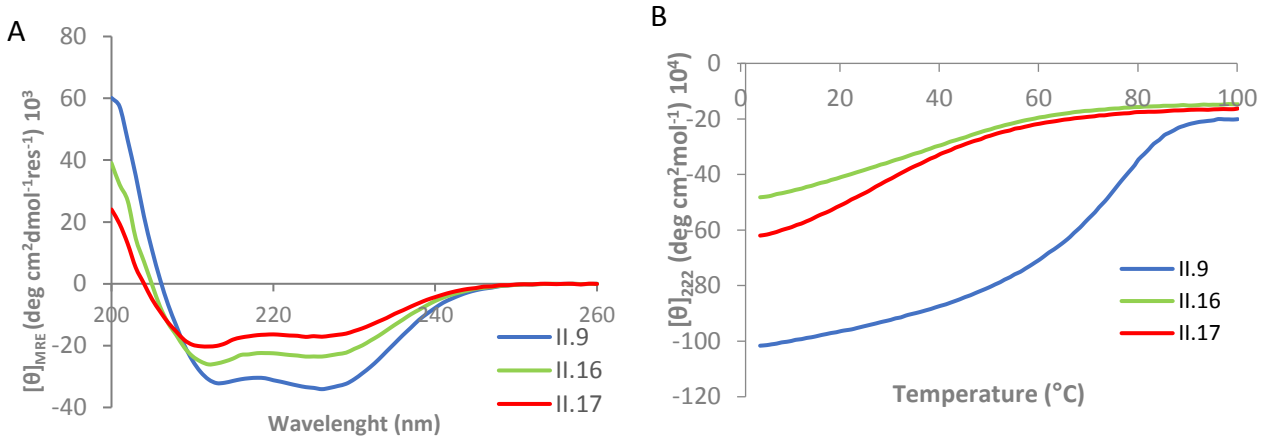


Figure 50: (A) CD curves expressed as mean residue molar ellipticity [θ] for the reference peptide **II.9**, chimeras **II.19** and **II.17**, at 100 µM concentration in 10 mM pH7 phosphate buffer at 20°C and (B) thermal spectra. Variable temperature CD scans were performed from 4 to 100°C at 100 µM concentration in 10mM pH7 phosphate buffer at 222 nm.

The CD spectra of **II.16** and **II.17** showed that having replaced 3 consecutive α-amino acid residues (Leu12Ser13Lys14 segment) by 2 ureas residues (diad) seems to be partially tolerated by the helix bundle assembly. The overall shape of the CD curves is maintained and is in favor of a conserved helicity, but for both chimeras (**II.16** and **II.17**) the absolute value of the $[\theta]_{MRE}$ at 222 nm is significantly lower than for the native peptide **II.9** (-23, -17 and -32 x10^3 deg.cm^2.dmol^{-1}.res^{-1} respectively see Figure 50A). Surprisingly, these results suggest that chimeras **II.16** and **II.17** are less folded than **II.9**. Despite oligourea have been shown to promote and stabilize helicity in a peptide segment, the introduction of this diad within the peptide sequence seems to alter the backbone conformation even though most of the side chains remained conserved. This caused important changes and do not lead to a completely mimicry of the α-peptide. The decrease in ellipticity may also reflect a weaker stability of the bundle due to less favorable inter helix interactions. The two chimeras present a difference of $[\theta]_{RME}$ at 222 nm of 6 000 deg.cm^2.dmol^{-1}.res^{-1}. This illustrates that the position of the Lys^u side chain on the C_α (**II.16**) is less stable than when the side-chain is appended on the C_β (**II.17**). This shift impacts the folding and the stability of the resulting chimera.

This CD structural study was completed with the second series of mutations by analysis of chimeras **II.18**-**II.20** (Figure 51 A). The minima values for $[\theta]_{222}$ were higher than for the reference peptide **II.9** meaning that as previously, the helicity is conserved but the α-helix is a disrupted by the presence of the urea units. Interestingly, the residue mean ellipticity of **II.19** is in the same range than **II.17** even though there is one missing methylene on the side chain of the first urea residue (Val^u to Leu^u) illustrating here that another lateral chain can be well-tolerated for overall folding. The two other chimera oligomers **II.18** and **II.20** showed higher loss of the negative contribution at 222nm concomitant with a decrease in helix folding propensity.

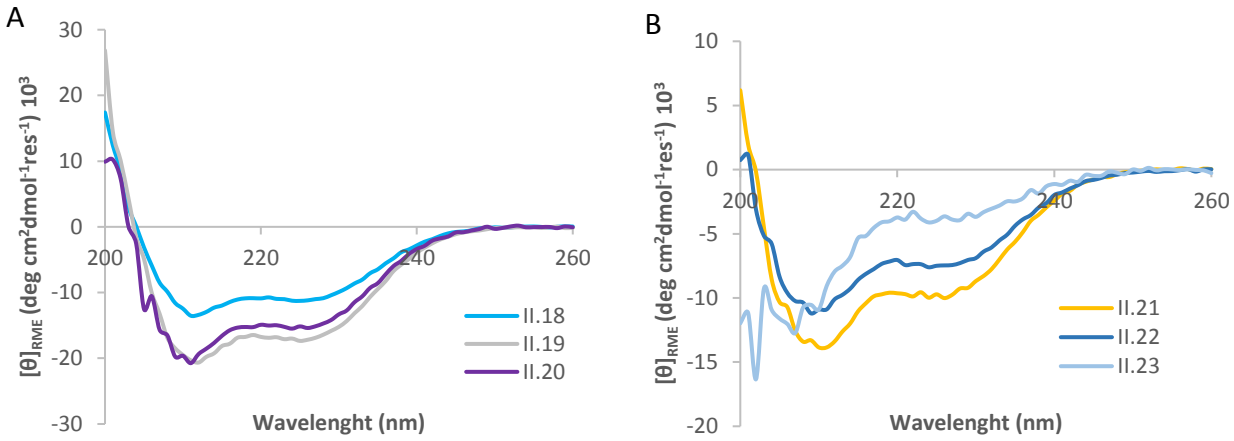


Figure 51: CD curves of chimeras with 3 amino acids replaced by 2 urea residues (A) inserted in LSK central position **II.18**-**II.20** and (B) involving hydrophobic position **II.21**-**II.23** at 20°C, at 100 µM concentration in 10 mM pH7 phosphate buffer.

We next recorded CD spectra from 260 to 200 nm at 20°C for the series of chimera oligomers where the diad was this time introduced at positions 12-14 and 23-25 in the peptide sequence (**II.21**-**II.23,** Figure 51 B). In these compounds, the diad replaced three consecutive α-residues one of them being involved in the hydrophobic core (in *a* position for **II.21** and in *d* position for **II.22** and **II.23** respectively). As previously, we focused our attention on the shape of the CD curves around 222 nm. The $[\theta]_{RME}$ value of the bands at 222 nm for these chimeras are weaker than for the reference peptide **II.9**, but also than for the previous chimeras. This result suggests that overall the helical folding is more disrupted by the insertion of the diads at these positions. The chimera exhibiting the highest molar ellipticity in this series is **II.21** with the diad replacing residues at positions 12-14, nearby the first zone of modification. However, with a $[\theta]_{RME222}$ = -10010. deg cm^{2}dmol^{-1}res^{-1}, the $[\theta]_{RME}$ absolute value is lower than for all the previous modifications (**II.16**-**II.20**) which suggests a negative contribution of the urea diad insertion on the hydrophobic core and therefore on the helical folding. For chimeras **II.22** and **II.23**, the diad was introduced to replace amino acids at positions 23-25 and again absolute ellipticity values are lower to that observed for **II.21**. Moreover, regarding **II.23** the shape of the curve does not suggest any helical folding meaning that this introduction of the urea diad at positions 23-25 is detrimental for folding and self-assembly.

Variable temperature CD analyses (from 4°C to 100°C) were next carried out in order to have access to the midpoint of the thermal transition (T_m) of each molecule. The molar ellipticity at

222 nm was plotted against the temperature and the T_m values were obtained by fitting the data with the Boltzmann sigmoid equation. As expected, thermal melts curve of **II.9** compared well to the one earlier published by Horne *et al.*[6] validating here our experimental conditions (T_m= 71 °C for **II.9** and 77 °C in the literature).

The results depicted in (Figure 50B) showed that the reference peptide presents an unfolding transition at a T_m= 71 °C and hybrids α/urea compounds present also this transition but the thermal unfolding midpoints for the chimeras are lower due to the introduction of two urea residues that may locally destabilize the helical folding. This observed behavior renders quantitative thermodynamic difficult to interpret. With the increase of the temperature, the observed $[\theta]_{222}$ absolute value for **II.16** and **II.17** is lower than the one of **II.9**, suggesting that the assemblies formed by these chimera oligomers are less stable than the assembly formed by the peptide itself. In contrast to what observed with CD spectra at 20°C, the best T_m for the chimeras was obtained for **II.17** (*i.e.* T_m = 33°C) whereas **II.16** presents an unfolding transition state at a T_m = 27°C. These results are not in line with the values of $[\theta]_{RME222}$ at 20 °C that was more favorable for **II.16** than for **II.17.** This suggests that the difference between the two T_m values of **II.16** and **II.17** may not be significant.

The second series of chimeras **II.18**-**II.20** present a less negative band at 222 nm (Figure 52 A), suggesting that the assemblies formed by **II.16** and **II.17** are likely to be a little bit more stable compared to the assemblies formed by **II.18**-**II.20**. Moreover, the variable temperature CD spectra show an even less negative band at 222 nm for **II.21**, **II.22** and a weak signal for

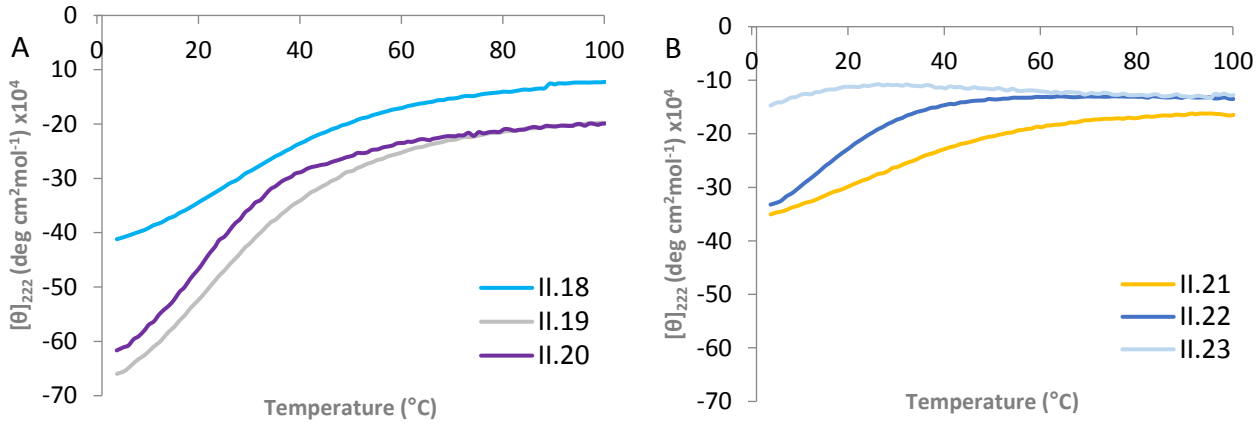


Figure 52: CD thermal spectra of chimeras with 3 amino acids replaced by 2 urea residues in LSK region for II.18-II.20 (A) and involving hydrophobic area II.21-II.23 (B) at 100 µM concentration in 10 mM pH7 phosphate buffer. Variable temperature CD was performed from 4 to 100°C at 222 nm.

II.23 (Figure 52 B). The T_m of the chimera **II.21**, bearing diad in its center, exhibits a much lower T_m than peptide **II.9**, but this T_m remains the best values in these series of three chimeras. Its value (≈ 11°C) is in the same range of order than that of **II.18** with the urea segment shifted of one residue. In regards to the two other chimeras, the T_m of **II.22** is extremely low (T_m = 16°C, below room temperature) and for **II.23**, the T_m could not be determined; the shape of the curve is not in favor of coiled coil-type folding. These CD results suggest that these second series of chimeras present a poor self-assembly propensity.

These results indicate that with the introduction of a urea segment, the T_m decreased substantially. The best T_m value was obtained for compound **II.17**, containing the Lys^U where the side-chain is branched on the C_α. Furthermore, we can observe that by decreasing the

number of carbons on the lateral chain of the urea residue mimicking the Leu13, the T_m decreased. This suggests that keeping the original side-chains in the urea diad seems to be the best compromise for helical secondary structure mimicry. Despite this Leu13 branched side-chain being solvent exposed, its nature seems to be important for folding and assembly. For all the chimeras screened in this study, temperature experiments have been performed from 4 - 100°C and from 100 - 4°C, and the all the curves were superimposable, highlighting here that the change of folding state is reversible.

We were next interested in gathering the thermodynamic parameters of the folding/unfolding process which in principle are accessible from these variable temperature CD experiments[25,26]. The peptide **II.9** and related chimeras undergo a two-state transition between a folded state (F) and an unfolded state (U) in function of the temperature. The fraction of oligomer folded at a given temperature is noted α.

$$\alpha = (\theta_t - \theta_U)/(\theta_F - \theta_t) \quad (1)$$

where θ_t is the ellipticity value observed at a given temperature, θ_U the ellipticity value of the unfolded state and θ_F the ellipticity value of the fully folded state. The Gibbs-Helmholtz equation can be used to describe the folding state as a function of temperature.

$$\Delta G = \Delta H(1-T/T_m) - \Delta C_p((T_m-T)+T\ln(T/T_m)) \quad (2)$$

T_m is the temperature at which K = 1 and α = 0.5

The constant of folding K is:

$$\Delta G = -RT\ln K \quad (3)$$

where R is the gas constant (1.98 cal.mol^{-1}), T the temperature (Kelvin) and K the constant of folding.

$$\alpha = K/(1+K) \quad (4)$$

The CD structural data have been reported in Table 10. Variable temperature CD experiments support the results obtained with the CD spectra of chimera oligomers **II.16**-**II.23** at 20°C on the destabilizing effect of urea diad on folding and self-assembly. We can observe in the table 8, that a low minimum of ellipticity at 222 nm does not necessary lead to a high T_m value. Indeed, **II.18** presents a molar ellipticity lower than **II.16** whereas the T_m is in the same range (≈ 27°C). This is quite surprising as **II.18** presents of the lowest negative band ($[\theta]_{222}$=-10 960 deg cm^{2}dmol^{-1}res^{-1}), whereas it leads the best T_m in this series. This result suggests that despite a poor α-helical content due to the introduction of a urea diad, the quaternary fold of chimera oligomers remains present in solution. This may also reflect a possible contribution of the urea chromophore to the CD spectrum even at 222 nm. The changes in T_m corresponds to a destabilization of about 1.9-2.2 kcal.mol^{-1} for chimeras **II.16** - **II.22**. The strongest destabilization occurs with **II.23** with 4.9 kcal.mol^{-1} which is in adequacy with previous results and confirms that this 23-25 position is not the best to substitute. The central position restores the most the desired conformation.

Table 10: Sequence and circular dichroism data for peptide II.9 and chimeras II.16-II.23. [a] $[\theta]_{222}$ determined from the CD curves at 20 °C and expressed in deg.cm^2.dmol^{-1}.res^{-1}. [b] the fraction of helicity was calculated with $[\theta]_{222}$ (see experimental part) [c] Melting temperature (T_m) of compound at 100 µM concentration in 10 mM phosphate buffer, pH 7. [d] Change of folding free energy with respect to II.9.

Compound	$[\theta]_{222}$[a]	% helicity[b]	T_m (°C)[c]	$\Delta\Delta G_{fold}$[d]
II.9	-32300	91	71±0.78	

II.16	-23032	67	27±0.12	1.9
II.17	-16764	49	33±0.23	2.0
II.18	-10965	33	26±0.61	1.6
II.19	-16796	49	19±0.52	2.1
II.20	-14972	44	16±0.99	2.2
II.21	-9881	30	23±0.34	1.9
II.22	-7361	23	16±0.23	4.9
II.23	-3286	12	n.m.	-

The insertion of oligourea segment in the center of the α-peptide appears to be better tolerated by the helix conformation (more than 33% of helicity) than when it is shifted to the C-terminal part of **II.9**. A drastic decrease of the ellipticity value was indeed observed for the two last chimeras **II.22** and **II.23** reflecting loss of helicity, in particular in case of **II.23** bearing an Ala^{u} side-chain at position 23. The observed effects appear to be cumulated. The shift of the mutation and the lack of branched side-chain are indeed detrimental for helix population and bundle formation. This last observation for **II.23** confirmed that the best results were obtained by keeping the native side chains on urea residues.

The destabilizing effect resulting from the introduction of a urea diad in the peptide backbone can be context dependent, *i.e*. it may be due to the location of the diad. Another hypothesis is that the number of urea residues is too low to sufficiently stabilize the helix (only two urea residues whereas 2.5 residues are required for one full helical turn). With only two urea residues the helical turn of the urea segment is not complete leading to a shift in the distribution of side-chains and to a decrease of the overall helicity in comparison with the reference peptide.

From these first results, we could consider that the introduction of an additional urea residue within the peptide backbone might offer a better 'lock' of the urea 2.5-helix and could have a beneficial effect on the stabilization of the secondary structure and quaternary folds.

b) CD for chimeras with a oligourea triad

The CD curves of chimeras encompassing a urea triad within the peptide sequence, were also recorded from 260 to 200 nm (Figure 53A). The CD spectra of **II.24** and **II.25** showed that the

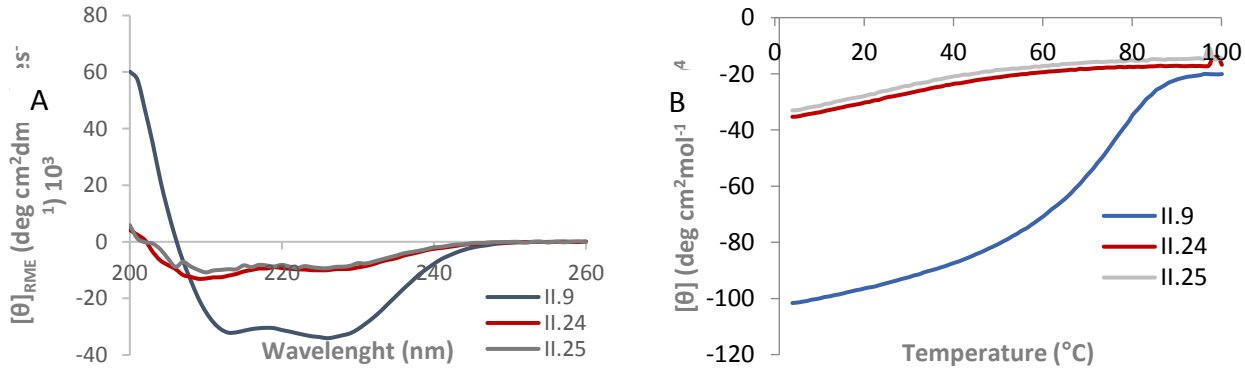


Figure 53: (A) CD spectra of chimeras II.24 and II.25 at 20°C and 100 µM concentration in 10 mM pH7 phosphate buffer and (B) Variable temperature CD scans recorded from 4 to 100°C at 100µM concentration in 10mM pH7 phosphate buffer at 222 nm.

replacement of four consecutive α-amino acids by three urea residues seems not to be really well-tolerated by the peptide backbone with poor helical folding. The overall shape of the CD curves is indeed in favor of a conserved helicity, however the helicity (30% for **II.24** and 29% for **II.25**) is much lower for both chimeras than for native peptide **II.9**. This suggests that the insertion of a urea triad at these two positions strongly disrupts the helical secondary structure.

This conclusion is further supported by variable temperature CD experiments recorded between 4°C and 100°C (and reversely from 100°C to 4°C) with the monitoring of the molar ellipticity at 222 nm for compounds **II. 24** and **II.25**. Curves of chimeras are quite similar to each other but showed a markedly different cooperative unfolding transition compared to **II.9** (Figure 53B). Indeed, chimeras **II. 24** and **II. 25** exhibited much lower T_m values (12 and 18 °C respectively) than peptide **II.9** indicating a low self-assembly propensity. For both chimeras, the T_m values are also lower than for chimeras with the introduction of two urea residues in place of three consecutive α-amino acids described in the previous section. Increasing the number of urea residues substitutions seems to further decrease the helix-bundle stability. This can be due to the position of the triad into the peptide sequence. The urea triad was inserted around the key residue Asn16, and as described before, the alteration of this residue can have a profound impact on the stoichiometry of the folding assembly. Furthermore, the triad involves the replacement of one polar residue at positions *a* and *d*, which has an impact on the packing arrangement in the hydrophobic core and can disrupt the folding as well as the assembly such as for chimeras **II.21**-**II.23**. The changes in T_m corresponds to 2.0 kcal.mol^{-1} destabilization for the introduction of triad in central positions 12-15 (**II.24**) and the introduction in positions 9-12 is a little bit more stable with 1.7 kcal.mol^{-1} (**II.25**) (Table 11). However, because we do not observe a true cooperative two-state unfolding process, these figures are just estimates and must be interpreted with caution.

Table 11: Sequence and circular dichroism data for peptide II.9 and chimeras II.24 and II.25. [a] $[\theta]_{222}$ determined with CD scans at 20 °C and expressed in deg cm^2 dmol^{-1} res^{-1}. .[b] the fraction of helicity was calculated with $[\theta]_{222}$ (see experimental part)[c] Melting temperature (T_m) of 100 µM compound in 10 mM phosphate buffer, pH 7. [d] Change of folding free energy with respect to II.9.

Compound	$[\theta]_{222}$[a]	% Helicity[b]	T_m (°C)[c]	$\Delta\Delta G_{fold}$[d]
II.9	-32300	91	71±0.78	
II.24	-9723	30	12±1.85	2.0
II.25	-9429	29	18±2.7	1.7

The CD data reported in Table 11 suggest that the introduction of a segment with one additional urea residues does not lead to a faithful mimic capable to lock and stabilize the helical folding. The insertion of triad decreases the coiled-coil assembly and hence the helicity. It is important to remind that as a parallel homodimer coiled-coil mimetic, the oligourea segment of each monomer interacts face to face in the self-assembly process, which could lead to increased destabilization. One hypothesis is that the size of triad, a bit higher than one urea helical turn (2.5 residues), could be too long and thus alter surrounding side chains encoded information, giving rise to visible changes in self-association behavior. It could be interesting, for future engineering efforts, to increase the number of urea residues to 5

(replacing 7 amino acids) to have two full helical turns in order to study the impact on the stability. Another point to figure out, is that in our design of triads, we change the nature of some side chains whereas we saw previously that best results were obtained with the native side chains. These results were not known when we first engineered this series of compounds and other results in the group had showed that to replace side chains of amino acids by shorter ones could be well tolerated (non-published results). Changing these side chains could have a positive impact for the stability but certainly not a colossal one because the backbone stills plays an important role also.

Despite a sufficient number of urea residues to make a full helix turn, results are quite disappointing. An explanation can be found by the lack of structural uniqueness after having inserted a three urea-residue segment. Indeed, the introduction of a urea segment can have a negative impact on the orientation of the side chain of Asn16 which is responsible for the specificity of a defined oligomeric state. The different CD signatures can reflect a tendency to adopt other stoichiometries than dimer but to confirm this hypothesis, other structural methods are required.

c) CD for chimeras with mutation of Asn16

CD experiments have been performed on compounds exhibiting mutation of Asn16 with the aim this time to study the importance of this residue on the oligomerization state. All the CD spectra of chimera oligomers with diad **II.26**, **II.28** and **II.29** and cognate peptide **II.27** were recorded from 260 to 200 nm at 20°C, at 100 µM concentration in phosphate buffer (10 mM, pH 7). We focused our attention on the residue mean ellipticity at 222 nm and we observed that θ_{222} for chimeras were lower than for the peptide **II.27** (Figure 54) that presents a higher helicity (83% see Table 12). These results show that, as discussed before, the introduction of urea diad in the peptide segment perturbs the helical folding.

Variable temperature CD experiments have also been performed from 4 to 100°C (and reversely) at 222 nm for this last series of compounds (Figure 54B). The shape of the curves for the peptide **II.27** (GCN4-pLI) and chimera compounds **II.26** and **II.28** is different than for previous compounds including the reference peptide **II.9**. Indeed, the sigmoidal shape is no more visible, but the shape obtained for the peptide **II.27**, more linear, is similar to what is reported in the literature for a tetrameric folding[5]. The absolute value of ellipticity for

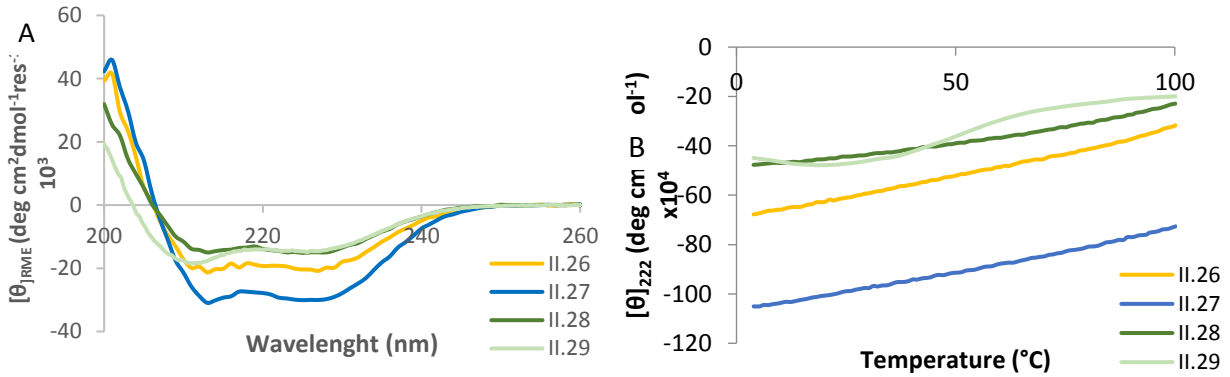


Figure 54: CD scans of peptide and chimeras II.26-II.29 at 20°C, at 100 µM concentration in 10 mM pH7 phosphate buffer (A) and thermal spectra (B) Variable temperature CD scans was performed from 4 to 100°C at 100µM concentration in 10mM pH7 phosphate buffer at 222 nm.

chimeras **II.26** and **II.28** is lower than for the peptide **II.27**, but for these three compounds, an exact determination of the melting temperature was not possible as we did not observe any transition state. The T_m obtained by fitting the data to a two-state model, were superior to 100 °C, suggesting the high stability of the coiled-coil or another oligomerization state.

Compound **II.26** that shares the same sequence than chimera **II.17,** at the exception of an Asn16 → Leu16 substitution, presents a much more stable coiled-coil assembly than previous chimera. This highlights that the substitution of this key residue (Asn16) has drastic structural consequences on the stability and self-assembly. The other chimera **II.28**, with Leu at *a* positions and Ile at *d* positions, presents a similar behavior meaning that the nature of the hydrophobic core is not essential here for the self-assembly propensity and coiled-coil stability. These results underline once again the key role of the Asn16 in the formation of dimer through polar contact.

Regarding chimera **II.29**, the shape of the CD curve reminds that of peptide **II.9** suggesting here a similar oligomerization behavior. However, the substitution of this Asn16 has a huge impact on the helicity (42% instead of 91%) highlighting the importance of this residue for the folding. The T_m value for **II.29** is lower than for peptide **II.9** (53 and 71°C respectively) meaning that the introduction of the urea residue destabilizes the assembly even though this fall in the value of T_m is less important than for the other studied chimera series. This result confirms that the change in the backbone composition has a direct impact on the helical structural folding and bundle self-assembly. Overall the GCN4 coiled-coil model has key structural constraints and any perturbation can have a huge impact on the folding, making any interpretation of these phenomena more complex.

Table 12: Sequence and CD data for peptide II.9 and chimeras II.26-II.29. [a] $[\theta]_{222}$ determined from the CD spectra recorded at 20 °C and expressed in deg cm^2 dmol^{-1} res^{-1}.[b] the fraction of helicity was calculated with $[\theta]_{222}$ (see experimental part). [c]Melting temperature (T_m) of chimeras at 100 µM, in 10 mM phosphate buffer, pH 7.

Compound	$[\theta]_{222}$[a]	% Helicity[b]	T_m (°C)[c]
II.9	-32300	91	71±0.78
II.26	-19223	56	_
II.27	-29291	83	_
II.28	-14503	43	_
II.29	-14366	42	53±0.79

These circular dichroism experiments gave us important structural information about the impact of the introduction of oligourea segment in a long model peptide. A main limitation of CD is the low-resolution structural information. Although a reliable estimation of the secondary structure of a peptide is obtained, the information on the quaternary structure are not detailed. To evaluate the stoichiometry of chimeras, other methods are necessary such as sedimentation equilibrium, differential scanning calorimetry, NMR or x-ray crystallography. In our case, we were interested in correlating these CD data with high-resolution structures of chimera oligomers assemblies. As GCN4 derivatives are known to be amenable to crystallization in aqueous conditions, we focused our efforts on x-ray crystallographic analysis.

4. Crystallographic assays

X-ray crystallography is a structural technique that requires crystals that diffract to high resolution. This technique allows the determination of the three-dimensional molecular structure at tomic resolution. The principle is to dissolve a pure sample in solution at high concentration and to find the correct conditions to induce its crystallization. The elucidation of these conditions is the rate limiting step of X-ray crystallography.

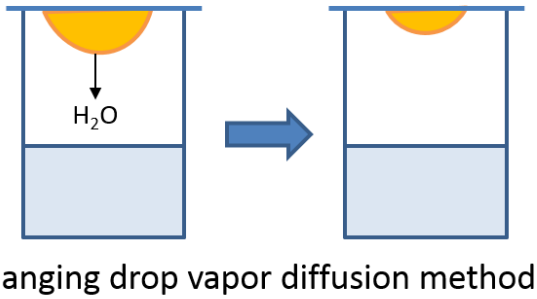


Figure 55: Hanging drop vapor diffusion method for crystallization with the inverted drop in yellow and the buffer solution in the well in blue.

All synthesized chimeras were subjected to crystallization assays through hanging drop vapor diffusion method (Figure 55) in aqueous conditions. This popular method has for principle a buffer solution in the crystallization well and a drop on the coverslip that is composed of the protein solution and the solubilizing buffer (1:1)[27]. The coverslip is returned and the drop stills suspended. The vapor pressure around the drop is higher than in the reservoir. With the pressure gradient across the vapor space, water is lost in the drop and precipitant concentration in the drop equilibrates with the precipitant concentration in the well.

Crystallization conditions were selected by variations around those reported for the reference peptide **II.9** (0.15 M sodium citrate tribasic pH 5.6, 20% v/v 2-propanol and 15% PEG w/v 4000). Several commercially available screening kits (JBS basic 1, basic 2 and classic ones) were also used to start and by learning this technique, I managed to grow crystals of good sizes for chimeras **II.16**, **II.18** and **II.24** (Figure 56). In contrast to **II.16**, we did not obtain crystals for **II.17** illustrating here that despite the best T_m of **II.17**, the switch of the lysine side chain on the carbon C_α, may destabilize the helical structure with an impact on the crystallization process. However, as the environment can have an influence on the stability of compounds, other conditions must be tried again to obtain good-quality crystals.

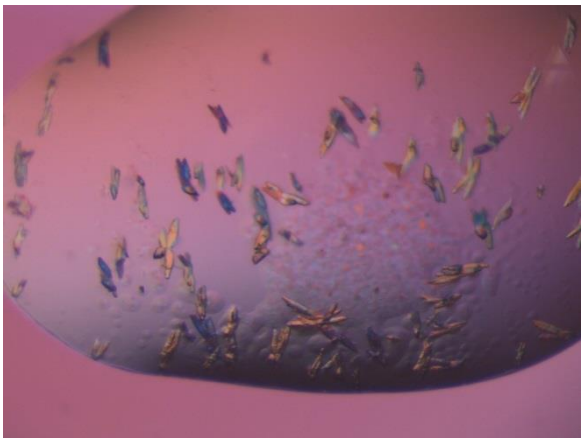


Figure 56: Example of crystals obtained for II.16 before optimization

These crystals are encouraging results because it indicates that some chimeras own a sufficient structural order to make crystallization possible. However, crystals were only obtained when the oligourea segment was introduced in the LLSK center position (12-15) suggesting that it could be a good area for modifications to obtain crystals. Crystals were frozen in N_2 and diffraction data of compounds **II.16** and **II.18** were obtained at 2.2 Å and 3.0 Å respectively by Jérémie Buratto (structural biologist in our team) who brought them to the synchrotron. For **II.16**, the space group P222 was found with

a unit cell (39.24, 47.31, 49.96, 90.00, 90.00, 90.00; a,b,c,α,β,γ). Concerning **II.24**, diffraction data were also obtained but the resolution was too low.

However, data obtained for **II.16** and **II.18** could not be used to solve the atomic structure as a phase problem appeared leading to the impossibility to exploit the data. Indeed, the phase angle between the diffracted angle and the amplitude of the reflection are related with the intensity of a diffracted spot but the determination of the phase angle is not direct in our case. To solve this problem, an indirect isomorpheous replacement method was used to incorporate heavy atoms into crystals. We thought to iodinate the tyrosine 17 in the chimera sequences. Indeed, using iodo-tyrosine as a phasing tool may be beneficial for structure determination using a single-wavelength anomalous diffraction[28]. Assays for the iodination of the tyrosyl side chain have then been performed during the crystallization process with the addition of an iodinating agent (I_2/KI) in the drop or directly in the well. However, we were not successful in reproducing crystals by using this iodination procedure. Another approach consists in performing the iodination post-synthesis by using N-iodosuccinimide as iodinating agent. This strategy is actually in progress in the group and we expect that it will help in solving the X-ray crystal structure of **II.16** and **II.18** chimeras.

Presently, our results with crystallography data did not allow the determination of atomic structures of our chimeras which is clearly missing. These crystal structures would indeed help us to understand and to study the influence of the introduction of the oligourea segment on the quaternary structure. Besides, it would be interesting to compare the structure of chimeras with diad or triad in order to optimize the design. Crystal structures would also allow us to check our hypothesis regarding the change in the orientation of Asn16 side-chain after having introduced the urea segment. Much effort on the crystallization method was done. First results are encouraging and new crystallogenesis efforts will be required to obtain crystals with heavy atoms, solve the phase problem and determine the structure with our data. New crystal assays will be necessary to find crystal conditions for other interesting compounds such as **II.17**, **II.21**, **II.26** - **II.29** and improve **II.24**. In short term, this would allow us to design new chimera sequences and to propose new modifications such as the increase of urea segment size. As mentioned before, to insert 5 urea residues within the peptide sequence would permit the installation of two full urea-helical turns which may potentially promote a better stabilization of the helical structure. Indeed, the proportion of α-residues is clearly higher compared to the urea residues (these chimeras are the longest synthesized in the group at the moment) and increasing the number of urea residues might be positive for structure and assembly.

In this project, we have performed a series of $\alpha \rightarrow$urea mutations with two different size of urea-segments with the aim to study the structural consequences on chimera folding and oligomerization state of the local insertion of a urea backbone in the main peptide chain. The nature of the side chains is also important for folding, in particular for residues involved in the hydrophobic core and inter-helix contacts. The position of the short urea segment is essential for the folding process and the chimeras equipped with urea blocks in the middle of the peptide sequence gave the best results. However, the introduction of this block in the center may induce a shift of the key residue Asn16 side-chain as well as other key side chains and is likely to have an impact on the oligomerization state. Unfortunately, all the structural

information gathered after CD experiments have not been yet confirmed by X-ray crystallography. This interpretation limitation precludes at this stage to define general rules to design effective mimics of α-helical peptides. However, the introduction of a urea segment in the center of a long peptide sequence (> 20 residues) is feasible via solid phase synthesis approaches earlier developed in the group, which is encouraging. Nevertheless, some of the chimeras described in this chapter have maintained significant helical folding even though the insertion of urea-diads or triads reduces overall stability, which is also a positive result. The results obtained for helix bundles can be extended to the rational design of α/urea hybrids as modulators of protein-protein interaction involving α-helix recognition. Indeed, GCN4 bundles required helical monomers and self-assembly that compare well to protein-protein interactions where α-helices are involved in the recognition process. We can now apply this strategy to our selected peptides involved in protein-protein interaction between histones H3-H4 and Asf1 proteins.

Bibliography

(1) Fremaux, J.; Mauran, L.; Pulka-Ziach, K.; Kauffmann, B.; Odaert, B.; Guichard, G. α-Peptide-Oligourea Chimeras: Stabilization of Short α-Helices by Non-Peptide Helical Foldamers. *Angew. Chem.* **2015**, *127* (34), 9954–9958.

(2) O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. X-Ray Structure of the GCN4 Leucine Zipper, a Two-Stranded, Parallel Coiled Coil. *Science* **1991**, *254* (5031), 539–544.

(3) Oshaben, K. M.; Salari, R.; McCaslin, D. R.; Chong, L. T.; Horne, W. S. The Native GCN4 Leucine-Zipper Domain Does Not Uniquely Specify a Dimeric Oligomerization State. *Biochemistry* **2012**, *51* (47), 9581–9591.

(4) O'Shea, E. K.; Rutkowski, R.; Kim, P. S. Evidence That the Leucine Zipper Is a Coiled Coil. *Science* **1989**, *243* (4890), 538–542.

(5) Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. A Switch between Two-, Three-, and Four-Stranded Coiled Coils in GCN4 Leucine Zipper Mutants. *Sci.-N. Y. THEN Wash.*- **1993**, *262*, 1401–1401.

(6) Oshaben, K. M.; Horne, W. S. Tuning Assembly Size in Peptide-Based Supramolecular Polymers by Modulation of Subunit Association Affinity. *Biomacromolecules* **2014**, *15* (4), 1436–1442.

(7) Malhotra, P.; Udgaonkar, J. B. How Cooperative Are Protein Folding and Unfolding Transitions: Cooperativity in Protein Folding/Unfolding Reactions. *Protein Sci.* **2016**, *25* (11), 1924–1941.

(8) Horne, W. S.; Price, J. L.; Keck, J. L.; Gellman, S. H. Helix Bundle Quaternary Structure from α/β-Peptide Foldamers. *J. Am. Chem. Soc.* **2007**, *129* (14), 4178–4180.

(9) Price, J. L.; Horne, W. S.; Gellman, S. H. Structural Consequences of β-Amino Acid Preorganization in a Self-Assembling α/β-Peptide: Fundamental Studies of Foldameric Helix Bundles. *J. Am. Chem. Soc.* **2010**, *132* (35), 12378–12387.

(10) Shin, Y.-H.; Mortenson, D. E.; Satyshur, K. A.; Forest, K. T.; Gellman, S. H. Differential Impact of β and γ Residue Preorganization on α/β/γ-Peptide Helix Stability in Water. *J. Am. Chem. Soc.* **2013**, *135* (22), 8149–8152.

(11) Arnold, U.; Hinderaker, M. P.; Nilsson, B. L.; Huck, B. R.; Gellman, S. H.; Raines, R. T. Protein Prosthesis: A Semisynthetic Enzyme with a β-Peptide Reverse Turn. *J. Am. Chem. Soc.* **2002**, *124* (29), 8522–8523.

(12) David, R.; Günther, R.; Baumann, L.; Lühmann, T.; Seebach, D.; Hofmann, H.-J.; Beck-Sickinger, A. G. Artificial Chemokines: Combining Chemistry and Molecular Biology for the Elucidation of Interleukin-8 Functionality. *J. Am. Chem. Soc.* **2008**, *130* (46), 15311–15317.

(13) Pendem, N.; Douat, C.; Claudon, P.; Laguerre, M.; Castano, S.; Desbat, B.; Cavagnat, D.; Ennifar, E.; Kauffmann, B.; Guichard, G. Helix-Forming Propensity of Aliphatic Urea Oligomers Incorporating Noncanonical Residue Substitution Patterns. *J. Am. Chem. Soc.* **2013**, *135* (12), 4884–4892.

(14) Claudon, P.; Violette, A.; Lamour, K.; Decossas, M.; Fournel, S.; Heurtault, B.; Godet, J.; Mély, Y.; Jamart-Grégoire, B.; Averlant-Petit, M.-C.; et al. Consequences of Isostructural Main-Chain Modifications for the Design of Antimicrobial Foldamers: Helical Mimics of Host-Defense Peptides Based on a Heterogeneous Amide/Urea Backbone. *Angew. Chem. Int. Ed.* **2010**, *49* (2), 333–336.

(15) Bansal, M.; Kumar, S.; Velavan, R. HELANAL: A Program to Characterize Helix Geometry in Proteins. *J. Biomol. Struct. Dyn.* **2000**, *17* (5), 811–819.

(16) Douat-Casassus, C.; Pulka, K.; Claudon, P.; Guichard, G. Microwave-Enhanced Solid-Phase Synthesis of N , N' -Linked Aliphatic Oligoureas and Related Hybrids. *Org. Lett.* **2012**, *14* (12), 3130–3133.

(17) García-Martín, F.; Quintanar-Audelo, M.; García-Ramos, Y.; Cruz, L. J.; Gravel, C.; Furic, R.; Côté, S.; Tulla-Puche, J.; Albericio, F. ChemMatrix, a Poly(Ethylene Glycol)-Based Support for the Solid-Phase Synthesis of Complex Peptides. *J. Comb. Chem.* **2006**, *8* (2), 213–220.

(18) Alper, P. B.; Hung, S.-C.; Wong, C.-H. Metal Catalyzed Diazo Transfer for the Synthesis of Azides from Amines. *Tetrahedron Lett.* **1996**, *37* (34), 6029–6032.

(19) Goddard-Borger, E. D.; Stick, R. V. An Efficient, Inexpensive, and Shelf-Stable Diazotransfer Reagent: Imidazole-1-Sulfonyl Azide Hydrochloride. *Org. Lett.* **2007**, *9* (19), 3797–3800.

(20) Nuyts, K.; Ceulemans, M.; Parac-Vogt, T. N.; Bultynck, G.; De Borggraeve, W. M. Facile Azide Formation via Diazotransfer Reaction in a Copper Tube Flow Reactor. *Tetrahedron Lett.* **2015**, *56* (13), 1687–1690.

(21) Schneider, J. P.; Lear, J. D.; DeGrado, W. F. A Designed Buried Salt Bridge in a Heterodimeric Coiled Coil. *J. Am. Chem. Soc.* **1997**, *119* (24), 5742–5743.

(22) Akey, D. L.; Malashkevich, V. N.; Kim, P. S. Buried Polar Residues in Coiled-Coil Interfaces † , ‡. *Biochemistry* **2001**, *40* (21), 6352–6360.

(23) Kelly, S. M.; Price, N. C. The Use of Circular Dichroism in the Investigation of Protein Structure and Function. 36.

(24) Fischer, L.; Guichard, G. Folding and Self-Assembly of Aromatic and Aliphatic Urea Oligomers: Towards Connecting Structure and Function. *Org. Biomol. Chem.* **2010**, *8* (14), 3101.

(25) Greenfield, N. J. Using Circular Dichroism Collected as a Function of Temperature to Determine the Thermodynamics of Protein Unfolding and Binding Interactions. *Nat. Protoc.* **2007**, *1* (6), 2527–2535.

(26) Greenfield, N. J. Applications of Circular Dichroism in Protein and Peptide Analysis. *TrAC Trends Anal. Chem.* **1999**, *18* (4), 236–244.

(27) Smyth, M. S.; Martin, J. H. J. X Ray Crystallography. *Mol. Pathol.* **2000**, *53* (1), 8–14.

(28) Collie, G. W.; Pulka-Ziach, K.; Guichard, G. In Situ Iodination and X-Ray Crystal Structure of a Foldamer Helix Bundle. *Chem Commun* **2016**, *52* (6), 1202–1205.

Chapter III: Conception of urea-based foldamers to target the interaction between Asf1 and H3-H4

We have previously demonstrated that the insertion of a urea-diad or triad in a peptide sequence is feasible and that overall the helical structure is maintained with nonetheless local distortions. We were interested in applying this chimera's concept to a more biologically relevant target by modulating PPIs involved in cancer and in particular the Asf1/H3-H4 complex. As discussed in chapter I.B, Asf1 is over-expressed in several cancers, including breast cancer and constitutes a valuable biological target for modern anticancer treatment[1]. The following work aims to design urea-based foldamers able to mimic the C-terminal part of histone H3 in order to modulate the natural Asf1/H3-H4 complex interaction with synthetic molecules that would potentially display better drug properties and possibly pharmaceutical interest.

This work was performed in close collaboration with the team of Françoise Ochsenbein (CEA of Saclay, University of Paris Sud) which has a unique expertise in the structural biology of the Asf1/H3-H4 interaction, including the characterization of peptides engineered to bind ASf1 using a combination of NMR and X-ray crystallography as well as molecular modeling. During my PhD training, I had the chance to be hosted in the Ochsenbein's team to perform some experiments. Remarkably, Ochsenbein *and coll.* have indeed designed α-peptides of different length and amino acid composition capable of inhibiting the Asf1/H3-H4 interaction up to the nanomolar concentration range[2]. However, synthetic and/or natural peptides present weaknesses, such as instability to proteolysis, immunogenicity in some cases and rapid clearance from the body, that make them unsuitable for cellular and in vivo applications and for their use as therapeutics. These peptide candidates were therefore selected as model sequences to engineer and finely tune the first urea-based foldamers as inhibitors of the human Asf1 and H3-H4 interaction and chromatin regulation function with improved biostability compared to natural α-peptides.

A. Design of α/urea chimeras for Asf1/H3-H4 inhibition

1. A synthetic α-helical peptide with improved binding affinity to Asf1 (developed by M. Bakail, previous phD in Ochsenbein team)

Given the large and complex interface between Asf1 and H3-H4, the inhibitory peptide strategy appeared to be more suitable than the small molecule approach. This was indeed recently confirmed by the poor affinity and selectivity of library of small molecules targeting Asf1/H3[3,4]. The fewer number of reports about this histone chaperone compared to papers about p53 for example, highlights the novelty of this anticancer target. The approach reported by F. Ochsenbein *et al.* for the design of peptides binding to Asf1 was based on the C-terminal helix-α3 of histone H3, herein referred as peptide **p1** (residues 117 to 131 of H3 encompassing the key residues: Leu126, Arg129 and Ile130). **p1** showed a measurable binding affinity for Asf1 with a dissociation constant (K_d) of 8.70±0.3µM determined by isothermal titration calorimetry (ITC). The x-ray

crystal structure of Asf1-**p1** complex at 2.0Å confirms that the **p1** peptide interacts with Asf1 in the same manner as in the full protein. However, this pentadecapeptide is only helically folded upon Asf1 binding[5]. In order to decrease the conformational entropic cost associated with the "folding and binding process", various mutations were introduced for helical stabilization. After several loops of computational designs followed by experimental validations, a first "good" affinity peptide ligand herein named **p2** (see Table 13) was conceived. *(1)* A capping box (ASTE tetrapeptide) was introduced at the *N*-terminal position with a shift of one residue compared to the reference peptide **p1**. *(2)* in addition, the peptide was lengthened on its C-terminus in order to reach the hydrophobic binding pocket of the histone H4 Phe100 residue (described in Chapter I) by engineering an α-amino acid sequence to bridge the C-terminus of H3 peptide and the N-terminal part of histone H4 β-strand. (3) Computational design with Rosetta software was performed to explore individual mutations effects. A set of amino acid systematic mutations was performed to stabilize the helical conformation by salt bridges and to gain better hydrophobic contacts with the Asf1 surface. In particular, the initial Asp6 in **p1** was substituted by a Trp making new hydrophobic contact with Asf1-Val92 and π-π interaction with Asf1-Tyr112. A global affinity gain of 100-fold with respect to **p1** was measured but the interaction could further be optimized. The last round of peptide binding optimization has consisted in performing the last critical mutations. Careful inspection of the Asf1-**p2** structure solved at 2.3Å, has indeed revealed repulsive electrostatic interactions between **p2** and the surface of Asf1. Consequently, Arg4 and Glu8 in **p2** sequence were switched. Furthermore, the Arg-Gly dipeptide motif prone to degradation and already present at the C-terminus of **p1** was mutated into Ala-Gly. The last mutation has consisted in replacing the neutral Asn22 by its Asp22 acidic version thus making an additional salt bridge contact with surface of Asf1. All together these three mutations significantly increased the binding of resulting **p3** peptide to Asf1 with a K_D value reaching the single digit nanomolar range: 3 ± 1nm.[6]

Table 13: Sequence of peptide inhibitors used as reference in the context of this PhD work. A one letter code for the natural α-amino acids is used and the Napthylalanine is referred with a three letters code Nal

	Sequence	K_D (µM)
	-1 5	
p1	T I M P K D I Q L A R R I R G	8.700±0.330
p2	A S T E R K W A E L A R R I R G A G G V T L N G F G	0.055 ±0.020
p3	A S T E E K W A R L A R R I A G A G G V T L D G F G	0.003±0.001
p3h	A S T E E K W A R L A R R I A G	0.150±0.050
p4	Ac E K **W** A R L A R R I A G	0.310±0.050
p5	Ac E K **Nal** A R L A R R I A G	0.240±0.110

X-ray crystal structure of the complex Asf1-**p3** was obtained at 1.9 Å resolution. As expected from the round of computational design and experimental evaluation, it revealed two binding epitopes corresponding to the N-terminal α-helix that reproduces the binding mode of the C-terminal α-helix of histone H3 and a C-terminal β-strand mimicking the C-terminal segment of histone H4[5,7,8]. In the same manner than in the natural complex, the major contributions to the Asf1 surface interaction occur through the helical part of **p3**. In particular, three hydrophobic residues Trp6, Leu9 and Ile13 are located into hydrophobic groove of Asf1 (Figure 57). As already observed for **p2**, Trp6 packs against Asf1 Val92 and Tyr112 side chains. As hypothesized, this interaction is also stabilized thanks to several salt bridges between residues Arg145, Asp88, Glu49 and Asp54 of Asf1 and residues Asp22, Lys5, Arg8 and Arg12 of **p3** respectively. Finally, intramolecular hydrogen bond between the carbonyl of Gly25 main chain and NH of Phe24 stabilizes the β-strand secondary structure allowing Phe24 to be well-positioned into the hydrophobic pocket but this part will be more detailed in the next chapter. In this chapter, we will focus on the α-helical part of **p3** to design α/urea hybrid compounds.

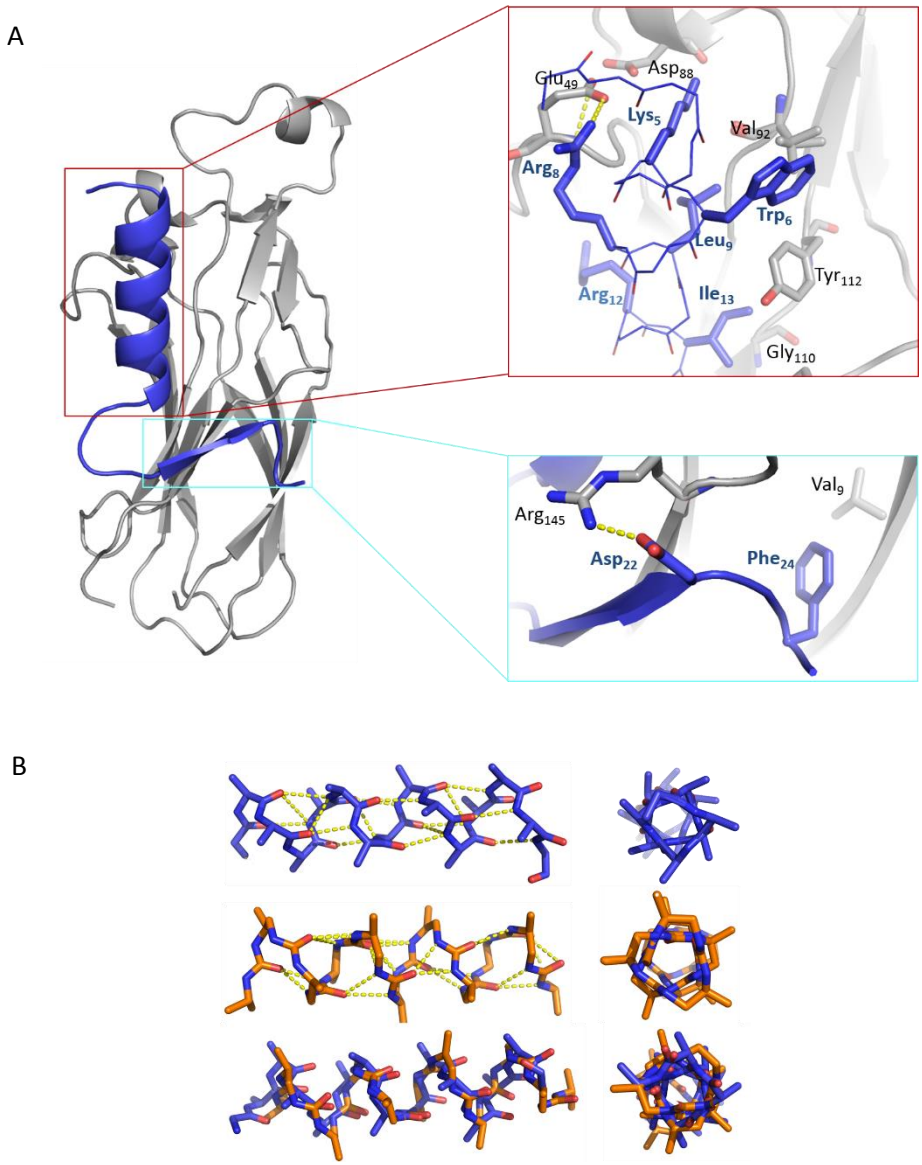


Figure 57: X-ray crystal structure of Asf1-p3 complex and similarities between the α-helical and urea foldamer backbone. A. Crystal structure of Asf1 (grey) in complex with p3 (blue). A) Cartoon representation of p3 (dark blue) in contact with the surface of Asf1 (light grey) with a zoom of the α-helical segment of p3 showing 3 key residues involved in the interaction. B) X-ray representation of peptide p3 (blue), a full oligourea (orange) and the superimposition of both.

Urea-based foldamers targeting the surface of the histone chaperone Asf1 were designed by relying on the crystal structure of Asf1-**p3** as a template. To simplify our design approach, we first focused our efforts on reproducing the key contact of the α-helix of **p3** with the surface of Asf1, herein referred as **p3h**. Importantly, as the N-terminal sequence of **p3h** (ASTE) was initially introduced as a capping box to stabilize the α-helical conformation of the peptide[9], these four residues were preliminary omitted as we sought that the urea contribution could be sufficient to lock a helix of this length. With the aim to best mimic the positioning of **p3h** α-helix, side-chain

orientation and spatial arrangement, we did a "simple" superimposition of its x-ray crystal structure from the Asf1/**p3** complex with the crystal structure of a homo-oligourea. This simple modeling quickly revealed the inherent difficulty to well mimic **p3h** with a homo-oligourea of the same length without observing a strong deviation of the urea-backbone helices and side-chain orientation that would have as major consequences steric clashes with the Asf1 surface. This first model prompted us to begin this project with an α-/urea hybrid approach in order to reproduce the folding of **p3h** and to maintain as much as possible the spatial arrangement of the key residue side chains (Trp5, Leu9, Arg12 and Ile13) at the surface of Asf1.

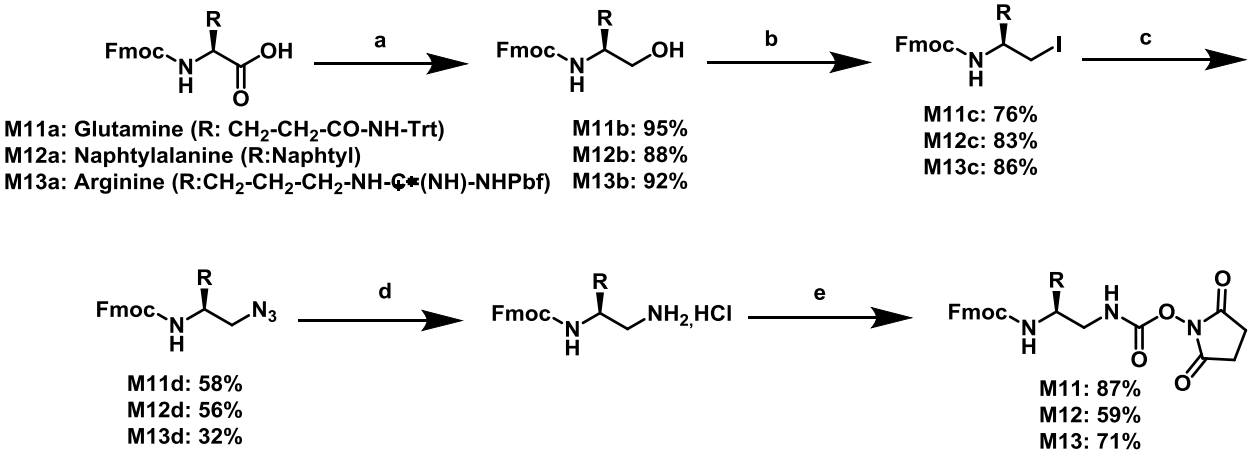
Based on our previous knowledge of the hybrid chimera approach, we selected **p3h** chimeras exhibiting three urea residues (triad) in place of four α-amino acids. As previously carried out for GCN4 Leucine zipper, we decided to perform a urea triad scan along the **p3h** backbone by each time keeping the side-chain composition. During this first round of design, we also observed that the **p3** Trp6 could potentially be substituted by non-proteinogenous aromatic residues such as a Naphtylalanine (Nal) that could better sit on an "empty" hydrophobic region of the Asf1 surface. Moreover, in order to gain additional binding information and an easy way of binding and structural comparisons between our chimeras and Asf1, two additional peptides were synthesized. Peptide **p4** that corresponded to an ASTE truncated version of **p3h** and its analogue **p5** equipped with Nal instead of Trp6 were synthesized by SPPS and their binding to Asf1 were determined by ITC (Table 1). When compared to **p3h**, K_D values are two and about 1.6 times higher for **p4** (K_d = 0.31 µM) and **p5** (K_d = 0.24 µM) respectively than for **p3h**, once again underlining here the α-helix stabilizing effect of the ASTE capping box during the binding to Asf1. Nonetheless, for peptides of this length (12-mer), **p4** and **p5** recorded K_D value seemed promising and especially the peptide **p5** can be viewed as a good starting candidate for urea-type residue installation. Additionally, installing a urea triad in this short peptide sequence may compensate the absence of the capping box by bringing new structural constraints such as the H-bond network between the carbonyl and the 2 NH of the urea function in *vis à vis*.

2. Urea triad scan of a short synthetic α-helical peptide

Urea triads were introduced along **p5** backbone at various positions in order to perform what we will call here a urea scan (by analogy to the Ala scan) aiming at identifying the best position of this short urea segment. Moreover, in order to reduce the number of positive charges meaning for **p5** the number of guanidinium moieties, the Arg8 residue which does not contribute to the Asf1 binding was replaced by a Gln one.

Firstly, activated carbamates deriving from Gln, Nal and Arg have been synthesized starting from the corresponding Fmoc-protected amino acids. The orthogonal protections, trityl for Gln and Pbf

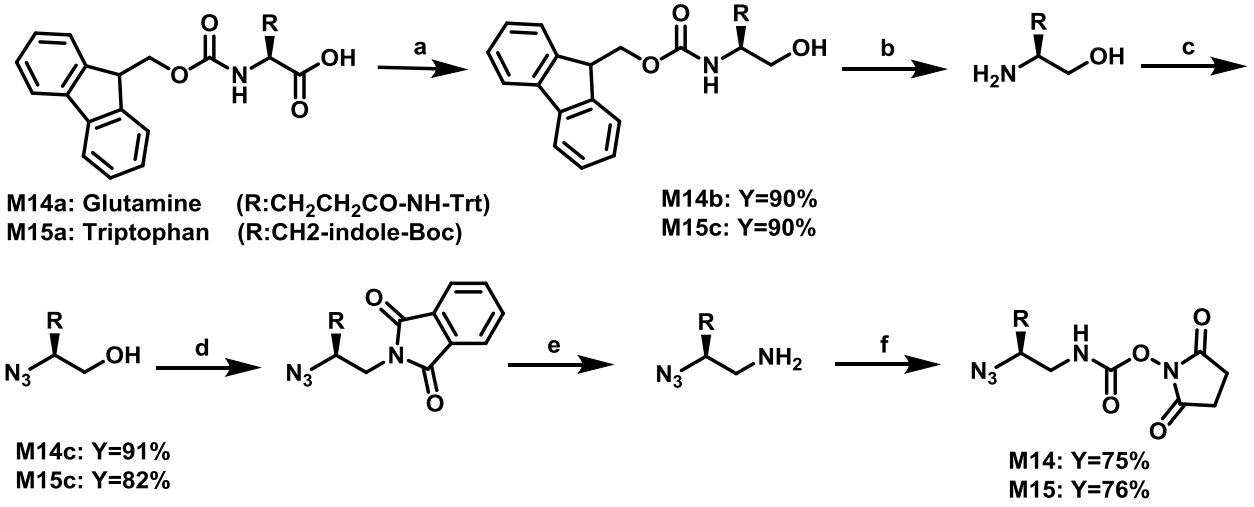
for Arg, were chosen for their compatibility with the final resin cleavage. For some chimera sequences in the group, low yields were obtained with the azide strategy due to partial reductions observed after the first urea residue. This difficulty in reducing the azide in chimera always occurs after the installation of the urea-residue on a peptide segment whereas the reverse was never observed. To overcome this problem, we thought of a new strategy that consisted to introduce the first urea residue in Fmoc strategy. Then, the activated monomer Fmoc-Gln(Trt)^{U}-OSu **M11**, Fmoc-(1)Nal^{U}-OSu **M12** and Fmoc-Arg^{U}(Pbf)-OSu **M13** have been synthesized in solution starting from the Fmoc-Gln(Trt)-OH, Fmoc-(1)Nal-OH and Fmoc-Arg(Pbf)-OH. After the formation of Fmoc-protected amino alcohol (step a), the alcohol was converted into the corresponding alkyl iodide (step b) then azide (Scheme 10). After a step of hydrogenation to obtain the corresponding amine (d), followed by the activation, the N-Fmoc protected activated carbamate monomers **M11**, **M12** and **M13** were recovered with 36%, 24% and 18% overall yield respectively.



Scheme 10: Multi-step synthesis of Fmoc-protected activated monomers M11-M13.

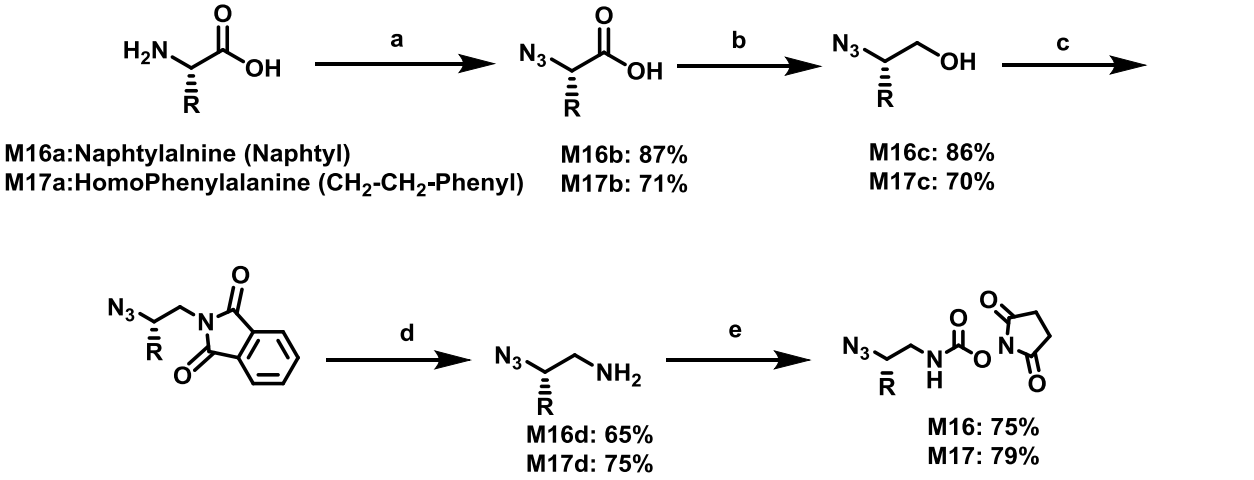
One main limitation of the use of the Fmoc strategy is the high number of equivalents required by coupling (2 x 5 eq.) and its total incompatibility with microwave irradiation[10] rendering the coupling steps much longer than the azide strategy. For these reasons, the urea residues following the first one were always introduced in azide strategy.

The activated monomer N_3-Gln(Trt)^{U}-OSu (**M14)**, was synthesized starting from the Fmoc-Gln(Trt)-OH. After the Fmoc-deprotection of the alcohol derivative and multi-step synthesis, **M14** was recovered with a yield of 23% ().



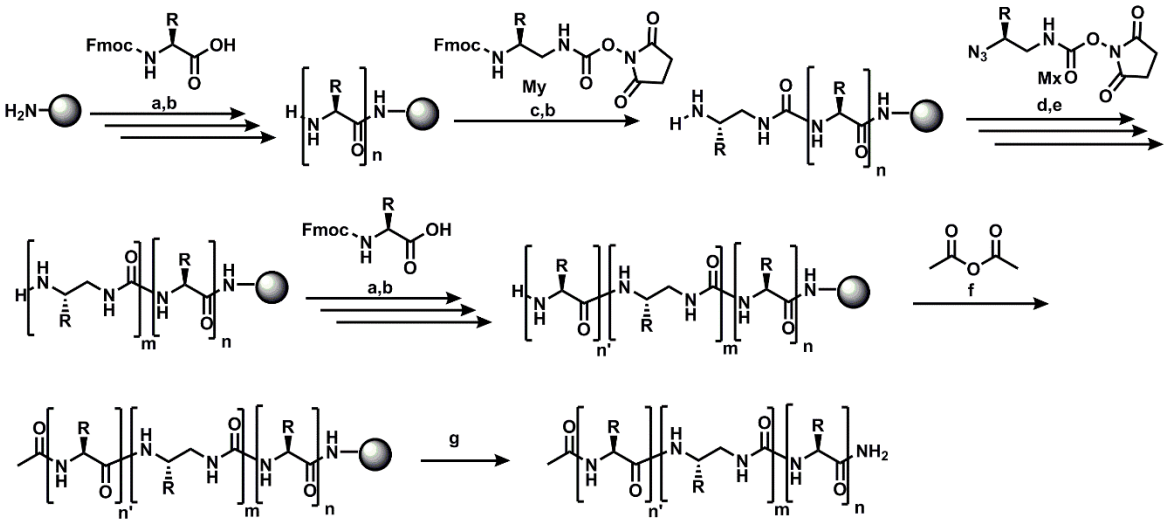
Scheme 11 : Multi-step synthesis of succinimidyl (2-azido-2-substituted ethyl) carbamate monomers, N_3-Gln(Trt)-OSu, (M14).

The activated carbamates N_3-Nal^U-OSu, **M16** and N_3-HPhe^U-OSu **M17**, have been synthesized starting from unprotected starting amino acids H-L-1Nal-OH and H-L-HPhe-OH. The side chain did not require protection, allowing to start with an unprotected amino acid and to save a deprotection-step during the multi-step synthesis. The activated monomers **M16** and **M17** were obtained with 36% and 30% global yield respectively (Scheme 12). The other activated monomers **M3** and **M6** were synthesized as described previously in Chapter II C.2.a)i).



Scheme 12: Multi-step synthesis of succinimidyl (2-azido-2-substituted ethyl) carbamate monomers N_3-Nal^u-OSu, (M16) and N_3-HPhe^u-OSu (M17).

For this series of compounds in urea scan, the solid phase synthesis (SPS) of the chimeras was performed on Rink amide MBHA resin. When the first urea residue introduced was derived from Gln, the first coupling was performed in Fmoc-strategy followed by the azide strategy (Scheme 13) whereas the azide strategy only was used for others first residue.



a: Fmoc -Xaa-OH (5 eq), DIC (5 eq), Oxyma (5 eq), 75°C , 25 W, 5 min x 2; b: 20% piperidine, DMF (75°C/25W/1'+3'); c: 5 eq My, 7 eq DIEA, 2 x 2h r.t.; d:1.5 eq Mx, 3 eq DIEA 50°C 50W 2 x 15 min; e:10 eq. PMe_3 in THF (1M), Dioxane, H_2O (7:3 v/v), 50°C, 50W, 2 x 15min; f: Acetic anhydride/ DCM 1:1 (15 min, r.t.); g: TFA/TIS/H_2O (92.5:2.5:2.5, v,v,v)

Scheme 13: Solid phase synthesis of chimeras with the introduction of the first urea residue in Fmoc-strategy

Chimeras **ch1**-**ch4** were synthesized and after purification on RP-HPLC expected compounds were identified by mass spectrometry and recovered in 15 to 30% overall yields (Table 14).

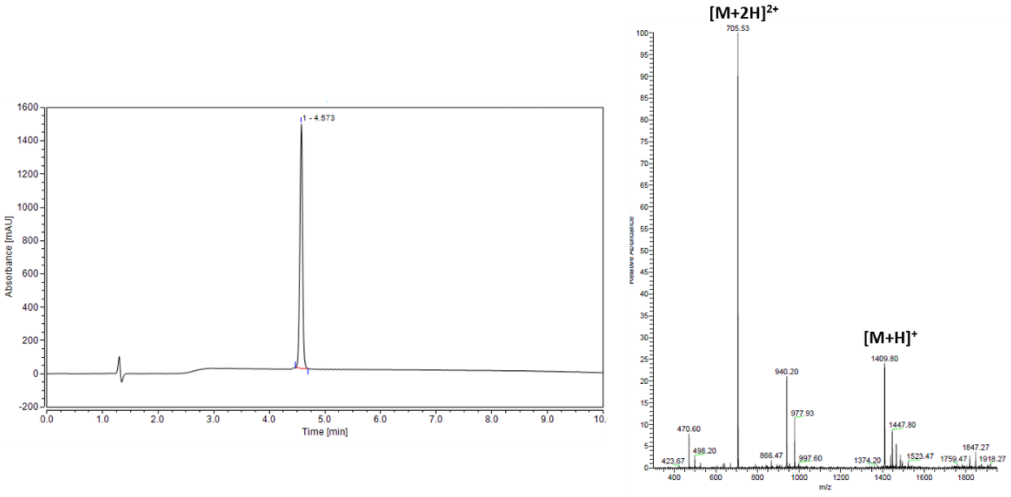


Figure 58: Example of HPLC chromatogram and ESI-MS of a pure chimera ch3

These four chimeras in hand, isothermal titration calorimetry (ITC) was used to investigate their binding to Asf1 and gain access to several important thermodynamic values including the dissociation constant K_D. For all the four chimeras, the measured values of K_d were comprised between 8.7 to 3.6 µM. The best sequence **ch3** exhibits a relatively good affinity for Asf1 (K_d= 3.6 µM), that nonetheless remains ten times higher than the binding affinity of **p5** (K_d= 0.2 µM). In this chimera, the central Leu-Ala-Arg-Arg α-tetrad of **p5** was swapped with Leu^{u}Gln^{u}Arg^u urea-triad. Moreover, as this binding affinity is more than twice better than the three other chimeras, this central position modification seems to be the most tolerated by **p5** α-helix when complexed to Asf1 and **ch3** foldamer was thus selected for a thorough structural study.

Table 14: Sequences of ch1-ch4 chimeras with urea triads and characterization by RP-HPLC and ESI-MS.

	Sequence													Yield (%)	t_R (min)	ESI-MS $[M+3H]^{3+}$
p5	Ac	E	K	Nal	A	R	L	A	R	R	I	A	NH_2	55	4.22	473.33
ch1	Ac	E	K	Nal	A	R	L	-	Q^u	R^u	I^u	A	NH_2	21	3.42	470.54
ch2	Ac	E	K	Nal	A	R	L^u	-	Q^u	R^u	I	A	NH_2	30	4.57	470.60
ch3	Ac	E	K	Nal	A	R^u	L^u	-	Q^u	R	I	A	NH_2	15	4.80	470.67
ch4	Ac	E	K	Nal^u	-	R^u	L^u	A	R	R	I	A	NH_2	28	4.37	478.96

Interestingly, ITC technique provides thermodynamic information and it is also possible to have access to the heats exchanged during the interactions between the two partners. The determination of other thermodynamic values can be obtained such as the change in free energy ΔG, the change of enthalpy ΔH and the change of entropy ΔS[11–15]. A given binding affinity can be the results of many combinations of enthalpy and entropy. These thermodynamic values were

calculated and reported for each chimera **ch1**-**ch4** and compared to those obtained for **p5** (Figure 59). Here, the analysis of these thermodynamic data reveals similar range of enthalpies and entropies for the four chimeras. The change of enthalpy is negative (meaning an exothermic mode of interaction takes place during the binding to Asf1) and varies between -3.1 and -4.5 kcal.mol^{-1} for chimeras **ch1** and **ch2** respectively. The binding enthalpy brings favorable contributions suggesting the formation of some interactions like H-bonds with Asf1. Concerning the entropies, they contribute between -2.3 and -3.7 Kcal.mol^{-1} to the binding energy. For **ch1**, the binding is mostly entropic, maybe driven by hydrophobicity or by possible rigidification of the backbone but coupled to a favorable enthalpy. In contrast, for chimeras **ch2**-**ch4**, the binding is predominantly enthalpic with favorable entropy suggesting that the hydrophobicity can be optimized.

In comparison to the peptide **p5**, the enthalpy values of the chimeras are lower whereas they are optimized entropically. These results highlight that the introduction of the urea-triad improves the order but concurrently decreases the enthalpy and therefore has a negative impact on the affinity to Asf1. Indeed, extremely good affinity is reached when both enthalpy and entropy are optimally contributing to the binding. This first series of α/urea chimeras are entropically well-suited but to optimize the affinity to Asf1, the next objective would be to improve their binding enthalpy. This unfavorable enthalpy is actually observed for many complexes[16].

Compound	K_D (µM)	N	ΔH (kCal.M^{-1})	-TΔS (kCal.M^{-1})	ΔG (kCal.M^{-1})
P5	0.21±0.23	1.14±0.01	-8.11±0.11	-0.84±0.15	-8.95±0.04
ch1	7.30±0.58	1.19±0.05	-3.14±0.17	-3.74±0.24	-6.88±0.06
ch2	8.77±1.41	1.09±0.01	-4.47±0.50	-2.31±0.59	-6.78±0.09
ch3	3.64±0.57	1.10±0.01	-4.57±0.24	-2.72±0.33	-7.29±0.09
ch4	8.20±1.80	1.30±0.05	-4.20±0.21	-2.61±0.05	-6.81±0.05

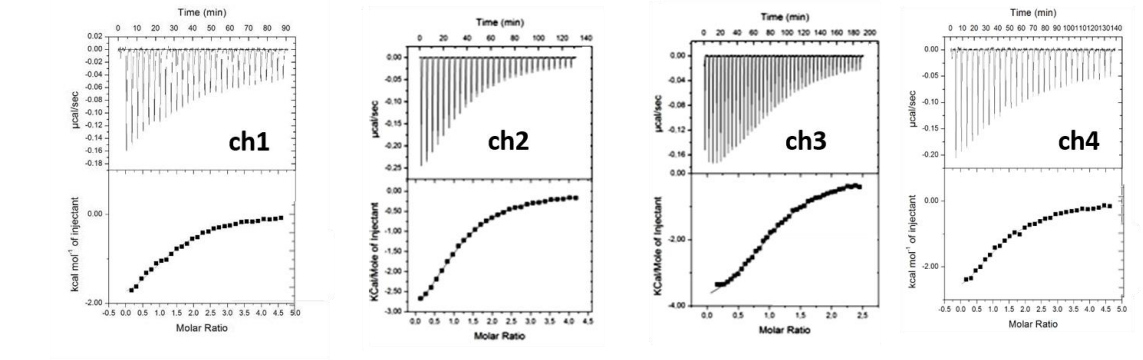


Figure 59: Thermodynamic parameters of compounds ch1-ch4 measured by ITC. ITC curves of ch1-ch4 and fitted curves are shown. Experiments were carried out at 20°C, in a 50mM Tris-HCl pH7.4

For some foldamer sequences, their interaction to the surface of Asf1 was further characterized by heteronuclear NMR techniques thus allowing us to compare their mode of binding to the one already studied in details by the group of F. Ochsenbein with their first generation of peptides[5,17]. NMR titration experiments were performed by the team of F. Ochsenbein on **ch2** and **ch3** with Asf1 uniformly ^{15}N labeled. The significant variation in the Asf1 ^{15}N-HSQC spectra reveals than **ch2** and **ch3** interact in the same region than **p3h**. However, during titration experiments, differences in the kinetic of exchange were noticed. Whereas a slow exchange rate is observed for the resonance of some protein residues (such as Asf1 Thr93) as for **p3h**, other residues are in rapid exchange (Asf1 Val95 and Tyr112) (Figure 60). Indeed, as illustrated for **ch3**, there are several crosspeaks for which two chemical shifts were observed, such as Val45 in a manner similar to **p3h**, reflecting here that the association/dissociation phenomenon is slow on NMR time scale (Figure 60B). In contrast, for some crosspeaks in the ^{1}H-^{15}N-HSQC spectrum such as Thr93, crosspeak shifting intermediates are observed reflecting a fast association/dissociation phenomenon. These observations indicate that an intermediate rate occurs for these chimeras.

All together these NMR results confirm that the original binding mode of **p3h** is partially restituted despite drastic backbone modifications with the introduction of the urea triad.

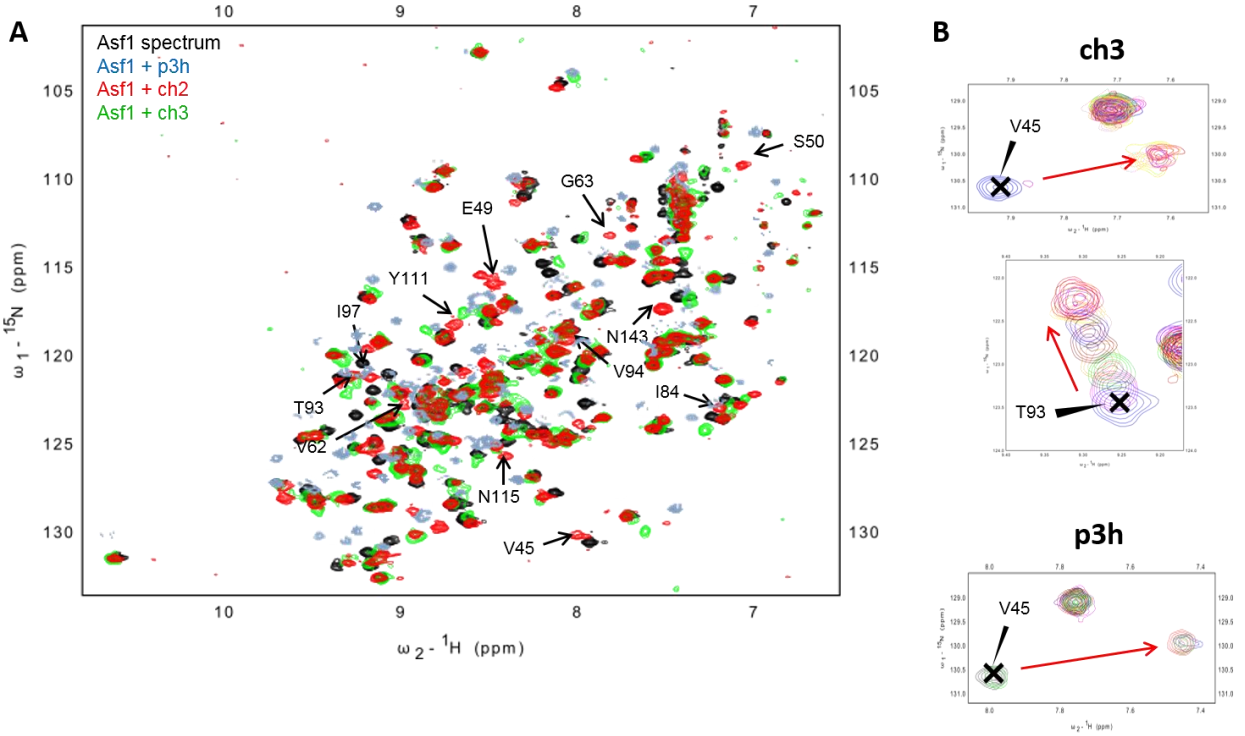


Figure 60: Comparative NMR analysis of Asf1, Asf1-ch2 and Asf1-ch3. A) Overlay of NMR ^{1}H-^{15}N HSQC spectra of free Asf1 (black), Asf1/**ch2** (red) and Asf1/**ch3** (green) complexes. B) Overlay of NMR ^{1}H-^{15}N HSQC spectra upon titration of Asf1 with **ch3** on the top and p3h on the bottom showing the NMR exchange regime. Experiments were carried out at 20°C, in Tris 10mM buffer pH7.5. Red arrows indicate the direction from the initial (free ASF1, blue with black cross) to the final (saturated ASF1, red) position of the resonance signal on the spectrum.

With the aim to get a structural snapshot of the mode of binding of this chimera with Asf1 protein, crystallogenesis screens have been performed for all compounds **ch1**-**ch4**. Remarkably, several crystals of the As1f/chimera complexes grown in different conditions were obtained for compounds **ch1** and **ch4**. After X-ray diffraction at the synchrotron (Soleil), the best data obtained were at 3.3 and 2.4 Å respectively. However, so far, the structures have not yet been solved but efforts are still devoted in this direction.

This first generation of chimera with a 4 $\alpha \rightarrow 3$ urea replacement strategy having given promising binding results, we decided to pursue this approach by increasing the number of urea residues to (1) improve the entropy of the systems and (2) to see if we could also gain in binding enthalpy. Adding extra urea-type units in the backbone of **ch3** may help to better lock the unbound helical conformation thus enhancing ΔS, the binding entropy. Indeed, the entropy of the system is

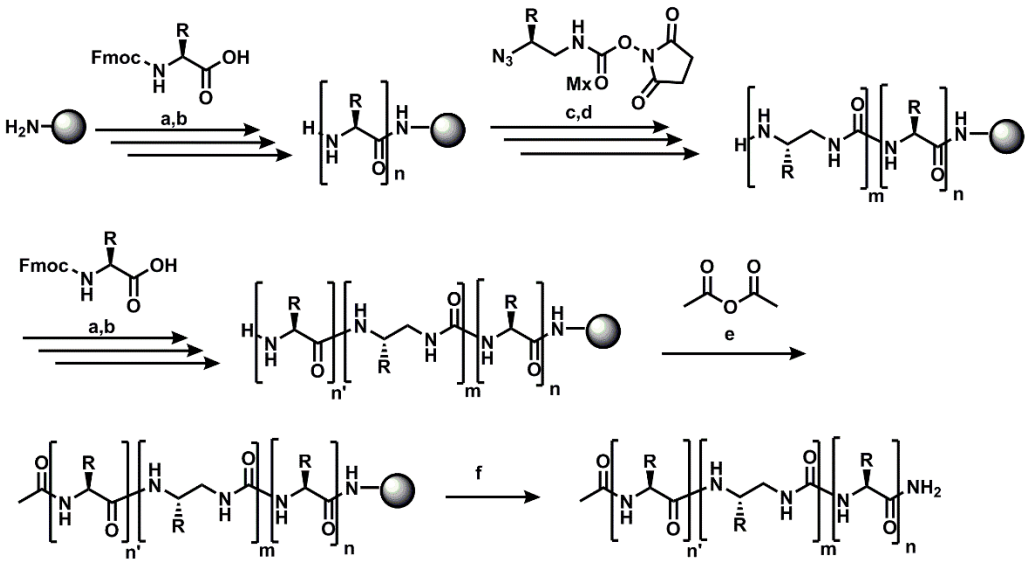
function of the chain length and the conformational entropy of each residue (α- or urea-) as it is related to the local folding and the degree of freedom[18]. Concurrently to improve the binding enthalpy (ΔH), the number of contacts (i.e. non-covalent interactions) at the interface between Asf1 surface and chimeras should be improved.

3. Different approaches to expand the oligourea region and first X-ray crystal structure of an α/urea chimera complexed to Asf1

a) Varying the number of urea residues in the chimera sequence

From the result gathered in the first series of chimera with a 4-α to 3 urea pattern we observed that the best installation of the urea triad was actually in the middle of peptide **p5**. With the aim to improve the Asf1 binding affinity, we synthesized several new sequences where one or even two extra urea residues were installed at N-terminus side of the urea triad of **ch3**. We did not want to modify the Arg9 in **p5** because from the X-ray crystal structure of **p3**, the guanidinium side chain was known to make a key salt bridge with the surface of Asf1 and its replacement could result in reduced binding affinity as previously observed for **ch2**.

Chimera **ch5** is indeed the version of **ch3** presenting a Nal^{u} urea-residue whereas in **ch6**, the naphthyl ring was substituted by a phenyl one with an additional methylene unit to bring additional flexibility to this side-chain and offer a potential better adaptation of the aromatic ring to the Asf1 surface. To compare with the previous SPS, these new chimeras were synthesized in full azide-chemistry (Scheme 14).



a: Fmoc -Xaa-OH (5 eq), DIC (5 eq), Oxyma (5 eq), 75°C , 25 W, 5 min x 2; b: 20% piperidine, DMF (75°C/25W/1'+3'); c: 1.5 eq Mx, 3 eq DIEA 50°C 50W 2 x 15 min; d:10 eq. PMe_3 in THF (1M), Dioxane, H_2O (7:3 v/v), 50°C, 50W, 2 x 15min; e: Acetic anhydride/ DCM 1:1 (15 min, r.t.); f: TFA/TIS/H_2O (92.5:2.5:2.5, v,v,v)

Scheme 14: Solid phase synthesis of chimera compounds in full azide-strategy

Chimeras **ch5** and **ch6** were only recovered with low yields (2 and 2.5% respectively). These poor yields result from a major limitation of this SPS using azide strategy introduced earlier: an incomplete reduction of the azide moiety of the newly installed Gln^u residue probably due to steric hindrance provided by the *Trt* protection on the amide side-chain. Indeed, the crude product of **ch5** obtained after cleavage was analyzed by RP-HPLC and several peaks were observed. They were isolated and independently identified by ESI mass spectrometry analysis and this partial azide reduction was mostly observed after the introduction of Gln^u but other truncated chimera sequences were also found after this urea residue. This issue of incomplete azide reduction was also observed in other projects developed in the team and seems to be sequence dependent. In all cases, the recovered quantities of **ch5** and **ch6** were sufficient to investigate their binding affinity to Asf1.

At this stage we were even tempted to enlarge the urea segment by one unit and to install a pentaurea "sandwiched" between two short peptide segments. In this context another compound was designed where Glu4 of **ch3** was conserved as a natural amino acid but Lys5 was replaced by its urea counterpart with the Lys side chain migration on the C_α carbon of the urea residue combined with an inversion of its absolute configuration ((*S*)→ (*R*)) to better reproduce the spatial orientation at the Asf1 surface. The canonical configuration would indeed result in an inappropriate orientation of the Lys side chain and consequently a steric clash with Asf1 Glu49 and would also lead to strong destabilization of the helical structure. The activated monomer N_3-

Lys(Boc)$_{inv}^{u}$-OSu **M8** was therefore prepared starting from corresponding *D*-amino acid and following the synthetic route detailed in chapter II C.2)a)ii). This activated monomer **M8** in hands, we synthesized the chimera **ch7** where a total of seven α-residues were replaced by five urea ones, which corresponds to two full α-helical turns (Table 15). Importantly, in this chimera side-chain composition, the HPhe^{u} urea residue was selected.

All the binding affinities to Asf1 were determined by performing ITC experiments and showed that the best affinity for Asf1 was obtained for chimera **ch5** equipped with tetraurea in its centre and a Naphtyl group as aromatic moeity (K_D= 2.7 µM in Table 15). This result was quite encouraging even though the affinity gain compared to **ch3** was of only 1.3-fold. The main interest here is that it confirmed that extanding the urea composition of the chimera of one unit does not have a negative impact on Asf1 binding and appears to be well-tolerated by the protein surface. Interestingly, the substitution of the Nal^{u} side chain by HPhe^{u} one induced a fall in the capacity of **ch6** to make key interactions with Asf1. Regarding the sequence **ch7** having a pentaurea segment, this chimera does not interact with Asf1 enough tightly to measure a K_D. This last result was a bit disappointing as the replacement of two full turn of helix (7 α→ 5 urea residues) is not enough efficient but this can be also due to the inversed configuration of the Lys^{U} side chain and/or to HPhe^{U} as aromatic that bring a negative contribution to the binding.

Table 15 : Second generation of chimeras equipped with urea segments composed by 4 to 5 unit and characterization by RP-HPLC and ESI-MS.

	Sequence													Yield (%)	t_R (min)	ESI-MS $[M+H]^+$
p5	Ac	E	K	Nal	A	R	L	A	R	R	I	A	NH_2	55	4.22	1420.84
ch3	Ac	E	K	Nal	A	R^{u}	L^{u}	Q^{u}		R	I	A	NH_2	15	4.80	1409.87
ch5	Ac	E	K	Nal^{u}		R^{u}	L^{u}	Q^{u}		R	I	A	NH_2	2	4.55	1367.73
ch6	Ac	E	K	HPhe^{u}		R^{u}	L^{u}	Q^{u}		R	I	A	NH_2	2.5	4.82	1331.60
ch7	Ac	E	$_{inv}$K^{u}	HPhe^{u}		R^{u}	L^{u}	Q^{u}		R	I	A	NH_2	1.7	4.72	1360.60

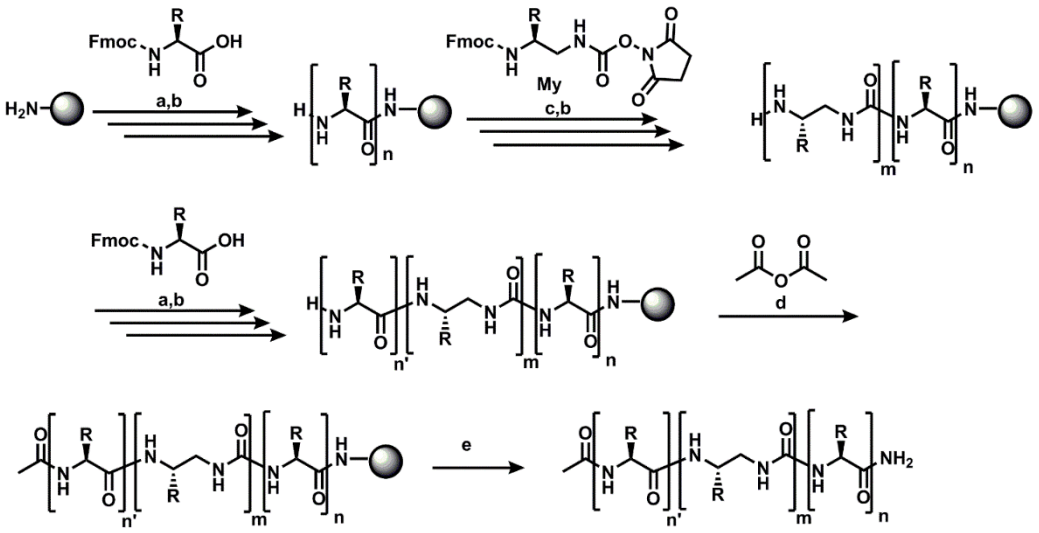
So far it is difficult to say that the observed negative effect on Asf1 binding was brought by the extra urea unit in **ch6** because the only sequence that we have evaluated did not contain the Nal^{U} and we did not investigate the consequence of repositioning the Lys side-chain on the regular Cβ of urea-residue on Asf1 binding. Nonetheless, we have acquired structural knowledge on the number of urea residues tolerated to maintain a correct binding to Asf1 and that their central location on **p5** is preferred and we can envision to modulate the nature and side chain compositions of the future generation of chimeras.

The thermodynamic data of this second generation of **p5** chimeras will be discussed in details in the next paragraph.

b) Effect of Replacing the aromatic residue in the chimeras

With the continuous aim to optimize the entropy, we designed sequences where the nature of the aromatic urea residue was modified to improve the interaction with Asf1 and its affinity. Before these modifications, I focused on the synthesis of our best chimera compounds, **ch5**, in order to compare different strategies for SPS. The low yield obtained in the previous paragraph can be explained by partial reductions of first azide residues. Firstly, we tried to increase the time of the reduction through the Staudinger reaction. Instead of twice 15 min, we tried to increase to 30 minutes twice or three times, but the yields were in the same range due to the same partial reduction problems. Secondly, **ch5** was synthesized in azide strategy except the first urea residue **M11** that was introduced in Fmoc strategy. The RP-HPLC analysis of the crude after cleavage showed a better quality of the chromatogram, in agreement with the improvement of the SPS. After purification by semi-preparative RP-HPLC, **ch5** was recovered this time in 7.5 % yield. To note, many problems occurred during this purification process leading to a loss of materials and pure **ch5**, letting us envisioning that the overall yield could have been much even better.

Finally, I have also synthesized the chimera **ch5** where all the urea-residues were installed by relying on the Fmoc-strategy (Scheme 15). The activated monomers Fmoc-(1)Nal^{U}-OSu (**M12**) and Fmoc-Arg^{U}(Pbf)-OSu (**M13**) were also required.

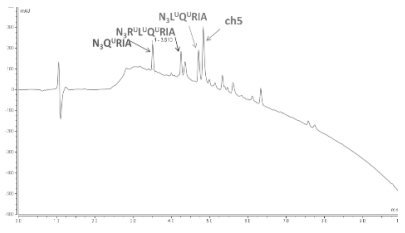


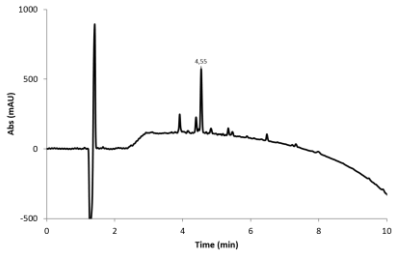
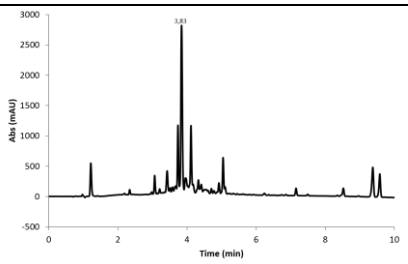
a: Fmoc -Xaa-OH (5 eq), DIC (5 eq), Oxyma (5 eq), 75°C , 25 W, 5 min x 2; b: 20% piperidine, DMF (75°C/25W/1'+3'); c: 3 eq My, 7 eq DIEA, 2 x 2h r.t.; d: Acetic anhydride/ DCM 1:1 (15 min, r.t.); e: TFA/TIS/H_2O (92.5:2.5:2.5, v,v,v)

Scheme 15: Solid phase synthesis of chimera ch5 in Fmoc strategy SIII.3

The SPS in Fmoc-strategy was next performed and **ch5** was recovered in 24% yield after semi-preparative purification. This improvement of the yield was quite encouraging but was mostly due to the ideal conditions of the purification. Indeed, the yields of crude were 56% and 49% for the first residue entered in Fmoc-strategy **SIII.2** and the full Fmoc-strategy **SIII.3** (Table 16). Because of the full Fmoc-strategy is time consuming and 10 equivalents of monomers are necessary compared to 3 eq for **SIII.2**, despite the better yield, we decided to use the Fmoc-strategy only to enter the Gln^U. We supposed that in the same condition of purification, the pure yield would be similar than what was obtained for the full Fmoc-strategy.

Table 16: Conditions of solid phase synthesis of urea segment of ch5

Synthesis	Coupling conditions	Deprotection conditions	Profile	Yields
SIII.1	DIC/Oxyma 2 x 75°C /25W/15min N_3-strategy	3 x 75°C/25W/30min		Crude Yield= 23% Pure Yield= 2.5%

SIII.2	DIC/Oxyma 2 x 50°C /50W/15min Gln^{u} in Fmoc strategy N_3-strategy	2 x 50°C/50W/15min		Crude Yield= 56% Pure Yield= 7.5%
SIII.3	DIC/Oxyma 2 x 50°C /50W/15min Fmoc-strategy	2 x 50°C/50W/15min		Crude Yield= 49% Pure Yield= 24%

After optimizing the SPS of our chimera, new sequences were synthesized through **SIII.2**. The nature of the side chain in third position of chimera **ch5** was varied with first the installation of Trp^{U} (**ch8**) and Ala^{U} (**ch9**) to evaluate the importance of the aromatic moiety for the Asf1 binding (Table 17). The determination of affinities by ITC experiment showed that the best K_D value remained the one measured for **ch5**. Having reintroduced the original Trp side-chain present in **p3h** induced a significant decreased in binding affinity (around 4-fold l increase of the K_D value compared to **ch5**; K_D=10.6 ± 4µM). The same trend was observed with the installation of Ala side-chain but this time it was anticipated (K_D=13.5±4 µM).

Table 17: Sequence of third generation of α/urea chimeras with variation of the side-chain composition at the third positions and characterization by RP-HPLC and ESI-MS.

	Sequence													Yield (%)	t_R (min)	ESI-MS $[M+H]^+$
p5	Ac	E	K	Nal	A	R	L	A	R	R	I	A	NH_2	55	4.22	1420.84
ch5	Ac	E	K	Nal^{u}	-	R^{u}	L^{u}	Q^{u}	-	R	I	A	NH_2	7.5	4.55	1367.73
ch6	Ac	E	K	HPhe^{u}	-	R^{u}	L^{u}	Q^{u}	-	R	I	A	NH_2	2.5	4.82	1331.60
ch8	Ac	E	K	W^{u}	-	R^{u}	L^{u}	Q^{u}	-	R	I	A	NH_2	4	5.30	$[M+2H]^{2+}$678.82
ch9	Ac	E	K	A^{u}	-	R^{u}	L^{u}	Q^{u}	-	R	I	A	NH_2	6.5	4.80	1241.60

The thermodynamic data were extracted from ITC experiments and are reported in Figure 61. The best combination between enthalpy and entropy was obtained for **ch5** that showed the best K_D.

The binding is mostly entropic coupled with a favorable enthalpy. The most important loss of binding affinity to Asf1 was observed for **ch8** and **ch9** with K_D values above 10 µM. These results are in favor of a better positioning of the naphtyl group at the surface of Asf1 and may better contribute to the Asf1 binding. In comparison with chimeras containing the triad oligourea segment, the combination of enthalpy and entropy differs in the sense that in this series, the entropy contribution to binding is predominant in most cases (**ch5**, **ch6** and **ch9**). However, values are in the same range of order, illustrating here that there is no thermodynamic penalty resulting from the presence of the additional urea unit. Similarly, the binding entropy is improved compared to the reference α-peptide **p5**, highlighting the robustness and the stability of the overall helical structure. However, for all the tested chimeras the absolute value of enthalpy remained much lower to that recorded for **p5** underlining that the contacts and noncovalent interactions between chimera and its protein partner are not optimal and that there is real room for improving the binding efficiency. To conclude the thermodynamic parameters gathered after ITC experiments told us that the weaker binding to Asf1 observed for this series of chimeras are the results of a modest enthalpy contribution. Nevertheless, we managed to optimize entropically our hybrid compounds.

Compound	K_D (µM)	N	ΔH (kCal.M^{-1})	-TΔS (kCal.M^{-1})	ΔG (kCal.M^{-1})
P5	0.21±0.23	1.14±0.01	-8.11±0.11	-0.84±0.15	-8.95±0.04
ch5	2.73±0.59	0.98±0.09	-2.22±0.03	-5.26±0.59	-7.47±0.12
ch6	4.63±0.27	0.98±0.08	-3.28±0.03	-3.50±0.06	-6.78±0.03
ch8	10.64±4.01	1.10±0.01	-4.01±0.21	-2.65±0.15	-6.66±0.14
ch9	13.29±3.51	1.08±0.18	-1.81±0.42	-4.73±0.57	-6.53±0.15

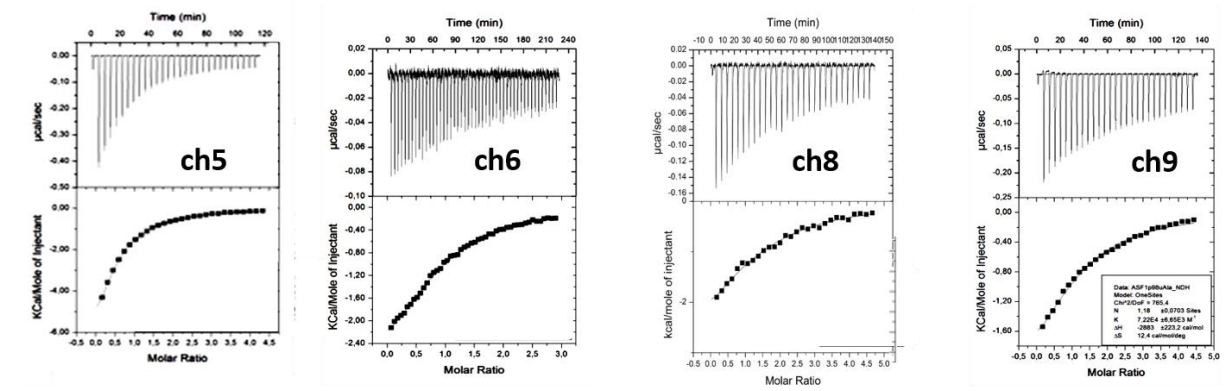


Figure 61: Thermodynamic data of compounds ch5-ch9 measured by ITC. ITC curves of ch5, ch6, ch8 and ch9 and fitted curves are shown. Experiments were carried out at 20°C, in a 50mM Tris-HCl pH7.4

As for the first series of chimeras, the structural analysis in solution was again performed using ^{15}N-HSQC NMR experiments with **ch5**, **ch6** and **ch9** to help localizing the binding region of the aromatic moiety**.** First, the best chimera **ch5** was compared to **p3h** (Figure 62). The addition of **ch5** induces significant variations of chemical shift in the Asf1 ^{15}N-HSQC spectrum that confirm its interaction with Asf1. Titration experiments revealed that a rapid exchange rate occurs, illustrating a fast dissociation and association mechanism between Asf1 and **ch5**. Furthermore, the affected residues of Asf1 (45-57, 90-97, 111 and 147) correspond to the ones involved in the binding with the natural H3 C-terminal helix[17]. In comparison with **p3h**, the Asf1 ^{15}N-HSQC spectra upon the addition of **ch5** show that the same resonances are affected but differences are observable. Indeed, some residues do not vary upon the addition of **ch5** whereas a strong variation of resonances are visible upon the addition of **p3h** (residues 60-63 and 71--77). These

affected residues upon the addition of **p3h** but not **ch5** is actually a hydrophobic patch corresponding to a non-specific binding zone (Figure 62B).

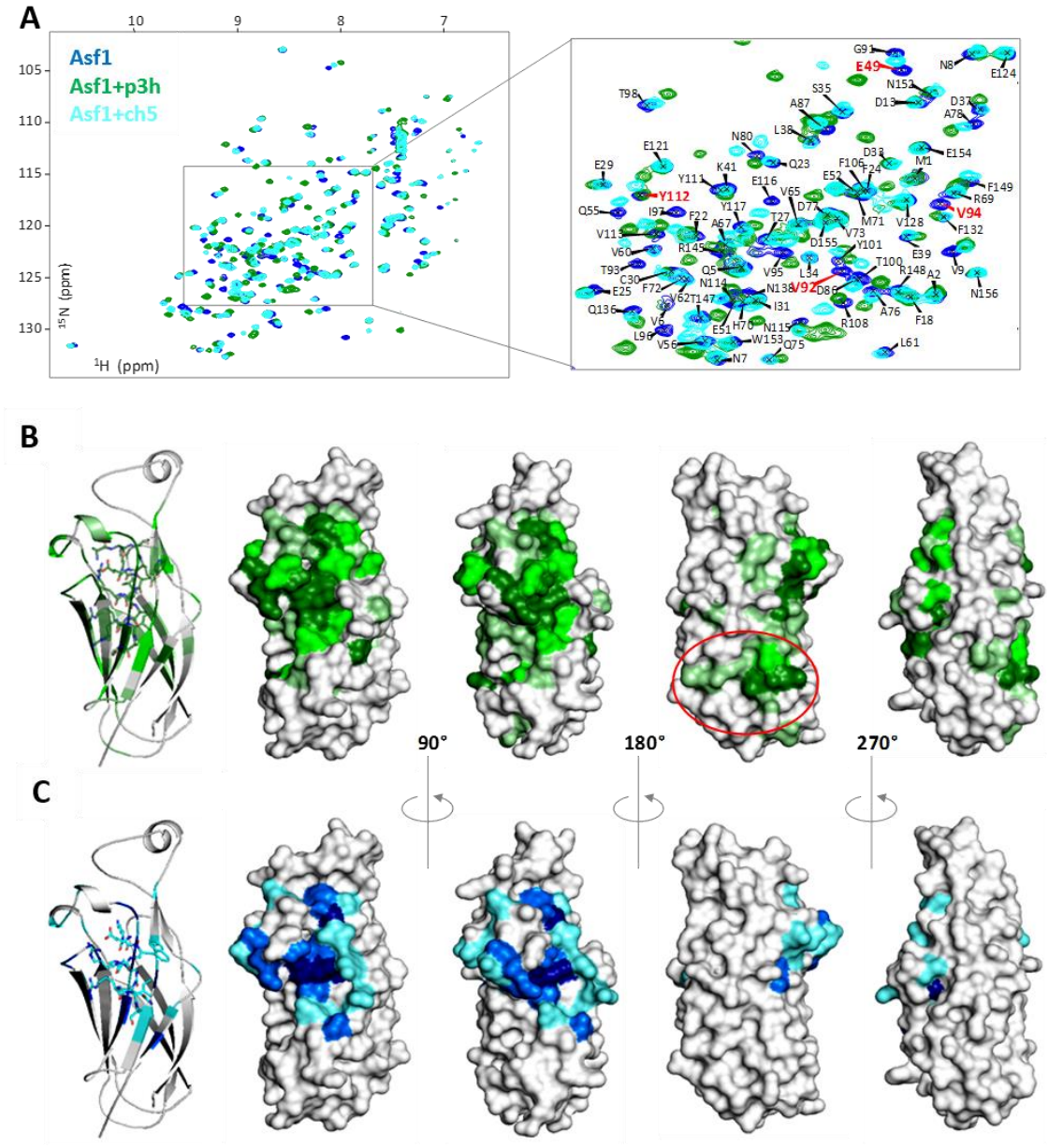


Figure 62: Comparative NMR analysis of Asf1 interacting with ch5 or p3h. A. overlay of NMR ^{1}H-^{15}N HSQC spectra corresponding to the free (blue), p3h-bound to Asf1 (green), or ch5-bound to Asf1 (cyan). Spectra were recorded at 25°C, in Tris 10mM buffer pH7.5, before and after addition of an excess of ligand (100 µM p3h or ch5). Asf1 hot spots are indicated in red. B. and C. Structure mapping of ASF1 chemical shift variations (CSV) upon p3h binding (B) or ch5 binding (C). The first representations show models of ASF1 (grey cartoon) bound to p3h (green sticks) or ch5 (blue sticks). In the second representations, surface color-coded representations of CSVs of ASF1 upon p3h or ch5 binding. From left to right, Asf1 is rotated by 90° each time with respect to the previous pose. CSV are reported as slightly (0.07-0.14, light color), moderately (0.14-0.21, moderate color), and highly (>0.21, dark color) affected Asf1 residues upon ligand binding. Variation below 0.07 ppm are considered negligible (grey). The coloration is green for p3h (B) or blue for ch5 (C) according to CSV cutoffs. The red circle represents the non-specific region.

In addition, upon the addition of **p3h**, a slow exchange rate is observed suggesting that a slow association/dissociation occur between Asf1 and **p3h**. These results prove that the referent peptide **p3h** and the chimera **ch5** present a different binding mode. However, the region of the interaction is the same and the chimera have the advantage to be more selective. Another sequence was synthesized and tested where ASTE was added in N-terminal of **ch5** (ASTE-ch5), used as a capping box for **p3h**. It did not change the binding properties of **ch5** as all data were comparable to **ch5**.

The ^{15}N-HSQC spectrum upon the addition of **ch5** and **ch6** shows a similar variation but the chimera **ch6** induces smaller chemical shift variations than **ch5**. In the case of **ch9**, the difference is stronger as the chemical shifts variation are more important. For all these chimeras, a rapid exchange rate is occurring.

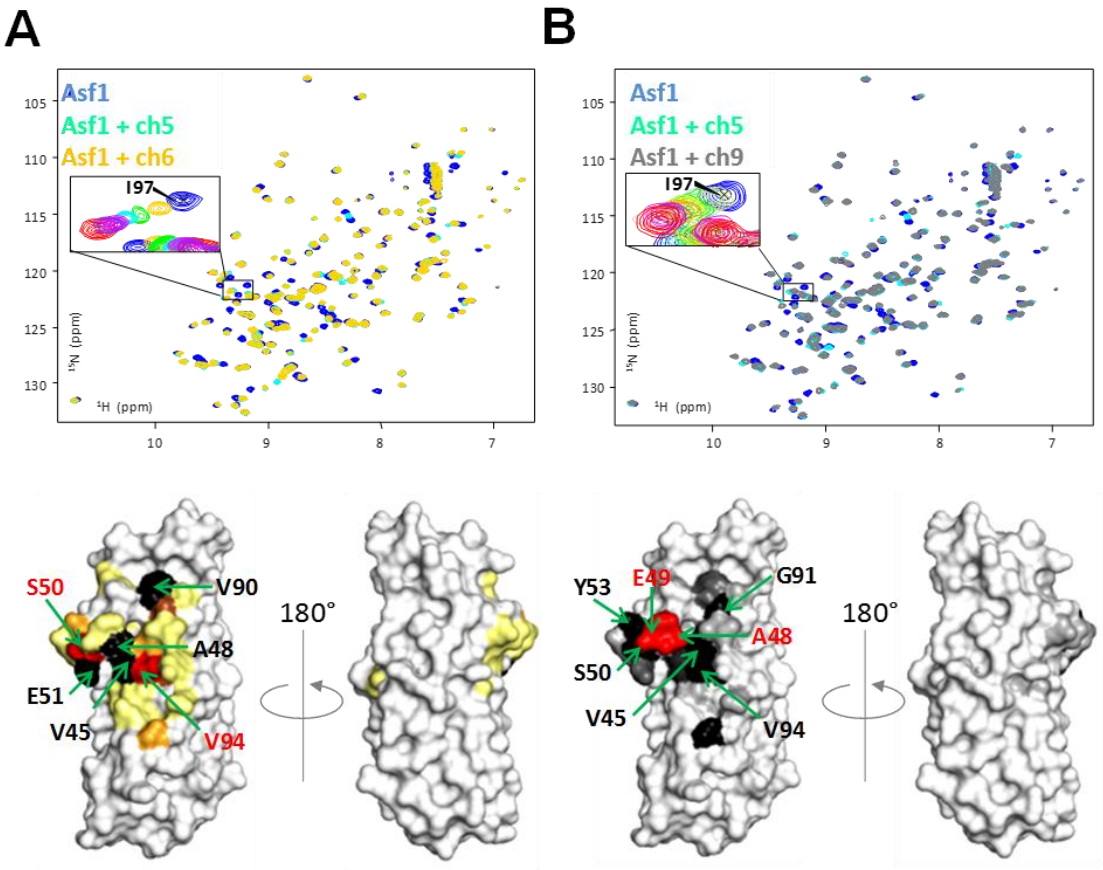


Figure 63: Comparative HSQC analysis of ch5, ch6 and ch9. A and B, upper panels, overlay of NMR ^{1}H-^{15}N HSQC spectra corresponding to the free Asf1 (blue), **ch5**-bound (cyan), ch6-bound (A) (orange), or **ch9**-bound (B) (red) forms of ASF1. Spectra were recorded at 25°C, in Tris 10mM buffer pH7.5, before and after addition of an excess of chimera (100 µM). The inset shows ASF1 Ile$_{97}$ signal variation upon titration with **ch6** or **ch9**. **ch6** and **ch9**, like ch5, show a rapid exchange rate. A and B, lower panels, surface color-coded representations of CSVs of Asf1 upon ch6 (A) or **ch9** (B) binding. Asf1 is in surface representation. In black and red, residues more or less affected upon **ch5** binding, respectively. Two orientations of ASF1 are shown (0-180°) (adapted from May Bakail's PhD thesis manuscript)

These results show the importance of naphthyl group for the interaction with Asf1 and confirms that the nature of the aromatic has an impact on the binding. As **ch5** is the best chimera, we used its sequence to check our hypothesis concerning the removal of the N-capping box of **p3h**. Indeed, we supposed that in presence of oligourea segment this cap should not be necessary. Then, the sequence **ch5** with ASTE added in N-terminal (**ASTE-ch5**) was synthesized and tested in solution. This cap did not change the binding properties of **ch5** as all data were comparable. Indeed, ^{15}N-HSQC spectra are overlapping considerably (Figure 64A) and the measure of affinity by ITC is comparable (K_D=4.27±0.52 µM). Otherwise, a negative peptide control was synthesized in order to confirm the importance of the oligourea segment in **ch5**. The peptide **p6** (= Ac-EK-Nal-RLQ-RIA-NH_2) was synthesized on Liberty Blue and recovered after RP-HPLC with 65% yield. ^{15}N-HSQC spectra show that there is no impact of **p6** in the region where **ch5** interacts (Figure 64B). However, the non-specific site of Asf1 is affected upon the addition of **p6**. This result was supported by ITC experiments where no interaction could be measured. It confirms the importance of oligourea segment and that our design allowed the conception of a chimera able to mimic the natural peptide.

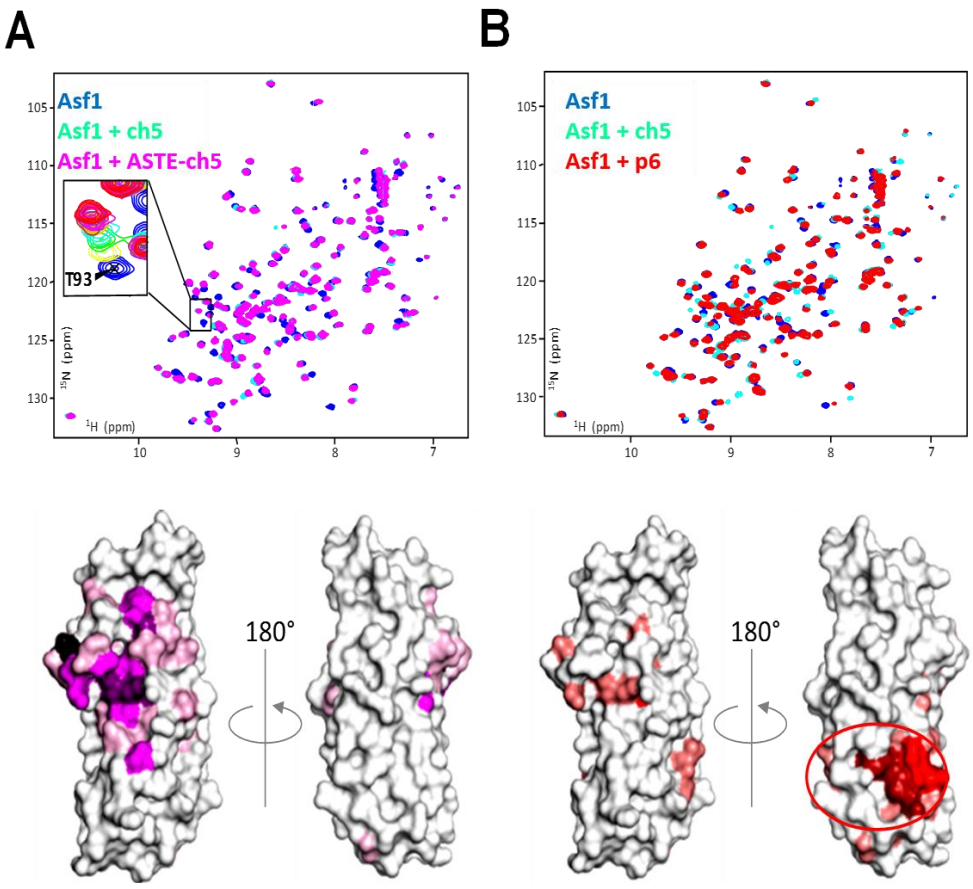


Figure 64: Comparative HSQC analysis of ch5, ASTE-ch5 and p6. A and B, upper panels, overlay of NMR ^{1}H-^{15}N HSQC spectra corresponding to the free Asf1 (blue), **ch5**-bound (cyan), ASTE-**ch5**-bound (A) (magenta), or **p6**-bound (B) (red) forms of Asf1. Spectra were recorded at25°C, in Tris 10mM buffer pH7.5, before and after addition of an excess

of chimera (100µM). The zoomed box shows Asf1 T93 signal variation upon titration with **ASTE-ch5**. As **ch5**, **ASTE-ch5** shows a rapid exchange rate. A and B, lower panels, surface color-coded representations of CSVs of Asf1 upon ASTE-ch5 (A) or p6 (B) binding. Asf1 is in surface representation with hot spots residues colored in red (acidic residues), and black (hydrophobic residues). Two orientations of ASF1 are shown (0-180°).

Finally, a sequence inspired by **ch5** where the basic amino acid Lys_2 was replaced by an Arg was synthesized. This sequence **ch10**= ER-Nal^{u}-Arg^{u}-Leu^{u}-Gln^{u}-RIA was synthesized as described before and recovered with 10% yield after purification on RP-HPLC. ITC experiment led to the determination of a K_d=4.6±0.5 µM illustrating the importance of Lys2 for the specific binding to Asf1. All in all, **ch5** is the chimera with the best affinity for Asf1 in this series, and we decided to focus on this compound and try crystallization in the presence of Asf1.

c) Crystal structure of ch5 interacting with Asf1

For the best Asf1 binding α/urea hybrid foldamers, many efforts have been devoted in crystallogenesis trials in order to have access to the solid state structure of the resulting complexes. During her PhD thesis in the group of F. Ochsenbein, May Bakail managed to grow crystals of Asf1-**ch5** complex and solved the crystal structure at 1.8 Å resolution. She used the X-ray crystal structure of human Asf1a (PDB 2I32) as reference and after having performed molecular replacement she determined that the space group was P1.

*Table 18: Crystal data collection and refinement statistics for the Asf1 –**ch5** complex*

Asf1-ch5 complex	
Data collection	
Space group	*P1*
Cell dimensions	
a. *b*. *c* (Å)	36.73. 51.28. 56.08
a. b. g (°)	112.38. 104.87. 101.190
Resolution (Å)	48.29-1.8 (1.91-1.8)
R_{merge}	8.4(66.6)
I / s*I*	9.41(2.14)
Completeness (%)	96.1(70.5)
Redundancy	3.9(2.6)
Refinement	
Resolution (Å)	1.8
No. unique reflections	31049
R_{work} / R_{free} (%)	17.64/20.72
No. atoms	
Protein	2576
Ligand/ion	369
Water	180
B-factors	
Protein	35.2
Ligand	47.1
Ion/small molecules	87.8
Water	44.3
R.m.s. deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.10

*Number of xtals for the structure: 2. wavelength= 0.978570
*Values in parentheses are for highest-resolution shell.

The X-ray crystal structure of Asf1-**ch5** complex confirmed that **ch5** interacts in the same region than **p3h**, corresponding to the histone H3 binding domain of Asf1. This supports the results gathered in solution by heteronuclear NMR spectroscopy. At first glance, the overall helix conformation of the chimera **ch5** at the surface of Asf1 is maintained. However, when looking in more detail a distortion of the helix structure is clearly visible and highlights a more opened helical conformation which is maintained by a H-bond network comprising the urea backbone but also by an H-bond between Tyr112 and the carbonyl of Leu^{u}5 (Figure 65B and D). A water molecule is also present at the center of the helix inducing a bridge of H-bonds within the urea segment that may stabilize the helix. Several other interactions occur between **ch5** and Asf1 such as salt bridge between Lys^{u}2 and residues Glu49 and Asp88 of Asf1. Other salt bridges occur also between Arg7 and Arg^{u}4 of **ch5** with Asp54 and Glu51 of Asf1.An additional H-bond occurs between the side

chain of Arg7 and the carbonyl of its main chain. Remarkably, the two key residues Leu^{u}5 and Ile8 of **ch5** are well positioned into hydrophobic pocket formed by residues Val92, Val94, Gly110, Tyr111 and Tyr112 of Asf1 (Figure 65C).

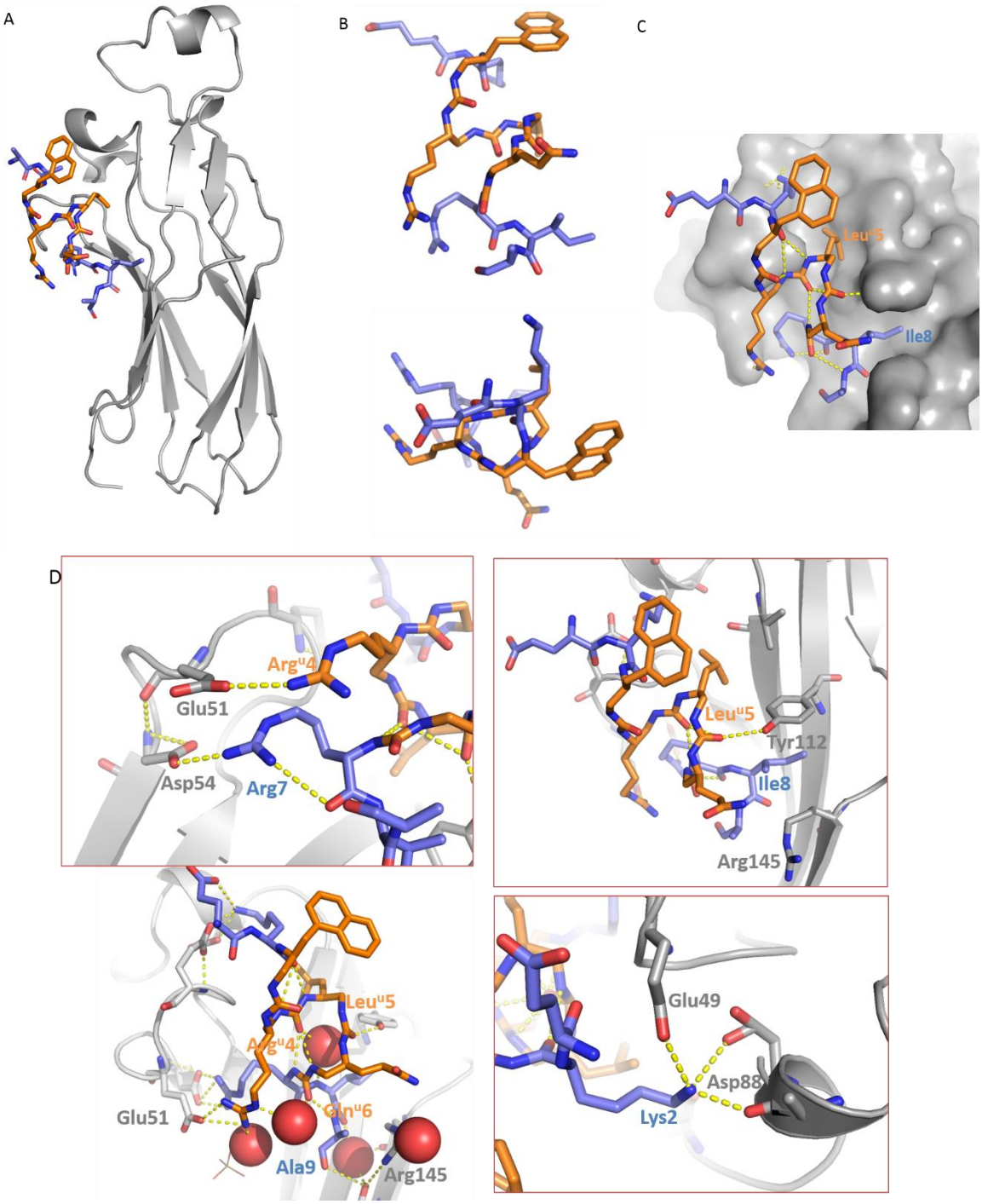


Figure 65: Crystal structure of the complex Asf1-ch5 at 1.8 Å. A. Overview of the structure of the complex Asf1 (grey) and ch5 with peptide segment in blue and oligourea segment in orange. B. Helical conformation of the chimera ch5 and top view. C. Representation of ch5 at the surface of Asf1 with H-bonds represented in yellow dots D. Zoom on the contacts between Asf1 and the lysine side chain at P2 of ch5. Red dots correspond to water molecule

However, some residues show a binding mode to Asf1 that differs from what would be expected. For example, $Arg^{U}4$ interacts with Asf1 by making a salt bridge with Glu51, thus mimicking the Arg12 of **p3h**. In addition, the surface of the guanidinium which is hydrophobic below the plane of the guanidinium[19] actually stacks over the alkyl chain of Arg7. Of particular interest, is the additional hydrogen bond that occurs between the carbonyl of $Leu^{U}5$ and the hydroxyl group of Tyr112. This interaction causes the oligourea/peptide helix to bend and to adopt a non-canonical conformation upon binding to Asf1. However, the napthyl ring of $Nal^{U}3$ that we were expecting to interact trough hydrophobic contacts with Asf1 Val92 and Tyr112 side-chains does not make any interaction and seems rather being oriented toward solvent. This result is quite surprising because we previously observed that this naphtyl group is involved in the binding to Asf1; its mutation to Phe^{U} or even Trp^{U} diminishes considerably the binding affinity (see Table 5). When looking more in details at the X-ray crystal structure, by adding asymmetrical units, it appeared that the $Nal^{U}3$ residue is actually positioned in the non-specific region of an asymmetrical molecule of Asf1. This was not observed in solution and may be due to crystal packing forces.

Overall, these results confirm the direct interaction between **ch5** and Asf1 and provide us with key structural information on the binding of this family of urea-based foldamers to Asf1. Moreover, the observed helix distortion suggests here that the urea-helix presents some plasticity to adapt itself to Asf1 surface, which is a main structural feature of these α/urea hybrid chimeras that, to the best of our knowledge, has not been reported so far for other foldamer architectures targeting PPIs. The results gathered with this first X-ray structure are valuable and should be practically useful for future designs of chimera sequences with improved binding affinity to Asf1.

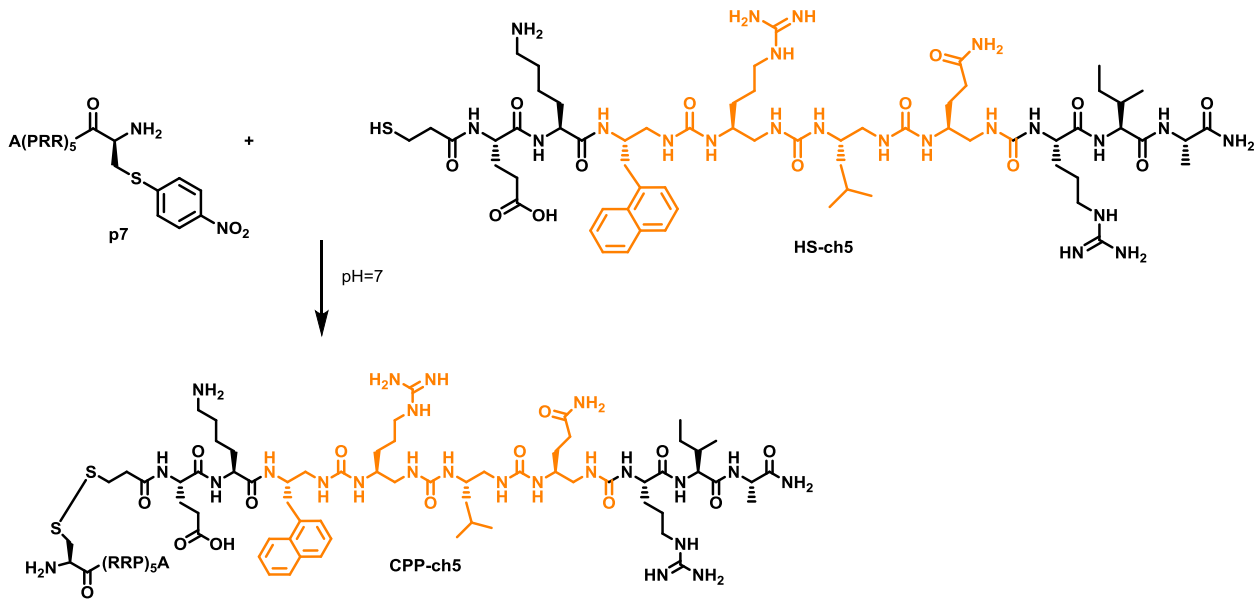
d) *In cellulo* activity of α/urea chimera ch5

i) Conjugation of **ch5** with a cell-penetrating peptide

As described in Chapter I, Asf1 is an intracellular target meaning that with the aim to inhibit its chaperone function, the synthesized α/urea chimeras must be able to cross over the cell membrane and reach the cytoplasm where Asf1 is located. In this context, we decided to append a cell-penetrating peptide (CPP) at the *N*-extremity of our lead compound, **ch5**. Cell-penetrating peptides brought major advances for the development of intracellular peptide-like drugs[20,21]. CPPs are capable to transport different cargos (oligonucleotides, proteins, plasmids) across the membrane. Depending on the nature and composition of CPP, the mechanism of cell penetration can vary: endosomal uptake is often encountered, but also direct membrane crossing can occur[20,22,23]. Usually, CPP sequences involve positively charged amino acids. Besides, Schepartz et al. have reported that the insertion of two or three Arg amino acids at judicious position into an α-helix can lead to cell permeability[24]. Active inhibitory peptides can be equipped at either N- or

C-terminus with a variety of CPP sequences to help their entrance in mammalian cells[25]. In the context of this project, we have decided to lengthen **ch5** sequence at its N-terminal end with the well-known sequence $(Arg\text{-}Arg\text{-}Pro)_5$, a proline rich CPP highly efficient for the delivery of peptide cargos within cells[26].

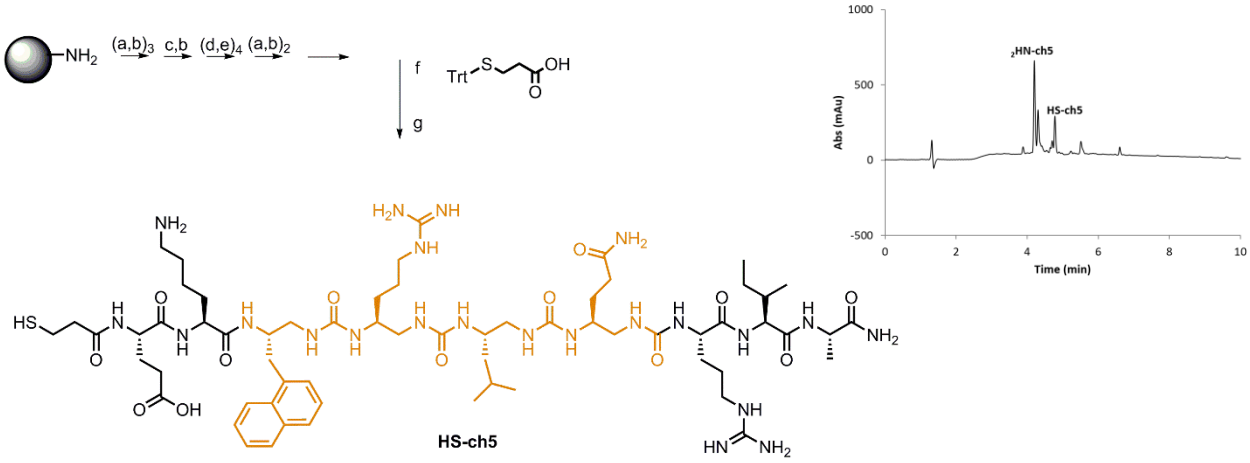
Regarding the synthesis of this CPP adduct, to circumvent any SPS issues related to the chain length of this peptide/urea construct, we decided to introduce a linker at the N-terminal part of **ch5** and to attach the CPP $C(RRP)_5A$ by relying on disulfide bridge formation (Scheme 16). Moreover, incorporating, a bioreducible disulfide linker would also enhance the chimera **ch5** delivery and release within the cytoplasm because the CPP segment would be cleaved by the presence of excess of glutathione inside the cells once the CPP adduct would have been internalized. Several synthesis trials have been performed to find how to conveniently couple the $H\text{-}Cys\text{-}(Arg\text{-}Arg\text{-}Pro)_5\text{-}Ala\text{-}NH_2$ to thiol-modified **ch5** (**HS-ch5**). First, the peptide **p7** (= $(NPys\text{-}S)\text{-}C\text{-}(RRP)_5\text{-}A\text{-}NH_2$) was synthesized on the automatic synthesizer *Liberty Blue* onto polystyrene rink amide resin. The commercially available Boc-L-Cys(NPys)-OH was used for the last coupling performed manually. After purification on RP-HPLC, the vector peptide was recovered with 40% yield.



Scheme 16: Conjugation approach for the synthesis of CPP-ch5

Second, we wanted to incorporate a S-trytil-3-mercaptopropionic acid on the resin bound **ch5** free amine to give rise to compound **HS-ch5.** (Scheme 17) However, the HPLC profile and MS analysis of the crude material recovered after TFA cleavage showed that this linker has not been

efficiently installed (as the desired product represented 24% of the crude) and the main product obtained (47% of crude) corresponded to the free N-terminus amine of **ch5**.



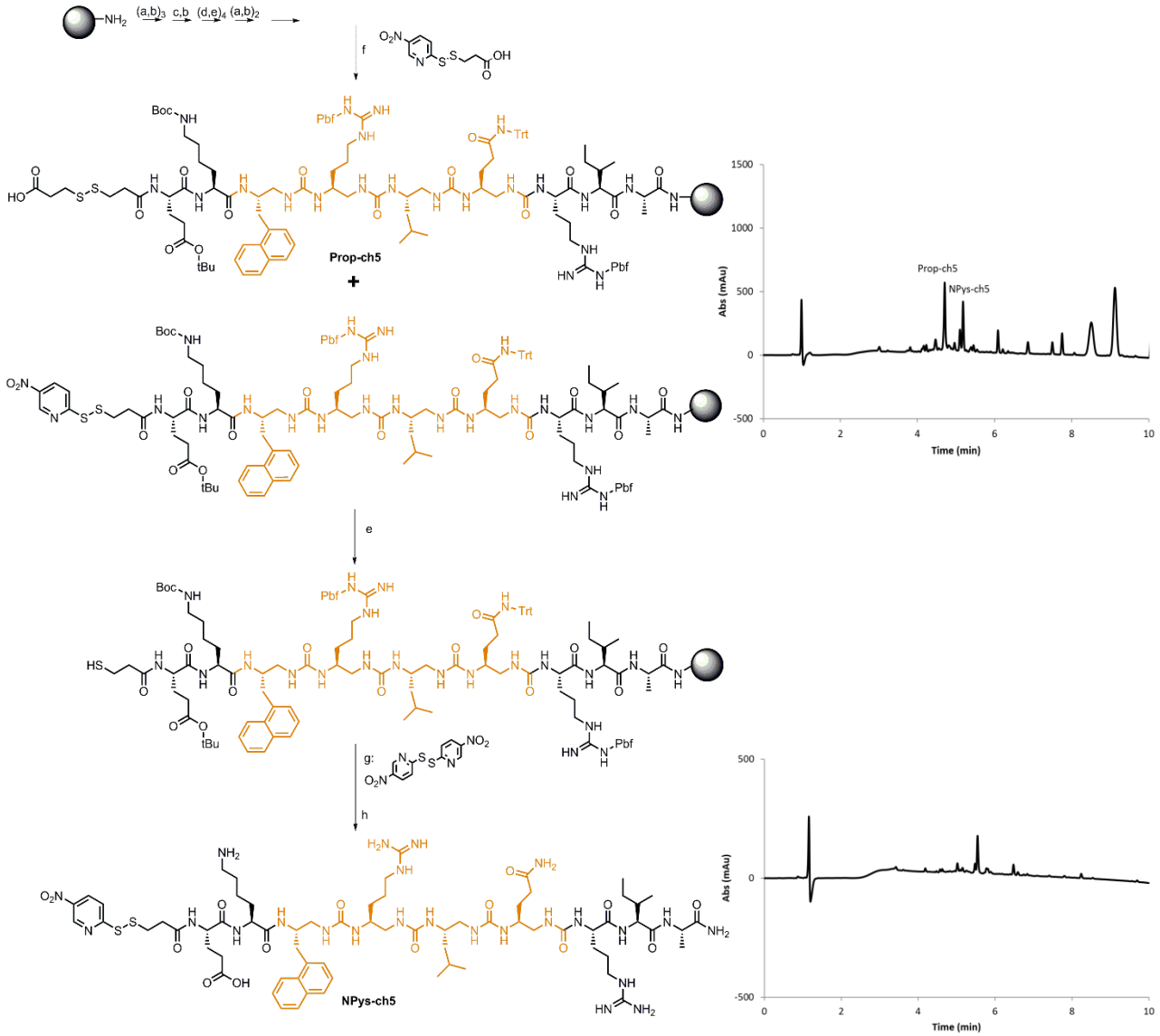
a: Fmoc -Xaa-OH (5 eq), DIC (5 eq), Oxyma (5 eq), 75°C , 25 W, 5 min x 2; b: 20% piperidine, DMF (75°C/25W/1'+3'); c: 5 eq My, 7 eq DIEA, 2 x 2h r.t.; d:1.5 eq Mx, 3 eq DIEA 50°C 50W 2 x 15 min; e:10 eq. PMe_3 in THF (1M), Dioxane, H_2O (7:3 v/v), 50°C, 50W, 2 x 15min; f: DIC (5eq), Oxyma (5eq), overnight, r.t.; g:TFA/TIS/H_2O/EDT (92.5:2.5:2.5:2.5, v,v,v,v)

Scheme 17 : Solid phase synthesis of HS-ch5 and its HPLC chromatogram

We hypothesized that the low coupling efficiency of this linker was the consequence of a steric hindrance induced by the Trityl protection on the thiol function. Furthermore, the conjugation with the CPP did not work properly, but was performed on few quantities. To circumvent this problem, we thought to change our strategy and to introduce directly the activated version of 3-Mercaptopropionic acid equipped with 3-nitro-2-pyridinesulfenyl (NPys) moiety (Scheme 18). Indeed, this activating group being TFA-resistant, it can be also easily removed in presence of tri-alkylphosphine and water[27]. The synthesis was performed following **SIII.2**, except that the coupling of NPyS-S$(CH_2)_2$COOH was performed at r.t. To confirm the efficiency of the last coupling properly, a microcleavage was performed after washings of the resin. Surprisingly again, two major peaks were observed by RP-HPLC and after LC-MS analysis corresponded to the desired product **NPys-ch5** bound to the resin and to a side-product **Prop-ch5** resulting of the addition of the free 3-mercaptopropionic acid to the resin-bound activated thiol (see Scheme 18).

This LC-MS results obtained from the crude material suggests here that once loaded on the free-amine bound resin, the NPys activated linker has probably reacted with the free-thiol linker remained in solution. This side-reaction took place due to an extended time of coupling thus letting for the excess of S-Npys-3-mercaptoproprionic acid to decompose in basic media and release the corresponding thiol-free 3-mercaptoproprionic acid. Even though the recovered compound was not entirely the expected one, this result underlines that this time the coupling

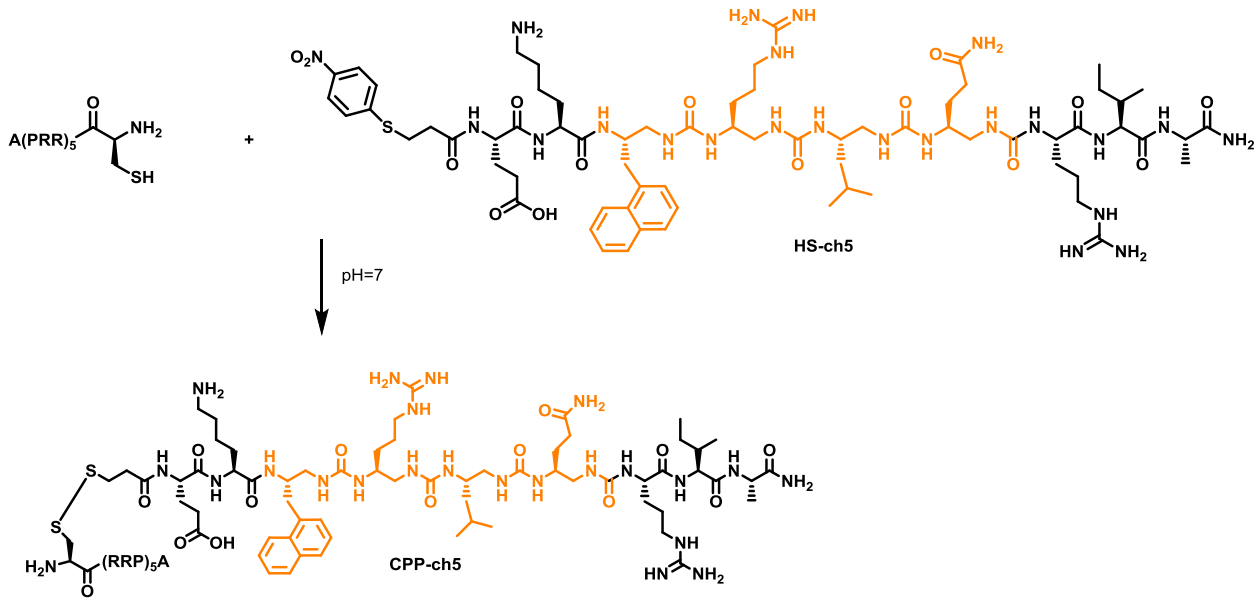
reaction was efficient. In order to save the synthesis time as well as to overcome the by-product formation on the remaining resin, we decided to reduce the disulfide bridge on resin by using trimethylphosphine and water on resin-bound **NPys-ch5** and **Prop-ch5** (both being present on the same batch of resin). Then the activating NPys group was re-installed by reaction in presence of 2,2'-dithiobis(5-nitropyridine) in a basic medium at room temperature. After final TFA cleavage and after semi-preparative RP-HPLC purification, the expected oligomer **NPys-ch5** was recovered with 6.5 % overall yield.



a: Fmoc -Xaa-OH (5 eq), DIC (5 eq), Oxyma (5 eq), 75°C , 25 W, 5 min x 2; b: 20% piperidine, DMF (75°C/25W/1'+3'); c: 5 eq My, 7 eq DIEA, 2 x 2h r.t.; d:1.5 eq Mx, 3 eq DIEA 50°C 50W 2 x 15 min; e:10 eq. PMe_3 in THF (1M), Dioxane, H_2O (7:3 v/v), 50°C, 50W, 2 x 15min; f: DIC (5eq), Oxyma (5eq), 3-(2-(5-nitropyridin-2-yl)disulfanyl)propanoic acid (5eq), overnight, r.t.; g: DIEA (1.5 eq), 2,2'-Dithiobis (5-nitropyridine) (1.5 eq); h: TFA/TIS/H_2O (95:2.5:2.5, v,v,v)

Scheme 18: SPS of NPys-ch5 with removal of the side product

The **NPys-ch5** compound in hands, the CPP vector peptide could be conjugated. As the peptide **p7** was already synthesized and stocked, the NPys was removed in presence of DTT to couple the resulting peptide to **NPys-ch5** in highly concentrated media at neutral pH (Scheme 19). The compound **CCP-ch5** was purified on semi-preparative HPLC, recovered with 42% yield and could be used for cellular test.



Scheme 19 : Chemical structure of ch5 conjugated with the CPP H-Cys-(Arg-Arg-Pro)$_5$-Ala-NH$_2$ with the urea segment depicted in orange, peptide extremities are depicted in black and the CPP is in black.

Once this conjugate between CPP and **ch5** in hands, cellular experiments were performed in order to evaluate the biological activity of the α/urea hybrid foldamer **ch5** *in cellulo* by targeting cancer cell lines. Previous study in the group of F. Ochsenbein have first revealed that **CPP-p3** exhibited a direct action on Asf1 activity *in cellulo* with regression of cell growth presumably by directly inhibiting the replication process leading to cell death (paper submitted). To monitor the cell viability in presence of our **ch5** conjugate **CPP-ch5**, a series of cellular assays such as MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assays (colorimetric tests used to assess the metabolic activity of cells) were performed on U20S cancerous cell lines[28].

Another peptide where the 3 key residues of **p2** were replaced by Ala side chain was used as negative control **(p8**=ASTERKAAEAARRARGAGGVTLNGFG). Then, natural peptide, **CPP-p8**, **CPP-p2** and **CPP-ch5** were tested into cell. Because of the poor biostability of natural peptides, each conjugated inhibitor was delivered twice a day for 48h. During this cell assay, **ch5** was also used free of conjugation in order to determine if this chimera could self-penetrate cancer cells.

Actually, the results show that **ch5** alone behaves like **CPP-p8**, meaning that there is no observed cytotoxic effect on U20S cell line whereas the **CPP-p2** decreases considerably cell viability (Figure 66A). This confirms the necessity to use a vector peptide for a cellular entry and activity of **ch5**.

Fluorescence microscopy experiments were performed at three different concentrations (5, 10, 20 µM) to efficiently count the living cells. The chimera **CPP-ch5** presents a dose-dependent effect such as **CPP-p2** (the number of cells decrease when the inhibitor concentration increases) whereas **CPP-p8** do not present any toxicity effect (Figure 66B). Furthermore, DNA was measured with fluorescence cell imaging and we could determine that **CPP-ch5** and **CPP-p2** show a similar accumulation of cells in S phase which is not the case for control peptide **p8** (Figure 66C). Finally, flow cytometry experiments, based on the staining of nuclei with propidium iodide (PI) followed by a bromodeoxyuridine (BrdU), were performed to distinguish different phase cycle (G0, G1, G2, S and M). This fluorescence technic shows that when cells are in contact with **CPP-ch5**, the cell progression into S phase is stopped (Figure 66D). This result highlights that α/urea hybrid **CPP-ch5** has perturbed the proliferation cycle of cancer cells and exhibits cytotoxicity for cancer cells.

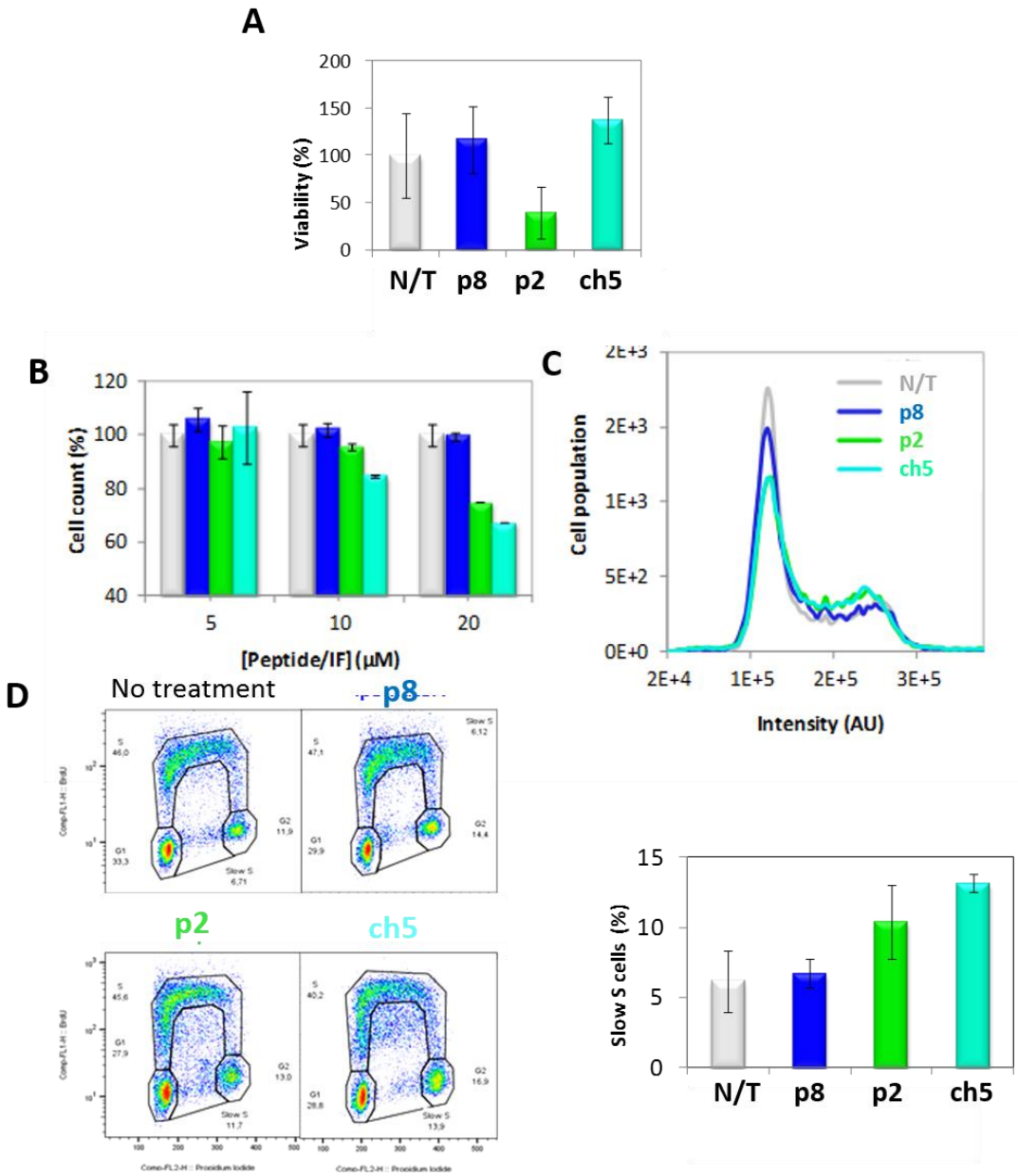


Figure 66: Cellular experiments performed on chimera CPP-ch5. A. when not coupled to a CPP, ch5 has no effect on U2OS cells compared to CPP-p8 and CPP-p2. Cells were treated and cell viability was assessed by MTT assay. Data are expressed as percentages of control cells (N/T) washed with DPBS/Pyrene butyrate solution. B. absolute cell count assessed by fluorescence microscopy reveals a significant loss of U2OS cells induced by $(RRP)_5$-SS-ch5. Vectorized peptides and ch5 were added directly in the culture medium every 12h during 48H at final concentrations of 5, 10 or 20µM. Data are expressed as a percentage of N/T control cells. C. U2OS cell cycle profiled by fluorescence cell imaging. Curves represent cell population (nuclei frequency) in intervals of Hoechst fluorescence intensity. D. Flow cytometry of cell cycle distribution according to PI (x-axis) and BrdU (y-axis) fluorescence intensities (left). $(RRP)_5$-peptides and $(RRP)_5$-SS-ch5 were added directly in the culture medium at final concentrations of 10 and 5µM, respectively. Data were acquired after 3h cells incubation. Right, Quantification of slowed S cells. In B, C and D, N/T refers to non-treated control cells. Error bars indicate s.d. in two independent experiments (experiments performed by May Bakail)

To confirm that the results of the cellular assays are really the reflect of the *in cellulo* action of the chimera **ch5** conjugate with the CPP (**CPP-ch5**), further negative control experiments in the presence of the vector peptide (*i.e.* **p9**=H-$(RRP)_5$-NH_2) has to be performed and complemented with two control chimera sequences where the two key anchoring residues Leu^U5 and Ile8 of **ch5** were replaced by Ala side chains (**CPP-ch11**= A$(RRP)_5$-C(-S-S-linker)-EK-Nal^U-Arg^U-Ala^U-Gln^U-RAA-NH_2). Peptide **p9** has been synthesized on the *Liberty Blue* peptide synthesizer and the adduct **CCP-ch11** was prepared according to the procedure described previously. However, the cellular assays have not been performed yet.

ii) Enzymatic degradation study with a selection of enzymes

Although peptides are promising candidates as therapeutics, their major weakness is their modest stability in biological fluids[29,30]. In our study, we introduced urea-type residues in order to improve the stability of natural peptides inhibitors such as **p2** and **p3**. To investigate the enhanced biostability brought by the urea segment, we performed a series of enzyme degradation assays. Trypsin was first selected as model enzyme. This enzyme is a pancreatic serine protease involved in the digestion of food proteins[31]. It specifically cleaves proteins on the C-terminal side of positively charged lysine and arginine residues. The degradation study was performed in a 96-well plate and followed by RP-HPLC (Figure 67).

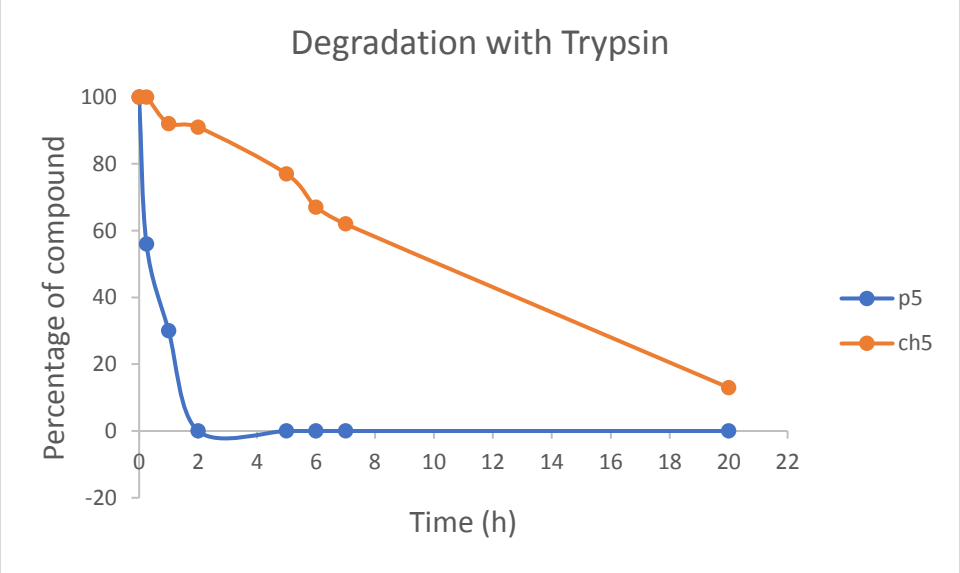


Figure 67: Comparison of the Trypsin enzymatic degradation of p5 and ch5

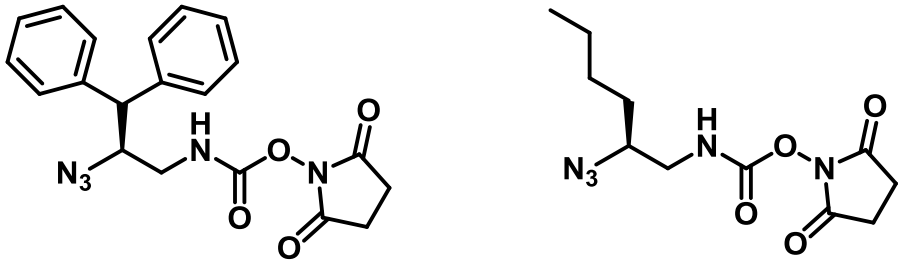
The curves of degradation study over a period of 20 hours show that **p5** is quickly digested by Trypsin whereas **ch5** is stable over the two first hours and is half degraded after 12 hours. Two main peptide **p5** segments are observed by RP-HPLC during the first 1 hours, and correspond to the predicted Ac-EK-Nal-ARLAR, and Nal-ARLARRIA fragments. Regarding **ch5** after 20 hours of Trypsin digestion, the main degraded segment was identified as being Ac-EK-Nal^{U}-Arg^{U}-Leu^{U}-Gln^{U}-R. This result highlights that having installed a urea segment dramatically improve the stability of the resulting chimera in biological environment. Interestingly, no cleavage after the Lys_2 occurred during the degradation assay suggesting here that connecting a urea residue on the C-terminal part of a positively charged residue prevent Trypsin's degradation.

This preliminary enzymatic degradation results are encouraging because they confirm our hypothesis that the introduction of an oligourea segment in a peptide sequence may increase its biostability. Indeed, in the presence of Trypsin, the chimera **ch5** presents a stability highly improved compared to **p5**. Others degradation studies in plasma for instance are necessary to determine half life time of **p5** and chimeras. This work is currently under progress.

e) Consequences of replacing Leu5 anchoring residue on Asf1 binding affinity

With these encouraging results obtained with **ch5**, we have designed new α/urea hybrids in order to improve the affinity to Asf1. For this new series of compounds, modeling studies have been performed by Seydou Traore, a post-doctoral researcher in the team of Raphaël Guerrois (@ CEA Saclay). The crystal structure of Asf1-**ch5** complex was used as model of modeling studies and single mutations were introduced in the sequence of **ch5** using the Rosetta software, a structure prediction tool developed by Baker and coll.[32,33]. A force field from Rosetta, called 'mm_std' was used for the modelling of sequences with the substitution of Leu^{U}5.

A new series of chimera hits were gathered during this modeling study where the best compounds corresponded to the substitution of the key residue Leu^{U}5 by various hydrophobic residues. Indeed, mastering the insertion of the aliphatic side-chain in the Asf1 hydrophobic pocket occupied by the α-Leu residue could have a huge impact on Asf1 binding affinity. To access to this new series of Asf1 inhibitors, two new azide-type monomers deriving from 3,3-Diphenyl-L-Alanine and Norleucine have been synthesized. Regarding the activated monomer N_3-DiPhe^{U}-OSu, the synthesis route described in Scheme 12 was followed and the final monomer **M18** was obtained with 31% global yield. The azide carbamate monomer N_3-Nle^{U}-OSu, **M19**, obtained from Ureka (young innovative enterprise from ImmuPharma group), synthesized as described in Scheme 3 from Chapter II and was obtained with 30% yield.



M18 M19

Figure 68: Chemical structure of azide activated monomers N_3-DiPhe^u-OSu M18 and N_3-Nle^u-OSu M19

These new monomers as well as **M12** and **M13** were next used to replace Leu$^u_5$ in the sequence of **ch5**. New sequences reported in Table 19 were synthesized as described in Scheme 14 and recovered after RP-HPLC purification in ≈ 10% yields. Chimeras **ch10**-**ch12** were designed thanks to the modelling study described above whereas meanwhile I suggested the last one encompassing the Nle^U urea-residue to evaluate the impact of branched versus linear aliphatic side chains on the binding affinity.

Table 19: Sequences of α/urea chimeras with variation of the side-chain composition at the 5th position and characterization by RP-HPLC and ESI-MS.

	Sequence											Yield (%)	t_r (min)	$[M+H]^+$
ch5	Ac	E	K	Nal^u	Arg^u	Leu^u	Gln^u	R	I	A	NH_2		4.55	1367.73
ch12	Ac	E	K	Nal^u	Arg^u	HPhe^u	Gln^u	R	I	A	NH_2	12	3.60	1416.69
ch13	Ac	E	K	Nal^u	Arg^u	Nal^u	Gln^u	R	I	A	NH_2	12	4.16	1451.75
ch14	Ac	E	K	Nal^u	Arg^u	DiPhe^u	Gln^u	R	I	A	NH_2	15	4.19	1478.62
ch15	Ac	E	K	Nal^u	Arg^u	Nle^u	Gln^u	R	I	A	NH_2	27	3.89	1367.90

ITC experiments allowed the determination of affinity constants and we can notice that the K_D values measured for this new series of compounds are actually less good than for **ch5**. Especially, the replacement of Leu^{U}5 by HPhe^U leads to a K_D almost 10-fold higher and it is even worse for the replacement by Nal^U as the interaction is no more measurable. This suggests that the length of the side chains and its aromaticity disturb the interaction with Asf1 leading to a loss or even an absence of affinity. These results highlight that a single side-chain modification can have a huge impact on the binding. Surprisingly, the substitution of Leu^U by DiPhe^U present a better K_D than the previous substitutions suggesting that replacing Leu side chain by a wider one is more tolerated but the binding affinity remains below that of **ch5**. In this second generation of chimeras, the best compound **ch15** possesses a Leu^U → Nle^U substitution meaning from a

branched to a linear aliphatic side-chain. However, the K_D value measured remains slightly higher to that of **ch5**.

Compound	K_D (µM)	N	ΔH (kCal.M^{-1})	-TΔS (kCal.M^{-1})	ΔG (kCal.M^{-1})
ch5	2.73±0.59	0.98±0.09	-2.22±0.03	-5.26±0.59	-7.47±0.12
ch12	20.90±0.71	0.91±0.27	-2.01±0.73	-4.26±0.89	-6.27±0.16
ch13	undectetable	ND	ND	ND	ND
ch14	5.10±0.33	1.72±0.09	-2.09±0.15	-5.01±0.24	-7.09±0.08
ch15	3.30±0.80	1.35±0.05	-1.72±0.09	-5.61±0.16	-7.33±0.07

ND= not determined

Figure 69: Thermodynamic data of compounds ch10-ch13 measured by ITC. Experiments were carried out at 20°C, in a 50mM Tris-HCl pH7.4

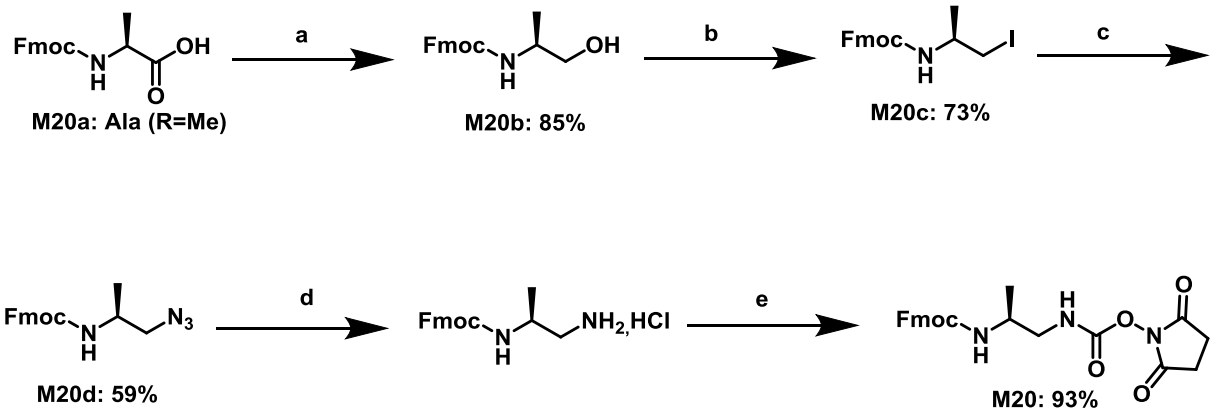
The thermodynamic data of **ch5**, **ch14** and **ch15** are in the same range of order although **ch15** presents a slightly better entropy than **ch5** whereas the enthalpy is lower as a consequence of the substitution of Leu^u (Figure 69). However, due to the margin errors, this observation has to be taken with care. However, the ratio enthalpy/entropy obtained for **ch12** was lower to that recorded for **ch5** resulting to an important loss of binding.

Despite promising modelling studies, none of these new chimeras present an improved binding affinity compared to **ch5**. This suggests that despite virtual predictions of improved contacts with Asf1 surface, binding experiments are mandatory to confirm or not the models and, in our case, it showed that the interaction with Asf1 is actually disrupted.

4. Effect of introducing a urea cap at the *N*-terminus of Asf1 binding peptide p4

In α-peptide with helix propensity, the lack of hydrogen bonds at the end of the helix can diminish the overall helical conformation. To circumvent this drawback, specific sequences, termed cap,

can be introduced at one helix end to stabilize the first or last helical turn[9]. These capping sequences contain appropriate residues that allows the formation of additional non-covalent interations (e.g. H-bonds, salt bridges) involving both main chain and side chains that will stabilize the helix. Concerning **p3h**, the ASTE sequence installed in the N-terminal end brings additional H-bonds. Another approach to develop Asf1 inhibitors was to stabilize the α-helix by replacing the ASTE sequence with a N-terminal urea capping box. In order to determine the number of urea residues required to replace ASTE capping box, two compounds containing a urea cap composed of three or four Ala^{u} were conceived. Due to the large number of α-amino acids to couple before to enter the first urea residue, we decided to follow the same *N*-Fmoc-strategy described in Scheme 15 to prepare these two **p3h** mimics. Hence, Fmoc-Ala^{u}-OSu (**M20**) was first synthesized in solution in 34% overall yield (Scheme 20) before to be manually coupled on the growing peptide sequence attached to the LL Rink amide MBHA resin.



a.1) NMM, IBCF 2) $NaBH_4$, THF; b.PPh_3, I_2, Imidazole, DCM; c. NaN_3, DMF; d. Pd/C, H_2, MeOH,, Chloroform (10%); e.1) NMM, IBCF 2) NaBH4, 0°C, THF

Scheme 20: Multi-step synthesis of Fmoc-Ala^{u}-OSu (M20).

After TFA cleavage and semi-preparative RP-HPLC purification, **ch17** was recovered in 25% yield. Although **ch17** was the major product identified in the crude after TFA cleavage, a second product was also present and actually corresponded to **ch18** (one additional Ala^{U} residue installed during the SPS process: over-insertion issue) (Figure 70). We took advantage of this issue and purified this “by-product” by semi-preparative RP-HPLC and recovered in 10% yield.

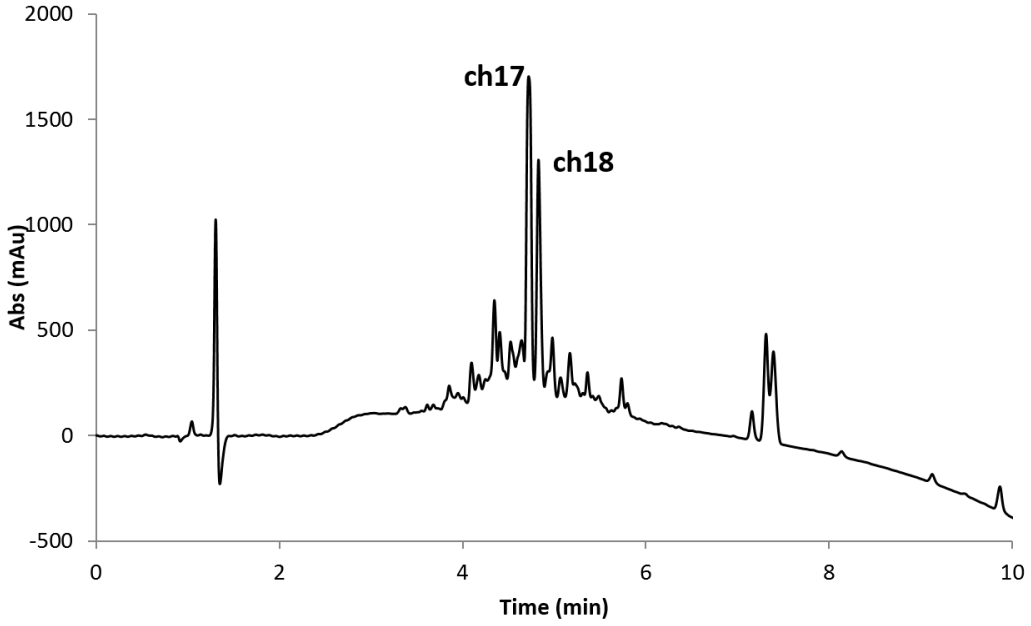


Figure 70: RP-HPLC of the crude ch17 with the by-product ch16. 10-100% ACN in 10 min at 50°C

Table 20: Sequences of chimeras ch17-ch19 with yield after purification, characterization by RP-HPLC and ESI-MS.

	Sequence																	Yield (%)	t_R (min)	$[M+2H]^{2+}$
ch17	Ac		A^u	A^u	A^u	E	K	W	A	R	L	A	R	R	I	A	NH_2	25	4.7	856.53
ch18	Ac	A^u	A^u	A^u	A^u	E	K	W	A	R	L	A	R	R	I	A	NH_2	10	4.8	906.47
ch19	Ac	A^u	A^u	A^u	A^u	-	K	W	A	R	L	A	R	R	I	A	NH_2	20	4.9	842.0

ITC experiments showed that the insertion of a urea cap led to a slight decrease of Asf1 binding affinity compared to **p3h**. This loss of affinity is more important for **ch18** (K_D= 0.50 ± 0.02 µM) encompassing 4 urea residues than for **ch17** equipped 3 Ala^u (K_D= 0.26 ± 0.02 µM). Thermodynamic data revealed that a loss of enthalpy occurs with four Ala^u residues leading to the loss of binding affinity (Figure 71). This can be due to steric clashes that could occur in the N-terminal region in contact with Asf1 surface. The loss of binding affinity can also be due to the absence of side chains in the capping box that could interact with the surface of Asf1 and make additional contacts. Regarding the entropy, in comparison with **p3h**, the -T∆S value decreased for **ch17** and even more for **ch18** and **ch19** which is in agreement with previous observations on introduction of urea residues in the middle of the peptide sequence (see paragraph 3a). Interestingly, the entropy values are better when 4 urea residues are present such as for **ch18**

and **ch19** and both chimeras shown similar values. This result highlights that in the central or terminal positions, the introduction of an oligourea segment brings an improvement of the entropy contribution thus suggesting a more constrained/stable helical conformation interacting with Asf1 surface than in the case of the regular peptide **p3h**.

Compound	K_D (µM)	N	ΔH (kCal.M^{-1})	-TΔS (kCal.M^{-1})	ΔG (kCal.M^{-1})
p3h	0.15±0.01	1.11±0.02	-9.15±0.02	1.47±0.32	-8.68±0.02
ch17	0.26±0.02	1.07±0.01	-9.38±0.06	0.56±0.10	-8.82±0.03
ch18	0.50±0.02	1.04±0.01	-8.26±0.11	-0.18±0.16	-8.45±0.05
ch19	0.64±0.02	1.1±0.02	-8.10±0.10	-1.75±0.06	-8.27±0.06

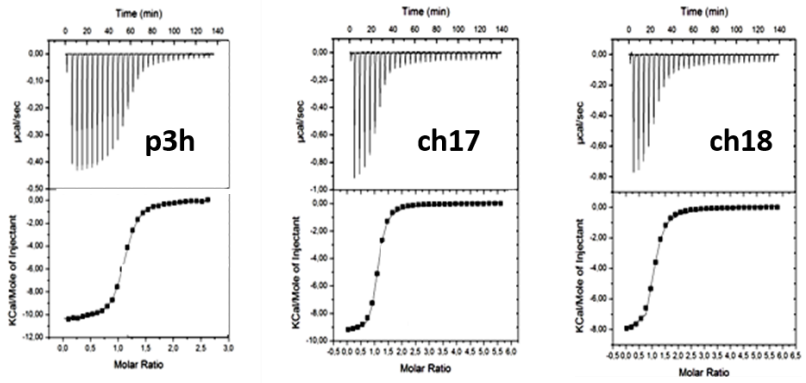


Figure 71: Thermodynamic data of compounds ch17-ch19 measured by ITC experiments. ITC curves of p3h, ch17-ch18 and fitted curves are shown. Experiments were carried out at 20°C, in a 50mM Tris-HCl pH7.4

Concurrently, ^{15}N-HSQC NMR experiments revealed a strong similarity between the spectra of Asf1 upon the addition of **p3h**, **ch17** or **ch18** (Figure 72). Mapping CSVs on the model of **p3**-Asf1 complex confirmed the common behavior between the three inhibitors. However, some differences are visible in the N-terminal part of **p3h** with residues Glu49 and Gln55 that are more affected in presence of the urea cap.

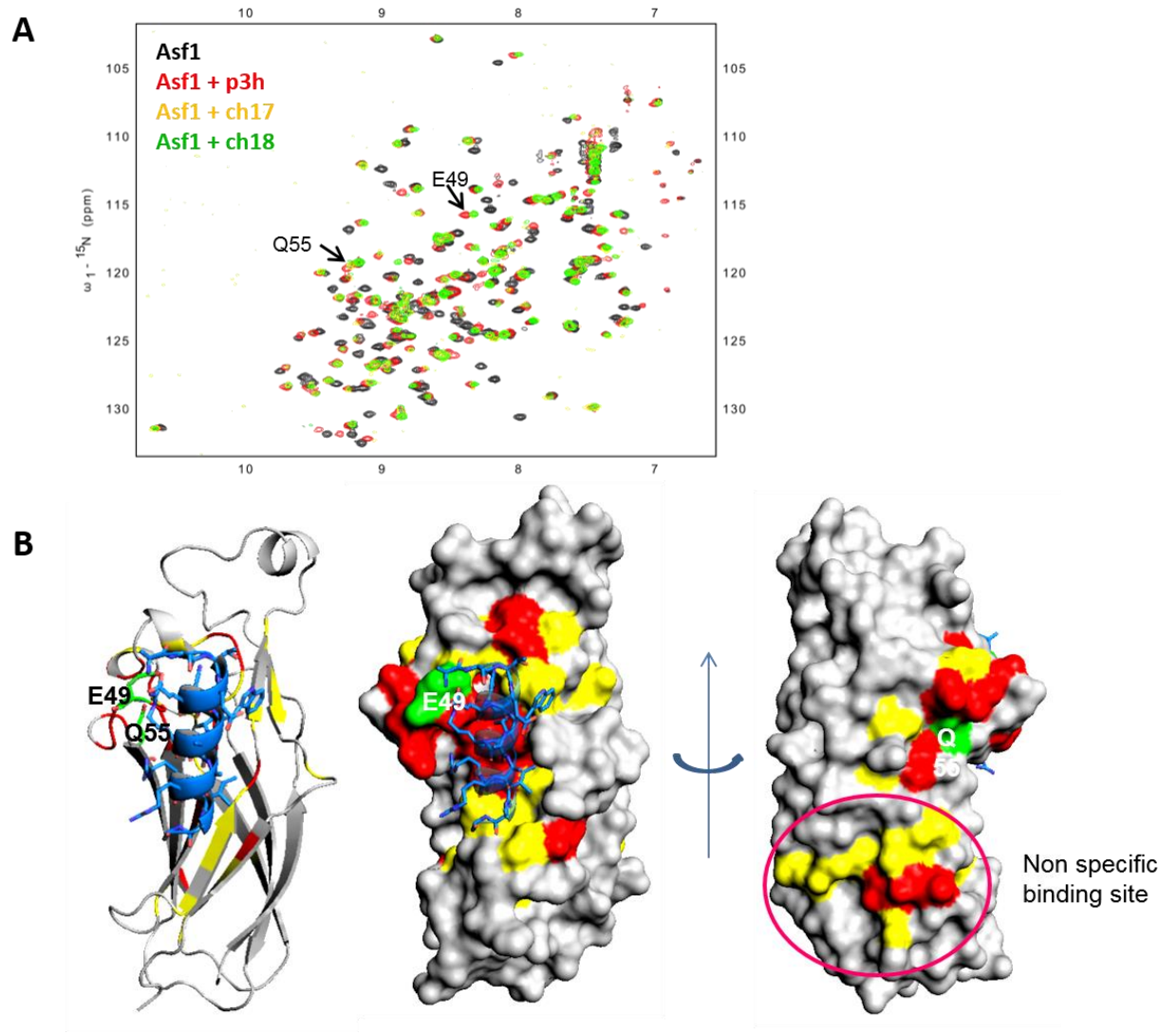


Figure 72 : ^{1}H-^{15}N HSQC NMR experiments of Asf1 bound to p3h, ch17 and ch18 *A.* overlay of ^{1}H-^{15}N HSQC NMR spectra corresponding to the free Asf1 (black), p3h- Asf1 (red) at 25 °C, ch17- Asf1 (orange), or ch18- Asf1 (green) at 20°C. *B.* structure mapping of Asf1 chemical shift variations upon **ch17** or **ch18** binding. Given the high similarity, a unique set of representations is shown. The first representation shows model of Asf1 (grey cartoon) bound to p3h (blue, cartoon/sticks) on which CSV are reported from yellow to red: slightly (yellow), moderately (orange), highly (red) affected Asf1 residues upon if binding. In green, residues less affected by p3h. For the surface representation, the same colors were used. From left to right, two side representations of Asf1, 0° and rotated by 180°.

To summarize this study of urea cap installation, the replacement of the ASTE peptide capping box by 3 or 4 urea residues is feasible without dramatically disturbing the binding affinity to Asf1. In this study, only Ala^{u} residues were used to form the capping box. Capping box with appropriate

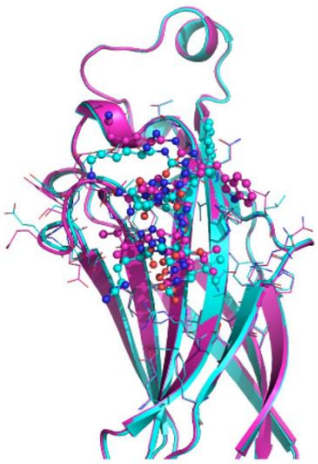
side-chains for additional H-bond formation could be envisioned to improve the affinity to Asf1. Moreover, having in hands a high-resolution X-ray structures of **ch17**-**ch19** would definitely be useful for successful engineering of this urea cap.

Overall during the course of these chimera designs, we have observed that the introduction of a urea segment at the central position or N-terminal part improves the binding entropy but decrease the enthalpy leading to a loss of binding affinity. However, these chimeras seem to be promising candidates for further design PPI modulators in particular in the context of Asf1, as they can accommodate their conformation to the surface of the protein, they target while exhibiting improved biostability. The best chimera **ch5** presents interesting properties including antiproliferative effect on cancer cells which is quite encouraging. Despite the fact that we were successful in getting one X-ray crystal structure of **ch5**/Asf1 complex, designing new chimeras based on this structure appeared to be trickier than anticipated. Having new X-ray crystal structures (i.e. urea scans and capping box) would help to better understand the binding and guide new designs. To pursue our efforts to design potent inhibitors of histone/Asf1 interactions, we were next interested in developing a full oligourea sequence. This has not been investigated so far in the group and remains particularly challenging.

B. Design and synthesis of homo-oligourea sequences mimicking $H3_{118-131}$ peptide

1. Molecular modeling approach for pure oligourea finding

To design homo-oligourea sequences mimicking the C-terminal part of H3 histone, molecular modeling has been performed by S. Traore (@ CEA Saclay). The X-ray crystal structure of Asf1-**ch5** complex was used to perform the computational modeling. The crystal structure of an oligourea was placed manually in the binding region of Asf1 and **ch5**. To satisfy an optimal positioning of side chains, Cα and Cβ of the three key residues (Nal, Leu, Ile) were fitted manually. The Rosetta software was used to determine a possible sequence.



Prediction sequence

Ideal $K^u\ Ar^u\ N^u\ L^u\ S^u\ R^u\ I^u$

Design + correction $K^u\ W^u\ N^u\ L^u\ N^u\ I^u\ I^u$

Figure 73: Prediction of the binding mode for a full oligourea

Several sequences of homo-oligourea were predicted for Asf1 binding (Figure 73). The obtained sequence after design and correction seems well-appropriate to interact with Asf1. As Asn^u3 and Asn^u5 may be solvent-exposed, several charged side-chains such as Arg or Gln can be used to replace Asn one.

2. Preparation of oligourea sequences and determination of the affinity to Asf1

To begin with this strategy of applying homo-oligourea to recognize the surface of Asf1, a first sequence was synthesized on LL Rink amide MBHA resin. We chose to replace the Asn^u side-chain from the model prediction by Gln and Arg ones because these urea residues were already present in the sequence of **ch5**. Regarding the synthesis of this first full oligourea, **ol1**, the SPS was performed by using the azide strategy and all the required azide-type monomers were available in the laboratory. After purification by semi-preparative RP-HPLC, the sequence **ol1** was obtained in 13% yield (Table 21).

Table 21: Sequence of ol1 with yield after purification, characterization by RP-HPLC and ESI-MS.

	Sequence											Yield (%)	t_R (min)	$[M+H]^+$
ch5	Ac	E	K	Nal^u	Arg^u	Leu^u	Gln^u	R	I	A	NH_2	8	4.55	1367.73
ol1	iPr^u	-	Lys^u	Trp^u	Arg^u	Leu^u	Gln^u	Arg^u	Ile^u	-	NH_2	13	4.32	1286.66

ITC experiments revealed that **ol1** interacts with Asf1 with a quite good K_D= 4.6 µM. This result is really encouraging for a first attempt to mimic the C-terminal part of H3 with a pure urea oligomer. The thermodynamic data determined by ITC experiment showed actually some similarities between **ol1** and **ch5** (Figure 74). Indeed, the **ol1** and **ch5** exhibit similar value of enthalpy, and more surprisingly, the binding entropy of **ol1** is also similar to that of **ch5**. With a full oligourea, we could expect that the entropy would increase as the 2.5-helix is expected to be more stable and already formed before the binding to Asf1. This result reflects that a wide range of enthalpy/entropy combinations can lead to similar affinity. The entropic contribution does not necessary increase despite the increasing number of urea residues suggesting that the only way to improve the affinity is to improve the enthalpy of binding. To do so, optimizing one side-chain can have a huge impact on the enthalpy and consequently on the binding affinity.

Compound	K_D (µM)	N	ΔH (kCal.M^{-1})	-TΔS (kCal.M^{-1})	ΔG (kCal.M^{-1})
ch5	2.73±0.59	0.98±0.09	-2.22±0.03	-5.26±0.59	-7.47±0.12
ol1	4.49±0.69	1.11±0.02	-2.30±0.07	-4.86±0.43	-7.16±0.36

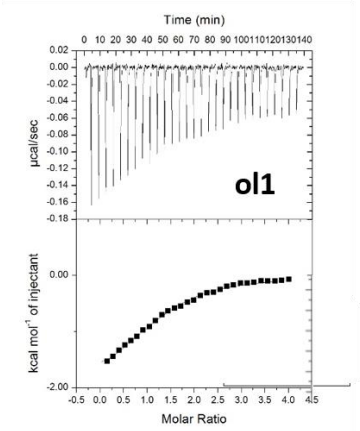


Figure 74. Thermodynamic data of compound ol1 measured by ITC experiments. ITC curves of ol1 and fitted curves are shown. Experiments were carried out at 20°C, in a 50mM Tris-HCl pH7.4

NMR titration and mapping experiments should be soon performed and will give access to precious information regarding the mode of binding and interaction zone on Asf1 surface. Besides, crystallogenesis assays of Asf1/**ol1** complex have been initialized and several crystals were obtained. Unfortunately, these recent data have not been treated yet to solve the crystal structure of the complex. The resolution of this new structure would be really interesting to

understand the binding of **ol1** to Asf1 and this would guide the design of new homooligoureas for optimizing the interaction with Asf1 surface.

In this chapter, we have demonstrated that the design of urea-based foldamers for the inhibition of PPIs, by focusing on the interaction with the chaperon protein Asf1, is within reach. The results obtained by ITC and NMR experiments are encouraging with some synthetic inhibitors presenting affinities in the micromolar range and interacting in the same region than the natural H3 C-terminal end. In particular, two compounds are promising: the chimera **ch5** and now the full oligourea **ol1**. Indeed, the first one presents the best affinity for Asf1 and is cytotoxic for a cancer cell line. Furthermore, we have demonstrated that the introduction of a urea segment in a peptide sequence improves the biostability which is a real advantage for pharmaceutical development. In continuation of this part, the full oligourea **ol1** showed promising data as the affinity for Asf1 is also in the micromolar range and mainly, the biostability should be considerably improved. Indeed, previous study in the group on full oligoureas confirmed their "super high" stability and integrity even after 24h in presence of a set of proteases[34]. However, the affinities of these compounds have to be improved to make them fully specific to the Asf1 target. We have been faceing during this PhD thesis the difficulty of improving the affinity of our lead compounds. One main concern remains how to get compounds presenting higher binding enthalpy t. After several rounds of designs and sequence optimization, the team of F. Ochsenbein managed to conceive an α-peptide with a final affinity to Asf1 100-fold higher to the starting peptide by reaching a fourth hydrophobic pocket located in another face of the Asf1 surface. We therefore thought that we could also improve our best chimera **ch5** by enlarging the C-terminal extremity.

Bibliography

(1) Brown, C. Targeted therapy: An elusive cancer target https://www-nature-com.inc.bib.cnrs.fr/articles/537S106a (accessed Mar 24, 2018).
(2) 387381a0.pdf http://www.nature.com.inc.bib.cnrs.fr/articles/387381a0.pdf (accessed Mar 15, 2018).
(3) Miknis, G. F.; Stevens, S. J.; Smith, L. E.; Ostrov, D. A.; Churchill, M. E. A. Development of Novel Asf1–H3/H4 Inhibitors. *Bioorg. Med. Chem. Lett.* **2015**, *25* (4), 963–968.
(4) Seol, J.-H.; Song, T.-Y.; Oh, S. E.; Jo, C.; Choi, A.; Kim, B.; Park, J.; Hong, S.; Song, I.; Jung, K. Y.; et al. Identification of Small Molecules That Inhibit the Histone Chaperone Asf1 and Its Chromatin Function. *BMB Rep.* **2015**, *48* (12), 685–690.
(5) Agez, M.; Chen, J.; Guerois, R.; van Heijenoort, C.; Thuret, J.-Y.; Mann, C.; Ochsenbein, F. Structure of the Histone Chaperone Asf1 Bound to the Histone H3 C-Terminal Helix and Functional Insights. *Structure* **2007**, *15* (2), 191–199.
(6) Scorgie, J. K.; Donham, D. C.; Churchill, M. E. A. Analysis of Histone Chaperone Antisilencing Function 1 Interactions. *Methods Enzymol.* **2012**, *512*, 223–241.
(7) English, C. M.; Adkins, M. W.; Carson, J. J.; Churchill, M. E. A.; Tyler, J. K. Structural Basis for the Histone Chaperone Activity of Asf1. *Cell* **2006**, *127* (3), 495–508.
(8) Natsume, R.; Eitoku, M.; Akai, Y.; Sano, N.; Horikoshi, M.; Senda, T. Structure and Function of the Histone Chaperone CIA/ASF1 Complexed with Histones H3 and H4. *Nature* **2007**, *446* (7133), 338–341.
(9) Aurora, R.; Rose, G. D. Helix Capping. *Protein Sci. Publ. Protein Soc.* **1998**, *7* (1), 21–38.
(10) Douat-Casassus, C.; Pulka, K.; Claudon, P.; Guichard, G. Microwave-Enhanced Solid-Phase Synthesis of N , N' -Linked Aliphatic Oligoureas and Related Hybrids. *Org. Lett.* **2012**, *14* (12), 3130–3133.
(11) Grolier, J.-P. E.; del Río, J. M. Isothermal Titration Calorimetry: A Thermodynamic Interpretation of Measurements. *J. Chem. Thermodyn.* **2012**, *55*, 193–202.
(12) Velazquez-Campoy, A.; Freire, E. Isothermal Titration Calorimetry to Determine Association Constants for High-Affinity Ligands. *Nat. Protoc.* **2006**, *1* (1), 186–191.
(13) Burnouf, D.; Ennifar, E.; Guedich, S.; Puffer, B.; Hoffmann, G.; Bec, G.; Disdier, F.; Baltzinger, M.; Dumas, P. KinITC: A New Method for Obtaining Joint Thermodynamic and Kinetic Data by Isothermal Titration Calorimetry. *J. Am. Chem. Soc.* **2012**, *134* (1), 559–565.
(14) Freire, E.; Mayorga, O. L.; Straume, M. Isothermal Titration Calorimetry. *Anal. Chem.* **1990**, *62* (18), 950A–959A.
(15) Dutta, A. K.; Rösgen, J.; Rajarathnam, K. Using Isothermal Titration Calorimetry to Determine Thermodynamic Parameters of Protein–Glycosaminoglycan Interactions. *Methods Mol. Biol. Clifton NJ* **2015**, *1229*, 315–324.
(16) Freire, E. Isothermal Titration Calorimetry: Controlling Binding Forces in Lead Optimization. *Drug Discov. Today Technol.* **2004**, *1* (3), 295–299.
(17) Mousson, F.; Lautrette, A.; Thuret, J.-Y.; Agez, M.; Courbeyrette, R.; Amigues, B.; Becker, E.; Neumann, J.-M.; Guerois, R.; Mann, C.; et al. Structural Basis for the Interaction of Asf1 with Histone H3 and Its Functional Implications. *Proc. Natl. Acad. Sci.* **2005**, *102* (17), 5975–5980.
(18) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. A Field Guide to Foldamers. *Chem. Rev.* **2001**, *101* (12), 3893–4012.

(19) Armstrong, C. T.; Mason, P. E.; Anderson, J. L. R.; Dempsey, C. E. Arginine Side Chain Interactions and the Role of Arginine as a Gating Charge Carrier in Voltage Sensitive Ion Channels. *Sci. Rep.* **2016**, *6*, 21759.

(20) Bechara, C.; Sagan, S. Cell-Penetrating Peptides: 20years Later, Where Do We Stand? *FEBS Lett.* **2013**, *587* (12), 1693–1702.

(21) Copolovici, D. M.; Langel, K.; Eriste, E.; Langel, Ü. Cell-Penetrating Peptides: Design, Synthesis, and Applications. *ACS Nano* **2014**, *8* (3), 1972–1994.

(22) Gump, J. M.; Dowdy, S. F. TAT Transduction: The Molecular Mechanism and Therapeutic Prospects. *Trends Mol. Med.* **2007**, *13* (10), 443–448.

(23) Herce, H. D.; Garcia, A. E.; Cardoso, M. C. Fundamental Molecular Mechanism for the Cellular Uptake of Guanidinium-Rich Molecules. *J. Am. Chem. Soc.* **2014**, *136* (50), 17459–17467.

(24) Smith, B. A.; Daniels, D. S.; Coplin, A. E.; Jordan, G. E.; McGregor, L. M.; Schepartz, A. Minimally Cationic Cell-Permeable Miniature Proteins via α-Helical Arginine Display. *J. Am. Chem. Soc.* **2008**, *130* (10), 2948–2949.

(25) Nadia, C.; Planel, S.; Salomon, A.; Jalinot, P.; Feige, J.-J. A novel concept in anti-angiogenic and anti-tumoral therapy: multi-target destabilization of short-lived mRNAs by the zinc finger protein ZFP36L1. *Bull. Cancer (Paris)* **2010**, *97* (1), S72.

(26) Daniels, D. S.; Schepartz, A. Intrinsically Cell-Permeable Miniature Proteins Based on a Minimal Cationic PPII Motif. *J. Am. Chem. Soc.* **2007**, *129* (47), 14578–14579.

(27) Matsueda, R.; WALTER, R. 3-NITRO-2-PYRIDINESULFENYL (Npys) GROUP. *Int. J. Pept. Protein Res.* **1980**, *16* (5), 392–401.

(28) Niforou, K. M.; Anagnostopoulos, A. K.; Vougas, K.; Kittas, C.; Gorgoulis, V. G.; Tsangaris, G. T. The Proteome Profile of the Human Osteosarcoma U2OS Cell Line. *Cancer Genomics Proteomics* **2008**, *5* (1), 63–78.

(29) Fosgerau, K.; Hoffmann, T. Peptide Therapeutics: Current Status and Future Directions. *Drug Discov. Today* **2015**, *20* (1), 122–128.

(30) Lau, J. L.; Dunn, M. K. Therapeutic Peptides: Historical Perspectives, Current Development Trends, and Future Directions. *Bioorg. Med. Chem.* **2018**, *26* (10), 2700–2707.

(31) Brown, W. E.; Wold, F. Alkyl Isocyanates as Active-Site-Specific Reagents for Serine Proteases. Reaction Properties. *Biochemistry* **1973**, *12* (5), 828–834.

(32) Simons, K. T.; Bonneau, R.; Ruczinski, I.; Baker, D. Ab Initio Protein Structure Prediction of CASP III Targets Using ROSETTA. *Proteins* **1999**, *Suppl 3*, 171–176.

(33) Parmeggiani, F.; Huang, P.-S.; Vorobiev, S.; Xiao, R.; Park, K.; Caprari, S.; Su, M.; Seetharaman, J.; Mao, L.; Janjua, H.; et al. A General Computational Approach for Repeat Protein Design. *J. Mol. Biol.* **2015**, *427* (2), 563–575.

(34) Antunes, S.; Corre, J.-P.; Mikaty, G.; Douat, C.; Goossens, P. L.; Guichard, G. Effect of Replacing Main-Chain Ureas with Thiourea and Guanidinium Surrogates on the Bactericidal Activity of Membrane Active Oligourea Foldamers. *Bioorg. Med. Chem.* **2017**, *25* (16), 4245–4

Chapter IV: Targeting two remote epitopes on Asf1 by combining two binding motifs in a single molecule

A. From micro to subnanomolar affinity to Asf1 by epitope tethering

1. Improved H3-H4 mimicry by integrating the histone H4 binding epitope

The X-ray crystal structure of the natural Asf1-H3-H4 complex allowed the determination of four key residues localized on H3 and H4 histones respectively that concentrate most of the binding interaction with Asf1: residues Leu126, Arg129, Ile130 for H3 and Phe100 for H4 (Figure 75)[1]. Designing an inhibitor capable of integrating these four binding epitopes was quite challenging because a distance of 21.6 Å separates the farthest anchor residues Leu126 and Phe100 (pdb 2io5). Thanks to a step by step approach evoked earlier in the chapter III,a.,1, the team of F. Ochsenbein and more particularly during the PhD thesis of M. Bakail succeeded in conceiving the peptide **p3** exhibiting a K_D value in the nanomolar range (K_D=3 ± 1 nM) reflecting a gain of affinity about 100-fold with respect to the helical part of **p3h**. This α-peptide can be decomposed in three peptide fragments: (1) the α-helical part that reproduced the C-terminal three key contacts of H3 (numbered in the peptide Leu10, Arg13 and Ile14) combined with extra interacting side-chain making additional hydrophobic and polar contacts, (2) a second α-peptide fragment mimicking the C-terminal β-strand of H4 and comprising the anchor residue Phe100 and (3) a peptide linker aimed to bridge the two fragments. For the design of this bridge, due to a distance of 8.7 Å between the Cα of Arg131 in H3 and the Cα of Arg95 in H4 (Figure 75B), a length of 4-residues (GAGG) was chosen. This fragment adopts a turn allowing the β-strand fragment mimicking H4 (in red) to be well positioned at the surface of Asf1 and restored the Phe side chain into the hydrophobic pocket as shown in the X-ray crystal structure of the complex between Asf1 and **p3** solved at 1.98 Å (Figure 75C).

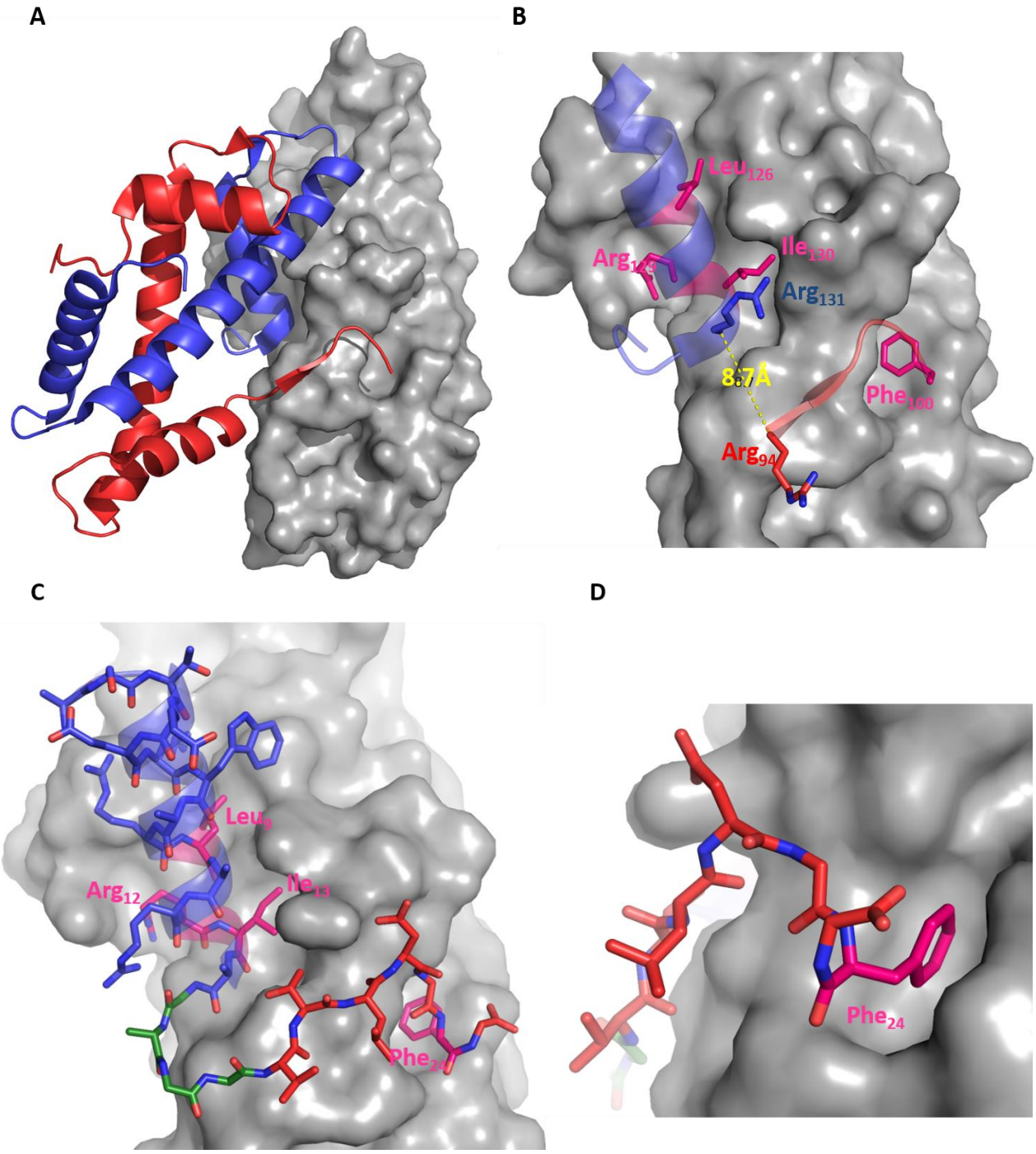


Figure 75 : Design method of peptide p3 and X-ray crystal structure with Asf1. A. X-ray crystal structure of Asf1-H3-H4 complex (pdb= 2io5)[1] with Asf1 in grey, H3 in blue and H4 in red. B. Binding region of Asf1 with the C-terminal part of H3 in blue and the C-terminal part of H4 in red. Binding epitopes are represented in pink. C. X-ray crystal structure of Asf1 in complex with **p3**. The helical part is depicted in blue, the H4 mimicking segment in red and the linker is represented in green. Binding epitope are represented in pink. D. Zoom on Phe24 of p3 into the hydrophobic pocket of Asf1.

This crystal structure confirmed that as expected the helical part of **p3** (**p3h**) is positioned in the same manner than the C-terminal part of H3 (described in Chapter III.A.1). Moreover, the four binding epitopes are well located within Asf1 hydrophobic pockets allowing to restore the contacts found in the natural complex between H3-H4 and Asf1. Indeed, in the same manner than the Phe100 of histone H4, the Phe24 of **p3** is well-positioned in the hydrophobic pocket of Asf1 (Figure 75D). This suggests that the choice and composition of the peptide linker was appropriate and nicely bridge the two α-peptide secondary structure (α-helix and β-strand). However, the structure also revealed that some of the residues constituting the loop of the linker (GAGG) do not interact directly with Asf1 revealing here some opportunities in terms of peptidomimetic design for further optimization. Interestingly, salt bridges are observed between residues Glu49 and Arg145 of Asf1 and residues Arg8 and Asp22 of **p3** (Figure 76). Concurrently, several hydrogen bonds are present all along the extension of **p3** such as Asf1 Arg108 with the backbone of **p3**-Arg12 or Asf1 Arg148 that interacts with OH function of Thr20 side-chain and the backbone of Phe24.

A

Asf1	p3
CO (side chain) of Glu49	NH (side chain) of Arg8
CO (side chain) of Asp54	NHs (side chain) of Arg12
NH (side chain) of Arg108	CO of Arg12
NH (side chain) of Arg108	CO of Ile13
CO of Arg148	NH of Val19
NH of Arg148	CO of Val19
OH of Thr147	OH of Thr20
CO of Val146	NH of Leu21
NH of Val146	CO of Leu21
NH (side chain) of Arg145	CO (side chain) of Asp22
CO of Pro144	NH of Gly23
NH (side chain) of Arg148	CO of Phe24

B

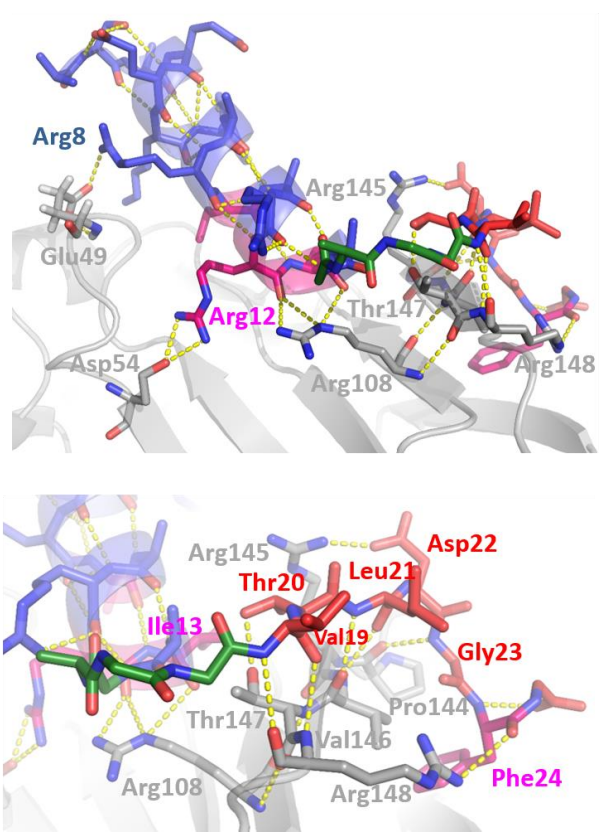


Figure 76: Interactions observed into the complex between Asf1 and p3. A List of Asf1 residues interaction with p3 residues. B X-ray crystal structure of the complex with zoom on H-bonds involved for the stabilization of the complex.

The lack of interactions between the linker of **p3** with the surface of Asf1 as well as its length prompted us to propose some structural optimizations. In order to master the synthesis of this

second generation of Asf1 modulators, we first decided to investigate the solid phase synthesis of **p3** because this peptide was previously bought by the group of F. Ochsenbein to different peptide suppliers.

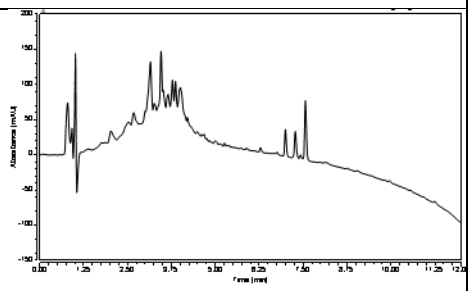
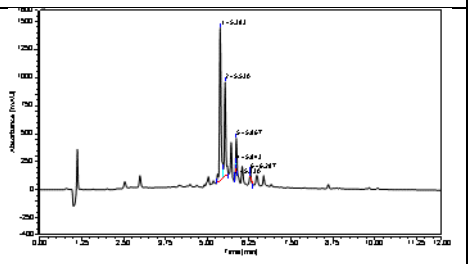
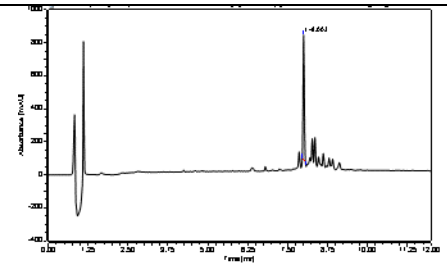
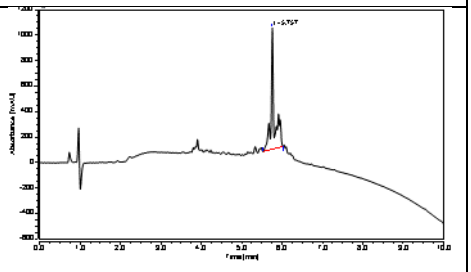
2. Optimized solid phase synthesis of elongated Asf1 binding peptides combining two remotes binding sites

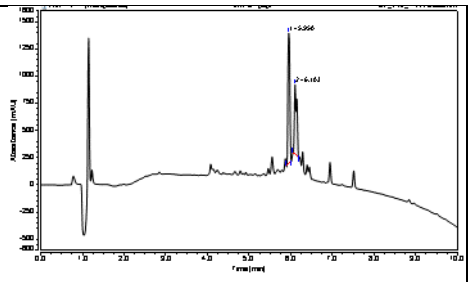
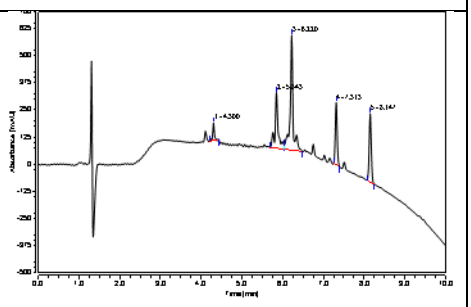
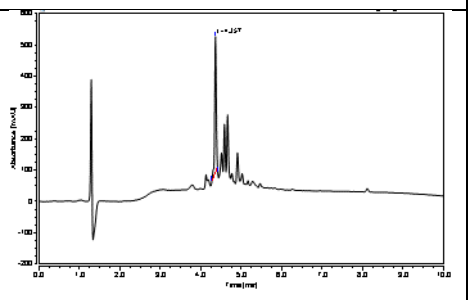
In order to unveil potential difficulties during the synthesis of full length α/urea hybrid foldamer, we first decided to synthesize the C-terminal part of **p3**, that will be necessary for the further modification of the **p3** segment linker (GAGG) as well as for the elongation of **ch5** at its C-terminal end. We focused on the synthesis of peptide **p10** = RIAGAGGVTLDGFG which was performed on a TFA-labile resin using classical *N*-Fmoc chemistry. As for the other peptide syntheses described earlier, all the SPPS were performed under microwave assistance to increase the rate of coupling and deprotection steps while improving the crude purity. **p10** was synthesized on the automated microwave peptide synthesizer *Liberty Blue* and a series of different conditions was screened.

In a first attempt, the synthesis of **p10** was carried out on Rink amide NOVAPEG resin (loading 0.46 mmol.g^{-1}) on a 0.050 mmol scale. Iterative couplings of *N*-Fmoc-protected amino acids were performed once in presence of 1,3-Diisopropylcarbodiimide (DIC) and Oxyma, coupling reagents known to decrease epimerization (compared to DIC/HOBt) and avoid side reactions occurring with high levels of base (Table 22)[2]. The peptide was then released from the resin through a TFA-cleavage using a "cocktail mixture" containing TFA/TIS/H_2O (95:2.5:2.5, v/v/v).

After precipitation of the peptide with cold Et_2O, the crude material was directly analyzed by analytic RP-HPLC. Unexpectedly, the RP-HPLC chromatogram revealed a crowded profile presenting multiple peaks including the expected **p10** (t_r=5.9 min for gradient 10-100% ACN in 10min) further identified by mass spectrometry (ESI positive mode). Consequently, the poor apparent purity combined with the presence of so many impurities surrounded the expected peptide rendered its purification extremely tricky (**SIV.1** Table 22).

Table 22: Different conditions to optimize SPPS of p10

Synthesis	Resin	Coupling conditions	Profile
SIV.1	NOVAPEG (loading 0.46 mmol.g^{-1})	Single coupling DIC/Oxyma 75°C /25W/15min Arg: Double coupling	
SIV.2	Polystyrene MBHA Rink amide (loading 0.37 mmol.g^{-1})	Single coupling DIC/Oxyma 50°C /50W/15min Arg: Double coupling	
SIV.3	Chemmatrix (loading 0.45 mmol.g^{-1})	Double coupling DIC/Oxyma 2 x 50°C /50W/15min Arg: Double coupling	
SIV.4	NOVAPEG (loading 0.46 mmol.g^{-1})	Double coupling DIC/Oxyma 2 x 75°C /25W/15min Arg: Double coupling +manually	

SIV.5	Polystyrene MBHA Rink amide (loading 0.37 mmol.g^{-1})	Double coupling DIC/Oxyma 2 x 75°C /25W/15min Arg: Quadruple coupling	
SIV.6	Polystyrene MBHA Rink amide (loading 0.37 mmol.g^{-1})	Double coupling BOP/DIEA 2 x 75°C /25W/15min Arg: Double coupling +manually	
SIV.7	Polystyrene MBHA Rink amide (loading 0.37 mmol.g^{-1})	Single coupling DIC/Oxyma 2 x 75°C /25W/15min Arg: Double coupling Fmoc-Val-Thr(psiMe-MePro)-OH	

During the second and third synthesis, we have replaced the NOVAPEG resin by either a Rink amide MBHA resin or ChemMatrix resin (100% polyethylene glycol, PEG, *Iris Biotech)*. Regarding the synthesis **SIV.2**, the coupling/deprotection cycle conditions remained identical to those applied for **SIV.1** whereas in the case of **SIV.3** a double coupling of each residue was systematically performed on ChemMatrix resin (Table 22). After TFA-cleavage and Et_2O precipitation, the RP-HPLC analysis showed this time the presence of a major peak identified by ESI-MS as being the expected peptide **p10**. Nonetheless and despite the visible improvement of the SPPS quality, many impurities were still surrounded the major peak that limited the purification step and more importantly, this was totally unfavorable with further peptide elongation to give rise to entire **p3** peptide and hybrid sequence development. A common chemical feature of these two syntheses

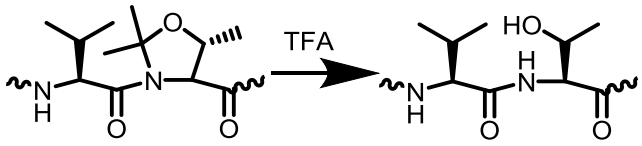
was the difficulty to introduce the last N-Fmoc-Arginine residue. The mass spectrometry analysis indeed revealed the presence of the peptide deleted of this residue (t_r= 6.1 min).

Nevertheless, in comparison with the SPPS on Rink amide MBHA, the SPPS on ChemMatrix was considerably improved. This was probably due to the double coupling performed for **SIV.3** whereas a single coupling was performed for **SIV.2**. To validate this hypothesis, two more syntheses with all the coupling steps repeated twice were performed on NOVAPEG and Rink amide MBHA resins (**SIV.4** and **SIV.5** in Table 22). Moreover, the final N-Fmoc-*L*-Arginine(Pbf)-OH was coupled twice with the automated synthesizer and once more manually for **SIV.4** and was even coupled four times with the *Liberty blue* synthesizer during **SIV.5** synthesis. After TFA cleavage and Et_2O precipitation of this two new batches of resin, the RP-HPLC chromatograms revealed a net improvement of the crude profiles with the presence of a major peak and a purity of 51 and 37% for **SIV.4** and **SIV.5** respectively. Even though impurities remained around the main peak, the SPS carried out on the NOVAPEG resin and accompanied by double coupling steps was considerably improved compared to **SIV.1**. Regarding the polystyrene resin (**SIV.5**), the quality of the SPS appeared better than the one of **SIV.1** but still contained the deleted sequence missing the last Arg residue.

Finally, a last synthesis was again performed on Rink amide MBHA resin by this time employing another coupling reagent system: Benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate (BOP) with N, N'-Diisopropylethylamine (DIEA) (**SIV.6**). Unfortunately, the RP-HPLC analysis of the crude reaction revealed a similar profile than those obtained in presence of DIC/Oxyma (the two last peaks at 7.3 and 8.1 min corresponds to Tris).

This study allowed us to determine the best conditions to do this part of peptide but in regards to our objective of elongating the C-terminal part of **ch5**, the Rink amide MBHA polystyrene resin was selected as this support gave the best results for the synthesis of **ch5**. However, further investigations to identify by-products were performed in order to improve at best this peptide sequence for the future introduction of urea-type units. First, considering the size of the peptide, we supposed that aggregation may occur rendering the chain elongation difficult, in particular to enter the final Arg. To overcome this problem, as our peptide sequence permitted it, we incorporated a pseudo-proline residue substituting the Val-Thr dipeptide. These units Fmoc-Xaa-Yaa-$\psi^{Me,Me}$pro are known to avoid aggregation as they removed a H-bond donor during the synthesis thanks to a cyclocondensation reaction of the side chain (Ser, Thr or Cys) linking it to the nitrogen in N-terminal[3]. Indeed, Mutter *et. al.* developed pseudo-prolines (oxazolidine) to overcome the problems of poor solvation and solubility that can occurred during the SPPS of large peptide [4,5]. The introduction of this pseudo-proline (oxazolidine) dipeptides brings kinks into the peptide backbone decreasing aggregation during chain grow. During the course of TFA-cleavage, the pseudo-proline dipeptide is reopened to give back to the initial dipeptide (Scheme 21). In the

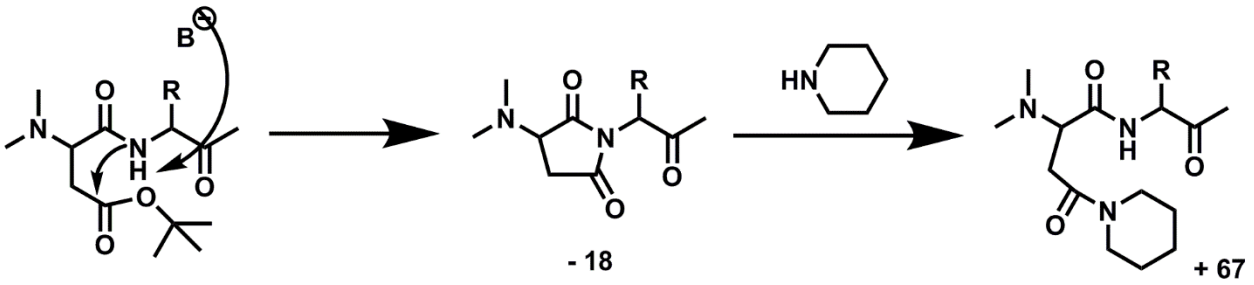
sequence of **p3**, the dipeptide Val20Thr21 was therefore substituted by its commercially available Fmoc-Val-Thr(psiMe-MePro)-OH version.



Scheme 21 : Opening of the pseudo-proline Val-Thr(psiMe-MePro) during TFA cleavage

To evaluate the effect of this pseudo-proline, the SPPS was performed on Polystyrene Rink amide MBHA resin (0.37 mmol/g), and the coupling steps were performed only once to see if the use of the N-Fmoc-Val-Thr(psiMe-MePro)-OH can have real impact on the efficiency of the SPPS. Furthermore, the Arg residues were coupled only twice on automated *Liberty Blue* synthesizer (**SIV.7** in Table 1). After release of the peptide from the support, the RP-HPLC chromatogram showed an optimized profile comparable with the double coupling synthesis (**SIV.5**). The major peak was identified by ESI mass spectrometry as being the expected product and interestingly we could not identify this time the truncated peptide missing an Arg underlining here that the introduction of a pseudo-proline has a beneficial effect on the automated **p10** synthesis on solid support.

Extensive investigations by ESI mass spectrometry were done to identify the two peaks at 4.65 min and 4.72 min respectively with mass corresponding to a loss of 18 and an addition of + 67. These two masses indicate that a side reaction occurred and they are actually characteristics of an aspartimide formation during the Fmoc deprotection in presence of piperidine. Indeed, the presence of Asp-Gly dipeptide in the **p10** sequence is prone to aspartimide formation[6,7]. First the cyclic aspartimide is formed and then piperidine can react on the carbonyl functions of aspartimide to give rise to piperidides and thus leading to mass excess of +67 (Scheme 22). It is clear, that this aspartimide formation was surely occurring since our first attempts of SPS of **p10** but due to the presence of numerous side-products, combined with the deletion of the Arg residue, its identification and characterization were rendered more difficult.



Scheme 22: Mechanism of aspartimide formation

To avoid the aspartimide formation we decided to introduce the commercially available *N*-Fmoc-Asp(*t*Bu)-(Dmb)Gly-OH dipeptide that contains 2,4-dimethoxybenzyl (Dmb), a backbone amide protecting group[8,9]. It will prevent aspartimide formation during the course of the synthesis and will be removed during the final TFA cleavage.

All these clues concerning the synthesis allowed us to synthesize the full **p3** peptide on WANG resin because we needed an acidic function at the C-terminus. The Fmoc-Val-Thr(psiMe-MePro)-OH (4 equiv.) was introduced with HATU (3.8 equiv.)/DIEA (0.8 equiv.) at 75°C/25 W for 20 minutes, the Fmoc-Asp(tBu)-(Dmb)Gly-OH (4 equiv.) with HATU (3.8 equiv.) and DIEA (8 equiv.) at 75°C/25 W for 5 minutes twice and all the *N*-Fmoc-protected amino acids were coupled once. RP-HPLC chromatogram of the crude confirms the presence of a major peak corresponding to the expected compound (purity of 44%). After RP-HPLC purification (gradient 20-50% ACN in 20min), the pure product was obtained with an overall yield of 10% (Figure 77).

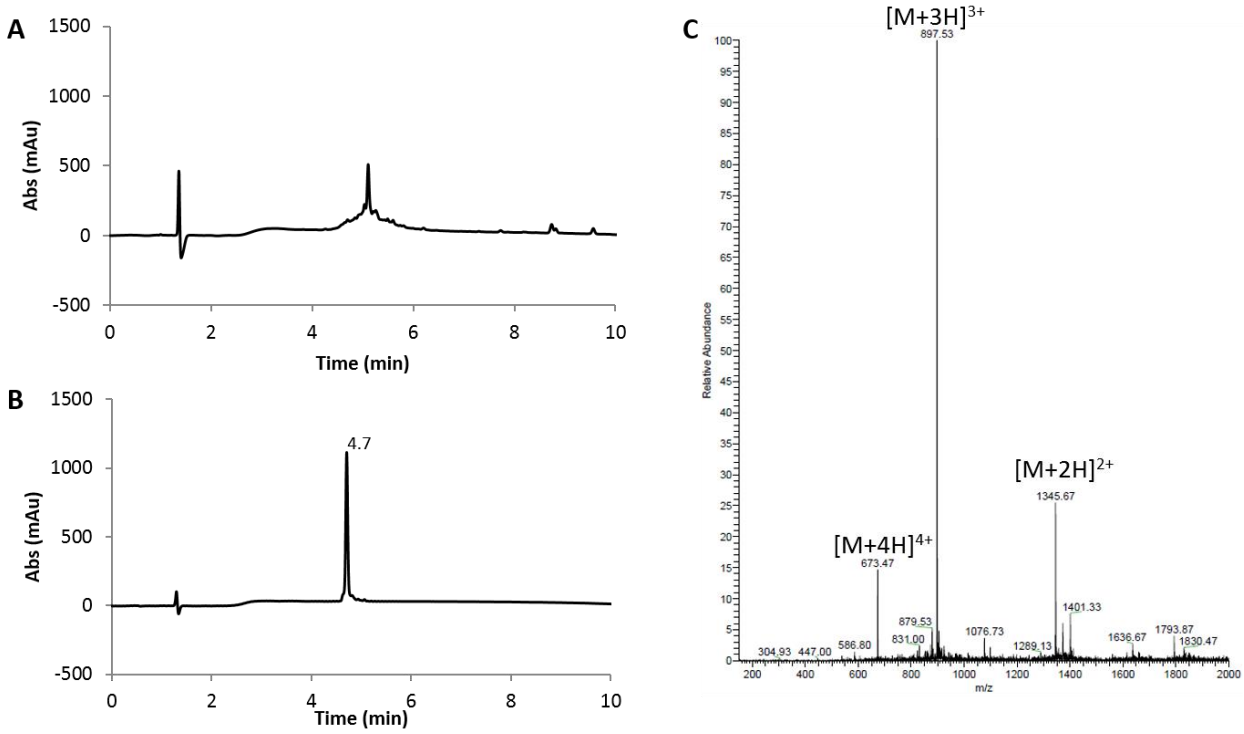


Figure 77: Characterization of the peptide p3 by RP-HPLC with the crude (A) and the pure product after purification with a gradient of 20-50% ACN in 20 min (B) and ESI-MS characterization (C)

The synthesis of **p3** was hence optimized which was essential for future modifications of its sequence.

B. Structure based variations of the spacer unit between the two peptide binding motifs in p3

1. Replacing the GAGG peptide segment by an isosteric peptidomimetic segment

The crystal structure of **p3** in complex with Asf1 confirms that GAGG peptide segment stops the helix propagation[10,11] and does not directly interact with the protein surface. However, its loop

conformation allows a good positioning of the Phe24 in the hydrophobic pocket occupied by the same residue (Phe100) of histone H4.

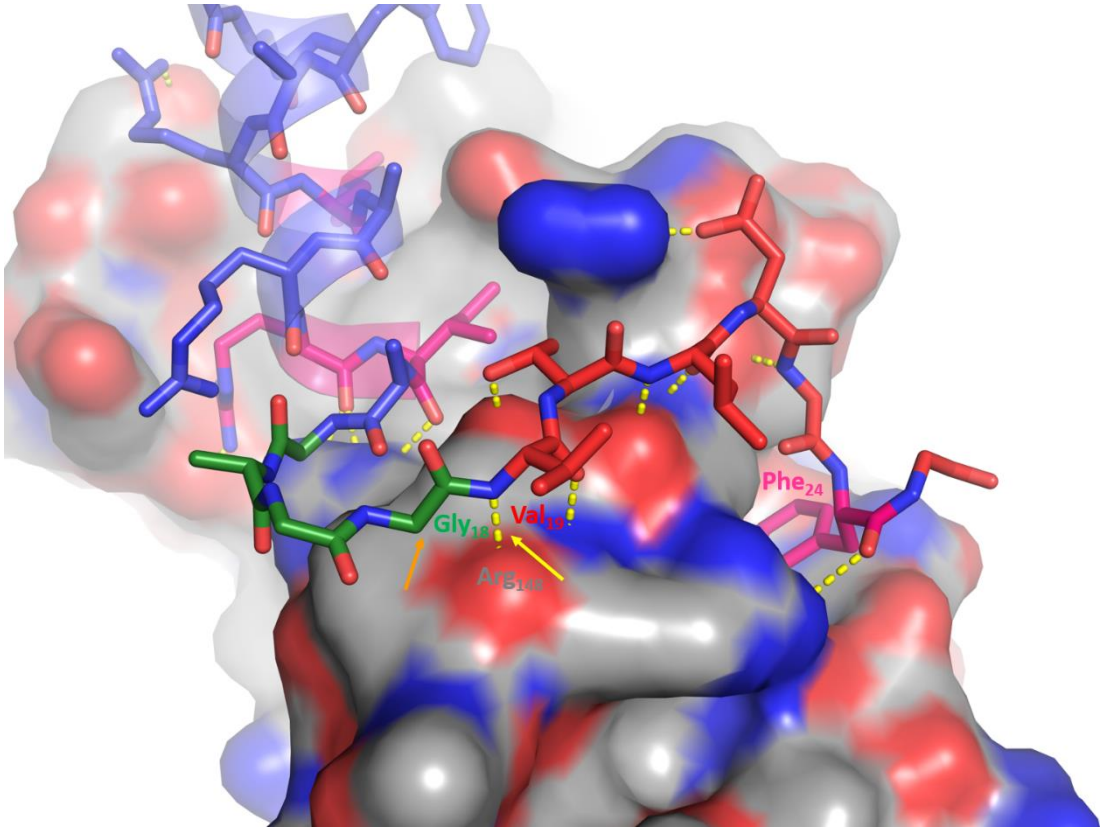
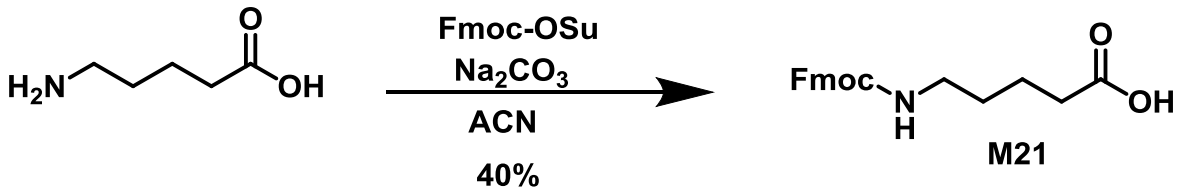


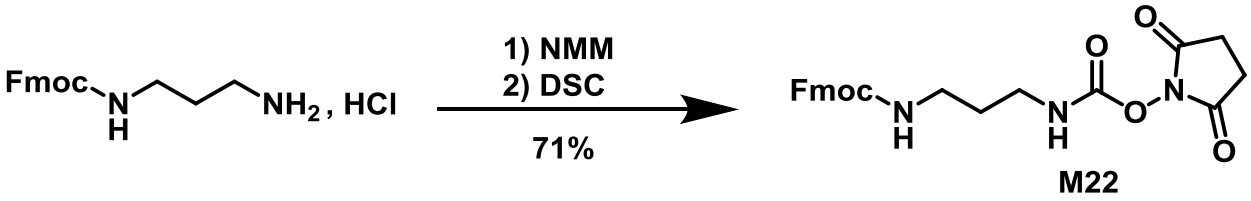
Figure 78: View at the GAGG interface of the X-ray crystal structure of p3/Asf1 complex. The α-helical part of **p3** is represented in blue, the spacer in green, the mimicking part of H4 in red and key hot-spot residues in pink. The orange arrow indicates the C_α of Gly18 and the yellow one indicates the H-bond between Val19 of **p3** and Arg148 of Asf1

We first reasoned on the possibility to directly substitute this GAGG segment because it does not make any contact with Asf1 surface. A basic study of **p3**/Asf1 X-ray crystal structure suggested us that this hydrophobic segment could be replaced by a non-peptidic aliphatic segment. Moreover, this simple substitution could also have a positive impact on the SPS of the peptide (see chapter IV paragraph A2). To replace this segment, we first envisioned to use an aliphatic linker containing five carbons such as aminovaleric acid. Hence, *N*-Fmoc-aminovaleric acid **M21** (N-Fmoc-AVA-OH) was synthesized from its commercially available unprotected aminovaleric acid following regular Fmoc protection conditions and was recovered in moderate yield (40%, see Scheme 23).



Scheme 23: Synthesis of Fmoc-AVA-OH (M21) in Acetonitrile

Additional structural information was gathered when looking deeply at this contact region of the **p3**/Asf1 complex. The NH amide of **p3** Val19 makes indeed an H-bond with Arg148-Asf1 side-chain and the C_α of Gly18-**p3** is oriented toward this Arg148 (Figure 78). This suggested the possibility of introducing a urea bond at this C_α position in order to form one additional H-bond with the Asf1 surface trough the guanidinium moiety of Arg148 side-chain and thereby improving (1) the well-positioning of the valeric-type linker on Asf1 surface and (2) the way it orients the rest of the peptide sequence. To do so, the commercially available *N1*-Fmoc-1,3-diaminopropane hydrochloride (*N*-Fmoc-Dap,HCl) was activated as a carbamate of succinimide to obtain Fmoc-Dap^u-OSu, **M22** in satisfactory yield (71%, see Scheme 24).



Scheme 24: Synthesis of Fmoc-DAP^u-OSu (M22) in CH_3Cl_2

These two monomers **M21** and **M22** in hands, the SPPS of **p11** (ASTEEKWARLARRIA-**AVA-DAP^u**-VTLDGFG see Figure 79) was performed on a preloaded *N*-Fmoc-Gly-Wang resin (loading 0.78 mmol.g^{-1}, 0.05 mmol, 64 mg). The unmodified peptide parts of **p3** were synthesized as previously described whereas *N*-Fmoc-Ava-OH was coupled twice in presence of BOP (5 equiv.)/DIEA (7 equiv.) and *N*-Fmoc-Dap^u-OSu was coupled three times at room temperature with DIEA (7 equiv.) manually. After cleavage in TFA conditions, the desired compound was recovered, purified on RP-HPLC and recovered with 9% yield.

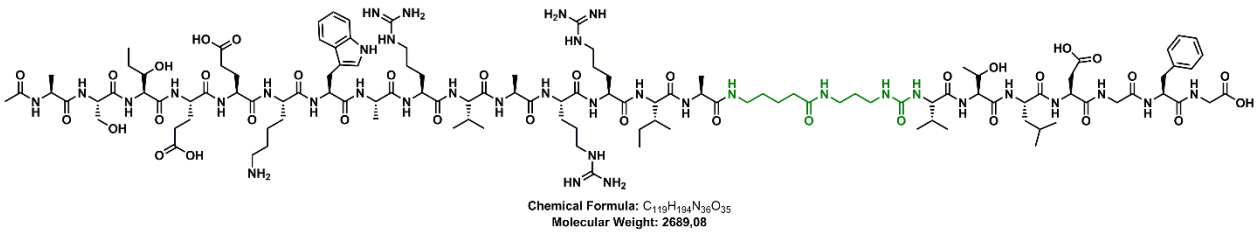


Figure 79: Chemical structure of p11 with the modified linker in green

ITC experiments were performed by adding the peptide **p11** (0.05 mM) into the cell sample containing the protein Asf1 (0.005 mM) in Tris buffer (50 mM, pH 7.4) (Figure 80). The binding affinity of **p11** for Asf1 was in the same range than **p3** with K_D=0.013 µM and a stoichiometry N=1 similar to **p3** was determined. This illustrates that this aliphatic modification allows the conservation of the positioning of key residues at the Asf1 surface. Thermodynamics data will be discussed in details in the following section.

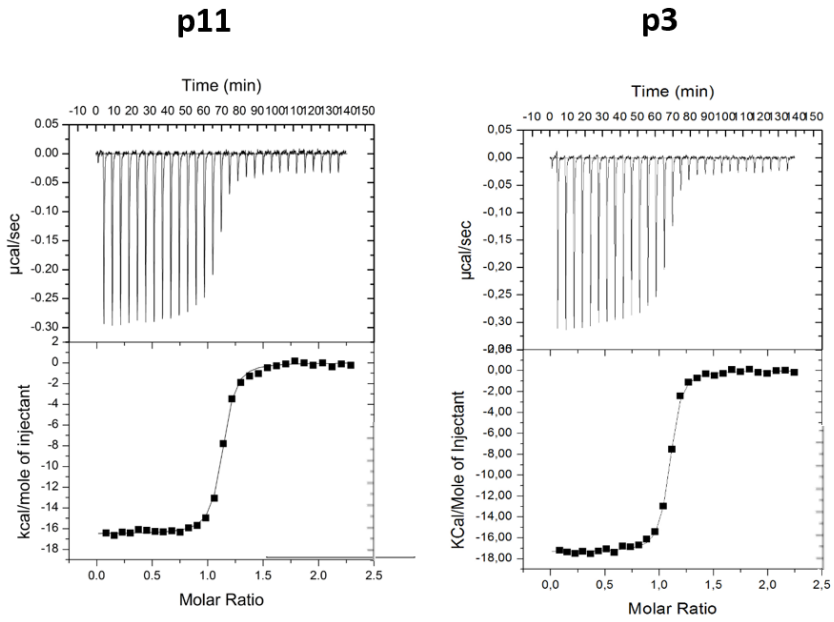


Figure 80: ITC curves and corresponding fitted curves giving access to the thermodynamics data of the binding. Experiments were carried out in Tris-HCl pH7.4 buffer at 20°C with the peptide in the syringe at 0.05 mM and Asf1 protein in the sample cell at 0.005 mM.

Besides, NMR experiments as well as X-ray crystallography assays are currently in progress in the lab of F. Ochsenbein.

2. Simplification of the connection between the two Asf1 binding motifs by removing the GAGG segment

Another design approach has consisted in shortening the length of the peptide spacer installed to reach the two binding zones of Asf1 surface. Based on the crystal structure of **p3**/Asf1 complex, we designed a shorter spacer bypassing the loop formed by the GAGG segment (Figure 81). We thought that the side-chain of Thr21 pointing toward the Ala15, could be substituted by a short aliphatic side-chain connected to Ala15 in order to maximize the contacts with Asf1 surface and to well-position the two peptide segments (the helical and the β-strand ones) of **p3**. For this study,

we decided to modulate the number of methylene units so as to determine the optimal length of this lateral spacer.

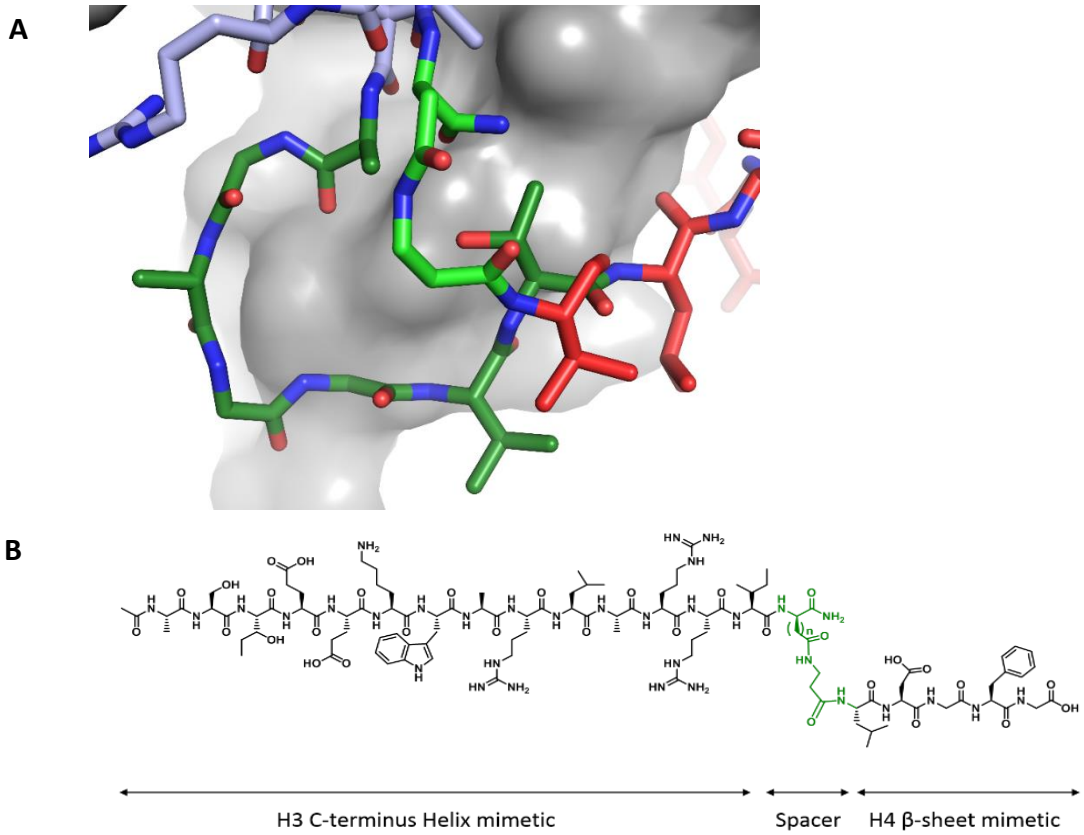
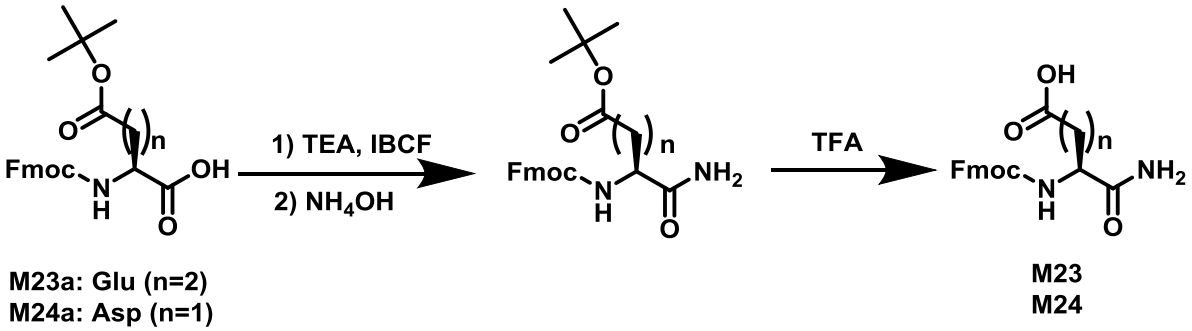


Figure 81: Design of a shorten linker to bypass the GAGG segment. A. Model design of the spacer (light green) on the X-ray crystal structure of p3-Asf1. B. Chemical structure of compound designed with the variable number of methylene units n= 1 or 2.

In this context and with the aim to reproduce the good spatial orientation of the Thr side chain, we decided to use a γ-amino acid with two methylene units coupled to the side-chains of an acidic residue (i.e. Glu or Asp residue) as connection point with the β-sheet part of peptide **p3**. While Fmoc-γ-aminobutyric acid (Fmoc-GABA-OH) is commercially available, monomers Fmoc-Glu(OH)-NH_2 (**M23**), and Fmoc-Asp(OH)-NH_2 (**M24**) were synthesized from *N*-Fmoc-Glu(O*t*Bu)-OH and *N*-Fmoc-Asp(O*t*Bu)-OH respectively by following the synthetic route depicted in Scheme 25. After the installation of the primary amide in the presence of ammoniac, the side-chain *tert*-Butyl protection was easily removed in the presence of TFA and desired products were recovered in 52 and 45% yield respectively.



Scheme 25: Synthesis of monomers M23 and M24

The solid phase syntheses of peptides **p12** and **p13** was performed on Fmoc-Gly-Wang resin LL (0.30 mmol.g^{-1}) and monomers Fmoc-GABA-OH, Fmoc-Glu-NH_2 or Fmoc-Asp-NH_2 were manually coupled with BOP/DIEA as coupling reagents whereas the unmodified peptide parts of **p3** were installed with the *Liberty Blue* synthesizer. After RP-HPLC purification, peptides **p12** and **p13** were recovered in 18% and 14% yield respectively (Table 23). These SPPS appeared to be much simpler compared to that of **p3** probably due to a different behavior of the growing peptide chain on the resin. In addition, this time we did not have to resort to pseudoproline (Fmoc-Val-Thr(ΨMe-MePro)-OH) residue, the $Val_{19}Thr_{20}$ dipeptide being absent of the **p12** and **p13** peptide sequences.

Table 23: Sequences of peptides p11-p13 and characterization by HPLC and mass spectrometry

	Sequence	Yield (%)	tr (min)	$[M+4H]^{4+}$
p3	A S T E E K W A R L A R R I A G A G G V T L D G F G	10	4.71	673.47
p11	A S T E E K W A R L A R R I AVA DAPu V T L D G F G	9	5.65	672.30
p12	A S T E E K W A R L A R R I Glu GABA L D G F G	18	3.35	598.24
p13	A S T E E K W A R L A R R I Asp GABA L D G F G	14	3.75	594.82

Remarkably, the determination of binding affinities to Asf1 of this new series of **p3** analogues carried out by ITC measurements showed that **p3**, **p11** and **p12** exhibited a similar K_D with a value of ≈ 0.01 µM.

Interestingly, if **p11** and **p12** showed comparable binding affinity than the one recorded for **p3**, **p13** missing one methylene unit on the central bridge presented a K_D value diminished by a factor of about three. It is likely that in the case of **p13** the aliphatic bridge is a little bit too short to perfectly position the resulting peptide backbone on the surface of Asf1 and that some deviations of the spatial orientation of the key side-chains occurred.

Compound	K_D (µM)	N	ΔH (kCal.M^{-1})	-TΔS (kCal.M^{-1})	ΔG (kCal.M^{-1})
p3	0.012±0.002	1.1±0.002	-17.300±0.060	6.630±0.100	-10.690±0.040
p11	0.013±0.001	1.1±0.003	-16.500±0.070	5.950±0.130	-10.560±0.050
p12	0.014±0.002	1.1±0.003	-15.370±0.060	4.860±0.110	-10.510±0.040
p13	0.029±0.002	1.1±0.002	0.013±0.001	3.440±0.160	-10.100±0.070

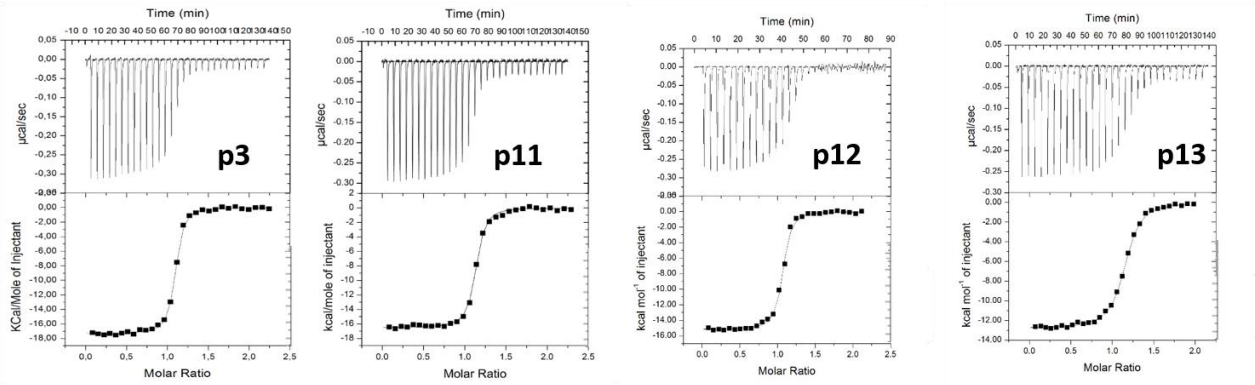


Figure 82: The binding parameters (affinity constant K_D, stoichiometry N, binding enthalpy ΔH, entropy -TΔS and free energy ΔG) of p11-p13 determined during ITC measurements

The binding signatures for the three designed peptides **p11**-**p13** were extracted from the ITC experiments and are reported in Figure 82. The enthalpy and entropy obtained for this series of peptides (*i.e.* **p11**-**p13**) are in the same range of order than **p3**, illustrating here that the central modifications do not lead to thermodynamic penalties. Similarly, the high absolute values of enthalpy of **p11**, **p12** and to a lesser extent of **p13** combined with unfavorable entropy factors reveal that binding to Asf1 surface occurs thanks to strong polar interactions with (i.e. H-bonds and Van der Waals interaction) with conformational changes of the peptides. To note, **p12** shows a value of entropic contribution that is more favorable to the binding that those measured for **p3** and **p11** suggesting here a higher contribution of the hydrophobic interactions and possibly a better anchoring of the Phenyl side-chain into the hydrophobic pocket of Asf1.

However, and even though showing the lowest entropy factor (-TΔS = 3.5), peptide **p13** also presents the lowest absolute enthalpy of binding value and consequently the lowest binding affinity to Asf1 due to the probable loss of polar interactions during the course of the binding.

This study confirmed the feasibility of mutating or even by-pass the GAGG spacer in **p3** leading to a simplified SPPS and maintaining a high affinity for Asf1 protein. Additional structural analysis in solution (^{15}N-HSQC NMR experiments) have been performed on this series of peptides but unfortunately all the data have not been retrieved and analyzed yet. Moreover, several crystals have been obtained for **p11** and **p12** at more than 10 Å and 2.2 Å resolution respectively. The crystals obtained for **p11** were really small and data set was of poor quality to be refined. An optimization of crystals conditions will be required. However, a rapid observation of the electron density for **p12** suggested that the helical part was well-localized at the surface of Asf1 whereas the C-terminal part was in contact with another Asf1 protein. This unexpected result can find an explanation in a crystallography artefact. Crystal conditions should be optimized soon in order to obtain higher resolution crystal data sets.

3. Extension of the urea/amide peptidomimetic by adding the histone H4 binding epitope at the C-terminus

The huge improvement of binding affinity to Asf1 brought by the presence of the second binding epitope Phe25 in **p3**, motivated us to investigate the potential increase of binding affinity after lengthening the C-terminal part of **ch5** chimera. Considering the binding affinity results gathered with the peptides **p3**, **p11** and **p12**, several peptides sequences were appended at the C-terminus of **ch5** foldamer with the aim to establish their incidence on the binding to Asf1.

a) Lengthening of ch5 C-terminus with the p3 peptide end and affinity determination by ITC experiment

First, we were interested to elongate the **ch5** foldamer with the native C-terminal sequence of **p3**. The incorporation of this C-terminal segment of **p3** including the GAGG loop at the C-terminal end of **ch5**, would indeed integrate the second binding epitope (Phe residue) and thereby should have a positive impact on the binding affinity to Asf1 protein. A basic superimposition of the X-ray crystal structures of **p3** and **ch5** confirmed the feasibility of this elongation strategy even though a small shift of 1.8 Å between the **p3** Ala15 and **ch5** Ala9 was noticed (Figure 83).

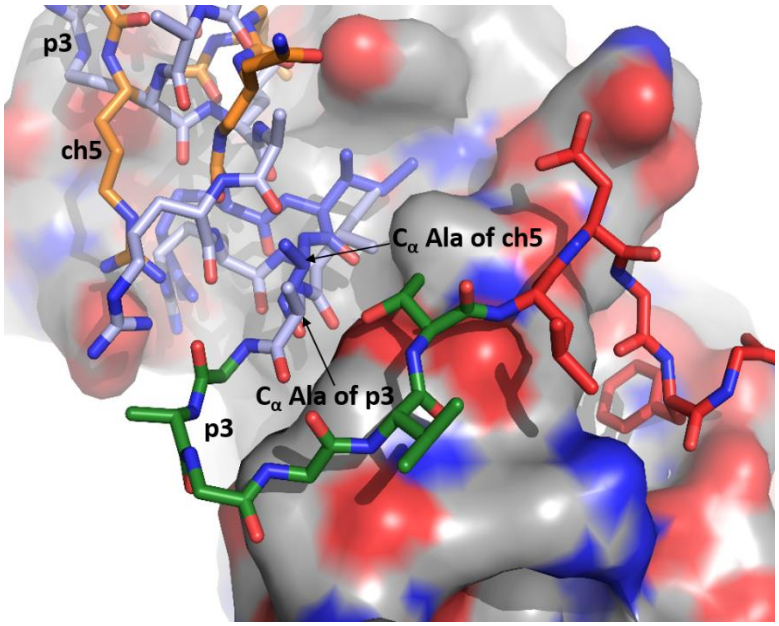


Figure 83 : Superimposition of ch5 and p3 X-ray crystal structures aiming at tethering the C-terminal segment of p3 and reach the second binding epitope. Urea residues of **ch5** are in orange whereas the peptide part is in dark blue. The helical part of **p3** is represented in light blue, the loop in green and the H4 β-sheet mimicking segment is represented in red.

The synthesis of the tethered chimera, **ch20**, was carried out on *N*-Fmoc-Gly-WANG resin LL. The C-terminal peptide segment was automatically introduced with the Liberty Blue microwave synthesizer using the same conditions and peptide strategy to those optimized for **p3**, whereas the urea part and the last α-amino acid residues were introduced manually as previously described during the synthesis of **ch5**. After RP-HPLC semi-preparative purification, the expected α/urea chimera **ch20** was recovered with an overall yield of 7%.

This new hybrid sequence was next assayed for its ability to bind better to Asf1. ITC experiments performed under the same conditions unveiled that chimera **ch20** actually exhibits a K_D value remaining in the micromolar range (K_D=1.2 ± 0.2 µm). This result illustrates a really slight improvement of the affinity compared to **ch5** (K_D=2.7 ± 0.6 µm). Compared to the drastic improvement in binding affinity observed between **p3h** (K_D=150 ± 50 nm) and **p3** (K_D= 3 ± 1 nm), the binding affinity obtained for **ch20** is quite disappointing and underlines that the targeted insertion of the phenyl side-chain within the hydrophobic pocket of Asf1 is probably not occurring.

Crystallography assays are currently in progress in the lab of F. Ochsenbein but we do not have data in hands yet. Having the crystallographic data of the **ch20**/Asf1 complex would obviously give us precious structural information in order to better understand the binding mechanism of **ch20** and optimize the chemical connection between the chimera **ch5** and the rest of **p3**.

b) Synthesis of ch5 lengthened at C-terminus with the short-cut strategy and affinity determination by ITC experiment

Another approach to lengthen **ch5** C-terminus was based on the insertion of the "by-pass linkers" developed for peptides **p12** and **p13** respectively. Using these shorter linkers would considerably reduce the size and the molecular weight of the foldamers making them more valuable for potential future drug development. Despite the fact that the best affinity results were obtained for **p12**, we decided to equip **ch5** with both linkers. Indeed, the superimposition of **ch5** and **p3** crystal structure in complex with Asf1 showed a shift of the position of C_α of Ala in C-terminus, suggesting that one linker potentially optimal for **p3** could be less efficient to bridge the chimera **ch5**. Hence, both linkers were installed and the synthesis of the two new chimeras **ch21** and **ch22** was performed (Figure 84).

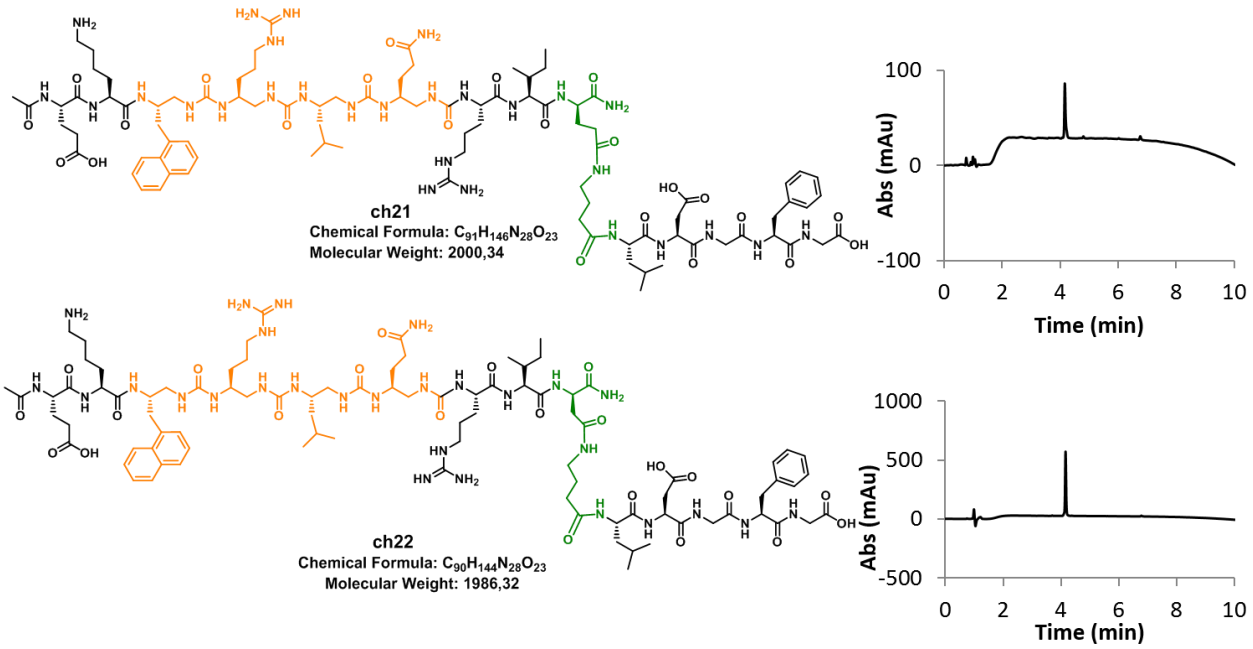


Figure 84: Chemical structure of ch5 elongated with "by-pass" spacers leading to ch21 and ch22. Natural amino acids are in black, urea residues in orange and spacer in green. The corresponding HPLC chromatograms of pure products with a gradient of 20-100% ACN in 10 min are represented in the right.

The solid phase synthesis of these two chimeras, **ch21** and **ch22** was performed on *N*-Fmoc-Gly-WANG-resin LL. The peptide part was introduced as described in the paragraph 2 and the urea part was introduced in the same way than for the synthesis of **ch5** (scheme 4, chapter 3, A, 3, b). After RP-HPLC semi-preparative purification, the expected α/urea hybrid foldamers **ch21** and **ch22** were both obtained with an overall yield of 13% (Table 24). Again, the recovered yields

underline that having removed the GAGG segment considerably simplified the SPS of the new chimera sequences.

Table 24: Sequences of chimera ch5 lengthened at C-terminal end characterization by HPLC and mass spectrometry

	Sequence																									Yield (%)	tr (min)	$[M+3H]^{3+}$
p3	A	S	T	E	E	K	W	A	R	L	A	R	R	I	A	G	A	G	G	V	T	L	D	G	F	G	10	4.71	673.47
ch5					E	K	Nal^u		R^u		L^u		Q^u		R	I	A										8	4.55	456.67
ch20					E	K	Nal^u		R^u		L^u		Q^u		R	I	A	G	A	G	G	V	T	L	D	G	F	G	7	5.41	767.75
ch21					E	K	Nal^u		R^u		L^u		Q^u		R	I	Glu			GABA			L	D	G	F	G	13	4.13	667.71
ch22					E	K	Nal^u		R^u		L^u		Q^u		R	I	Asp			GABA			L	D	G	F	G	13	4.16	663.05

The binding affinities to Asf1 were again determined by performing ITC experiments and chimeras **ch21** and **ch22** exhibit dissociation constant of 6.2 ± 0.9 µM 2.2 ± 0.8 µM respectively (Figure 85). The shapes of the fitting curves are not as good as for the peptide and more similar to **ch5**. In contrast with α-peptides **p12** and **p13**, the best affinity was recorded for the shortest linker equipped with Asp residue of **ch22** which is better tolerated than with the additional carbon from Glu residue spacer residue in **ch21**. This confirms our wish to test both linkers due to the shift of Ala residue in C-terminus of **ch5**. However, these binding affinity values remain less good than the one recorded after tethering with the native C-terminal part of **p3**. Despite these encouraging results the best affinity for lengthened chimera remains the one observed with **p3** (**ch20**). This suggests that the bridge between **ch5** with the H4 mimicking segment is not optimal and considering the K_D values there are big chance that the Phe side-chain does not reach its hydrophobic pocket.

At the opposite of **ch5** that showed a favorable entropic factor (-TΔS negative), all the new chimera sequences exhibit penalty from the entropic contribution that diminishes the binding to Asf1 (Figure 85). Concurrently, the absolute enthalpy value of each new sequence appears to be considerably enhanced in comparison to **ch5** highlighting an optimization of the polar contacts and interactions during the association process of the chimeras with Asf1.

ITC measurements reveal that sequences **ch21** and **ch22**, which are equipped with the "by-pass" linker, have an improved binding enthalpy compared to **ch5**, exhibiting an absolute ΔH (14 and around 11) value close to that of **p3** (17). However, the penalty brought by the entropic factor considerably diminish the capacity of the resulting sequences **ch20**-**ch22** to tightly bind to Asf1, this even though the ΔS values recorded for chimeras **ch20-ch21** appear better than that of **p3**.

Compound	K_D (µM)	N	ΔH (kCal.M^{-1})	-TΔS (kCal.M^{-1})	ΔG (kCal.M^{-1})
p3	0.012±0.002	1.11±0.01	-17.30±0.06	6.63±0.10	-10.69±0.04
ch5	2.73±0.59	0.98±0.09	-2.22±0.03	-5.26±0.59	-7.47±0.12
ch20	1.16±0.21	1.20±0.01	-11.00±0.12	3.04±0.15	-7.95±0.03
ch21	6.19±0.90	1.11±0.01	-13.97±1.00	6.99±1.04	-6.98±0.04
ch22	2.19±0.80	1.10±0.01	-10.69±0.11	3.12±0.12	-7.58±0.02

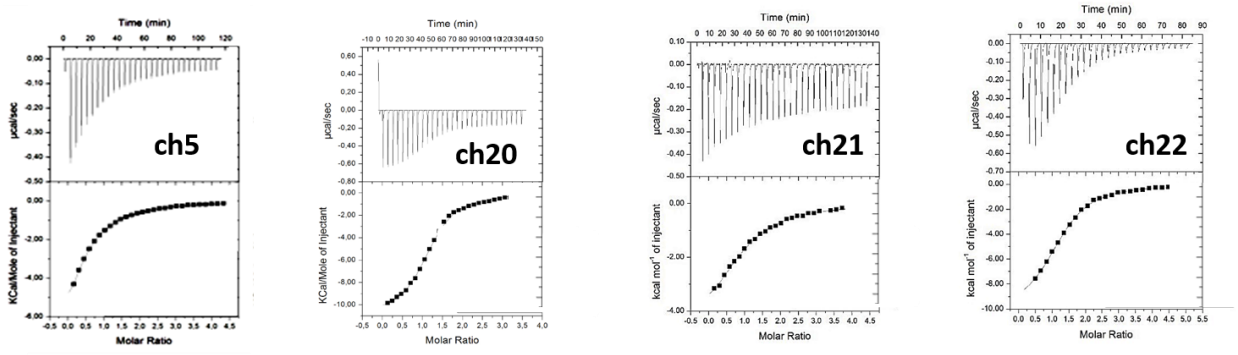


Figure 85: The binding parameters (affinity constant K_D, stoichiometry N, binding enthalpy ΔH, entropy -TΔS and free energy ΔG) of ch20-ch22 determined during ITC measurements

As for the other series of chimeras discussed in the previous chapters, structural analyses in solution will be soon performed (^{15}N-HSQC NMR experiments) and the X-ray crystallography assays as well. Intense efforts are currently done in this direction in the group of Françoise Ochsenbein.

In this chapter, we showed that the binding affinity of chimeras could be improved by increasing the absolute value of the binding enthalpy. The lengthening of **ch5** in C-terminus has been performed and the synthesis has been optimized allowing the preparation of several sequences that have been screened for their binding affinity to Asf1. All these chimeras showed affinities for the protein Asf1 in the same range or slightly better than **ch5**. These results highlight that the fragment approach can be highly efficient to gain binding affinity to Asf1. Despite the fact that crucial structural information is still missing (in solution and at solid state), these results are extremely encouraging for future peptidomimetic constructions with the aim to improve the affinity of chimeras or even to access homo-oligoureas as H3 helical mimicking fragments. The determination of the X-ray crystal structure would be also beneficial for these new designs even though molecular modeling studies can be helpful in finely tune the engineered sequences.

Bibliography

(1) Natsume, R.; Eitoku, M.; Akai, Y.; Sano, N.; Horikoshi, M.; Senda, T. Structure and Function of the Histone Chaperone CIA/ASF1 Complexed with Histones H3 and H4. *Nature* **2007**, *446* (7133), 338–341.

(2) SubirÃ³s-Funosas, R.; Prohens, R.; Barbas, R.; El-Faham, A.; Albericio, F. Oxyma: An Efficient Additive for Peptide Synthesis to Replace the Benzotriazole-Based HOBt and HOAt with a Lower Risk of Explosion [1]. *Chem. - Eur. J.* **2009**, *15* (37), 9394–9403.

(3) Abedini, A.; Raleigh, D. P. Incorporation of Pseudoproline Derivatives Allows the Facile Synthesis of Human IAPP, a Highly Amyloidogenic and Aggregation-Prone Polypeptide. *Org. Lett.* **2005**, *7* (4), 693–696.

(4) Haack, T.; Mutter, M. Serine Derived Oxazolidines as Secondary Structure Disrupting, Solubilizing Building Blocks in Peptide Synthesis. *Tetrahedron Lett.* **1992**, *33* (12), 1589–1592.

(5) Mutter, M.; Nefzi, A.; Sato, T.; Sun, X.; Wahl, F.; Wöhr, T. Pseudo-Prolines (Psi Pro) for Accessing "Inaccessible" Peptides. *Pept. Res.* **1995**, *8* (3), 145–153.

(6) Lauer, J. L.; Fields, C. G.; Fields, G. B. Sequence Dependence of Aspartimide Formation during 9-Fluorenylmethoxycarbonyl Solid-Phase Peptide Synthesis. *Lett. Pept. Sci.* **1995**, *1* (4), 197–205.

(7) Yang, Y.; Sweeney, W. V.; Schneider, K.; Thörnqvist, S.; Chait, B. T.; Tam, J. P. Aspartimide Formation in Base-Driven 9-Fluorenylmethoxycarbonyl Chemistry. *Tetrahedron Lett.* **1994**, *35* (52), 9689–9692.

(8) Abdel-Aal, A. M.; Papageorgiou, G.; Raz, R.; Quibell, M.; Burlina, F.; Offer, J. A Backbone Amide Protecting Group for Overcoming Difficult Sequences and Suppressing Aspartimide Formation. *J. Pept. Sci.* **2016**, *22* (5), 360–367.

(9) Behrendt, R.; Huber, S.; White, P. Preventing Aspartimide Formation in Fmoc SPPS of Asp-Gly Containing Peptides - Practical Aspects of New Trialkylcarbinol Based Protecting Groups: New Protecting Groups Preventing Aspartimide Formation in Fmoc SPPS. *J. Pept. Sci.* **2016**, *22* (2), 92–97.

(10) Scholtz, J. M.; Baldwin, R. L. The Mechanism of Alpha-Helix Formation by Peptides. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21* (1), 95–118.

(11) Bhattacharjee, N.; Biswas, P. Helical Ambivalency Induced by Point Mutations. *BMC Struct. Biol.* **2013**, *13*, 9.

Conclusion and perspectives

Oligourea peptidomimetics combine several interesting properties such as a stabilized and predictable helical conformation and resistance to proteolysis, making them interesting candidates for the development of drugs. In this thesis, we showed that their urea-based backbone could be combined with the α-peptide backbone to form α/urea chimeras retaining helical conformation. Short urea segments were successfully introduced at different positions of water-soluble peptides of various sizes. In several cases, the solid phase synthesis required optimization cycles to efficiently prepare the target chimera sequences. Remarkably, the desired α/urea hybrid foldamers were all recovered in honorable yields. During this SPS optimization time, it also clearly appeared that the synthesis of this family of chimera foldamers could only be mastered if the peptide component was accessible in high purity and yield.

Regarding the design of α/urea hybrid foldamers mimicking GCN4-p1 peptide, the selection of the side-chains appeared to be a key parameter to mimic as much as possible the native peptide. The mimicry appeared to be even more complex when mutations were introduced at the vicinity of the Asn16residue in GCN4-p1 sequence due to its key role in the self-assembly process. In this project, we showed that the insertion of a short urea-based segment within a long peptide has an impact on the overall structural conformation of the resulting chimera. We studied the conformational behavior of a series of GCN4-p1 mimics in solution while focusing on getting a high-resolution crystal structure essential for the characterization of the quaternary structure of the resulting chimera assemblies. However, despite having obtained crystals suitable for X-ray diffraction and high-resolution data sets for several compounds, unfortunately no X-ray structure have yet been solved. Many efforts are currently devoted in the group to overcome the phasing problem and to grow new crystals in different crystallization conditions. Indeed, solving one structure of a GCN4-p1 mimic at atomic resolution would give us crucial structural information on the impact of the α→urea substitution pattern and would facilitate new designs for an efficient chimera copy of native GCN4-p1 assembly. The length of the urea segment might be increased to have a stronger effect of the urea segment for optimal stabilization and propagation of the helical conformation. Another perspective would be to study for example the self-assembly in solution by performing differential scanning calorimetry experiments. In the absence of high resolution structural data, it remains difficult to establish a general and precise rule for replacement of a peptide segment by a urea backbone but nevertheless, we have shown here that introducing a urea segment into a peptide sequence known to adopt α-helix conformation is a feasible concept.

The second part of this thesis was devoted to the design of α-helix mimics for the inhibition of Asf1/H3-H4 complex by resorting to the urea-based foldamer approach. We essentially focused our work on the design and the synthesis of α/urea chimeras able to mimic the C-terminal part of histone H3. The side chain selection was determinant to reach a good binding affinity to Asf1, in particular to find a way to recapitulate the main hot-spot interactions. However, the choice of the other side-chains was also important in order to improve the binding enthalpy with the formation of additional salt bridges or H-bonds between the urea-based foldamer and the protein surface. To do so, basic molecular modelling methods or more sophisticated ones were used and, in some cases, revealed their efficiency. Indeed, a simple superimposition of the X-ray crystal structure of reference **p3** peptide with the backbone of an oligourea (for which the crystal structure was available) led to the design of chimera **ch5** which after synthesis and binding assays revealed to be really promising. More sophisticated methods such as Rosetta were also used to optimize key isobutyl side chain of $Leu^{U}5$ in **ch5**. However, the results gathered from this new optimization cycle were quite disappointing. Nonetheless and to our delight, when the same software was used to design the first full homooligourea **ol1,** this sequence showed promising binding affinity to Asf1.

The thermodynamic data characterizing the interactions between urea-based foldamers and Asf1 were determined by ITC experiments and showed that the introduction of the urea segment results in a more favorable entropy term but the loss of binding enthalpy impacts the overall binding affinity. However, we managed to obtained short chimeras (9 residues) with micromolar range affinities. Efforts have been done to characterize at the solid state the protein-foldamer interactions and we were pleased to obtain in the group the first X-ray crystal structure of a urea-based foldamer (**ch5**) in complex with a protein (Asf1). This result confirmed that **ch5** interacts with Asf1 in the same region than the C-terminal part of H3. The X-ray crystal structure also highlighted the adaptability/plasticity of **ch5** to adapt itself to the surface of Asf1. Furthermore, this compound was tested *in cellulo* and showed a toxicity for cancer cells which is really encouraging for future drug development. However, additional experiments are necessary to confirm that the observed toxicity was induced by engagement of the target (by the chimera) and not from the cell penetrating peptide used to deliver the chimera within cells in a non-specific manner. Nevertheless, this preliminary cellular study unveils the potential of peptide backbone modifications with urea segment. Additionally, the enzymatic stability of **ch5** was clearly improved compared to the corresponding peptide which offers a significant advantage over α-helical backbones. For this series of α/urea chimera, ^{1}H-^{15}N HSQC NMR experiments have been performed and the data are currently being treated to give access to additional structural information in solution. A number of new crystallization trials have been performed and current efforts focus on the resolution of structures or improvement of crystallization conditions, in particular for the series of α/urea hybrid foldamers encompassing three urea-type units.

Finally, by targeting a second binding site located on another face of Asf1 surface, we managed to slightly improve the affinity of our lead compound **ch5**, in particular, the binding enthalpy was clearly improved. However, compared to peptide **p3**, the affinity of the lengthened version of **ch5** remains significantly lower. An important design work on the GAGG linker has been performed but further improvements are still mandatory to optimally position the Phe side-chain into its Asf1 hydrophobic pocket. As a first possibility, the aliphatic linker composed of the aminovaleric acid and the DAP^{u} (Figure 86) could be inserted in **ch5** sequence as the corresponding full peptide version presented an affinity similar to **p3**.

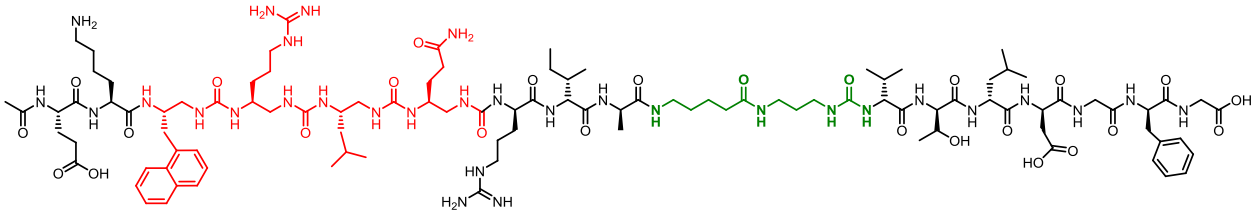


Figure 86: Chemical structure of ch5 lengthened in C-terminal with aliphatic linker

Besides, to improve the binding enthalpy of such elongated chimeras, S. Traore in the team of R. Guerois has done a simple molecular modelling study using Pymol to build new linkers (Figure 87). Indeed, as unnatural backbones are present into the sequence, the use of Rosetta appeared to be too much complicated as new parameters were requested to be first implemented. Despite this limitation, this model suggests us new sequences to synthesize. In contrast to already prepared and discussed chimeras, the Val19Thr20 residues of **p3** are conserved in this model, increasing the size of the linker. This could be beneficial for the binding to Asf1 as it could restore the H-bond occurring between side chains Thr20 of **p3** and Thr147 of Asf1.

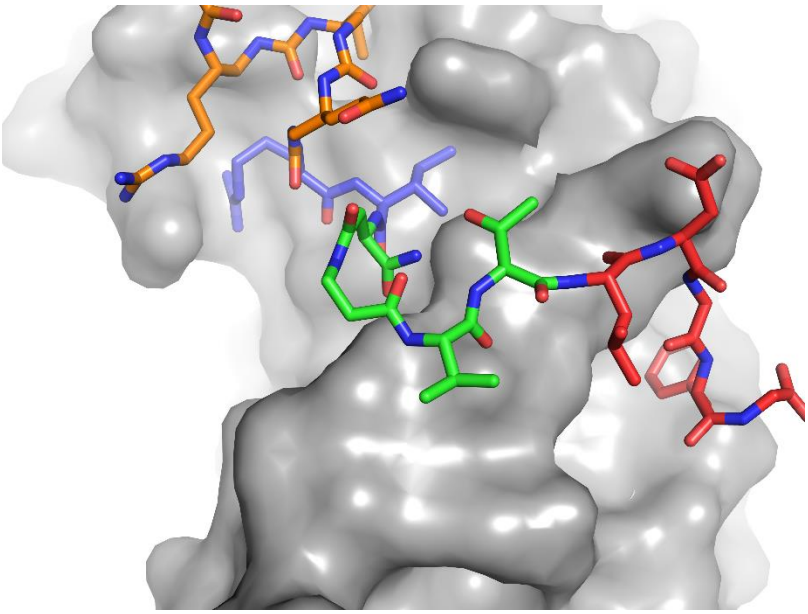


Figure 87: Design of a short linker to remove GAGG segment in ch5 elongation.

This basic model could allow the design of future sequences to be synthesized and evaluated for their ability to associate with Asf1. Another sequence that could be synthesized corresponds to **ch5** with the aliphatic spacer found in **p11**. This illustrates a room for improving the binding affinity of chimeras to Asf1. An alternative approach could be envisioned to design a linker with oligourea residues.

Finally, efforts are needed to obtain additional structural information in solution by HSQC-NMR as well as in the solid state with crystallogenesis trials. Crystals have already been obtained for peptides derived from **p3** and efforts to improve the crystallogenesis conditions are underway.

This thesis was supported by *La Ligue contre le Cancer* which allowed the initiation of new projects that will be pursued, in particular the project targeting Asf1 in collaboration with the team of F. Ochsenbein as new funding was obtained at the end of my PhD in the group. Importantly, a collaboration with the team of G. Almouzni at the Curie institute would provide increased knowledge about the mode of action of compounds in this series *in cells* and also *in vivo.*

Moreover, all the results gathered during this thesis will be useful for other projects in particular for those involving PPIs in general. The development of the chimera synthesis on solid support has been optimized during my PhD work, in particular for water-soluble sequences and for the introduction of urea segments into large peptides.

Finally, the potential of full oligoureas used as PPI inhibitors is only at its first stage and will be deeper explored because this peptidomimetic approach could be highly valuable for future drug design. These urea-based foldamers represent an interesting synthetic tool for extracellular and intracellular targets highlighting their clear potential for pharmaceutics. In addition, this foldamer-based toolbox could be combined with other orthogonal chemical approaches developed by other groups to improve the biological activities of α-helical peptides.

Experimental part

A. General methods

Commercially available reagents were used throughout without purification. The solvents were purchased from Sigma-Aldrich. THF, DCM and Toluene were preliminary dried by passing through solvent drying system (SPS 800 Manual from MBraun). TFA was purchased from Alfa Aesar. 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 semi-preparative purification. *N*-protected amino acids were purchased from Polypeptide Laboratories or Iris Biotech. NovaPEG Rink amide resin, Rink amide MBHA resin and Fmoc-Gly-Wang resin were purchased from Novabiochem.

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 chromatography was carried out on silica gel (40-63 µm, Merck). Purifications by automated Combiflash were performed with the Teledyne Isco Combiflash Rf^+.

Solid phase peptide syntheses with microwave irradiation were carried out on Liberty Blue System, from CEM (CEM µWaves S.A.S., Orsay, France). Solid phase oligourea syntheses with microwave irradiation were carried out on Discover Bio system, from CEM (CEM µWaves S.A.S., Orsay, France).

Analytical RP-HPLC analyses were performed on a Dionex ultimate U3000 using a Macherey-Nagel Nucleodur column (4.6 × 100 mm, 3 µm) at a flow rate of 1 mL/min. The mobile phase was composed of 0.1% (v/v) TFA-H_2O (Solvent A) and 0.1% (v/v) TFA-CH_3CN (Solvent B). Semi-preparative purifications of oligoureas were performed on a Dionex ultimate U3000 using a Macherey-Nagel Nucleodur VP250/10 100-16 C18ec column (10 × 250 mm, 16 µm) at a flow rate of 4 mL/min. Preparative purifications of oligoureas were performed on a Gilson GX-281 using a Macherey-Nagel Nucleodur VP250/21 100-5 C18ec column (21 × 250 mm, 5 µm) at a flow rate of 20 mL/min.

^1H NMR and ^{13}C NMR spectra were recorded on three different NMR spectrometers: (1) an Avance II NMR spectrometer (Bruker Biospin) with a vertical 7.05T narrow-bore/ultrashield magnet operating at 300 MHz for ^1H 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 (Bruker Biospin) with avertical 9.4T narrow-bore/ultrashield magnet operating at 400 MHz for ^1H observation by means of a 5-mm direct QNP $^1H/^{13}C/^{31}P/^{19}F$ probe with gradient capabilities; (3) an Avance III NMR spectrometer (BrukerBiospin) with a vertical 16.45T narrowbore/ultrashield magnet operating at 700 MHz for ^1H observation by means of a 5-mm TXI $^1H/^{13}C/^{15}N$ probe with Z gradient capabilities. Chemical shifts are reported in parts per million (ppm, δ) relative to the ^1H or ^{13}C residual signal of the deuterated solvent used. ^1H NMR splitting patterns with observed first-order coupling are designated as singlet (s), broad singlet (brs),

doublet (d), triplet (t), or quadruplet (q). Not defined ^{1}H NMR splitting patterns are designated as multiplet (m). Coupling constants (J) are reported in hertz (Hz).

ESI-HRMS analyses were carried out on a Thermo Exactive with an ion trap mass analyzer from the Mass Spectrometry Laboratory at the European Institute of Chemistry and Biology (UMS 3033 - IECB), Pessac, France. ESI-MS analyses were carried out on a ThermoElectron LCQ Advantage spectrometer equipped with an ion trap mass or on a Agilent 6230 with a time of flight analyzer with an UHPLC Agilent 1290 Infinity (Agilent Eclipse Plus C18 RRHD 1.8µM 2.1*50 mm) from the Mass Spectrometry Laboratory at the European Institute of Chemistry and Biology (UMS 3033-IECB).

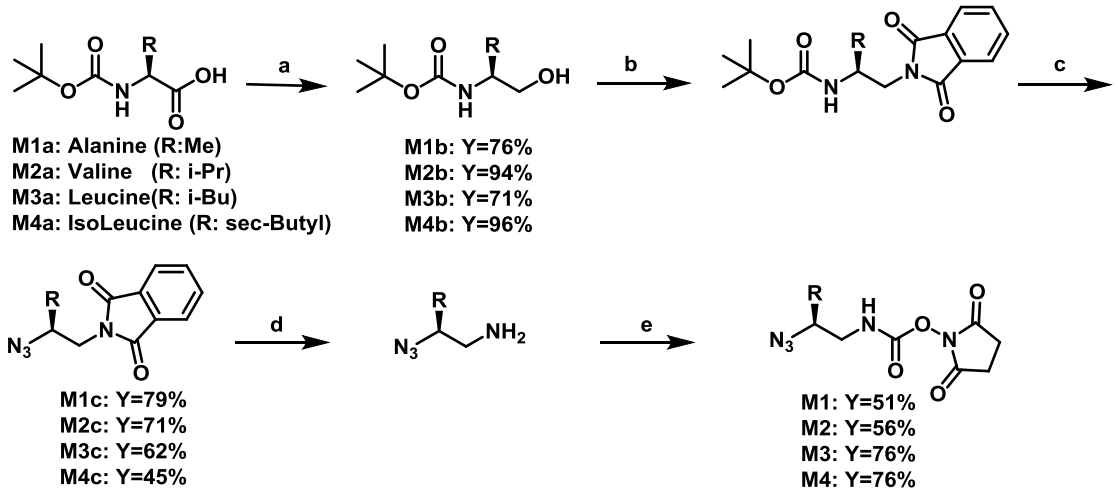
All Circular dichroism (CD) spectra were recorded on a J-815 Jasco dichrographe (Jasco France, Nantes, France) using a quartz cell with a path length of 1 mm or 2mm (Hellma, Paris, France). Samples temperature was regulated at 20°C. Data were collected in continuous scan mode with a data pitch of 1 nm, a scanning speed of 50 nm/min, an integration time of 5 s, 2 nm bandwidth and two accumulations per sample. Samples Data were collected as raw ellipticity (ψ in mdeg) and converted to per residue mean ellipticity (PRME with [θ] = ellipticity) in deg.cm^{2}.dmol^{-1}.residue^{-1} using the following equation: $[\theta] = (\psi \times 10^{-3}) / (res \times l \times c)$

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

B. Solution synthesis of Monomers

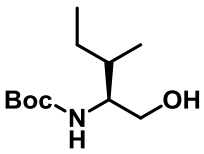
1. Preparation of azide protected *O*-succinimidyl carbamate monomers

Activated monomers equipped with Ala- (**M1**), Val- (**M2**), Leu- (**M3**) and Ile- (**M4**) side chains were synthesized as previously reported[1].



a: 1) NMM, IBCF, THF, -20°C; 2) $NaBH_4$, H_2O; b: PPH_3, Phthalimide, DIAD, THF, r.t; c: 1) TFA; 2) $N_3SO_2Im.HCl$, K_2CO_3, $CuSO_4$:$5H_2O$, CH_3CN:H_2O(1:1), r.t; d: $N_2H_4.H_2O$, MeOH, 70°C,e: DSC, CH_2Cl_2,r.t.

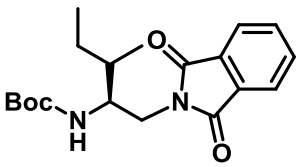
Procedure for conversion into alcohol



The N-Boc-Ile OH*$2H_2O$ (10g, 10.4mmol) was dissolved in THF under N_2 and cooled to -20°C. After addition of NMM (1.35 mL, 12.24 mmol) and IBCF (1.6 mL, 12.24 mmol) the mixture reaction was stirred at -20°C for 40 min. The resulting white suspension was filtered off and a solution of $NaBH_4$ (462 mg, 12.24 mmol) in water was added drop by drop to the filtrate cooled to -20°C. The reaction was left to react overnight. After the addition of water and THF removal under reduced pressure, the aqueous solution was extracted with EtOAc. The organic layers were combined and washed two times with 1M $KHSO_4$, two times with solution of saturated $NaHCO_3$, once with Brine solution and dried over $MgSO_4$. After concentration under vacuum, and silica gel flash chromatography, the desired pure product was furnished as a colorless oily product with 96% yield.

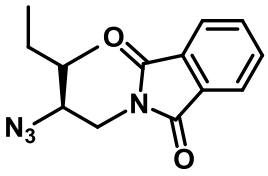
tert-butyl (2S,3S)-1-hydroxy-3-methylpentan-2-ylcarbamate M4b: ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.65 (s, 1H), 3.75 – 3.55 (m, 2H), 3.50 (d, *J* = 4.7 Hz, 1H), 1.45 (s, 9H), 1.42 (s, 2H), 1.20 – 1.09 (m, 1H), 0.93 (d, *J* = 2.4 Hz, 3H), 0.91 (d, *J* = 1.9 Hz, 3H).

Procedure for conversion of alcohol into phtalimide- Mitsunobu reaction:



N-Boc-Ile-alcohol (2.16g, 10 mmol) was dissolved in anhydrous THF under N_2 atmosphere pressure. PPh_3 (3.15g, 12 mmol) in THF and phthalimide (1.76g, 12 mmol) were added, followed by dropwise addition of DIAD (2.4mL, 12 mmol). The reaction mixture was stirred overnight at r.t. After concentration under reduced pressure, a flash silica gel chromatography was performed to remove PPh_3 and the material was directly used to the next step.

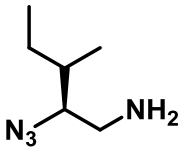
Boc removal and azide formation on Boc-Ile-Phthalimide:



N-Boc-Ile-Phthalimide was dissolved in pure TFA (10mL) and left to react without stirring for 2 h. The TFA was removed under reduced pressure and the TFA salt was dried overnight on vacuum line. The TFA salt was directly engaged in the next step without any purification. To a solution containing the TFA salt (10 mmol) in CH_3CN/H_2O (1:1), were successively added K_2CO_3 (2.04g, 12.84 mmol), imidazole-1-sulfonyl azide hydrochloride (N_3SO_2Imidazole.HCl) (2.46 g, 11.77 mmol) and $CuSO_4.5H_2O$ (15 mg, 0.1 mmol). The mixture was stirred at r.t. overnight. The reaction mixture was then filtered off to remove the remaining K_2CO_3. After concentration under reduced pressure, EtOAc was added. The organic phase was washed two times with $KHSO_4$ solution (1M), once with Brine, dried over magnesium sulfate and concentrated under reduced pressure. After silica gel flash chromatography, the desired pure product was furnished as a colorless oily product with 45% yield.

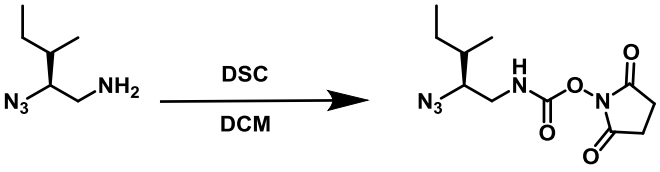
(S)- (2-azido-3-methylpentyl)phtalimide M4c: ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.98 – 7.33 (m, 4H), 3.64 (s, 1H), 2.27 (dd, *J* = 58.3, 15.9 Hz, 1H), 1.57 (s, 2H), 1.40 (d, *J* = 15.0 Hz, 2H), 1.18 – 1.00 (m, 3H), 0.98 – 0.72 (m, 3H).

Procedure for azide Ile amine derivative:



To a solution of azido phthalimide derived of Ile (4.8 mmol, 1.3g) in MeOH, hydrazine hydrate (690 µL, 14 mmol) was added. The reaction mixture was heat to reflux and the reaction was stirred 4h with apparition of a white precipitate. The reaction mixture was filtered off, washed with MeOH and the filtrate was concentrated under reduced pressure. The crude material was dissolved in EtOAc and the organic layer was washed twice with 1M HCl solution. The combined aqueous phases were again washed with EtOAc and were neutralized by addition of K_2CO_3 until pH8. The aqueous phase was finally extracted with DCM and the combined organic layers were dried over MgSO4, concentrated under reduced pressure but not until dryness to avoid the loss of azido Ile derivative which is quite volatile. Resulting azido amine was directly engaged for activation step without further purification.

Synthesis of N_3-Ile-OSu:

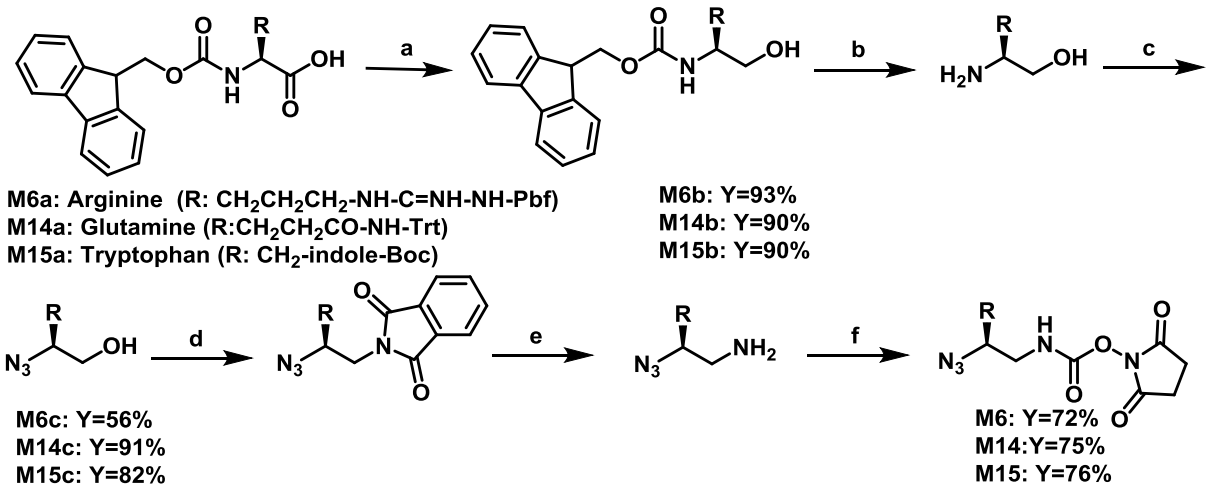


To a stirred suspension of disuccinimidyl carbonate (DSC) (1.23g, 4.8 mmol) in distilled DCM, was added drop by drop a solution of the derivative amine in DCM. The reaction mixture was stirred 4 h and was concentrated under reduced pressure. The crude material was dissolved in EtOAc and the organic phase was washed two times with 1M $KHSO_4$, once with Brine solution and dried over MgSO4

before to be concentrated under reduced pressure. The product was precipitated in a cold mixture pentan/Et_2O (3:7, v/v) and furnished as solid with a yield of 76%.

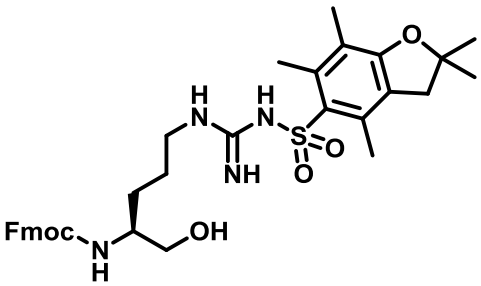
(S)-2,5-dioxopyrrolidin-1-yl (2-azido-3-methylpentyl)carbamate M4 : solid, 45%, ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm) = 5.46 (s, 1H), 3.58 – 3.40 (m, 1H), 3.17 – 3.04 (m, 2H), 2.83 (s, 4H), 1.77 – 1.58 (m), 1.61 – 1.44 (m), 0.99 – 0.92 (m, 6H); HRMS (ESI-TOFMS) *m/z* calcd for $C_{11}H_{21}N_6O_4$ $[M+NH_4]^+$ 301.1618, found 301.1618

The activated monomers containing Lys- (**M5**), Arg- (**M6**), Ser- (**M7**), Glu- (**M9**), Asn- (**M10**) and Gln- (**M11**) side chains were synthesized as previously reported[1], following path A, starting from the *N*-Fmoc protected amino acid for Lys- and Arg- derivatives and the *N*-Cbz protected (L) amino acid for Ser- and Glu- derivatives. The Asn type monomer was synthesized following the path B, starting from *N*-Fmoc protected amino acid. Fmoc deprotection was performed using DBU, as previously described[2]. Activated monomers **M5**, **M7**, **M9** and **M10** were previously described[3].

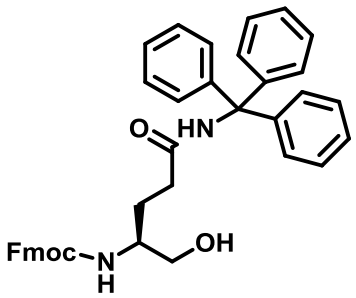


General procedure for N-Fmoc-protected alcohol formation

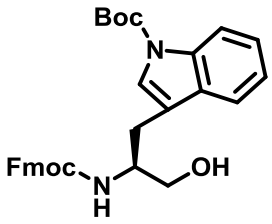
The N-Fmoc protected amino acid (10 mmol) was dissolved in THF under N_2 and cooled to -20°C. After addition of NMM (1.35 mL, 12.24 mmol) and IBCF (1.6 mL, 12.24 mmol) the mixture reaction was stirred at -20°C for 40 min. The resulting white suspension was filtered off and a solution of $NaBH_4$ (462 mg, 12.24 mmol) in water was added drop by drop to the filtrate cooled to -20°C. The reaction was left to react overnight. After the addition of water and THF removal under reduced pressure, the aqueous solution was extracted with EtOAc. The organic layers were combined and washed two times with 1M $KHSO_4$, two times with solution of saturated $NaHCO_3$, once with Brine solution and dried over $MgSO_4$. After concentration under vacuum, and silica gel flash chromatography, the desired pure product was furnished.



M6b: solid, 93%^{1}H NMR (300 MHz, $CDCl_3$) δ 7.75 (t, *J* = 7.1 Hz, 2H), 7.65 – 7.54 (m, 2H), 7.48 – 7.32 (m, 4H), 6.24 (s, 1H), 5.29 (t, *J* = 14.5 Hz, 1H), 4.42 (d, *J* = 6.1 Hz, 2H), 4.20 (d, *J* = 5.9 Hz, 2H), 3.63 (dd, *J* = 21.4, 10.6 Hz, 2H), 3.23 (d, *J* = 21.9 Hz, 2H), 2.95 (d, *J* = 7.0 Hz, 3H), 2.54 (dd, *J* = 20.5, 9.1 Hz, 6H), 2.27 (d, *J* = 20.3 Hz, 1H), 2.10 (d, *J* = 5.3 Hz, 3H), 1.48 – 1.40 (m, 6H).



M14b: White solid, 90% yield**,** ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.73 (t, *J* = 9.5 Hz, 2H), 7.61 – 7.54 (m, 2H), 7.38 (dd, *J* = 14.6, 7.1 Hz, 2H), 7.33 – 7.26 (m, 10H), 7.21 – 7.14 (m, 8H), 7.01 (s, 1H), 5.85 (d, *J* = 6.1 Hz, 1H), 4.39 (p, *J* = 10.4 Hz, 2H), 4.20 (dd, *J* = 12.3, 5.7 Hz, 2H), 2.67 – 2.39 (m, 2H), 2.07 (dd, *J* = 7.8, 4.2 Hz, 2H).



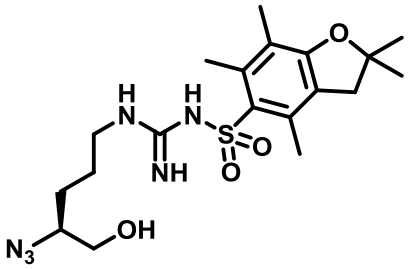
(S)-tert-butyl2-(2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-hydroxypropyl)-1Hindole-1-carboxylate M15b: Solid, 90%; ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm)= 8.13 (d, *J* = 8.0 Hz, 1H), 7.79 (d, *J* = 7.5 Hz, 2H), 7.69-7.25 (m, 10H), 5.21 (d, *J* = 6.6 Hz, 1H), 4.45 (d, *J* = 6.9 Hz, 2H), 4.23 (t, *J* = 6.8 Hz, 1H), 4.10 (brs,1H), 3.72 (d, *J* = 4.7 Hz, 2H), 3.02 (d, *J* = 6.0 Hz, 2H), 2.05 (brs, 1H), 1.67 (s, 9H).

General procedure for Fmoc removal

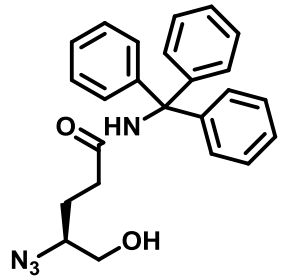
N-Fmoc protected amino alcohol was dissolved in EtOAc and DBU was added. The, mixture reaction was left to react at r.t. for 1h. EtOAc was removed under reduced pressure and the resulting crude product was directly engaged in the next step without any further purification.

General procedure for amine to azide conversion

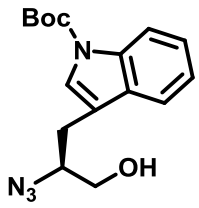
To a solution containing the TFA salt (10 mmol) in CH_3CN/H_2O (1:1), were successively added K_2CO_3 (2.04g, 12.84 mmol), imidazole-1-sulfonyl azide hydrochloride (N_3SO_2Imidazole.HCl) (2.46 g, 11.77 mmol) and $CuSO_4.5H_2O$ (15 mg, 0.1 mmol). The mixture was stirred at r.t. overnight. The reaction mixture was then filtered off to remove the remaining K_2CO_3. After concentration under reduced pressure, EtOAc was added. The organic phase was washed two times with $KHSO_4$ solution (1M), once with Brine, dried over magnesium sulfate and concentrated under reduced pressure. After silica gel flash chromatography, the desired pure product azides was furnished as pure compounds.



(S)-4-azido-5-hydroxypentyl (3-N-(2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl guanidine M6c: white solid, 76%; ^{1}H NMR (300 MHz, $CDCl_3$) δ 6.27 (s, 1H), 6.11 (s, 1H), 5.30 (s, 1H), 3.75 – 3.52 (m, 2H), 3.52 – 3.38 (m, 1H), 3.34 – 3.13 (m, 2H), 2.96 (s, 3H), 2.53 (d, *J* = 20.4 Hz, 6H), 2.15 – 2.07 (m, 3H), 1.61 (tt, *J* = 16.7, 8.5 Hz, 3H), 1.44 (d, *J* = 10.9 Hz, 6H).



(S)-4-azido-5-hydroxy-N-tritylpentanamide M11c: white solid, 91%; ^{1}H-RMN ($CDCl_3$, 300MHz): δ 7.16-7.42 (m, 15H) , 6.64 (s, 1H), , 3.92 (s, 1H), 3.39-3.62 (m, 2H), 2.36-2.54 (m, 2H), 1.71-1.82 (m, 2H), 1.25 (s, 1H),



(S)-tert-butyl2-(2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-hydroxypropyl)-1Hindole-1-carboxylate M15c: oil, 82%; ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm)= 8.13 (d, *J* = 8.0 Hz, 1H), 7.69-7.25 (m, 4H), 3.95-3.62 (m, 3H), 3.02 (m, 2H), 1.72 (s, 9H).

General procedure for conversion of alcohol into phthalimide- the Mitsunobu reaction

To a solution of PPh_3 in anhydrous THF at 0°C and under positive N_2 atmosphere pressure, were successively added DIAD and phthalimide. Azido alcohol dissolved in anhydrous THF was next added dropwise and the reaction mixture was stirred and allowed to reach the r.t. After 4 hours, the reaction was completed. THF was removed under reduced pressure and the crude material was directly engaged in the next step.

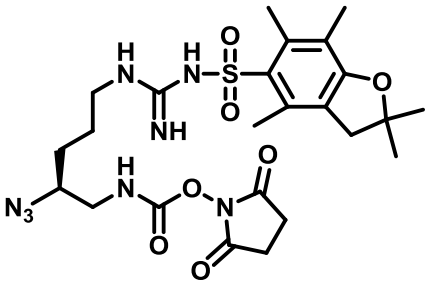
General procedure for azido amine synthesis

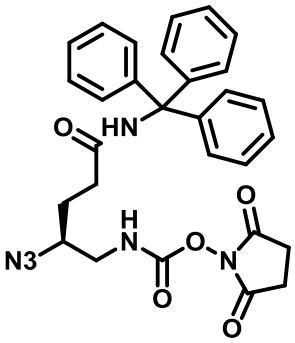
To a solution of azido phthalimide derived of Ile (1eq) in MeOH, hydrazine hydrate (3eq) was added. The reaction mixture was heat to reflux and the reaction was stirred 4h with apparition of a white precipitate. The reaction mixture was filtered off, washed with MeOH and the filtrate was concentrated under reduced pressure. The crude material was dissolved in EtOAc and the organic layer was washed twice with 1M HCl solution. The combined aqueous phases were again washed with EtOAc and were neutralized by addition of K_2CO_3 until pH8. The aqueous phase was finally extracted with DCM and the combined organic layers were dried over MgSO4, concentrated under reduced pressure but not until dryness to avoid the loss of azido Ile derivative which is

quite volatile. Resulting azido amine was directly engaged for activation step without further purification.

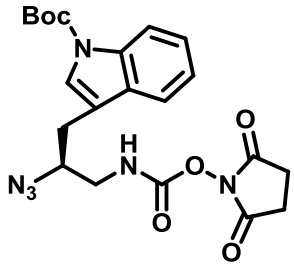
General procedure for the preparation of azide protected *O*-succinimidyl carbamate monomers
To a stirred suspension of disuccinimidyl carbonate (DSC) (1eq) in distilled DCM, was added drop by drop a solution of the derivative amine in DCM. The reaction mixture was stirred 4 h and was concentrated under reduced pressure. The crude material was dissolved in EtOAc and the organic phase was washed two times with 1M $KHSO_4$, once with Brine solution and dried over MgSO4 before to be concentrated under reduced pressure to furnish activated carbamate as solid.

(S)-2,5-dioxopyrrolidin-1-yl (2-azidopentyl-3-N-(2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)guanidine) carbamate M6: solid with 72% yield; ^{1}H NMR (300 MHz, $CDCl_3$) δ 6.12 (s, 2H), 5.91 (s, 1H), 5.15 – 5.01 (m, 1H), 3.65 (s, 2H), 3.44 (d, *J* = 6.9 Hz, 1H), 3.37 – 3.27 (m, 2,), 2.96 (s, 2H), 2.85 (s, 4H), 2.57 (d, *J* = 20.3 Hz, 6H), 2.10 (s, 3H), 1.63 (m, 2H), 1.46 (s, 6H), 1.32 – 1.24 (m, 2H).^{13}C NMR (75 MHz, $CDCl_3$) δ 170.45, 156.19, 152.28, 138.67, 132.61, 132.39, 132.28, 128.80, 124.92, 117.78, 86.69, 77.58, 77.36, 77.16, 76.74, 60.95, 44.60, 44.14, 43.33, 28.74, 28.14, 25.67, 25.28, 19.43, 18.10, 12.61. HRMS (ESI-TOFMS) *m/z* calcd for $C_{24}H_{34}N_8O_7S$ $[M+H]^+$ 578.2271, found 579.2327

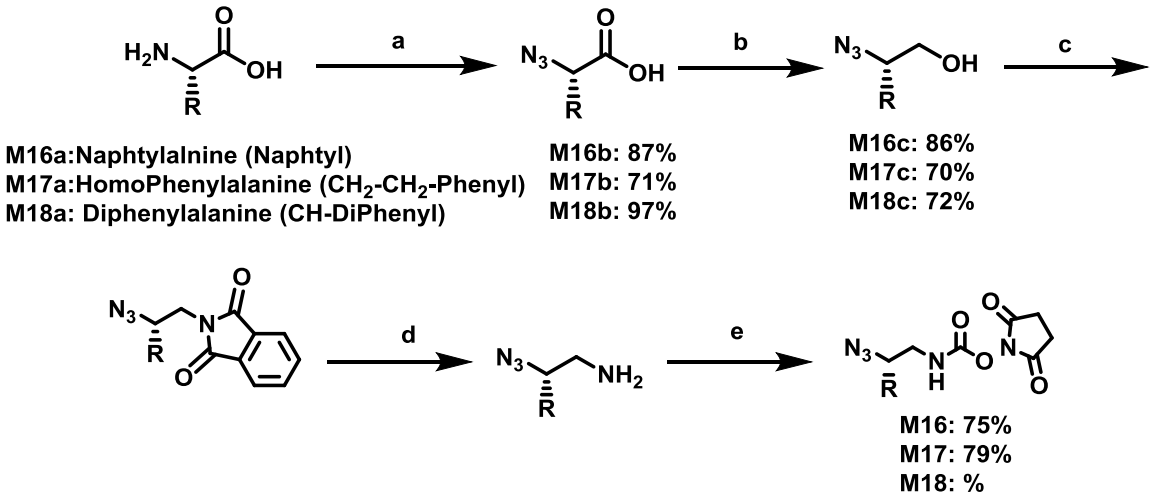




(*S*)-2,5-dioxopyrrolidin-1-yl(-4-azido-N-tritylpentanamide) carbamate M11 : solid, 75% yield ; ^{1}H-RMN ($CDCl_3$, 300 MHz): δ (ppm) = 7.16-7.42 (m, 15H), 6.7 (s, 1H), 5.6 (s, 1H), 3.5 (m, 1H), 3.2 (m), 2.8 (s, 4H), 2.4 (t, 2H), 1.8 (q, 2H); ^{13}C-RMN ($CDCl_3$, 75 MHz): δ= 168.8, 143.5, 127.1, 126.18, 75.9, 69.5, 59.2, 30.6, 24.4, 23.8.



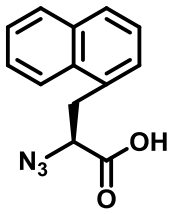
(S)-tert-butyl 2-(2-azido-3-((((2,5-dioxopyrrolidin-1-yl)oxy)carbonyl)amino)propil)-1Hindole-1-carboxylate M15[1]:solid, 76% yield; ^{1}H NMR (300 MHz, $CDCl_3$) δ (ppm)= 8.13 (s, 1H), 7.62 – 7.49 (m, 2H), 7.46 – 7.31 (m, 2H), 6.33 (d, *J* = 28.6 Hz, 1H), 3.70 – 3.53 (m, 2H), 3.45 (dt, *J* = 45.8, 23.1 Hz, 1H), 3.07 – 3.03 (m, 2H), 2.80 (s, 4H), 1.67 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.08, 124.90, 123.90, 122.85, 118.84, 115.61, 77.58, 77.58, 77.36, 77.16, 76.74, 52.80, 46.48, 31.57, 28.35, 25.60. HRMS (ESI-TOFMS) *m/z* calcd for $C_{21}H_{24}N_6O_6Na$ $[M+Na]^+$ 479.1649, found 479.1643



a: N_3SO_2-Im,HCl, K_2CO_3, $CuSO_4$:5H_2O, CH_3CN:H_2O(1:1), r.t.; b:1) NMM, IBCF, THF, -20°C; 2) $NaBH_4$, H_2O; c:PPh_3, Phthalimide, DIAD, THF, r.t.; d: N_2H_4.H_2O, MeOH, 70°C; e: DSC, CH_2Cl_2, r.t.

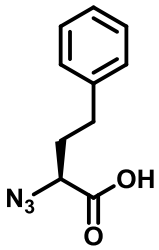
General procedure for amine to azide conversion

To a solution containing the *N*-free amino acid (10 mmol) in CH_3CN/H_2O (1:1), were successively added K_2CO_3 (2.04g, 12.84 mmol), imidazole-1-sulfonyl azide hydrochloride (N_3SO_2Imidazole.HCl) (2.46 g, 11.77 mmol) and $CuSO_4.5H_2O$ (15 mg, 0.1 mmol). The mixture was stirred at r.t. overnight. The reaction mixture was then filtered off to remove the remaining K_2CO_3. After concentration under reduced pressure, EtOAc was added. The organic phase was washed two times with $KHSO_4$ solution (1M), once with Brine, dried over magnesium sulfate and concentrated under reduced pressure. After silica gel flash chromatography, the desired pure product was furnished as pure compounds.



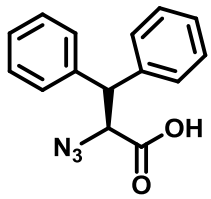
(S)-2-azido-3-(naphthalen-1-yl)propanoic acid

M16b: brown oil; 87% yield; ^{1}H-RMN ($CDCl_3$, 300 MHz): δ 3.18 (dd, *J* = 14.4, 9.5 Hz, 1H), 3.57-3.70 (m, 2H), 4.08-4.21 (m, 2H), 7.28-7.87 (m, 7H)



(S)-2-azido-4-phenylbutanoic acid

M17b: yellow oil, 71%; ^{1}H-RMN ($CDCl_3$, 300MHz): δ 2.04-2.28 ppm (m, 2H), 2.68-2.90 (m, 2H), 3.9 ppm (dd, J=8.8, 4.9 Hz, 1H), 7.21-7.35 (m, 5H), 8.56 (s, 1H)

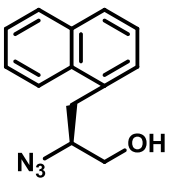


(S)-2-azido-3-diphenylpropanoic acid

M18b: 97%^{1}H NMR (300 MHz, $CDCl_3$) δ 7.67 (dd, *J* = 14.0, 7.3 Hz, 5H), 7.35 – 7.15 (m, 5H), 4.42 (d, *J* = 9.3 Hz, 1H), 2.30 (s, 1H).

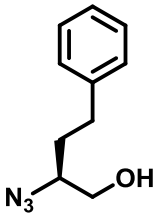
General procedure for azide alcohol formation

The azide amino acid (2g, 1 eq) was dissolved in THF under N_2 and cooled to -20°C. After addition of NMM (1.2 eq) and IBCF (1.2 eq) the mixture reaction was stirred at -20°C for 40 min. The resulting white suspension was filtered off and a solution of $NaBH_4$ (1.2 eq) in water was added drop by drop to the filtrate cooled to -20°C. The reaction was left to react overnight. After the addition of water and THF removal under reduced pressure, the aqueous solution was extracted with EtOAc. The organic layers were combined and washed two times with 1M $KHSO_4$, two times with solution of saturated $NaHCO_3$, once with Brine solution and dried over $MgSO_4$. After concentration under vacuum, and silica gel flash chromatography, the desired pure product was furnished.



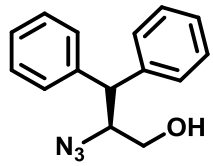
(S)-2-azido-3-(naphthalen-1-yl)propan-1-ol

M16c: brown oil**,** 86%; ^{1}H-RMN ($CDCl_3$, 300 MHz): δ 3.35 (qd, *J* = 14.1, 7.1 Hz, 2H), 3.56-3.82 (m, 2H), 3.86-3.94 (m, 1H), 7.42-8.09 (m, 7H)



(S)-2-azido-4-phenylbutan-1-ol

M17c: yellow oil, 70%; ^{1}H-RMN ($CDCl_3$, 300 MHz): δ 1.79-1.91 (m, 2H), 2.63-2.94 (m, 2H), 3.74 (ddd, *J* = 39.2, 12.9, 7.2 Hz, 2H), 3.96 (d, *J* = 6.7 Hz, 1H), 7.26-7.35 (m, 5H)



(S)-2-azido-3-diphenylpropan-1-ol

M18c: ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.00 – 7.64 (m, 10H), 4.94 – 4.77 (m, 2H), 4.42 (d, *J* = 9.3 Hz, 1H), ^{1}H NMR (300 MHz, $CDCl_3$), 2.30 (s, 1H).

General procedure for conversion of alcohol into phthalimide- the Mitsunobu reaction

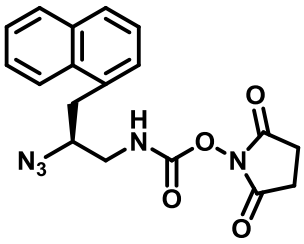
To a solution of PPh_3 in anhydrous THF at 0°C and under positive N_2 atmosphere pressure, were successively added DIAD and phthalimide. Azido alcohol dissolved in anhydrous THF was next added dropwise and the reaction mixture was stirred and allowed to reach the r.t. After 4 hours, the reaction was completed. THF was removed under reduced pressure and the crude material was directly engaged in the next step.

General procedure for azido amine synthesis

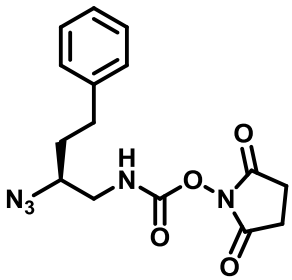
To a solution of azido phthalimide derived of Ile (1eq) in MeOH, hydrazine hydrate (3eq) was added. The reaction mixture was heat to reflux and the reaction was stirred 4h with apparition of a white precipitate. The reaction mixture was filtered off, washed with MeOH and the filtrate was concentrated under reduced pressure. The crude material was dissolved in EtOAc and the organic layer was washed twice with 1M HCl solution. The combined aqueous phases were again washed with EtOAc and were neutralized by addition of K_2CO_3 until pH8. The aqueous phase was finally extracted with DCM and the combined organic layers were dried over MgSO4, concentrated under reduced pressure but not until dryness to avoid the loss of azido Ile derivative which is quite volatile. Resulting azido amine was directly engaged for activation step without further purification.

General procedure for the preparation of azide protected *O*-succinimidyl carbamate monomers

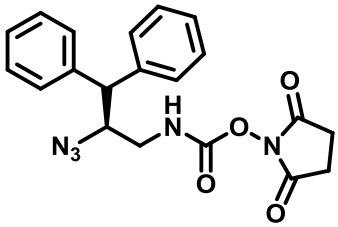
To a stirred suspension of disuccinimidyl carbonate (DSC) (1eq) in distilled DCM, was added drop by drop a solution of the derivative amine in DCM. The reaction mixture was stirred 4 h and was concentrated under reduced pressure. The crude material was dissolved in EtOAc and the organic phase was washed two times with 1M $KHSO_4$, once with Brine solution and dried over MgSO4 before to be concentrated under reduced pressure to furnish activated carbamate as solid.



(S)-2,5-dioxopyrrolidin-1-yl (2-azido-3-naphten-1-ylpropan)carbamate M16: solid, 75% yield; ^{1}H-RMN ($CDCl_3$, 300MHz): δ (ppm) =, 7.42-8.07 (m, 7H), 5.52 (s, 1H), 3.92-4.02 (m, 1H), 3.5 (ddd, *J* = 13.9, 6.5, 4.0 Hz, 2H), 3.30 (m, 2H), 2.82 (s, 4H); ^{13}C-RMN ($CDCl_3$, 75 MHz): δ= 170, 151.2, 134, 132.2, 128, 127.2, 126.5, 125.8, 122.5, 61.8, 44.8, 35, 25.1; HRMS (ESI-TOFMS) *m/z* calcd for $C_{18}H_{17}N_5O_4Na$ $[M+Na]^+$ 390.1173, found 390.1178.



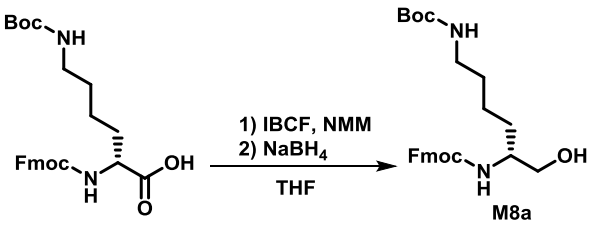
(S)-2,5-dioxopyrrolidin-1-yl (2-azido-4-phenylbutan) carbamate M17 : White solid, 79%; ^{1}H-RMN ($CDCl_3$, 300MHz): δ (ppm) = 7.18-7.34 (m, 5H), 5.51 (s, 1H), 3.39-3.58 (m, 2H), 3.17-3.28 (m, 1H), 2.82 (s, 4H), 2.73 (dd, *J* = 12.0, 4.4 Hz, 2H), 1.87 (dd, *J* = 15.9, 8.1 Hz, 2H); ^{13}C-RMN ($CDCl_3$, 75 MHz): δ= 170.8, 151.5, 141.2, 129, 127.2, 61.8, 46, 32.5, 31.8, 25.1; HRMS (ESI-TOFMS) *m/z* calcd for $C_{15}H_{17}N_5O_4Na$ $[M+Na]^+$ 354.1173, found 354.1180.



(S)-2,5-dioxopyrrolidin-1-yl (2-azido-3-diphenylpropan) carbamate M18: solid, 57% yield, ^{1}H NMR (300 MHz, DMSO) δ 7.47 (dd, *J* = 14.0, 7.3 Hz, 4H), 7.35 – 7.15 (m, 6H), 4.53 (t, *J* = 8.4 Hz, 1H), 4.11 (d, *J* = 10.7 Hz, 2H), 3.19 (dd, *J* = 10.5, 2.8 Hz, 1H), 3.10 – 2.96 (m, 1H), 2.77 (s, 4H);^{13}C NMR (75 MHz, $CDCl_3$) δ 169.80, 140.52, 129.07, 128.14, 127.47, 65.69, 55.01, 45.19, 30.59, 25.47.HRMS (ESI-TOFMS) *m/z* calcd for $C_{20}H_{23}N_5O_4$ $[M+H]^+$ 411.1775 found 411.1800

2. Preparation of azide protected O-succinimidyl carbamate monomer M8 encompassing the Lys-type side chain in □ position with an inversion of configuration

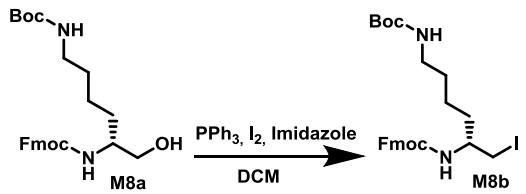
(*S*)-(9*H*-fluoren-9-yl)methyl(1-hydroxy-5-*N*-tert-butoxycarbonylaminohex-2-yl)carbamate M8a:



In a round bottom flask, *N*-Fmoc protected amino acid Fmoc-D-Lys(Boc)-OH (5 g, 10.7 mmol) was dissolved in anhydrous THF under positive N_2 atmosphere and cooled down to -20°C[4]. Isobutyl chloroformate (1.67 mL, 12.8 mmol) and NMM (1.4 mL, 12.8 mmol) were added and the reaction was stirred at -20°C for 25 min. After filtration of the resulting NMM salt, the filtrate was cooled down to 0°C and a solution of $NaBH_4$ (485 mg, 12.8 mmol) in water was slowly added. The reaction mixture was let under stirring for 10 min. The crude mixture was then concentrated under reduced pressure. A 1 M aqueous solution of $KHSO_4$ was next added then extracted twice with EtOAc. The organic layers were combined and washed twice with a saturated aqueous solution of $NaHCO_3$, once with brine and dried over sodium sulfate before to be concentrated under vacuum. Silica gel flash chromatography was performed in a mixture of cyclohexane / EtOAc (7: 3; v/v) to give pure **M8a** as a white solid in 92% yield.

M8a: ^{1}H-NMR ($CDCl_3$, 300MHz): δ 1.43 (s, 9H), 1.59 (d, *J* = 28.7 Hz, 2H), 2.04 (d, *J* = 2.9 Hz, 1H) 3.02-3.12 (m, 1H), 3.17 (d, *J* = 5.9 Hz, 2H), 3.63 (s, 2H), 4.21 (t, *J* = 6.7 Hz, 2H), 4.42 (d, *J* = 5.2 Hz, 4H), 4.56 (s, 1H), 5.02 (s, 1H), 7.29-7.82 (m, 8H)

(*S*)-(9*H*-fluoren-9-yl)methyl(1-iodo-5-*N*-tert-butoxycarbonylaminohex-2-yl)carbamate M8b:

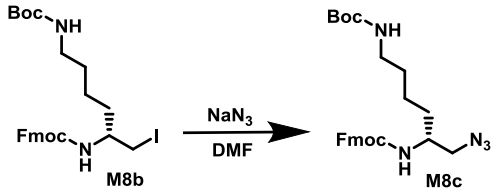


M8a was converted into iodine **M8b** according to the procedure described by Sureshbabu et al.[5]. To a solution of triphenylphosphine (2.9 g, 25.3 mmol) in anhydrous CH_2Cl_2 under N_2 atmosphere, at O°C, were successively added Imidazole (4.26g, 42.2 mmol) and Iodine (6.42g, 25.3 mmol). *N*-Fmoc protected amino alcohol derivative **M8a** was dissolved in anhydrous CH_2Cl_2 and added dropwise to the reaction mixture. The reaction mixture was allowed to reach room temperature and stirred overnight. After evaporation of CH_2Cl_2 under reduced pressure, the crude mixture was dissolved in EtOAc

and successively washed three times with a 0.5 M aqueous solution of $Na_2S_2O_3$, twice with a 1M aqueous solution of $KHSO_4$, once with brine and dried over sodium sulfate. The crude product was concentrated under reduced pressure. Silica gel flash chromatography was performed in cyclohexane/EtOAc (9 : 1, v/v) to obtain pure **M8b** as a white solid in 90% yield.

M8b: ^{1}H-NMR ($CDCl_3$, 300MHz): δ 1.36 (d, *J* = 7.6 Hz, 2H), 1.41 (s, 9H), 1.53 (d, *J* = 9.3 Hz, 4H), 3.12 (d, *J* = 5.5 Hz, 2H), 3.29 (d, *J* = 6.5 Hz, 1H), 3.42 (d, *J* = 8.9 Hz, 2H), 4.23 (t, *J* = 6.7 Hz, 1H), 4.31-4.38 (m, 2H), 4.51 (d, *J* = 10.3 Hz, 1H), 4.87 (d, *J* = 4.5 Hz, 1H), 7.30-7.79 (m, 8H)

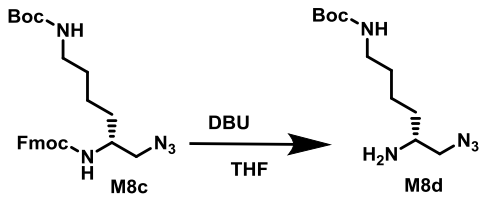
(*S*)-(9*H*-fluoren-9-yl)methyl(1-azido-5-*N*-tert-butoxycarbonylaminohex-2-yl)carbamate M8c:



The *N*-Fmoc protected amino iodine derivative **M8b** (3.92 g, 6.95 mmol) was dissolved in DMF. NaN_3 (1.36 g, 20.8 mmol) was added and the reaction mixture was stirred at room temperature overnight. EtOAc was added to this reaction mixture and the organic layer was washed six times with H_2O, three times with 1M aqueous solution of $KHSO_4$ solution, twice with brine and dried over sodium sulfate. The crude was concentrated under reduced pressure before to be purified with silica gel flash chromatography in cyclohexane/EtOAc (9: 1, v/v) to obtained **M8c** as a white solid, 53% yield.

M8c: ^{1}H-NMR ($CDCl_3$, 300MHz): δ 0.72 – 0.92 (m, 2H) 1.42 (s, 9H), 2.15 (s, 1H), 3.11 (d, *J* = 6.4 Hz, 2H), 3.41 (tt, *J* = 12.4, 6.1 Hz, 2H), 3.76 (s, 1H), 4.21 (t, *J* = 6.6 Hz, 1H), 4.41 (dd, *J* = 17.7, 8.9 Hz, 2H), 4.53 (s, 1H), 4.83 (d, *J* = 7.9 Hz, 1H), 6.93 (d, *J* = 14.3 Hz, 3H), 7.87 – 7.28 (m, 8H)

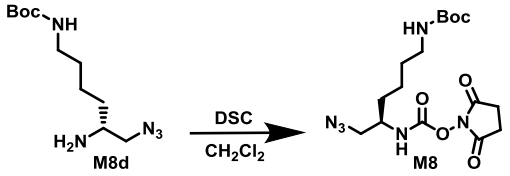
(R)-tert-butoxycarbonyl-6-azidohexyl-5-amine M8d:



The *N*-Fmoc protected azide derivative **M8c** was dissolved in EtOAc and DBU (3 eq, 1.6 ml, 11 mmol) was added. The mixture reaction was stirred at room temperature for 10 minutes. Precipitate was filtered and washed twice with EtOAc. The product was concentrated under reduced pressure and **M8d** was obtained as a white solid.

M8d: ^{1}H-NMR ($CDCl_3$, 300MHz): δ 1.43 (s, 9H), 1.51 (d, *J* = 6.2 Hz, 4H), 1.69 – 1.63 (m, 2H), 3.15 – 3.08 (m, 2H), 3.18 (dd, *J* = 10.1, 4.2 Hz, 2H), 3.39 – 3.27 (m, 2H), 4.63 (s, 1H), 5.29 (s, 1H)

(S)-2,5-dioxopyrrolidin-1-yl (2-azido-5-N-ter-butoxycarbonylaminohexyl) carbamate M8 :



To a stirred suspension of disuccinimidyl carbonate (795 mg, 3.1 mmol) in freshly distilled CH_2Cl_2 (50 mL) was added drop by drop a solution of **M8d** (3.1 mmol) in DCM (50 mL). After 4 hours, the reaction mixture was concentrated under reduced pressure and the crude material was dissolved in EtOAc (70 mL). The organic layer was washed twice with 1M aqueous solution of $KHSO_4$, once with brine, dried over Na_2SO_4 and concentrated under reduced pressure to furnish an oily product. The activated carbamate **M8** was successfully precipitated in a mixture of pentane/Et_2O (3:7, v/v) and obtained as a white solid with 75 % yield.

M8: overall yield of 30%. ^{1}H-NMR ($CDCl_3$, 300MHz): δ 1.41 (s, 9H), 1.50-1.58 (m, 2H), 1.6 (s, 2H), 1.64-1.7 (m, 2H), 2.88 (s, 4H), 3.09 (dt, *J* = 13.3, 8.3 Hz, 2H), 3.56 – 3.44 (m, 2H), 3.74 (s, 1H) 4.62 (s, 1H), 5.54 (s, 1H); ^{13}C-RMN ($CDCl_3$, 75 MHz): δ 23, 24.8, 27, 27.5, 30, 40, 51.2, 54.8, 79.2, 152, 156, 170; HRMS (ESI-TOFMS) *m/z* for $C_{16}H_{26}N_6O_6Na$ $[M+Na]^+$ 421.1812, found 421.1809.

3. Preparation of Fmoc-N-protected O-succinimidyl carbamates monomers

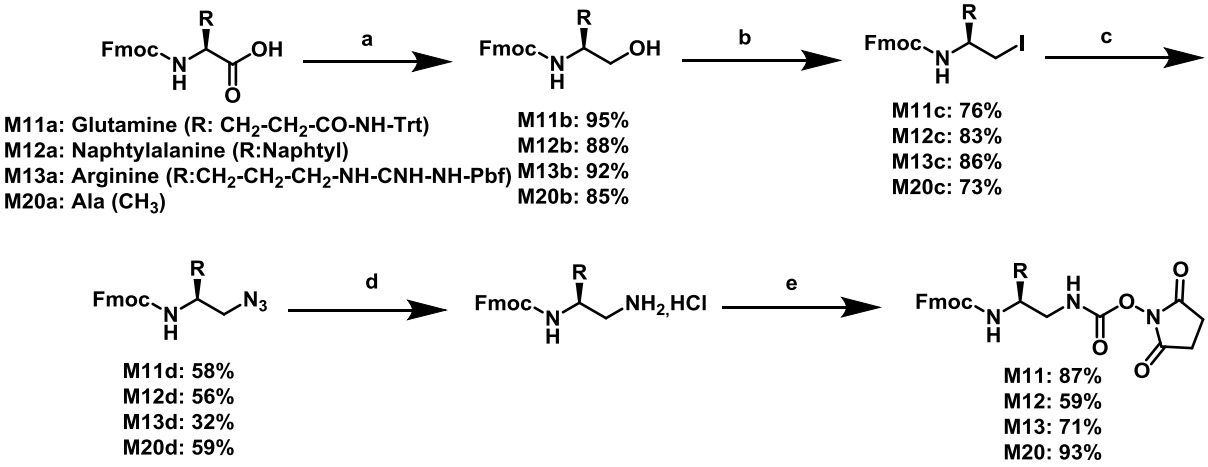
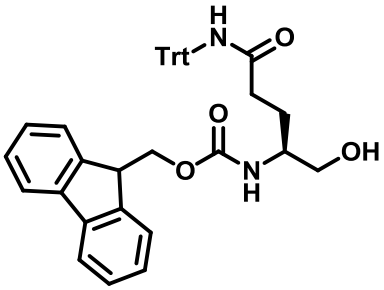


Figure 88: Multi-step synthesis of Fmoc-N-protected

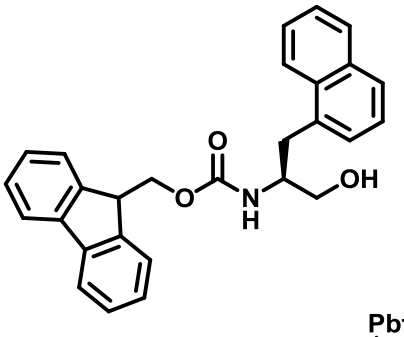
Fmoc-Gln(Trt)^{u}-OSu (**M11**), Fmoc-Nal^{u}-OSu (**M12),** Fmoc-Arg^{u}(Pbf)-OSu (**M13)** Fmoc-Ala^{u}-OSu (**M20)** were synthesized through the multi-step synthesis describe above.

<u>General procedure for N-Fmoc protected amino alcohol formation</u>

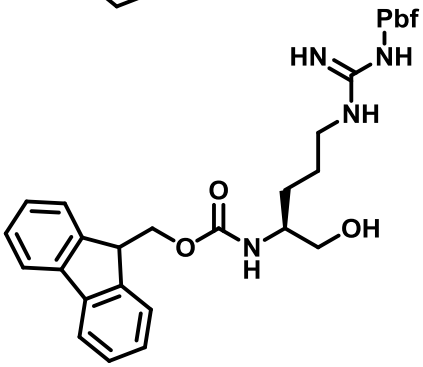
In a round bottom flask containing *N*-Fmoc protected amino acid (1 eq) dissolved in anhydrous THF cooled to -20°C, under N_2 atmosphere, NMM (1.2 eq) and IBCF (1.2 eq) were added. The reaction mixture was let stirring for 40 min. The resulted white precipitate was filtered off. $NaBH_4$ (30 mmol, 1 g, 1.2 eq) in solution in water was added dropwise to the filtrate at 0°C. The reaction was left stirring 10 min, quenched with water and then concentrated under reduced pressure. After adding ethyl acetate, organic layer was washed twice with $KHSO_4$ solution (1M) and twice with saturated solution of $NaHCO_3$, once with brine. After drying over sodium sulfate, the organic layer was concentrated. The product was purified on silica gel flash chromatography in EtOAc/ cyclohexane (from 1/9 to 6/4).



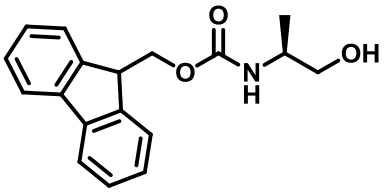
(S)-(9H-fluoren-9-yl)methyl (1-hydroxy N-tritylpentanamide-2-yl)carbamate M11a : White solid, 95% yield**,** ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.73 (t, *J* = 9.5 Hz, 2H), 7.61 – 7.54 (m, 2H), 7.38 (dd, *J* = 14.6, 7.1 Hz, 2H), 7.33 – 7.26 (m, 10H), 7.21 – 7.14 (m, 8H), 7.01 (s, 1H), 5.85 (d, *J* = 6.1 Hz, 1H), 4.39 (p, *J* = 10.4 Hz, 2H), 4.20 (dd, *J* = 12.3, 5.7 Hz, 2H), 2.67 – 2.39 (m, 2H), 2.07 (dd, *J* = 7.8, 4.2 Hz, 2H).



(S)-(9H-fluoren-9-yl)methyl (1-hydroxy-3-naphtylpropan-2-yl)carbamate M12a: White solid, 88% yield**,** ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.19 (t, *J* = 17.0 Hz, 1H), 7.95 – 7.83 (m, 1H), 7.77 (d, *J* = 7.7 Hz, 2H), 7.53 (dt, *J* = 15.1, 10.8 Hz, 4H), 7.43 – 7.29 (m, 4H), 5.11 (s, 1H), 4.41 (d, *J* = 6.8 Hz, 2H), 4.18 (d, *J* = 11.3 Hz, 1H), 3.67 (d, *J* = 15.7 Hz, 2H), 3.41 (s, 1H), 3.30 (dd, *J* = 13.8, 8.2 Hz, 1H).



(S)-(9H-fluoren-9-yl)methyl (1-hydroxy-3-N-(2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)guanidine)-2-yl)carbamate M13a:white solid**,** 92%1;H NMR (300 MHz, $CDCl_3$) δ 7.75 (t, *J* = 7.1 Hz, 2H), 7.65 – 7.54 (m, 2H), 7.48 – 7.32 (m, 4H), 6.24 (s, 1H), 5.29 (t, *J* = 14.5 Hz, 1H), 4.42 (d, *J* = 6.1 Hz, 2H), 4.20 (d, *J* = 5.9 Hz, 2H), 3.63 (dd, *J* = 21.4, 10.6 Hz, 2H), 3.23 (d, *J* = 21.9 Hz, 2H), 2.95 (d, *J* = 7.0 Hz, 3H), 2.54 (dd, *J* = 20.5, 9.1 Hz, 6H), 2.27 (d, *J* = 20.3 Hz, 1H), 2.10 (d, *J* = 5.3 Hz, 3H), 1.48 – 1.40 (m, 6H).

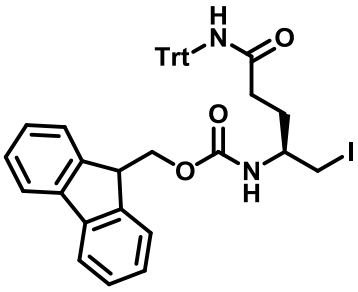


(S)-(9H-fluoren-9-yl)methyl (1-hydroxypropan-2-yl)carbamate M20a: white solid, 85 % yield; ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.4 Hz, 2H), 7.59 (d, *J* = 7.4 Hz, 2H), 7.46 – 7.29 (m, 4H), 4.83 (s, 1H), 4.44 (d, *J* = 6.3 Hz, 3H), 4.22 (t, *J* = 6.6 Hz, 2H), 3.82 (s, 1H), 3.65 (s, 1H), 3.54 (s, 2H).

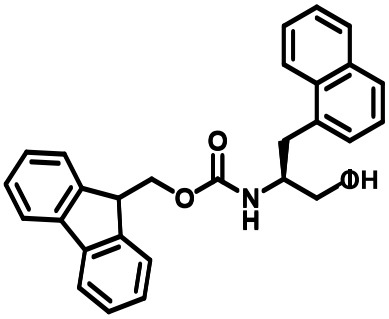
General procedure for conversion of alcohol into iodine derivative

Iodine derivatives were obtained from the corresponding alcohols according to the procedure of Sureshbabu et al.[5]. To a solution of triphenylphosphine (3 eq) in anhydrous CH_2Cl_2 under N_2 atmosphere, at O°C, were successively added Imidazole (5 eq) and Iodine (3 eq). N-Fmoc protected amino alcohol derivative (1 eq) was dissolved in anhydrous CH_2Cl_2 and added dropwise into the reaction mixture. The reaction mixture was allowed to reach the room temperature and stirred overnight. After evaporation of CH_2Cl_2 under reduced pressure. The crude was dissolved in EtOAc and washed three times with a 0.5 M aqueous solution of $Na_2S_2O_3$, twice with a 1M

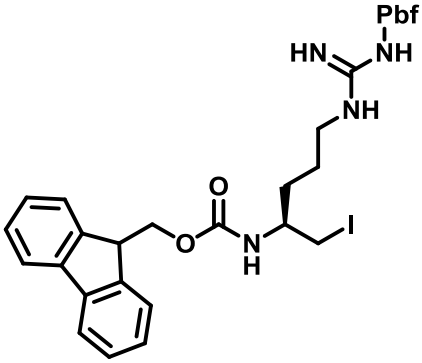
aqueous solution of $KHSO_4$, once with brine and dried over sodium sulfate. The crude product was concentrated under reduced pressure. Silica gel flash chromatography was performed in cyclohexane /EtOAc (9 : 1, v/v) to obtain pure **M11b** (95%), **M12b** (88%), **M13b** (92%), and **M19b** (85%).



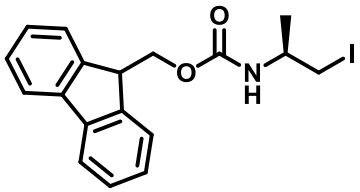
(S)-(9H-fluoren-9-yl)methyl (1-iodo N-tritylpentanamide-2-yl)carbamate M11b : white solid, 83% yield, ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.6 Hz, 2H), 7.63 – 7.56 (m, 2H), 7.48 – 7.28 (m, 11H), 7.25 – 7.16 (m, 8H), 6.73 (s, 1H), 5.04 (d, *J* = 8.0 Hz, 1H), 4.51 (dd, *J* = 10.5, 6.9 Hz, 1H), 4.37 (t, *J* = 8.4 Hz, 1H), 4.20 (dd, *J* = 18.5, 12.0 Hz, 1H), 3.47 (s, 1H), 3.29 (dt, *J* = 10.3, 7.7 Hz, 2H), 2.34 (t, *J* = 6.5 Hz, 2H), 1.87 (s, 2H).



(S)-(9H-fluoren-9-yl)methyl (1-iodo-3-naphtylpropan-2-yl)carbamate M12b: white solid, 83% yield, ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.23 (d, *J* = 8.0 Hz, 2H), 7.94 – 7.80 (m, 2H), 7.82 – 7.69 (m, 4H), 7.57 (dd, *J* = 15.4, 7.9 Hz, 4H), 7.49 (dd, *J* = 14.0, 6.9 Hz, 3H), 7.42 (dd, *J* = 15.3, 7.4 Hz, 4H), 7.40 – 7.28 (m, 3H), 5.02 (s, 1H), 4.49 – 4.34 (m, 2H), 4.21 (dd, *J* = 15.6, 9.9 Hz, 1H), 3.85 (s, 1H), 3.54 – 3.34 (m, 2H), 3.17 (dd, *J* = 18.0, 9.4 Hz, 2H).



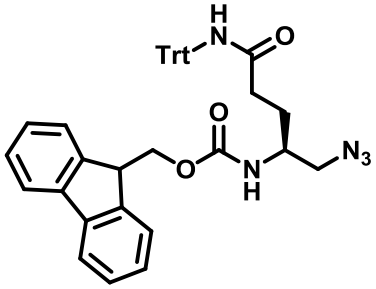
(S)-(9H-fluoren-9-yl)methyl (1-iodopentyl-3-N-(2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)guanidine)-2-yl)carbamate M13b: solid, 86% yield, ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.78 – 7.61 (m, 4H), 7.60 – 7.36 (m, 4H), 6.28 – 6.08 (m, 1H), 5.07 (d, *J* = 9.2 Hz, 1H), 4.45 (d, *J* = 5.6 Hz, 2H), 4.20 (t, *J* = 6.8 Hz, 1H), 3.74 (s, 1H), 3.48 – 2.73 (m, 1H), 2.99 (s, 3H), 2.95 (s 2H), 2.92 (s, 3H), 2.63 (s, 2H), 2.56 (s, 2H), 2.20 – 1.82 (m, 3H), 1.69 (s, 6H), 0.95 – 0.83 (m, 2H).



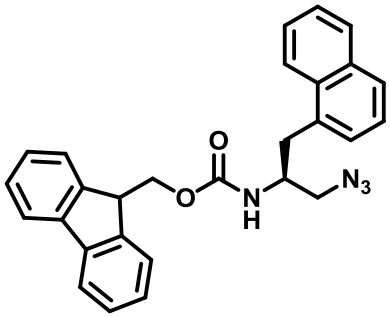
(S)-(9H-fluoren-9-yl)methyl (1-iodopropan-2-yl)carbamateM20b: white solid, 73% yield. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.80 (d, *J* = 7.4 Hz, 2H), 7.65 (t, *J* = 8.0 Hz, 2H), 7.51 – 7.32 (m, 4H), 4.86 (s, 1H), 4.44 (dt, *J* = 16.6, 10.0 Hz, 3H), 4.25 (dd, *J* = 13.0, 6.2 Hz, 2H), 3.63 (s, 1H), 3.49 (s, 2H), 1.48 (s, 1H).

General procedure for conversion of iodine into azide

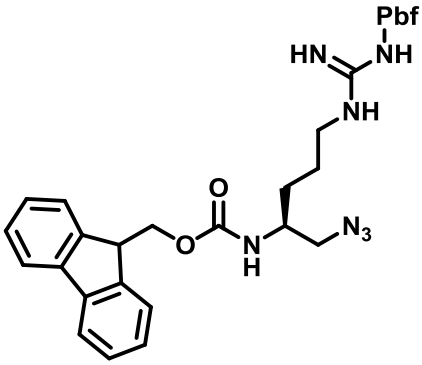
The *N*-Fmoc protected amino iodine derivative (1 eq) was dissolved in DMF. NaN_3 (3 eq) was added into the reaction mixture that was stirred at room temperature overnight. EtOAc was added to the reaction mixture that was washed with H_2O six times, three times with 1M $KHSO_4$ solution, twice with brine and dried over sodium sulfate. The crude was concentrated under reduced vacuum before to be purified with silica gel flash chromatography in cyclohexane/EtOAc (9: 1, v/v).



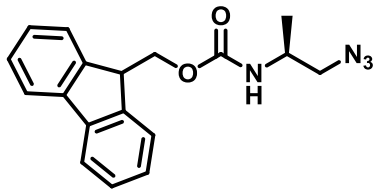
(S)-(9H-fluoren-9-yl)methyl (1-azido N-tritylpentanamide-2-yl)carbamate M11c: White solid, 56% yield, ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.74 (d, *J* = 7.5 Hz, 2H), 7.62 – 7.51 (m, 3H), 7.43 – 7.28 (m, 9H), 7.23 – 7.12 (m, 9H), 6.73 (s, 1H), 4.99 (d, *J* = 8.9 Hz, 1H), 4.54 – 4.33 (m, 3H), 4.20 (t, *J* = 6.5 Hz, 1H), 3.73 (d, *J* = 4.7 Hz, 1H), 3.36 (s, 2H), 2.30 (dd, *J* = 14.8, 7.9 Hz, 2H), 1.75 (d, *J* = 44.6 Hz, 2H).



(S)-(9H-fluoren-9-yl)methyl (1-azido-3-naphtylpropan-2-yl)carbamateM12c : White solid, 56% yield, ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.18 (d, *J* = 7.9 Hz, 1H), 7.93 – 7.67 (m, 4H), 7.61 – 7.27 (m, 10H), 5.02 (d, *J* = 7.1 Hz, 1H), 4.43 (d, *J* = 6.9 Hz, 2H), 4.22 (d, *J* = 6.2 Hz, 2H), 3.41 (t, *J* = 13.6 Hz, 3H), 3.22 (dd, *J* = 13.9, 8.6 Hz, 1H).



(S)-(9H-fluoren-9-yl)methyl (1-azidopentyl-3-N-(2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)guanidine)-2-yl)carbamate M13c :^{1}H NMR (300 MHz, $CDCl_3$) δ 7.63 (dd, *J* = 69.1, 53.9 Hz, 4H), 7.17 (dt, *J* = 20.5, 5.6 Hz, 4H), 6.91 (s, 1H), 6.52 (s, 1H), 4.85 (s, 2H), 3.53 (s, 3H), 2.27 (d, *J* = 60.3 Hz, 6H), 2.17 (s, 4H), 1.33 (s, 2H), 1.27 (d, *J* = 8.7 Hz, 7H), 1.03 (s, 2H), 0.92 – 0.74 (m, 6H).



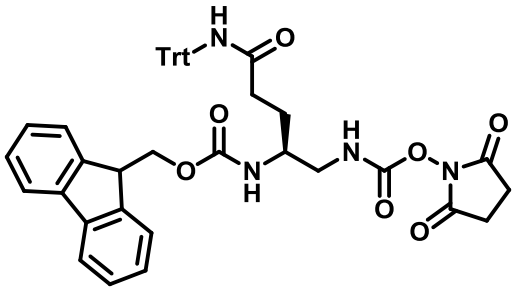
(S)-(9H-fluoren-9-yl)methyl (1-azido-propan-2-yl)carbamate M20c: White solid, 59% yield. ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.75-7.62 (m, 4H), 7.54 – 7.38 (m, 4H), 4.84 (s, 1H), 4.68 (d, *J* = 6.2 Hz, 3H), 4.12 (s, 2H), 3.62 (s, 1H), 3.41 (d, *J* = 6.6 Hz, 2H), 2.55- (s, 1H).

General procedure for conversion of azide into amine

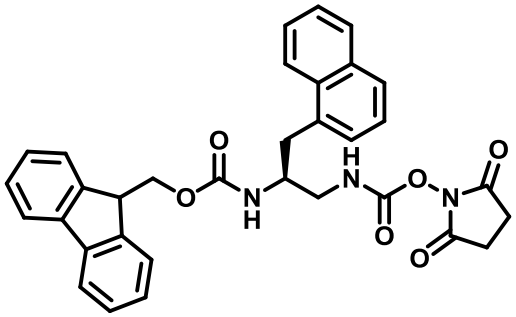
In a round bottom flask containing *N*-Fmoc protected amino azide derivative (1 eq) dissolved in a mixture MeOH/ chloroform (9:1; v/v), Pd/C (10% in mass, 300 mg) was gently added under N_2. The reaction mixture was stirred overnight at r.t. under H_2 atmosphere. Pd/C was filtered carefully over Millipore paper filter and the filtrate was concentrated under reduced pressure. The product was precipitated in ice bath with a mixture of EtOAc/Et_2O (5:5, v/v). The precipitate was filtered, dried and engaged to the following step without any purification.

General procedure for the preparation of *N*-Fmoc protected activated carbamate monomers

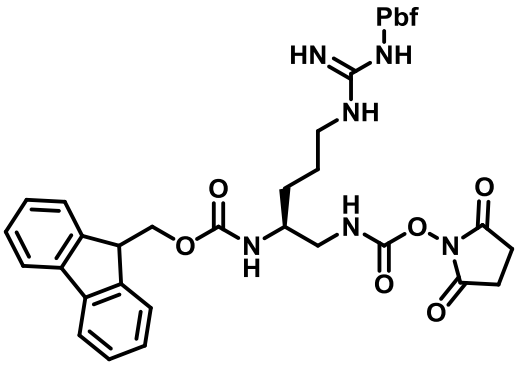
In a round bottom flask containing N-Fmoc amine derivative (1 eq) dissolved in anhydrous CH_2Cl_2, under N_2 atmosphere, at 0°C, NMM (1.2 eq) was added. The reaction was let under magnetic stirring for about 5 min and disuccinimidyl carbonate (1 eq) dissolved in anhydrous CH_2Cl_2 was added slowly. The reaction mixture was let under magnetic stirring for additional 4 hours after which the reaction mixture was quenched by adding a 1M aqueous solution of $KHSO_4$. The aqueous solution was then extracted twice with EtOAc and the combined organic layers were finally washed twice with brine. After drying over sodium sulfate, the product was concentrated under reduced pressure and vacuum to furnish the activated carbamate monomers.



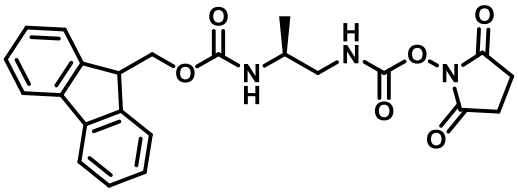
(S)-(9H-fluoren-9-yl)methyl (1-(2,5-dioxopyrrolidin-1-yl)-5-N-tritylpentanamide-2-yl)carbamate M11 : white solid, 91% yield, ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.74 (d, *J* = 7.7 Hz, 2H), 7.57 (dd, *J* = 16.0, 8.8 Hz, 2H), 7.33 (ddd, *J* = 11.0, 9.8, 4.5 Hz, 10H), 7.25 – 7.16 (m, 9H), 6.74 (s, 1H), 5.78 (d, *J* = 24.7 Hz, 1H), 5.33 – 5.11 (m, 1H), 4.57 – 4.27 (m, 3H), 3.66 (d, *J* = 13.4 Hz, 1H), 3.28 (d, *J* = 20.1 Hz, 1H), 3.08 (s, 1H), 2.74 (s, 4H), 2.45 (d, *J* = 37.3 Hz, 2H), 1.96 – 1.65 (m, 2H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 169.67, 144.45, 141.34, 128.64, 128.01, 127.69, 127.12, 119.95, 71.65, 69.70, 67.74, 65.78, 52.53, 50.35, 47.33, 46.13, 45.00, 33.39, 26.94, 25.43.; HRMS (ESI-TOFMS) *m/z* for $C_{44}H_{40}N_4O_7$ $[M+H]^+$ 736.2897, found 737.2973.



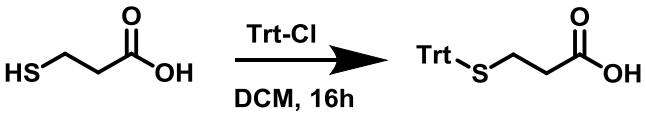
(S)-(9H-fluoren-9-yl)methyl (1-(2,5-dioxopyrrolidin-3-naphtylpropan-2-yl)carbamateM12 :White solid, 59% yield, ^{1}H NMR (300 MHz, DMSO) δ 8.57 (t, *J* = 5.5 Hz, 1H), 8.16 (d, *J* = 8.0 Hz, 1H), 7.84 (dd, *J* = 32.7, 7.4 Hz, 3H), 7.69 – 7.49 (m, 4H), 7.45 – 7.21 (m, 6H), 4.31 – 4.05 (m, 3H), 3.90 (t, *J* = 14.4 Hz, 1H), 3.51 – 3.41 (m, 1H), 3.26 (dd, *J* = 9.5, 7.2 Hz, 2H), 3.05 (dd, *J* = 13.9, 9.2 Hz, 1H), 2.83 – 2.72 (m, 4H), 2.35 – 2.21 (m, 1H). HRMS (ESI-TOFMS) *m/z* for $C_{23}H_{29}N_3O_6Na$ $[M+H]^+$ 563.2056, found 564.2131.



S)-(9H-fluoren-9-yl)methyl (1-(2,5-dioxopyrrolidin -3-N-(2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl)guanidine)-2-yl)carbamateM13 : White solid, 91% yield, ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.77 (d, *J* = 7.4 Hz, 2H), 7.59 (d, *J* = 8.2 Hz, 2H), 7.47 – 7.31 (m, 4H), 6.38 (s, 1H), 6.20 (s, 2H), 5.44 (s, 1H), 4.46 (d, *J* = 6.5 Hz, 2H), 4.23 – 4.15 (m, 1H), 3.77 (s, 2H), 3.35 (s, 1H), 3.26 (s, 2H), 2.94 (s, 2H), 2.79 (s, 4H), 2.57 (d, *J* = 21.1 Hz, 6H), 2.10 (s, 3H), 1.45 (s, 6H), 1.38 (d, *J* = 9.0 Hz, 2H), 1.29 (t, *J* = 4.4 Hz, 2H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.47 (s), 144.64 – 143.51 (m), 141.40 (s), 132.23 (d, *J* = 9.8 Hz), 128.71 (d, *J* = 12.1 Hz), 127.53 (d, *J* = 43.1 Hz), 120.05 (s), 87.81 – 85.02 (m), 67.56 – 66.20 (m), 47.32 (s), 43.26 (s), 28.70 (s), 25.63 (s), 19.50 (s), 18.12 (s), 12.61 (s).HRMS (ESI-TOFMS) *m/z* for $C_{39}H_{46}N_6O_9S$ $[M+H]^+$ 774.30, found 774.32



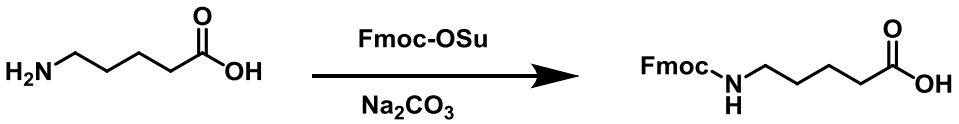
(S)-(9H-fluoren-9-yl)methyl (1-(2,5-dioxopyrrolidin -propan-2-yl)carbamate M20: White solid, 93% yield; ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.75 (t, *J* = 10.3 Hz, 2H), 7.67 – 7.56 (m, 2H), 7.35 (ddt, *J* = 18.1, 11.6, 9.1 Hz, 4H), 5.87 (s, 1H), 4.75 (s, 1H), 4.51 (dd, *J* = 19.1, 12.6 Hz, 2H), 4.41 (d, *J* = 6.3 Hz, 1H), 4.22 (t, *J* = 6.4 Hz, 1H), 3.88 (s, 1H), 3.32 (d, *J* = 21.3 Hz, 2H), 2.79 (s, 4H); HRMS (ESI-TOFMS) *m/z* for $C_{23}H_{23}N_3O_6$ $[M+H]^+$ 438.15, found 438.23



3-tritylsulfanyl-propionic acid[6]**:** The desired compound was synthesized as described previously[6].

4. Synthesis of monomers M21-M24

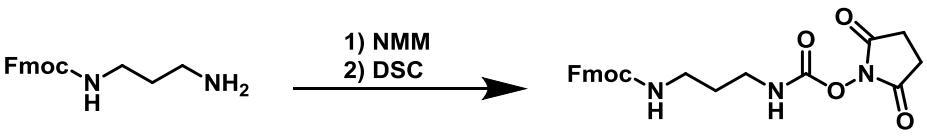
Synthesis of *N*-Fmoc-valeric acid M21:



In a round bottom flask containing 5-aminovaleric acid (8.5 mmol, 1 g) dissolved in and aqueous solution of Na_2CO_3 (0.15 M), a 0.1M solution of Fmoc-OSu (10.2 mmol, 3.45 g) in CH_3CN was added dropwise. The reaction was let under magnetic stirring overnight at r.t. After concentration under rotavapor, the crude product was purified on silica gel flash column chromatography and the expected *N*-Fmoc-valeric acid was isolated in 39% yield.

M21: ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.75 (t, *J* = 14.3 Hz,26H), 7.61 (d, *J* = 8.4 Hz, 2H), 7.47 – 7.30 (m, 4H), 4.83 (d, *J* = 6.6 Hz, 1H), 4.50 – 4.34 (m, 2H), 4.28 – 4.18 (m, 1H), 3.24 (d, *J* = 6.8 Hz, 2H), 1.26 (d, *J* = 5.6 Hz, 6H). HRMS (ESI-TOFMS) *m/z* for $C_{20}H_{21}NO_4$ $[M+H]^+$ 339.39, found 339.83

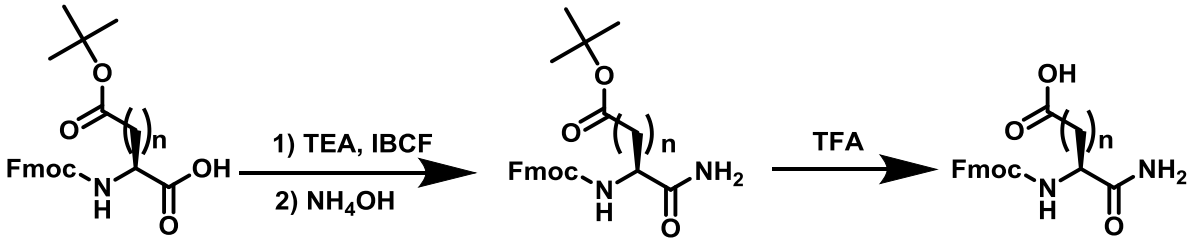
Synthesis of *N*-Fmoc-DAP-OSu M22:



In a round bottom flask containing *N*-Fmoc-NH-$(CH_2)_2NH_2$, HCl (1 g, 3.37 mmol) in solution in DCM, NMM (430 µL, 4 mmol,) was added. After 5 min, the resulting solution was added to DSC (0.863 g, 3.37 mmol, 1 eq) dissolved in DCM. The reaction was let under magnetic stirring at r.t. for 4 hrs. After concentration under reduced pressure, EtOAc was added and the organic phase was washed 2 times with 1M aqueous solution of $KHSO_4$ and once with brine. Recrystallisation was performed with a pentane/Et_2O (3:7, v/v) solvent mixture and compound **M22** was recovered a white solid in 71% yield.

M22: ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.76 (d, *J* = 7.4 Hz, 2H), 7.59 (d, *J* = 7.1 Hz, 2H), 7.44 – 7.27 (m, 4H), 5.99 (s, 1H), 4.98 (s, 1H), 4.47 (d, *J* = 6.5 Hz, 2H), 4.20 (t, *J* = 6.3 Hz, 1H), 3.26 (dd, *J* = 13.0, 6.1 Hz, 4H), 2.82 (s, 4H), 1.66 (dd, *J* = 15.9, 10.3 Hz, 2H). HRMS (ESI-TOFMS) *m/z* for $C_{23}H_{23}N_3O_6$ $[M+H]^+$ 437.15, found 437.93

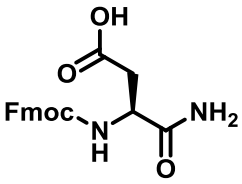
Synthesis of monomers M23 and M24



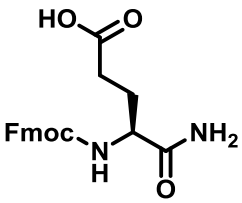
M23a: Glu (n=2)
M24a: Asp (n=1)

M23: Y=76%
M24: Y=72%

To a solution of *N*-Fmoc protected amino acid (1 eq) dissolved in anhydrous THF cooled to -20°C, under N_2 atmosphere, TEA (1.2 eq) and IBCF (1.2 eq) were added. The reaction mixture was let stirring for 40 min. NH_4OH (1.2 eq) in solution was added dropwise to the filtrate at 0°C. The reaction was left stirring 2h. After concentration under reduced pressure, ethyl acetate was added and the organic layer was washed twice with $KHSO_4$ solution (1M) and twice with saturated solution of $NaHCO_3$, once with brine. After drying over sodium sulfate, the organic layer was concentrated. The product was purified on silica gel flash chromatography in EtOAc/ cyclohexane (from 1/9 to 6/4). TFA was directly added to the pure compound and the reaction was le stirring for 2 hrs. After concentration under reduced pressure, the product was precipitated in cold Et_2O and the pure product was obtained as white solid.



(S)-(9H-fluoren-9-yl)methyl (1-amide-butanoic acid-2-yl)carbamate M23: white solid, 76% yield; ^{1}H NMR (300 MHz, MeOD) δ 7.83 (d, *J* = 7.3 Hz, 2H), 7.68 (dt, *J* = 24.4, 12.2 Hz, 2H), 7.60 – 7.23 (m, 4H), 4.72 – 4.32 (m, 1H), 4.27 (d, *J* = 6.7 Hz, 1H), 4.25 – 3.81 (m, 2H), 3.34 (dt, *J* = 3.2, 1.6 Hz, 2H), 2.34 (t, *J* = 7.5 Hz, 2H), 2.21 – 1.98 (m, 1H), 1.98 – 1.72 (m, 1H). HRMS (ESI-TOFMS) *m/z* for $C_{19}H_{18}N_2O_5$ $[M+H]^+$ 354.12, found 354.62

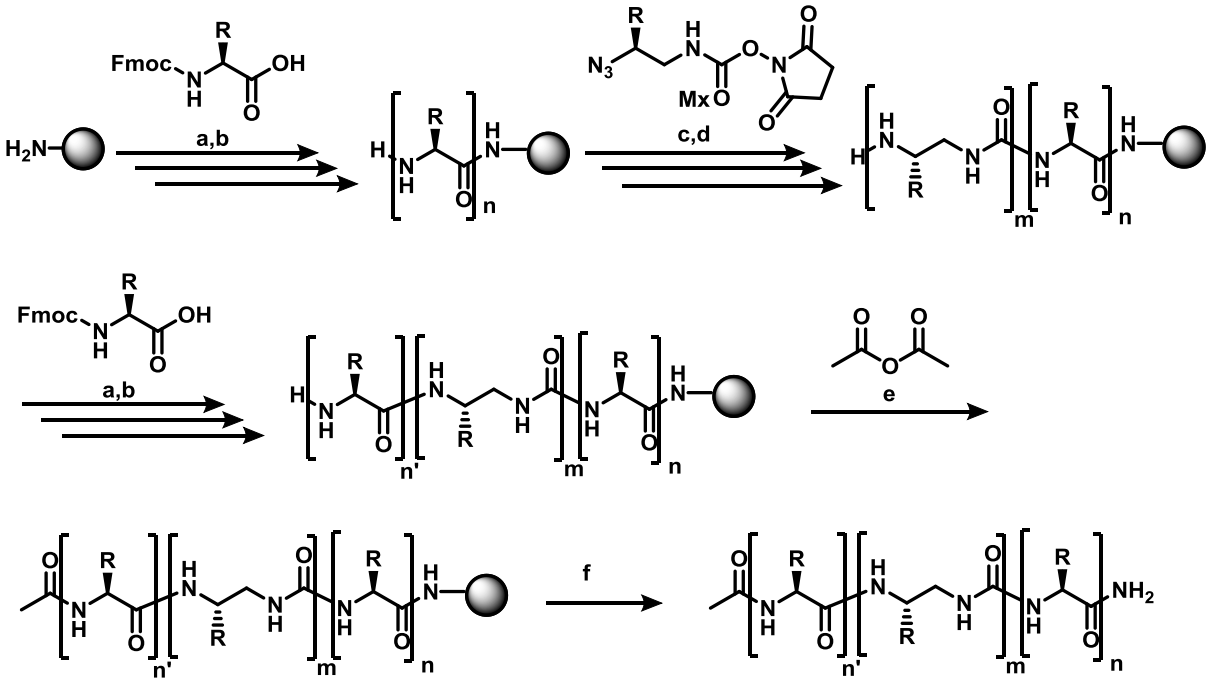


(S)-(9H-fluoren-9-yl)methyl (1-amide-pentanoic acid-2-yl)carbamate M24: white solid, 72% yield; ^{1}H NMR (300 MHz, DMSO) δ 7.91 (d, *J* = 7.5 Hz, 2H), 7.71 (dd, *J* = 14.2, 7.5 Hz, 2H), 7.40 (dt, *J* = 26.2, 7.4 Hz, 4H), 4.74 – 4.61 (m, 1H), 4.51 – 4.36 (m, 2H), 4.27 (t, *J* = 6.4 Hz, 2H), 2.87 (dd, *J* = 18.4, 6.2 Hz, 1H), 1.93 (s, 1H). ^{13}C NMR (75 MHz, DMSO) δ 170.94, 155.21, 145.65, 141.18, 128.14, 127.59, 125.74, 120.60, 67.44, 49.45, 48.09. HRMS (ESI-TOFMS) *m/z* for $C_{20}H_{20}N_2O_5$ $[M+H]^+$ 368.13, found 368.22

C. Solid phase synthesis of oligomers

1. Solid phase synthesis of oligomers in Chapter II

General procedure A for solid phase peptide synthesis under microwave irradiation using azide strategy



a: Fmoc -Xaa-OH (5 eq), HOBt (6 eq), HBTU (6 eq), DIEA (12 eq), 75°C (except for Fmoc-His-OH at 50°C), 30 W, 5 min; b: 20% piperidine, DMF (90°C/35W/50s); c:1.5 eq A, 3 eq DIEA 50°C 50W 2 x 15 min; d:10 eq. PMe_3 in THF (1M), Dioxane, H_2O (7:3 v/v), 50°C, 50W, 2 x 15min; e: Acetic anhydride/ DCM 1:1 (15 min, r.t.); f: TFA/TIS/H_2O/EDT (92.5:2.5:2.5:2.5, v,v,v,v)

Scheme 26 : Solid phase synthesis of compounds in chapter II

Peptides were prepared on a 50 to 100 µmol scale. Polystyrene rink amide MBHA resin (loading 0.45 mmol/g) was placed in the reaction vessel (CEM), and pre-swelled with DMF for 1-2h. All coupling steps were performed under inert atmosphere and microwave irradiation on the *Liberty Blue* system (CEM). The temperature was maintained by modulation of power and controlled with a fiber optic sensor. Stock solution of Fmoc-α-Xaa-OH (6 eq relative to the resin loading), HBTU (6 eq), HOBt (6 eq) and DIEA (12 eq) solutions were prepared in DMF except for DIEA that was prepared in N-methyl pyrolidone (from Carlo Erba). The mixture was added into the reaction vessel through *Liberty Blue* system. The vessel was then irradiated twice (75°C/30W/5min and 50°C for Fmoc-His-OH). The resin was then filtered off and washed with DMF. Fmoc removal was carried out with 20% of piperidine in DMF under microwave irradiation (155 W, 75°C, 15 sec) + (35W, 90°C, 50 sec). Finally, the resin was swelled in a mixture TFA/TIS/H_2O (92.5:2.5:2.5:2.5,

v/v/v/v) and let to react for 4hrs. Then, the resin was filtered off, washed with TFA (2 x 2mL) and DCM (2 x 2mL) into a falcon and cold Et_2O was added into the falcon. The filtrate in the falcon was centrifuged and the supernatant was removed. The crude oligomer was analyzed by RP-HPLC and lyophilized. The crude oligomer was next purified by preparative RP-HPLC using the appropriate gradient to a final purity ≥ 95% and lyophilized.

II.9: Ac-RMKQLEDKIEELLSKNYHLENEIARLKKLIGER-NH_2

The peptide **II.9** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 $mmol.g^{-1}$) with commercially available Fmoc-α-Xaa-OH following the general procedure A with coupling performed in the presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC with H_2O/0.1% TFA (solvent A) and ACN/0.1% TFA (solvent B) as mobile phases. A gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 10%. ESI-MS (MW = 4079.79: m/z 2040.6 $[M+2H]^{2+}$, 1360.7 $[M+3H]^{3+}$, 1020.8 $[M+4H]^{4+}$, 816.8 $[M+5H]^{5+}$,680.9 $[M+6H]^{6+}$; gradient 20-100% B, 5 min: t_R= 7.98 min.

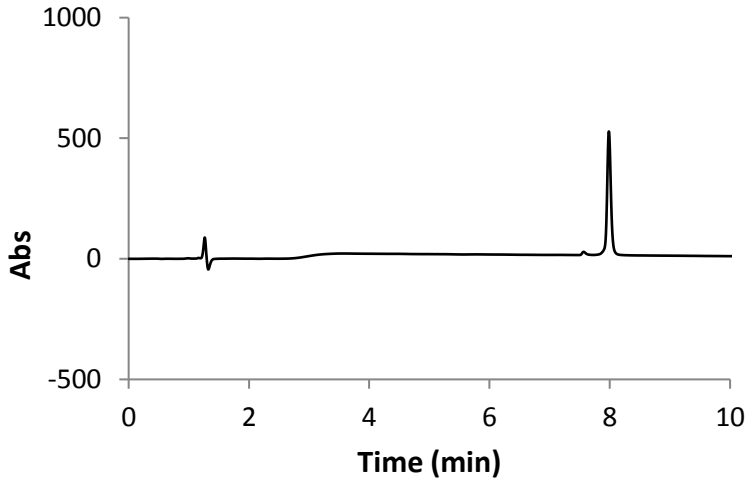


Figure 89: RP-HPLC chromatogram of II.9 (10-100% of B in 10 min)

II.16: Ac-RMKQLEDKIEELL^uK^uNYHLENEIARLKKLIGER-NH_2

Peptide **II.16** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 $mmol.g^{-1}$) with commercially available Fmoc-α-Xaa-OH and monomers **M3** and **M5** following the general procedure A with couplings performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 8%. ESI-MS (*Mw* 4048.8): m/z 1013.4 $[M+4H]^{4+}$, 810.9 $[M+5H]^{5+}$, 675.9 $[M+6H]^{6+}$, 579.5 $[M+7H]^{7+}$; gradient 20-100% B, 5 min: t_R= 7.56 min.

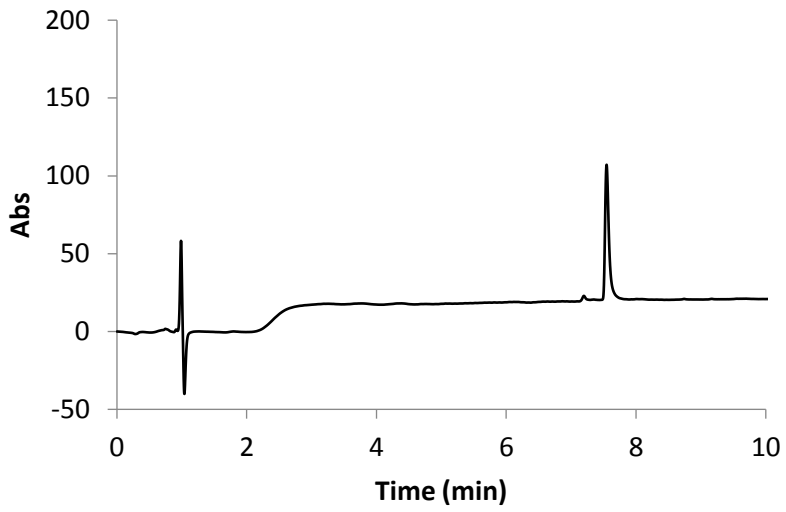


Figure 90: RP-HPLC chromatogram of II.16 (10-100% ACN in 10 min, 50°C)

II.17: Ac-RMKQLEDKIEEL$^{u}K_{inv}{}^{u}$NYHLENEIARLKKLIGER-NH_2

The peptide **II.17** was synthesized on a 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and monomers **M3** and **M8** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to 60% of B in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 8%. ESI-MS (*Mw* 4050.79): m/z 2025.9 $[M+2H]^{2+}$ 1013.4 $[M+4H]^{4+}$, 810.9 $[M+5H]^{5+}$, 675.9 $[M+6H]^{6+}$; gradient 10-100% B, 10 min, 50°C: t_R= 7.5 min.

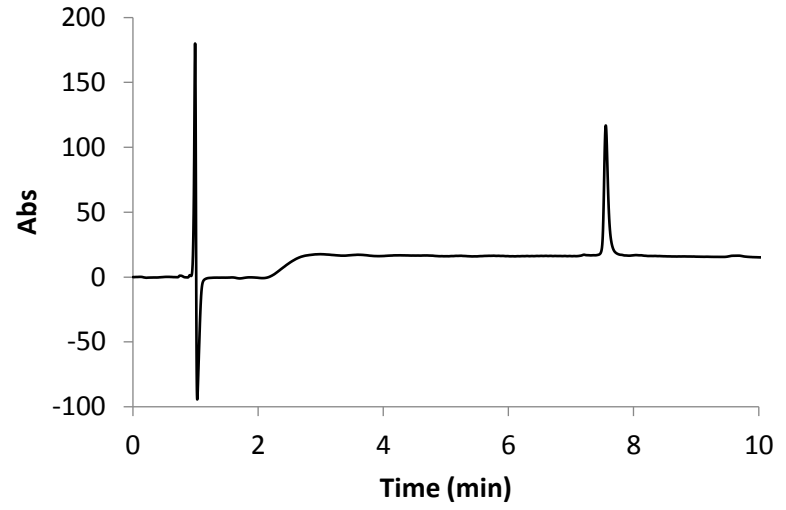


Figure 91: RP-HPLC chromatogram of II.17 (10-100% B in 10 min, 50°C)

II.18: Ac-RMKQLEDKIEELI^uK^uNYHLENEIARLKKLIGER-NH_2

The peptide **II.18** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and monomers **M4** and **M5** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC with H_2O (0.1% TFA) and ACN (0.1% TFA). A gradient 20-60% of ACN in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 9%. ESI-MS (*Mw* 4050.8): m/z 1013.4 $[M+4H]^{4+}$, 810.9 $[M+5H]^{5+}$, 675.9$[M+6H]^{6+}$;579.5 $[M+7H]^{7+}$; gradient 20-100% B, 5 min: t_R= 7 min.

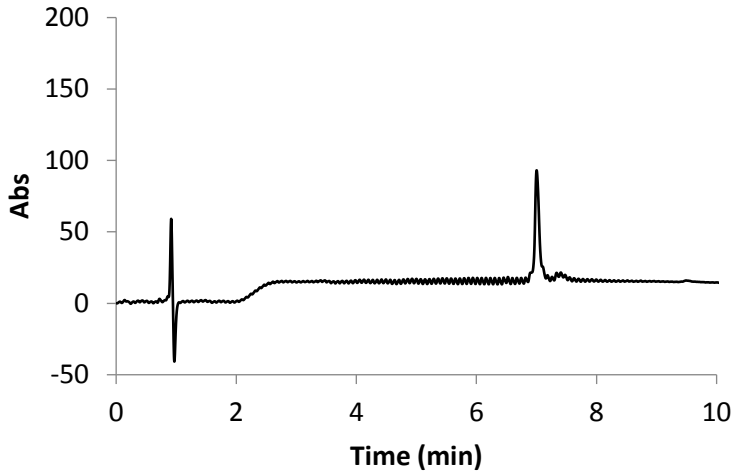


Figure 92: RP-HPLC chromatogram of II.18 (10-100% B in 10 min, 50°C)

II.19: Ac-RMKQLEDKIEELV^uK^uNYHLENEIARLKKLIGER-NH_2

The peptide **II.19** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and monomers **M2** and **M5** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 8%. ESI-MS (*Mw* 4036.7): m/z 1009.9 $[M+4H]^{4+}$, 808.1 $[M+5H]^{5+}$, 577.5 $[M+7H]^{7+}$; gradient 20-100% ACN, 5 min: t_R= 6.68 min.

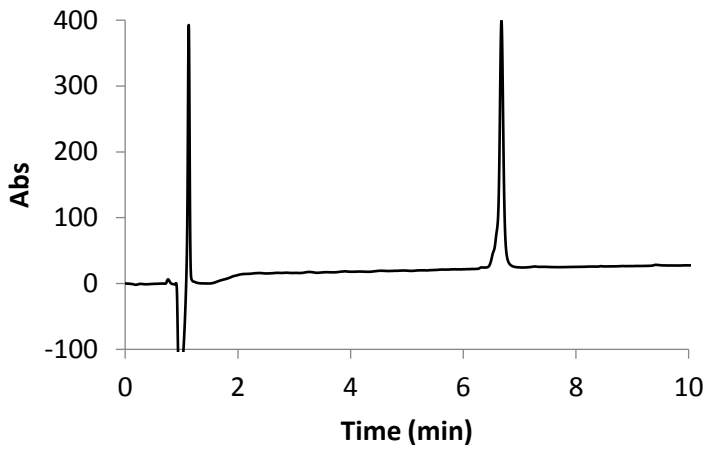


Figure 93: RP-HPLC chromatogram of II.19 (10-100% B in 10 min, 50°C)

II.20: Ac-RMKQLEDKIEELA^{u}K^{u}NYHLENEIARLKKLIGER-NH_2

The peptide **II.20** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and monomers **M1** and **M5** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC with H_2O (0.1% TFA) and ACN (0.1% TFA). A gradient 20-60% of ACN in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 11%. ESI-MS (*Mw* 4007.7): m/z 1336.9 $[M+3H]^{3+}$, 1002.9 $[M+4H]^{4+}$, 802.3 $[M+5H]^{5+}$; 668.9 $[M+6H]^{6+}$; 573.5 $[M+7H]^{7+}$; gradient 20-100% ACN, 5 min: t_R= 6.49 min.

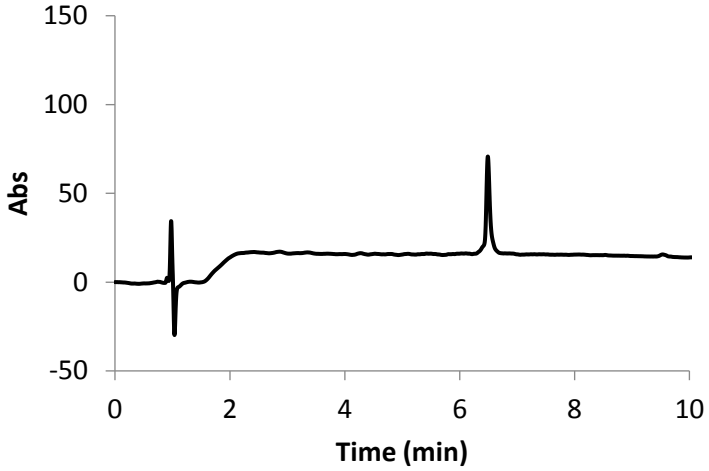


Figure 94: RP-HPLC chromatogram of II.20 (10-100% B in 10 min, 50°C)

II.21: Ac-RMKQLEDKIEEL^{u}S^{u}KNYHLENEIARLKKLIGER-NH_2

The peptide **II.21** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and monomers **M3** and **M7** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20-to 60% of B in 20 min with a flow of 20 mL/min

was used. The product was recovered with an overall yield of 11%. ESI-MS (*Mw* 4024.7): m/z 1006.9 $[M+4H]^{4+}$,805.7 $[M+5H]^{5+}$, 671.6 $[M+6H]^{6+}$; gradient 20-100% ACN, 5 min: t_R= 5.9 min.

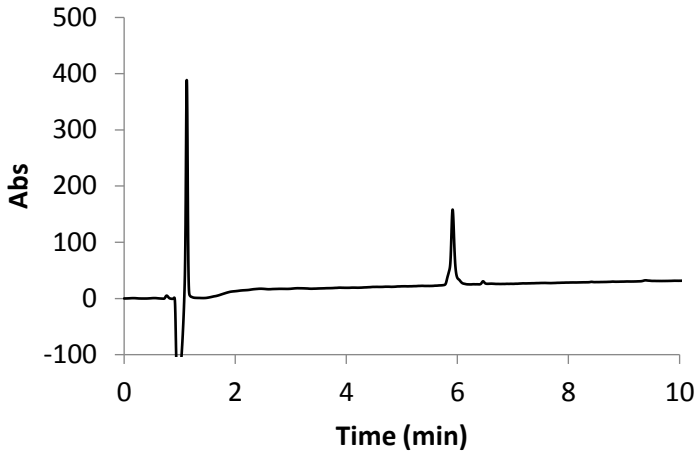


Figure 95: RP-HPLC chromatogram of II.21 (10-100% B in 10 min, 50°C)

II.22: Ac-RMKQLEDKIEELLSKNYHLENEI^uR^uLKKLIGER-NH_2

The peptide **II.22** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 $mmol.g^{-1}$) with commercially available Fmoc-α-Xaa-OH and monomers **M4** and **M6** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 16%. ESI-MS (*Mw* 4066.8): m/z 1356.3 $[M+3H]^{4+}$, 1017.4 $[M+4H]^{4+}$, 814.1 $[M+5H]^{5+}$, 581.8 $[M+7H]^{7+}$; gradient 20-100% ACN, 5 min: t_R= 7.23 min.

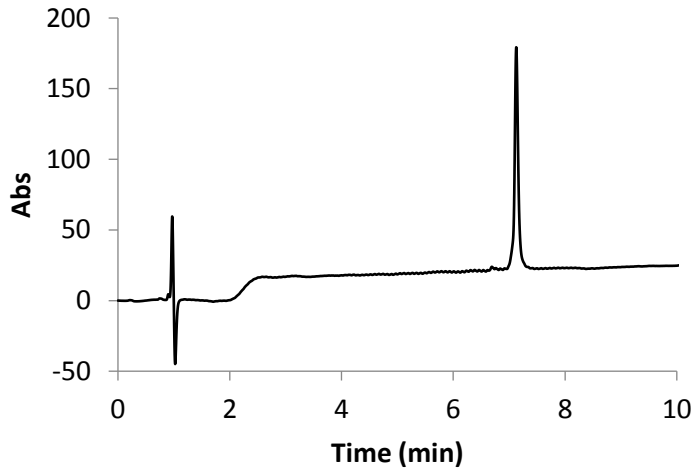


Figure 96: RP-HPLC chromatogram of II.22 (10-100% B in 10 min, 50°C)

II.23: Ac-RMKQLEDKIEELLSKNYHLENEA^uR^uLKKLIGER-NH_2

The peptide **II.23** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 $mmol.g^{-1}$) with commercially available Fmoc-α-Xaa-OH and monomers **M1** and **M6** following the general

procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to 60% of B in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 10%. ESI-MS (*Mw* 4028.7): m/z 1006.9 $[M+4H]^{4+}$, 805.7 $[M+5H]^{5+}$, 575.8 $[M+7H]^{7+}$; gradient 20-100% ACN, 5 min: t_R= 6.29 min.

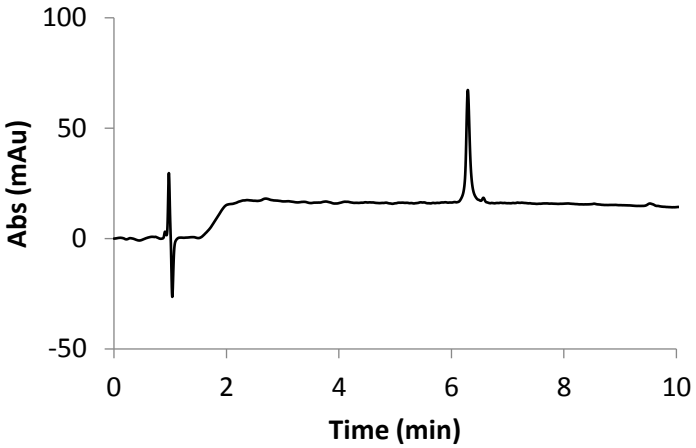


Figure 97: RP-HPLC chromatogram of II.23 (10-100% B in 10 min, 50°C)

II.24: Ac-RMKQLEDKIEE**L^{u}A^{u}K^{u}**NYHLENEIARLKKLIGER-NH_2

The peptide **II.24** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 $mmol.g^{-1}$) with commercially available Fmoc-α-Xaa-OH and monomers **M1**, **M3** and **M5** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to 60% of B in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 15%. ESI-MS (*Mw* 4037.7): m/z 2019.5 $[M+2H]^{2+}$, 1346 $[M+3H]^{3+}$, 1010.2 $[M+4H]^{4+}$, 808.3 $[M+5H]^{5+}$, 673.8 $[M+6H]^{6+}$; gradient 20-100% ACN, 5 min: t_R= 6.38 min.

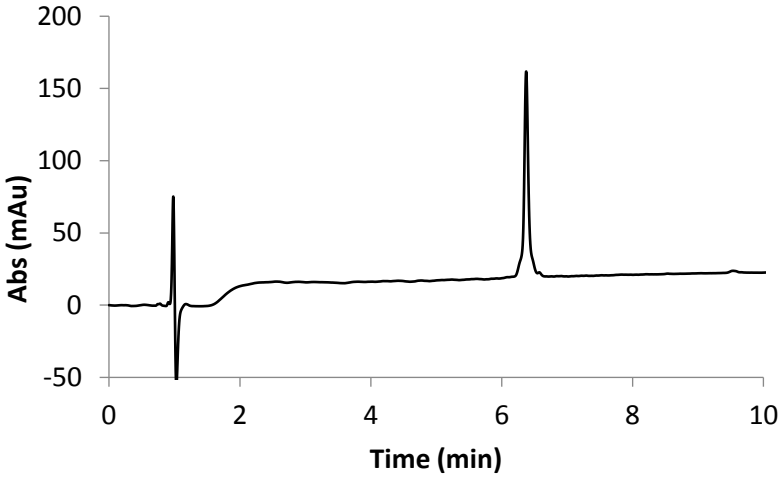


Figure 98: RP-HPLC chromatogram of II.24 (10-100% B in 10 min, 50°C)

II.25: Ac-RMKQLEDK$V^uE^uL^u$LSKNYHLENEIARLKKLIGER-NH_2

The peptide **II.25** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and monomers **M2, M3** and **M9** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to 60% of B in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 20%. ESI-MS (*Mw* 4009.7): m/z 1337.3 $[M+3H]^{3+}$, 1003.2 $[M+4H]^{4+}$, 802.7 $[M+5H]^{5+}$, 669.1 $[M+6H]^{6+}$; gradient 20-100% ACN, 5 min: t_R= 6.27 min.

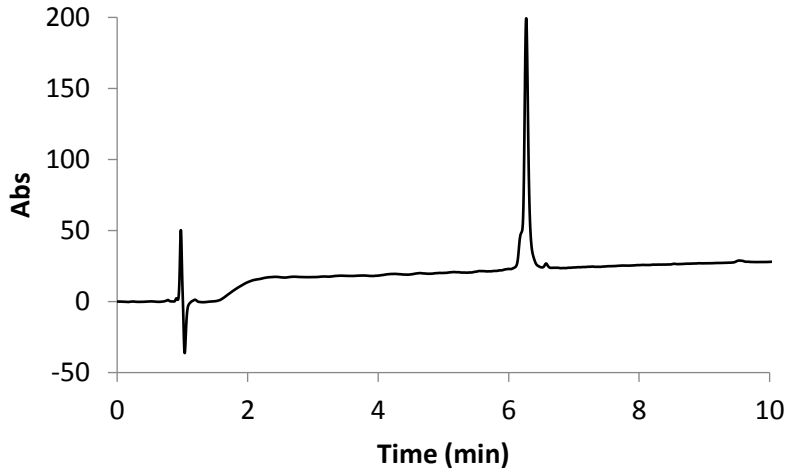


Figure 99: RP-HPLC chromatogram of II.25 (10-100% B in 10 min, 50°C)

II.26: Ac-RMKQLEDKIEEL$L^uK_{inv}{}^u$LYHLENEIARLKKLIGER-NH_2

The peptide **II.26** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and monomers **M3** and **M8** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to 60% of B in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 8%. ESI-MS (*Mw* 4049.8): m/z 1013.5 $[M+4H]^{4+}$, 810.8 $[M+5H]^{5+}$, 875.6 $[M+6H]^{6+}$; gradient 20-100% B, 5 min: t_R= 7.83 min.

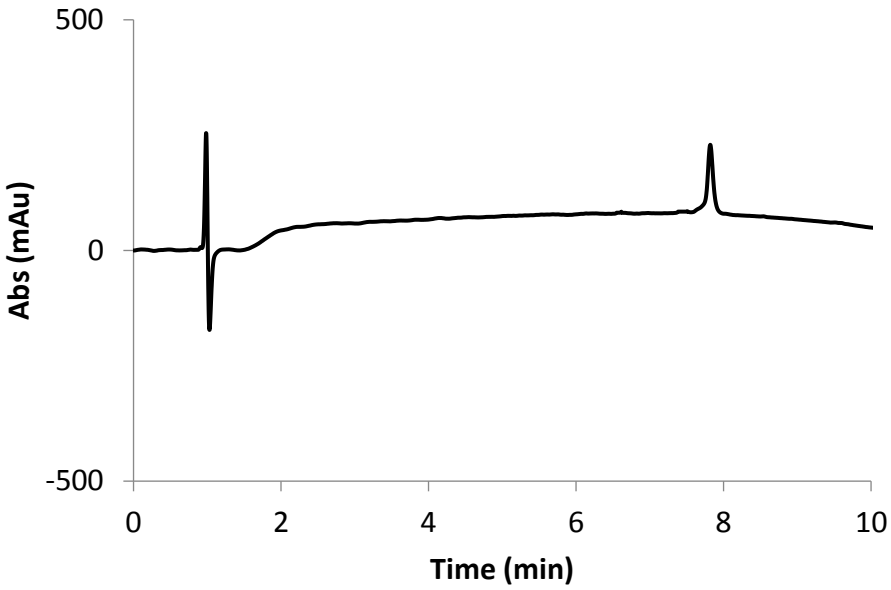


Figure 100: RP-HPLC chromatogram of II.26 (10-100% ACN in 10 min, 50°C)

II.27: Ac-RMKQIEDKLEEILSKLYHIENELARIKKLLGER-NH_2

The peptide **II.27** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC. A gradient 20 to 60% of B in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 11%. ESI-MS (*Mw* 4078.8): m/z 1019.3 $[M+4H]^{4+}$, 815.1 $[M+5H]^{5+}$, 679.7$[M+6H]^{6+}$; gradient 20-100% B, 5 min: t_R= 7.95 min.

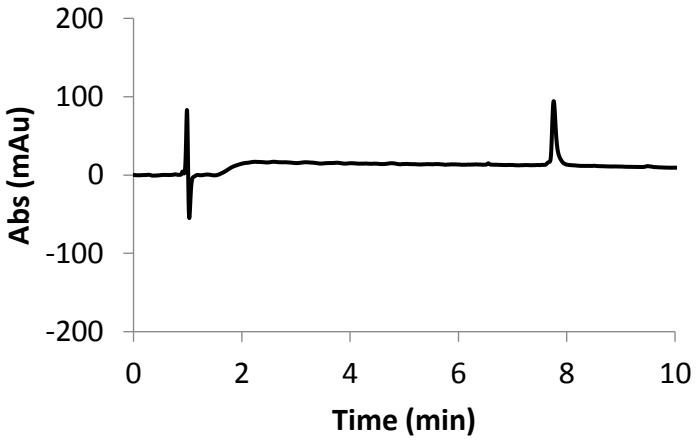


Figure 101: RP-HPLC chromatogram of II.27 (10-100% B in 10 min, 50°C)

II.28: Ac-RMKQIEDKLEEI $L^uK_{inv}{}^uL$YHIENELARIKKLLGER-NH_2

The peptide **II.28** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and monomers **M3** and **M8** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC with H_2O (0.1% TFA) and ACN (0.1% TFA). A gradient 20-60% of B in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 10%. ESI-MS (*Mw* 4049.8): m/z 1350.8 $[M+3H]^{3+}$, 1013.4 $[M+4H]^{4+}$, 810.9 $[M+5H]^{5+}$, 675.9 $[M+6H]^{6+}$, 579.8 $[M+7H]^{7+}$; gradient 20-100% B, 5 min: t_R= 7.76 min.

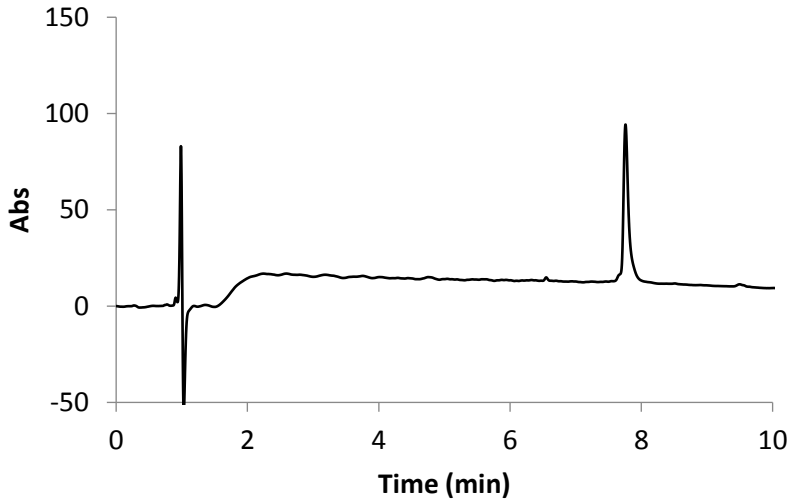


Figure 102: RP-HPLC chromatogram of II.28 (10-100% B in 10 min, 50°C)

II.29: Ac-RMKQLEDKIEELLSK**N^{u}**YHLENEIARLKKLIGER-NH_2

The peptide **II.29** was synthesized on 50 µmol scale, on Polystyrene resin (0.37 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-OH and the monomer **M10** following the general procedure A with coupling performed in presence of HOBt, HBTU and DIEA. The desired product was purified by preparative RP-HPLC with H_2O (0.1% TFA) and B (0.1% TFA). A gradient 20 to 60% of B in 20 min with a flow of 20mL/min was used. The product was recovered with an overall yield of 12%. ESI-MS (*Mw* 4066.7): m/z 1017.6 $[M+4H]^{4+}$, 814.2 $[M+5H]^{5+}$, 678.7 $[M+6H]^{6+}$, 581.9 $[M+7H]^{7+}$; HPLC (H_2O 0.1% TFA), (B 0.1%TFA); gradient 20-100% B, 5 min: t_R= 6.99 min.

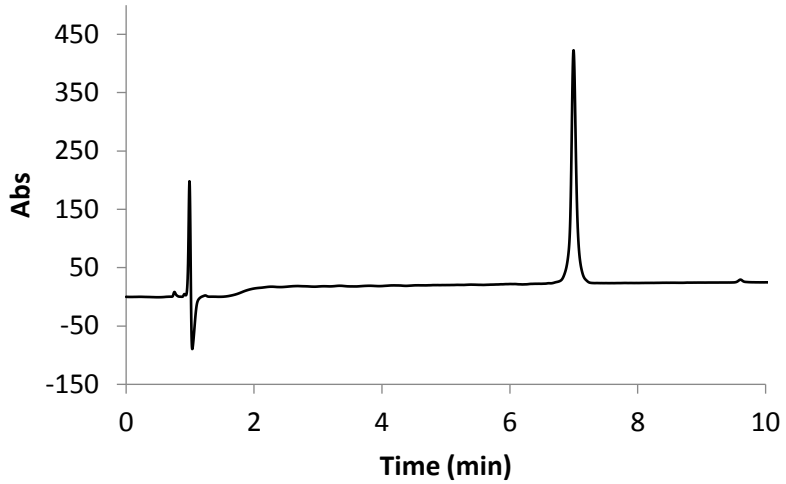


Figure 103: RP-HPLC chromatogram of II.29 (10-100% B in 10 min, 50°C)

Analyze with circular dichroism:

To calculate the percentage of helical structure, the minimum of ellipticity $[\theta]_{RME}$ was observed at 222nm.

$$\%helicity = ([\theta]_{RME222} - [\theta]c) + ([\theta]\infty 222 - [\theta]c)$$

where [θ]c is the molar ellipticity of a non-structured peptide and [θ]∞222 is the molar ellipticity of a fully α-helical structured peptide. They could be respectively defined as:

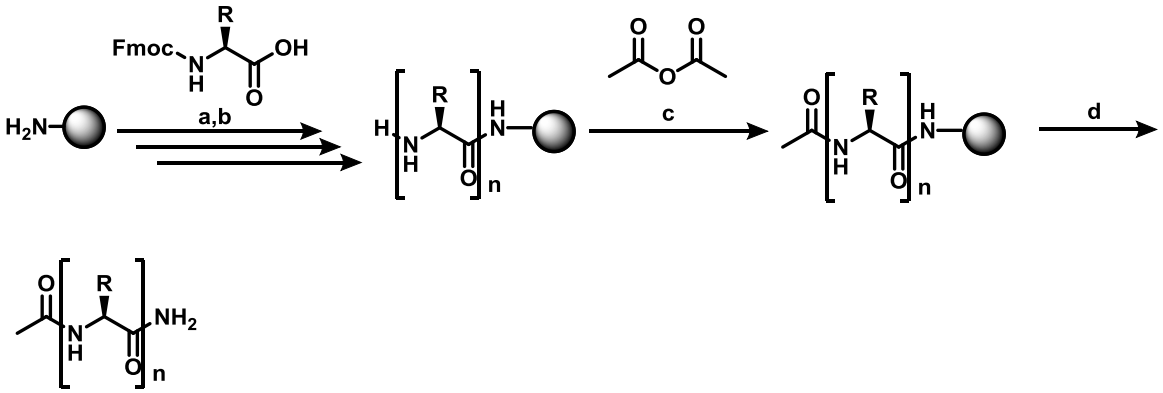
$[\theta]c = (2220 - 53T)$

$[\theta]\infty 222 = (-44000 + 250T) \times (1 - k / N)$

With T the temperature in degree Celsius and N the number of peptide units. (1- k / N) is a correction of terminal effect with k the number of carbonyl group not involved in H-bond network. For amide peptides in C-terminus, a value of k=3 was used[7].

2. Solid phase synthesis of oligomers in Chapter III

General procedure B for solid phase peptides synthesis under microwave irradiation with *Liberty Blue*



a: Fmoc -Xaa-OH (5 eq), DIC (5 eq) Oxyma (5 eq), (75°C, 170 W, 15sec) + (90°C,35W,110sec) (Double coupling for Arg); b: 20% piperidine, DMF (75°C,155W,15sec) + (90°C,35W,50sec) c: Acetic anhydride/ DCM 1:1 (15 min, r.t.); d: TFA/TIS/H_2O (95:2.5:2.5, v,v,v)

Scheme 27:General method of SPPS with *Liberty Blue*

Oligomers were prepared on a 50 to 100 µmol scale. Polystyrene Rink amide resin (loading 0.45 mmol/g) was placed in the reaction vessel, and pre-swelled with DMF for 1-2h. The vessel was then placed inside the microwave reactor of the *Liberty Blue*. The temperature was maintained by modulation of power and controlled with a fiber optic sensor. Fmoc-α-Xaa-OH (5 eq relative to the resin loading), DIC (5 eq) and Oxyma (5 eq), were dissolved in DMF. The vessel was then irradiated (75°C,170 W, 15 sec) + (90°C, 35 W, 110 sec). The resin was then filtered off and washed with DMF (3 × 3 mL). Fmoc removal was carried out with 20% of piperidine in DMF (2mL) under microwave irradiation (75°C, 155W, 15 sec) + (90°C, 35W, 50 sec). When Acetyl *N*-cap was needed, an additional step was performed manually with a mixture acetic anhydride/DCM (1:1, v/v). Finally, the resin was swelled in a mixture TFA/TIS/H_2O/ (95:2.5:2.5, v/v/v) and let to react for 4hrs under slight shaking. Then, the resin was filtered off, washed with TFA (2 x 2mL) and CH_2Cl_2 (2 x 2mL) and the filtrate was evaporated under reduced pressure. After precipitation in cold Et_2O, the crude oligomer was analyzed on RP-HPLC and lyophilized. The crude oligomer was purified by preparative or semi-preparative RP-HPLC using the appropriate gradient to a final purity ≥ 95% and lyophilized.

p4: Ac-EKWARLARRIA-NH_2

Peptide **p4** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- following the general procedure C. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 49%. ESI-MS (*Mw* 1410.7): m/z 705.8 $[M+2H]^{2+}$, 470.8 $[M+3H]^{3+}$, 353.4 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 3.88 min.

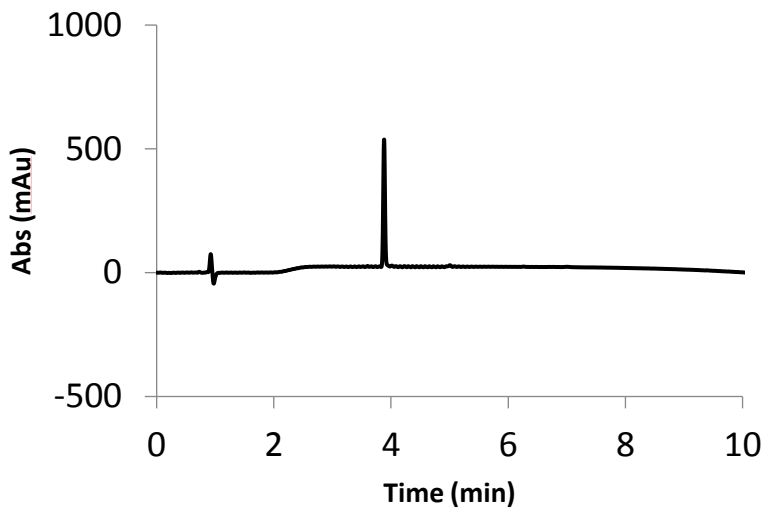


Figure 104: RP-HPLC chromatogram of p4 (10-100% ACN in 10 min, 50°C)

p5: Ac-EK-Nal-ARLARRIA-NH_2

Peptide **p5** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- following the general procedure B. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 55%. ESI-MS (*Mw* 1420.8): m/z 711.6 $[M+2H]^{2+}$, 474.7 $[M+3H]^{3+}$, 366.3 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 4.22 min.

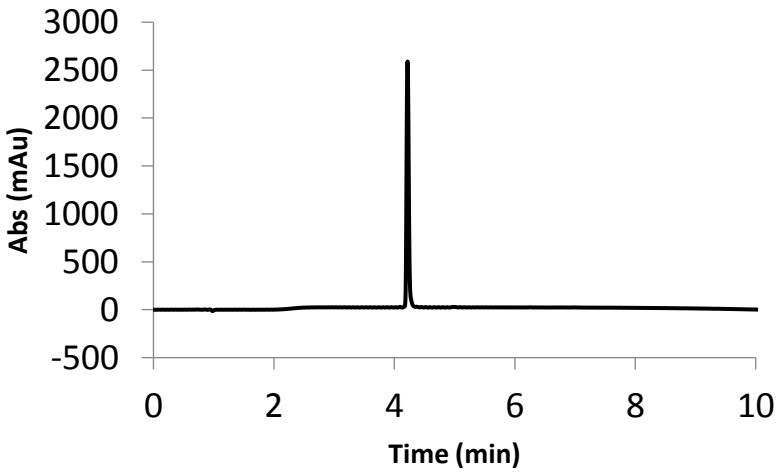


Figure 105: RP-HPLC chromatogram of p5 (10-100% ACN in 10 min, 50°C)

p7: (NPys-(S-S))-C-$(RRP)_5$-A-NH_2

Peptide **p7** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- following the general procedure C. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 15 to 30% of B in 15 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 40%. ESI-MS (*Mw* 2347.8): m/z 782.8 $[M+3H]^{3+}$, 587.8 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 3.34 min.

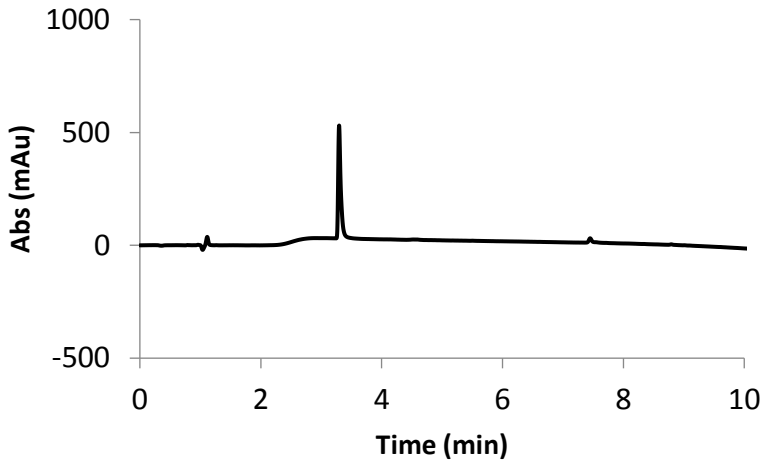


Figure 106: RP-HPLC chromatogram of p7 (10-100% ACN in 10 min, 50°C)

p9=H-$(RRP)_5$-NH_2

Peptide **p9** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- following the general procedure B. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 40% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 49%. ESI-MS (*Mw* 2064.5): m/z 687.9 $[M+3H]^{3+}$, 516.9 $[M+4H]^{4+}$, 344.9 $[M+6H]^{6+}$; gradient 20-100% B, 5 min: t_R= 3.51 min.

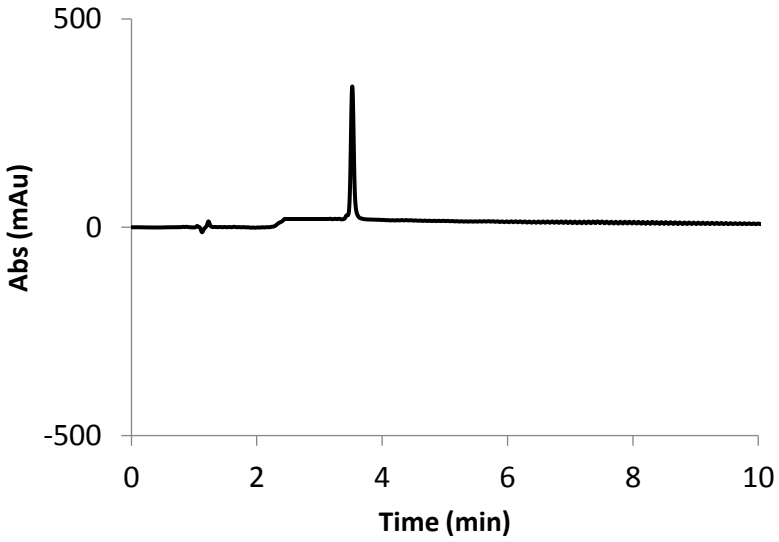
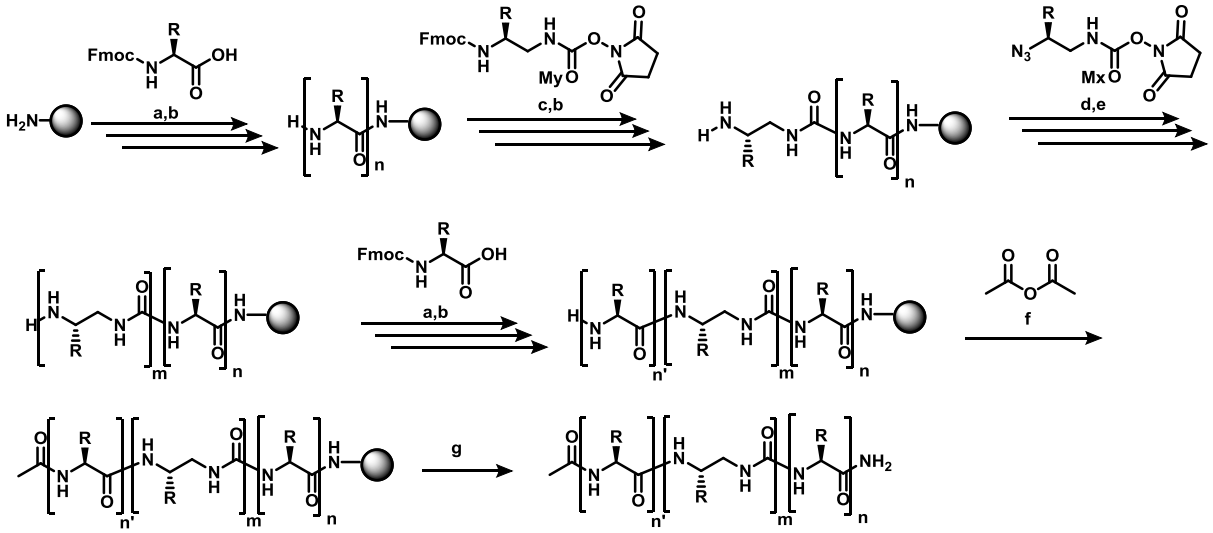


Figure 107: RP-HPLC chromatogram of p9 (10-100% ACN in 10 min, 50°C)

General procedure C for solid phase oligomers synthesis under microwave irradiation with *Discover Bio*



a: Fmoc -Xaa-OH (5 eq), DIC (5 eq), Oxyma (5 eq), 75°C, 25 W, 5 min (x 2 for Arg); b: 20% piperidine, DMF (75°C,25W,1'+3'); c: 5 eq My, 7 eq DIEA r.t. 2hrs x 2; d: 1.5 eq Mx, 3 eq DIEA 50°C, 50W, 15 min x 2; e:10 eq. PMe_3 in THF (1M), Dioxane, H_2O (7:3 v/v), 50°C, 50W, 2 x 15min; f: Acetic anhydride/ DCM 1:1 (15 min, r.t.); f: TFA/TIS/H_2O (95:2.5:2.5, v,v,v)

Scheme 28: General procedure of SPS of chimeras

Oligomers were prepared on a 50 to 100 µmol scale. Polystyrene Rink amide resin (loading 0.76 mmol/g) was placed in the reaction vessel, and pre-swelled with DMF for 1-2h. For sequences composed of less than four amino acids at the C-terminus, all steps were performed under inert atmosphere and microwave irradiation on the Discover Bio system (CEM). The vessel was then placed inside the microwave reactor. The temperature was maintained by modulation of power and controlled with a fiber optic sensor. Fmoc-α-Xaa-OH (5 eq relative to the resin loading), DIC (5 eq) and Oxyma (5 eq), were dissolved in DMF. The resulting mixture was manually added to the reaction vessel. The vessel was then irradiated (75°C,25 W, 5 min) once for all amino acids except Arg, repeated twice. The resin was then filtered off and washed with DMF (3 × 3 mL). Fmoc removal was carried out with 20% of piperidine in DMF (2mL) under microwave irradiation (75°C, 155W, 15 sec) + (90°C, 35W, 50 sec). These coupling and reduction steps were monitored with the Kaiser test[8]. In some sequences, the first urea residue (Fmoc-Gln(Trt)-OSu) **M11** was introduced in Fmoc-strategy with DIEA (7eq) at r.t. for 2h twice. Fmoc removal was performed as described before. Otherwise, the azide protected succinimidyl carbamate activated monomer (1.5 equiv. relative to the resin loading) were dissolved into DMF (1.5 mL) and were added to the reaction vessel (CEM), followed by DIEA (3 equiv.). The reaction vessel was then irradiated under microwave (50°C, 50W, 15 min). A double coupling was performed systematically. The resin was filtered and washed with DMF (3 x 3mL). After that, the resin was washed with a mixture of 1,4-dioxane:H_2O (7:3 v/v) in order to perform the reduction of the azide group (3 x 3mL). The

Staudinger reaction was performed under the microwave irradiation (50°C, 50W, 15 min) in the mixture 1,4- dioxane:H_2O (1.5mL) with 1M PMe_3 solution in THF (10 eq relative to the resin loading) used as reducing agent. The reduction step was systematically performed twice. The resin was then filtered and washed with DMF (4 x 3 mL). These coupling and reduction steps were monitored with the chloranil test[9].For adding the last amino acids, the Fmoc-α-Xaa-OH was added as described before. When an Acetyl *N*-cap was needed, an additional step was performed with a mixture acetic anhydride/DCM (1:1, v/v). Finally, the resin was swelled in a mixture TFA/TIS/H_2O/ (95:2.5:2.5, v/v/v) and let to react for 4hrs under slight shaking. Then, the resin was filtered off, washed with TFA (2 x 2mL) and CH_2Cl_2 (2 x 2mL) and the filtrate was evaporated under reduced pressure. After precipitation in cold Et_2O, the crude oligomer was analyzed on RP-HPLC and lyophilized. The crude oligomer was purified by preparative or semi-preparative RP-HPLC using the appropriate gradient to a final purity ≥ 95% and lyophilized.

ch1: Ac-EK-Nal-ARL$Q^{u}R^{u}I^{u}$A-NH_2

The compound **ch1** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M4**, **M6** and **M14** following the general procedure C without introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 21%. ESI-MS (*Mw* 1409.7): m/z 1409.7 $[M+H]^+$, 705.3 $[M+2H]^{2+}$, 470.5 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 3.42 min.

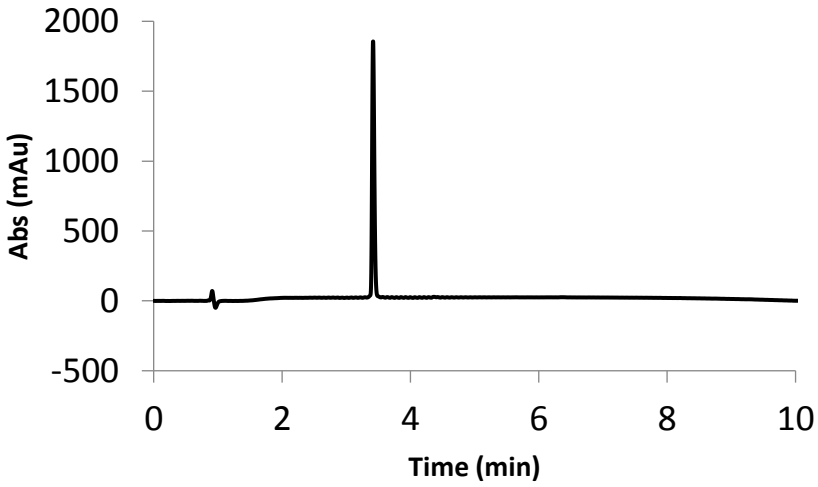


Figure 108: RP-HPLC chromatogram of ch1 (10-100% ACN in 10 min, 50°C)

ch2: Ac-EK-Nal-AR$L^{u}Q^{u}R^{u}$IA-NH_2

The compound **ch2** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M3**, **M6** and **M14** following the general procedure C without introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O

(0.1% TFA), ACN (0.1% TFA); a gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 30%. ESI-MS (*Mw* 1409.7): m/z 1409.8 $[M+H]^+$, 705.5 $[M+2H]^{2+}$, 470.6 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.57 min.

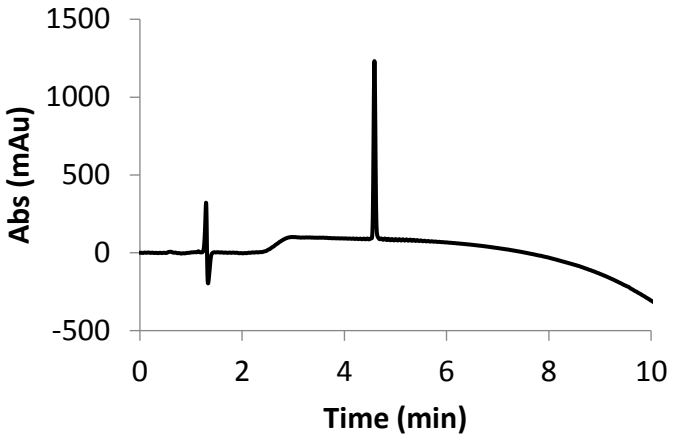


Figure 109: RP-HPLC chromatogram of ch2 (10-100% ACN in 10 min, 50°C)

ch3: Ac-EK-Nal-A-R^{u}L^{u}Q^{u}RIA-NH_2

The compound **ch3** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M3**, **M6** and **M11** following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 15%. ESI-MS (*Mw* 1409.7): m/z 1409.8 $[M+H]^+$, 705.5 $[M+2H]^{2+}$, 470.7 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.80 min.

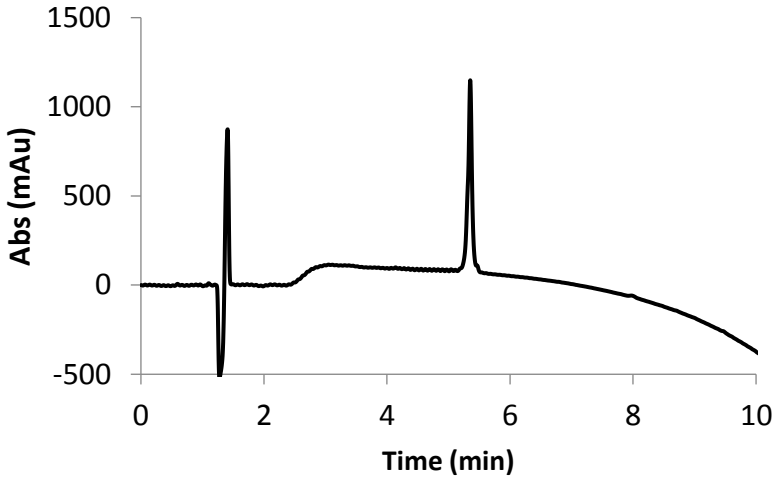


Figure 110: RP-HPLC chromatogram of ch3 (10-100% ACN in 10 min, 50°C)

ch4: Ac-EK-Nal^u-Arg^{u}Leu^{u}ARRIA-NH_2

The compound **ch4** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M3**, **M6** and **M16** following the general procedure C without introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 28%. ESI-MS (*Mw* 1436.8): m/z 1409.7 $[M+H]^+$, 705.3 $[M+2H]^{2+}$, 470.5 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.37 min.

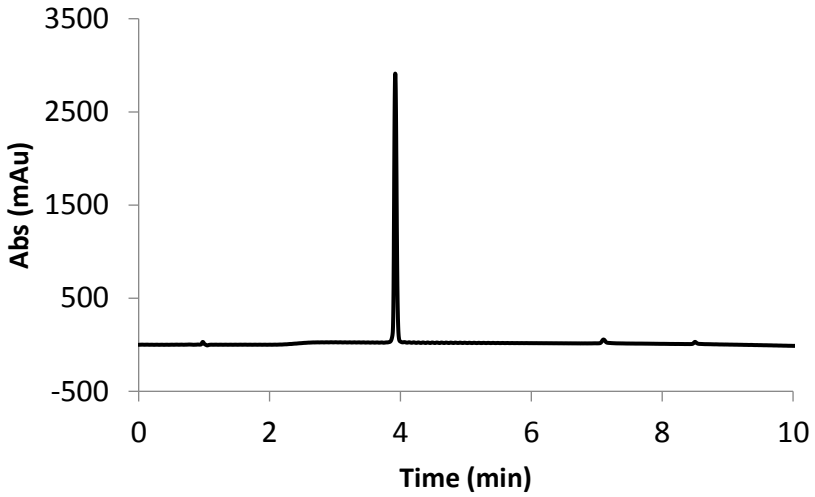


Figure 111: RP-HPLC chromatogram of ch4 (10-100% ACN in 10 min, 50°C)

ch5: Ac-EK-Nal^u-R^{u}L^{u}Q^{u}RIA-NH_2

The compound **ch5** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M3**, **M6**, **M14** and **M16** following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 7%. ESI-MS (*Mw* 1367.7): m/z 1367.7 $[M+H]^+$, 684.5 $[M+2H]^{2+}$, 456.7 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.55 min.

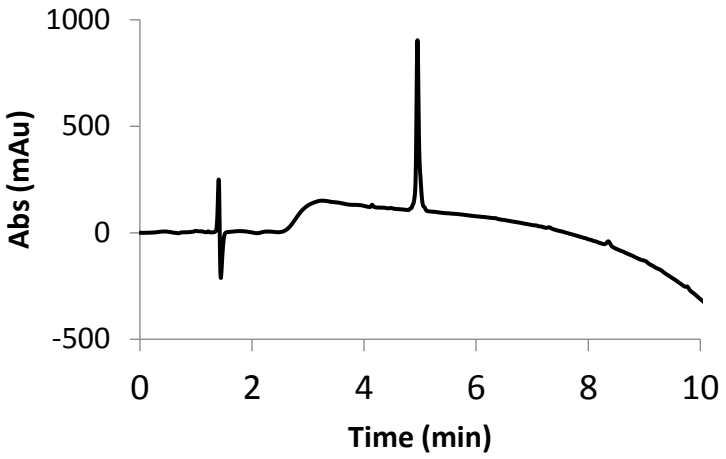
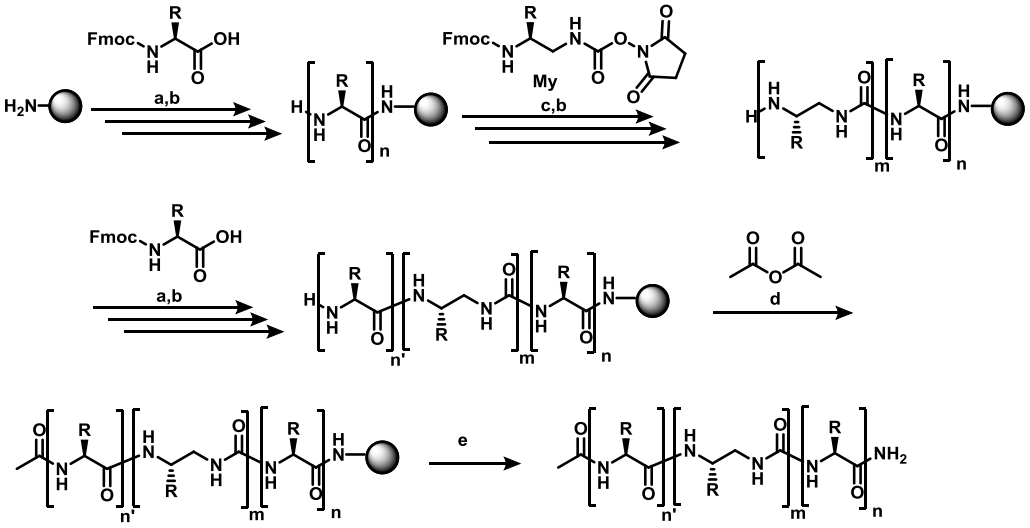


Figure 112: RP-HPLC chromatogram of ch5 (10-100% ACN in 10 min, 50°C)

Solid phase synthesis of **ch5** under microwave irradiation using Fmoc strategy



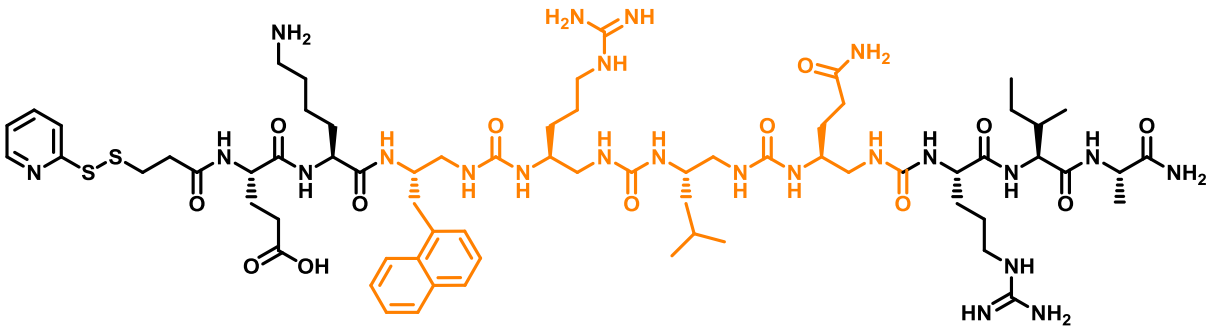
a: Fmoc -Xaa-OH (5 eq), DIC (5 eq), Oxyma (5 eq), 75°C , 25 W, 5 min x 2; b: 20% piperidine, DMF (75°C/25W/1'+3'); c: 3 eq My, 7 eq DIEA, 2 x 2h r.t.; d: Acetic anhydride/ DCM 1:1 (15 min, r.t.); e: TFA/TIS/H_2O (92.5:2.5:2.5, v,v,v)

Scheme 29: Solid phase synthesis of ch5 with the Fmoc-strategy

Oligomer was prepared on a 50 µmol scale. Polystyrene rink amide resin (loading 0.52 mmol/g, 0.05 mmol, 96 mg) was placed in the reaction vessel and pre-swelled with DMF for 1h. All steps were performed under inert atmosphere and microwave irradiation on the *Discover Bio* system (CEM). The vessel was then placed inside the microwave reactor. The temperature was maintained by modulation of power and controlled with a fiber optic sensor. Fmoc-α-Xaa-OH (5 eq relative to the resin loading), DIC (5 eq) and Oxyma (5 eq), were dissolved in DMF. The mixture was added into the reaction vessel manually. The vessel was then irradiated (170W, 75°C, 1 min) + (35W, 90°C, 3 min). The resin was then filtered off and washed with DMF (3 × 3mL). Fmoc

removal was carried out with 20% of piperidine in DMF (2 mL) under microwave irradiation (155 W, 75°C, 15 sec) + (35 W, 90°C, 50 sec). These coupling and reduction steps were monitored with the Kaiser test[8]. The *N*-Fmoc protected succinimidyl carbamate activated monomers (5 equiv. relative to the resin loading) were dissolved into DMF (2 mL) and were added to the reaction vessel (CEM), followed by DIEA (7equiv.). The reaction vessel was then let without microwave irradiation (r.t., 2hrs). A double coupling was performed systematically. The resin was filtered off and washed with DMF (3 × 3mL). Fmoc removal was carried out as described above. These urea-building block coupling and reduction steps were monitored with the chloranil test[9]. The installation of the last amino acids was performed as described in procedure B. Finally, the resin was swelled in a mixture TFA/TIS/H_2O/ (95:2.5:2.5, v/v/v) and let to react for 4hrs. Then, the resin was filtered off, washed with TFA (2 x 2mL) and CH_2Cl_2 (2 x 2 mL) and the filtrate was evaporated under reduced pressure. After precipitation in cold Et_2O, the crude oligomer was analyzed by RP-HPLC and lyophilized. The crude oligomer was purified by preparative RP-HPLC using the 20-50% ACN in 20 min gradient to obtain **ch5** with 21% yield and a final purity ≥ 95% and lyophilized.

NPys-ch5



The compound **NPys-ch5** was synthesized on 100 µmol scale with commercially available Fmoc-α-Xaa-and monomers **M3**, **M6**, **M14** and **M16** following the general procedure C with introduction of the first urea residue in Fmoc-strategy. 3-Nitro-2Pyridine-3-Mercaptopropionic acid (5eq) was coupled in presence of DIC (5eq) and Oxyma (5eq) at r.t., overnight. After microcleavage, and identification of a side produc by ESI-MS, disulfide bond was reduced with PMe_3 (10 eq) in Dioxane/H_2O (70:30) 50°C,50W, 15 min x 2. Finally, 3-Nitro-2-pyridinesulfonyle (1.5 eq) was coupled with 1.1 eq DIEA 1.1 eq 1hr atr.t. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 6%. ESI-MS (*Mw* 1522.9): m/z 507.7 $[M+H]^+$; gradient 20-100% B, 5 min: t_R= 5.35 min.

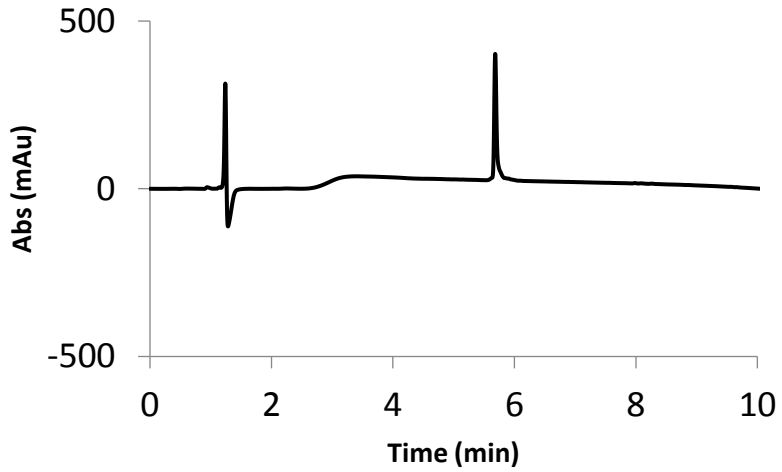
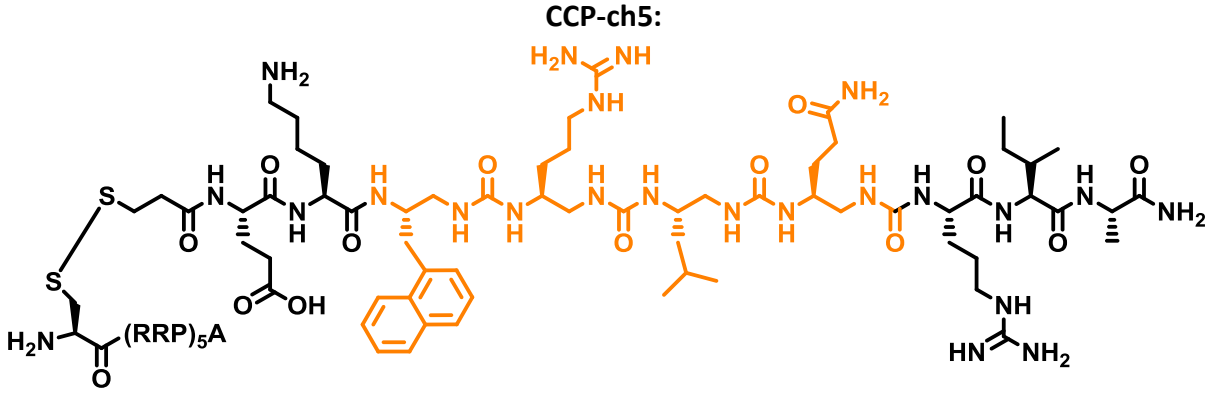


Figure 113: RP-HPLC chromatogram of NPys-ch5 (10-100% ACN in 10 min, 50°C)

CCP-ch5:



4 mg of **NPys-ch5** was dissolved into 100µL of water and C-$(PRR)_5$-A dissolved in 50 µL of water was added. 10 µL of NH_4CO_3 (0.1M) was added to reach a pH of 7. A yellow coloration was observed in solution. The reaction was left overnight at r.t. The desired compound **CCP-ch5** was purified by semi-preparative RP-HPLC (10-50 % ACN in 20min). The product was recovered with a yield of 48%. ESI-MS (*Mw* 3650.5): m/z 523.3 $[M+7H]^{7+}$; gradient 20-100% B, 5 min: t_R= 3.62 min.

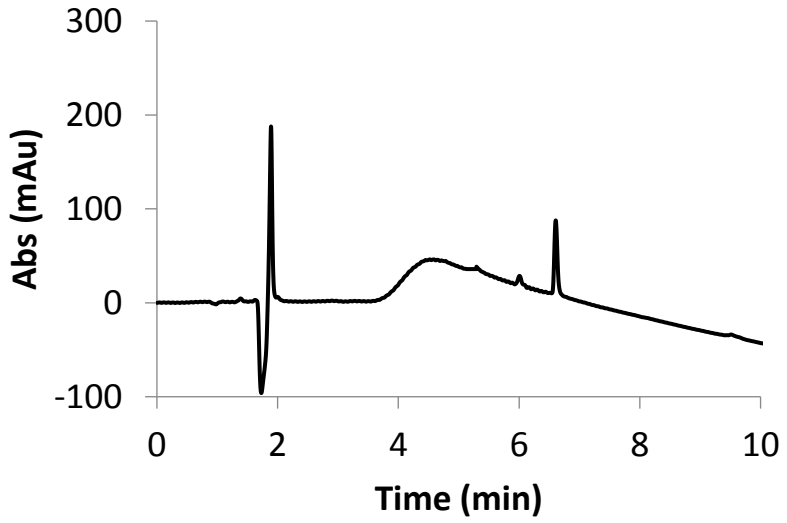


Figure 114: RP-HPLC chromatogram of CCP-ch5 (10-100% ACN in 10 min, 50°C)

ASTE-ch5: ASTEEK-Nal^{u}-Arg^{u}-Leu^{u}-Gln^{u}-RIA-NH_2

The compound **ASTE-ch5** was synthesized on 50 µmol scale on NOVAPEG resin (loading 0.46 mmol.g^{-1}) with commercially available Fmoc-α-Xaa-and monomers **M3**, **M6**, **M14** and **M16** following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 6%. ESI-MS (*Mw* 1756.1): m/z 1756.8 $[M+H]^+$; gradient 20-100% B, 5 min: t_R= 4.11 min.

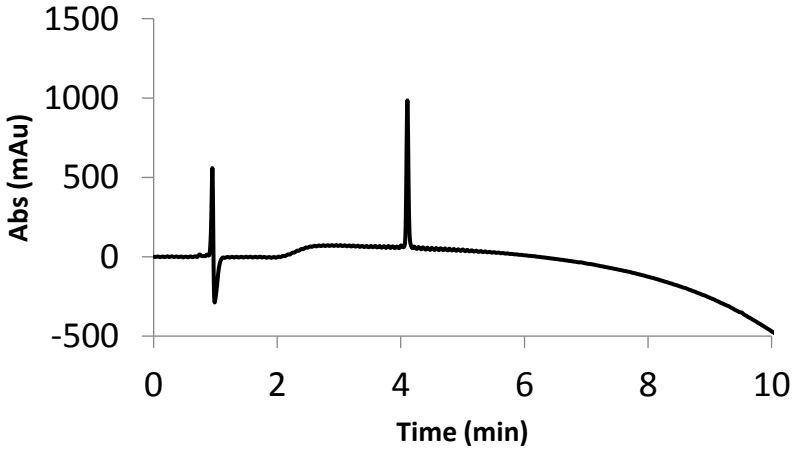


Figure 115: RP-HPLC chromatogram of ASTE-ch5 (10-100% ACN in 10 min, 50°C)

ch6: Ac-EK-HPhe^{u}-R^{u}L^{u}Q^{u}RIA-NH_2

The compound **ch6** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M3**, **M6**, **M14** and **M17** following the general procedure C without introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 2%. ESI-MS (*Mw* 1331.6): m/z 1331.6 $[M+H]^+$, 666.4 $[M+2H]^{2+}$, 444.9 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.82 min.

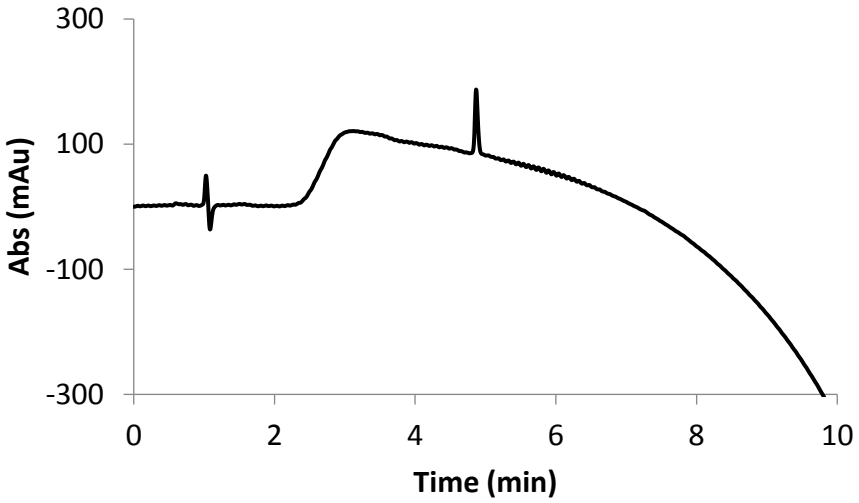


Figure 116: RP-HPLC chromatogram of ch6 (10-100% ACN in 10 min, 50°C)

ch7: Ac-EK^{u}-HPhe^{u}-R^{u}L^{u}Q^{u}RIA-NH_2

The compound **ch7** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M3**, **M6**, **M14**, **M16** and **M8**, following the general procedure C without introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 3%. ESI-MS (*Mw* 1346.7): m/z 1346.6 $[M+H]^+$, 673.0 $[M+2H]^{2+}$; gradient 20-100% B, 5 min: t_R= 4.55 min.

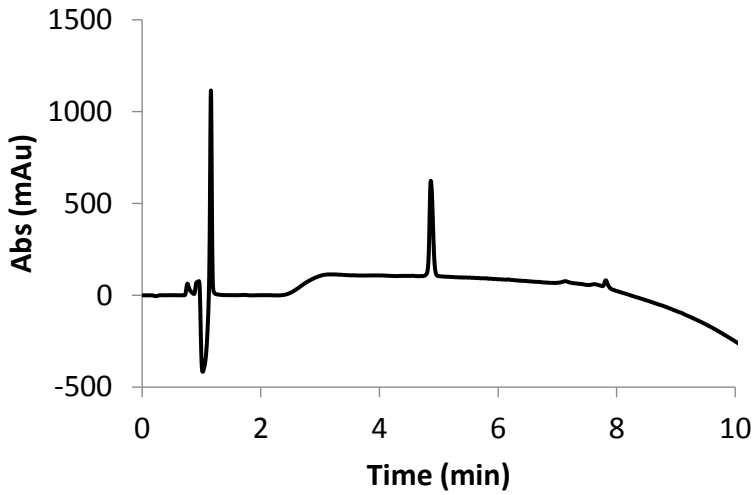


Figure 117: RP-HPLC chromatogram of ch7 (10-100% ACN in 10 min, 50°C)

ch8: Ac-EK-W^{u}-R^{u}L^{u}Q^{u}RIA-NH$_2$

The compound **ch8** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M3**, **M6**, **M14**, and **M15**, following the general procedure C without introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 40% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 4%. ESI-MS (*Mw* 1353.6): m/z 678.8 $[M+2H]^{2+}$, 452.8 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 5.3 min.

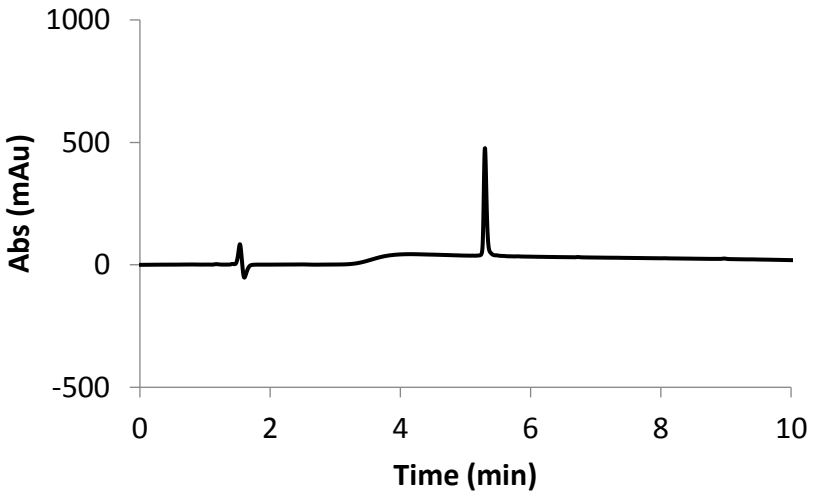


Figure 118: RP-HPLC chromatogram of ch8 (10-100% ACN in 10 min, 50°C)

ch9: Ac-EK-A^{u}-R^{u}L^{u}Q^{u}RIA-NH$_2$

The compound **ch9** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M1**, **M3**, **M6**, and **M14**, following the general procedure C without introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 3%. ESI-MS (*Mw* 1241.5): m/z 1241.6 $[M+H]^+$, 621.4 $[M+2H]^{2+}$; gradient 20-100% B, 5 min: t_R= 3.81 min.

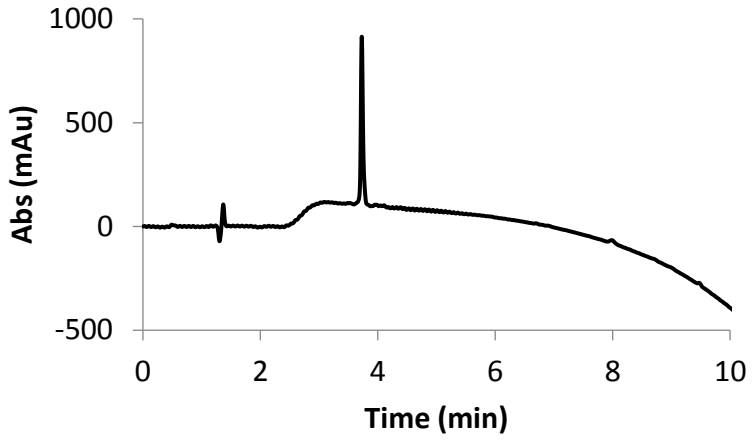


Figure 119: RP-HPLC chromatogram of ch9 (10-100% ACN in 10 min, 50°C)

ch10= ER-Nal^u-R^u-L^u-Q^u-RIA-NH_2

The compound **ch9** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M3**, **M6**, **M11** and **M16**, following the general procedure C without introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 10%. ESI-MS (*Mw* 1395.7): m/z 1395.8 $[M+H]^+$, 698.7 $[M+2H]^{2+}$; gradient 20-100% B, 5 min: t_R= 5.03 min.

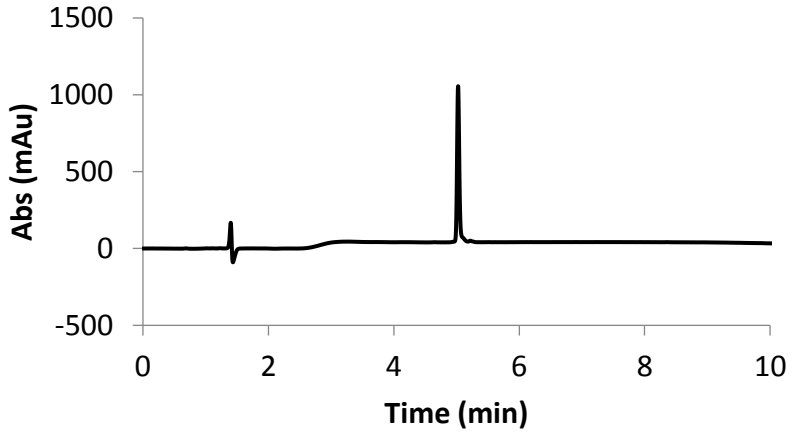


Figure 120: RP-HPLC chromatogram of ch10 (10-100% ACN in 10 min, 50°C)

CPP-ch11= A(RRP)$_5$-C(-S-S-linker)-EK-**Nal^u-Arg^u-Ala^u-Gln^u**-RAA-NH_2

ch11 was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa-and monomers **M1**, **M6**, **M11**, and **M16**, following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The last amino acid, Boc-L-Cys.NPyS (5eq) was coupled with 5 equivalents of DIC (5eq) and Oxyma (5 eq) at r.t. overnight The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. Then, 4 mg of activated **ch11** was dissolved into 100µL of water and C-(PRR)$_5$-A dissolved in 50 µL of water was added. 10 µL of NH_4CO_3 (0.1M) was added to reach a pH of 7. A yellow coloration was observed in solution. The reaction was left overnight at r.t. The desired compound **CCP-ch11** was purified by semi-preparative RP-HPLC (10-50 % ACN in 20min). The product was recovered with a yield of 52% for the conjugation. ESI-MS (*Mw* 3581.3): m/z 1193.6 $[M+3H]^{3+}$, 895.4 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 4.29 min.

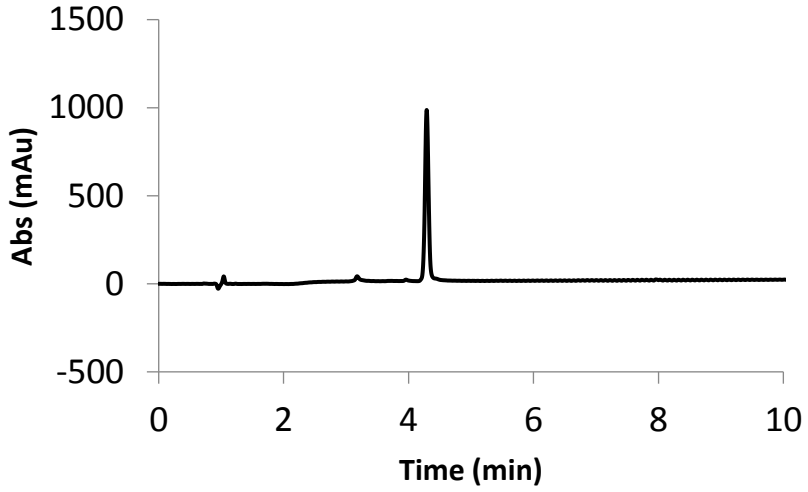


Figure 121: RP-HPLC chromatogram of CPP-ch11 (10-100% ACN in 10 min, 50°C)

ch12= EK-**Nal^u-Arg^u-HPhe^u-Gln^u**-RIA-NH_2

The compound **ch12** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa-and monomers **M6**, **M11**, **M16**, and **M17**, following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 12%. ESI-MS (*Mw* 1415.7): m/z 1416.7 $[M+H]^+$, 708.3 $[M+2H]^{2+}$, 472.5$[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 3.60 min.

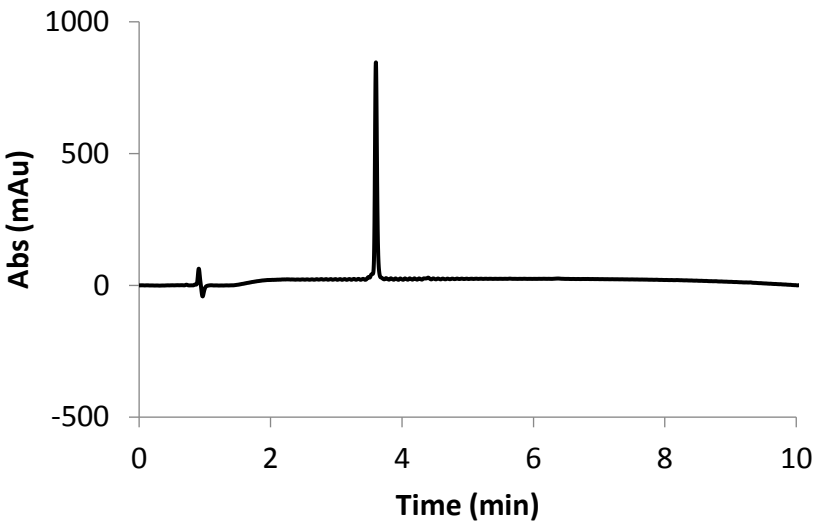


Figure 122: RP-HPLC chromatogram of ch12 (10-100% ACN in 10 min, 50°C)

ch13= EK-Nal^{u}-R^{u}-Nal^{u}-Q^{u}-RIA-NH_2

The compound **ch13** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M6**, **M11**, and **M16**, following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 12%. ESI-MS (*Mw* 1451.7): m/z 1451.7 $[M+H]^+$, 725.5 $[M+2H]^{2+}$, 483.7$[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.14 min.

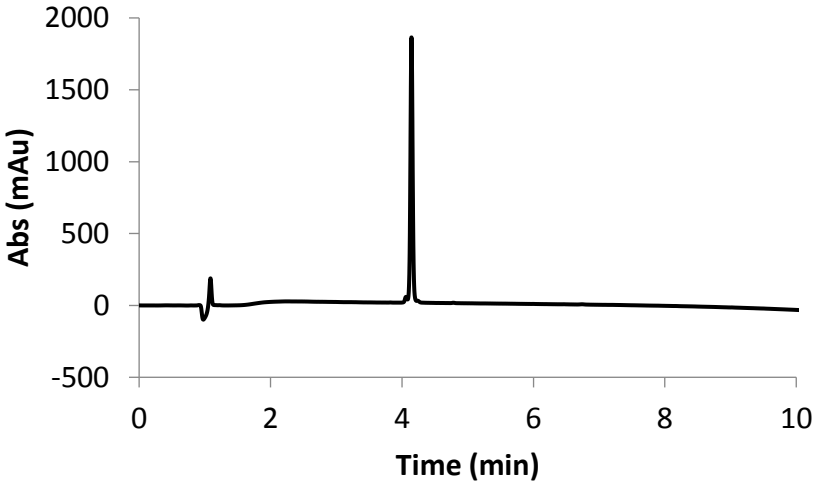


Figure 123: RP-HPLC chromatogram of ch13 (10-100% ACN in 10 min, 50°C)

ch14= EK-Nal^{u}-R^{u}-DiPhe^{u}-Q^{u}-RIA-NH_2

The compound **ch14** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M6**, **M11**, **M16**, and **M18**, following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 15%. ESI-MS (*Mw* 1477.8): m/z 1478.6 $[M+H]^+$, 739.3 $[M+2H]^{2+}$, 493.2$[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.19 min.

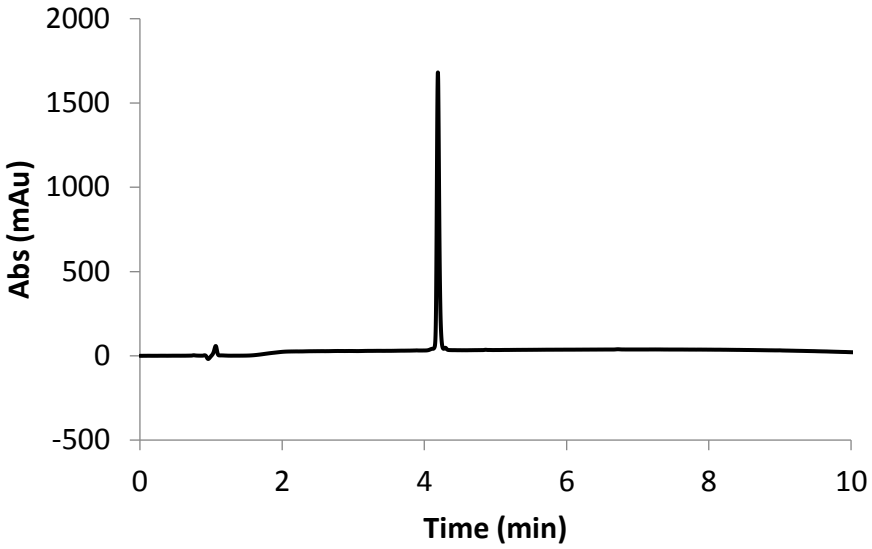


Figure 124: RP-HPLC chromatogram of ch14 (10-100% ACN in 10 min, 50°C)

ch15= EK-Nal^u-R^u-Nle^u-Q^u-RIA-NH_2

The compound **ch15** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomers **M6**, **M11**, **M16**, and **M19**, following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 27%. ESI-MS (*Mw* 1367.7): m/z 1367.9 $[M+H]^+$, 684.5 $[M+2H]^{2+}$, 456.6$[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 3.89 min.

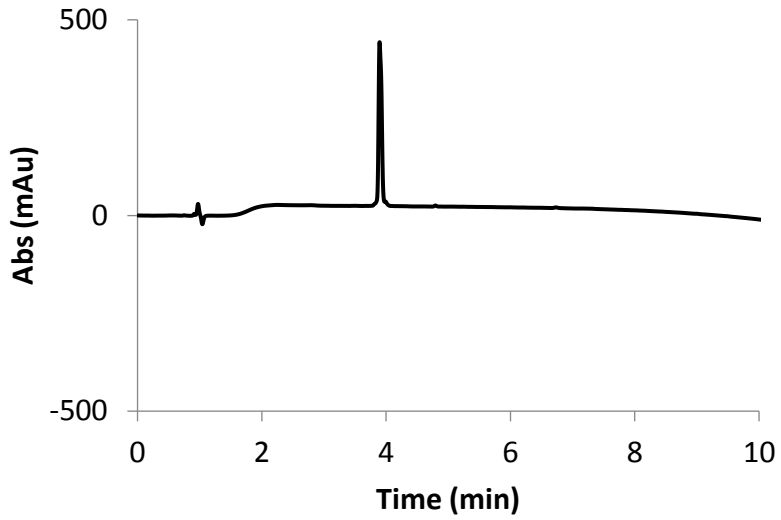


Figure 125: RP-HPLC chromatogram of ch15 (10-100% ACN in 10 min, 50°C)

ch16= $A^uA^uA^u$EKWARLARRIA-NH_2

The compound **ch16** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomer **M20**, following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 25%. ESI-MS (*Mw* 1711.1): m/z 856.5 $[M+2H]^{2+}$, 571.1 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.69 min.

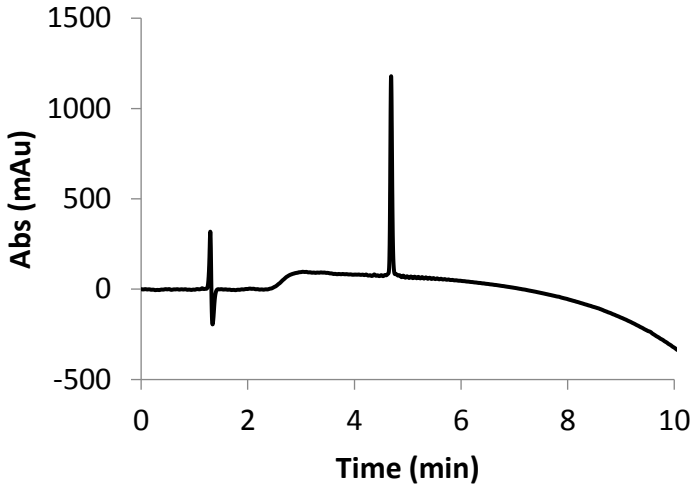


Figure 126: RP-HPLC chromatogram of ch16 (10-100% ACN in 10 min, 50°C)

ch17= $A^uA^uA^uA^u$EKWARLARRIA-NH_2

The compound **ch17** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomer **M20**, following the general procedure C with introduction of the first urea residue

in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 10%. ESI-MS (*Mw* 1811.1): m/z 906.5 $[M+2H]^{2+}$, 604.6 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.72 min.

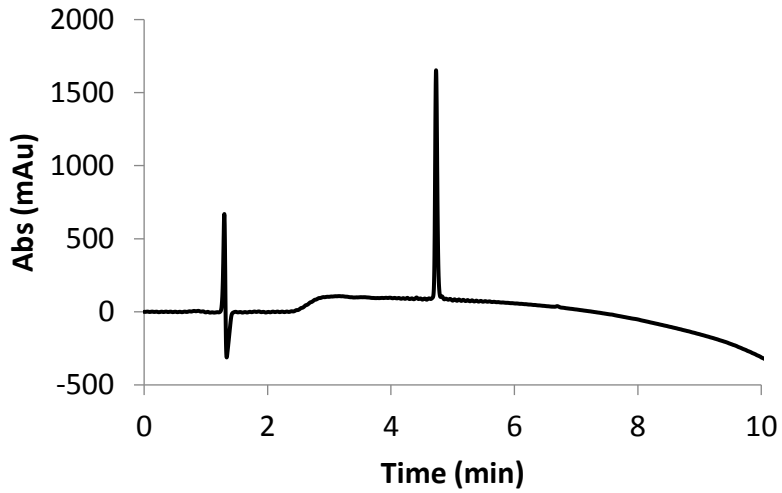


Figure 127: RP-HPLC chromatogram of ch17 (10-100% ACN in 10 min, 50°C)

ch18= $A^uA^uA^uA^u$KWARLARRIA-NH_2

The compound **ch18** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomer **M20**, following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 25%. ESI-MS (*Mw* 1682.1): m/z 842.0 $[M+2H]^{2+}$, 561.6 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 4.94 min.

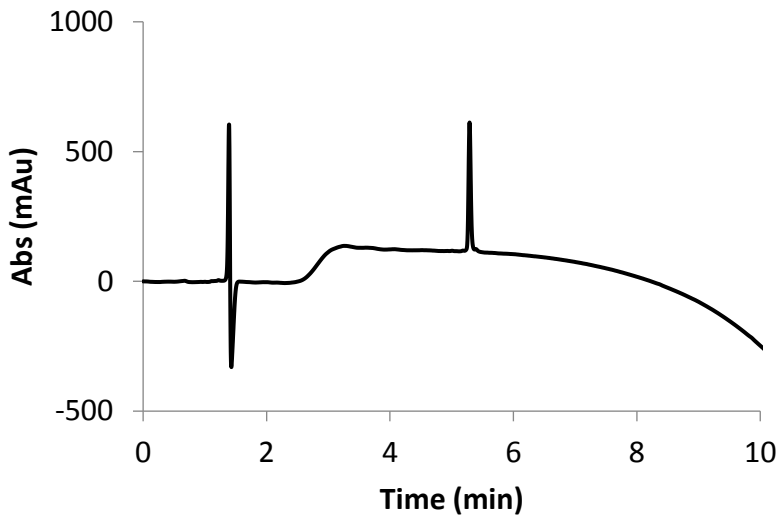


Figure 128: RP-HPLC chromatogram of ch18 (10-100% ACN in 10 min, 50°C)

ol1 : Ip-K^{u}W^{u}R^{u}L^{u}Q^{u}R^{u}I^{u}-NH_2

The compound **ol1** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa- and monomer **M20**, following the general procedure C with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 13%. ESI-MS (*Mw* 1285.9): m/z 1286.6 $[M+2H]^{2+}$, 643.8 $[M+3H]^{3+}$, 429.5 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 4.32 min.

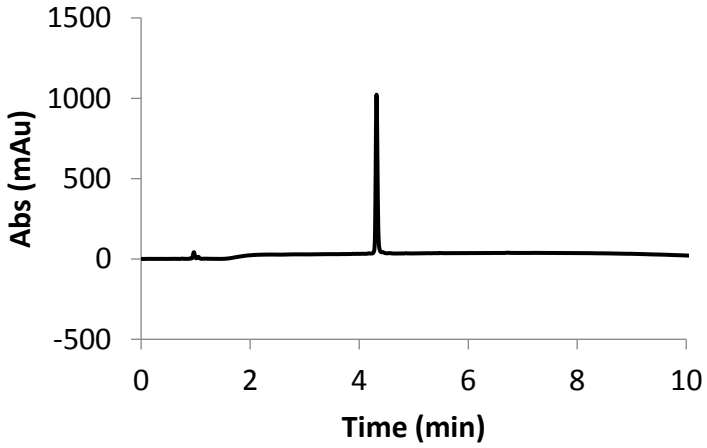


Figure 129: RP-HPLC chromatogram of ol1 (10-100% ACN in 10 min, 50°C)

3. Solid phase synthesis of oligomers in Chapter IV

General procedure D for solid phase oligomers synthesis under microwave irradiation using azide strategy

Oligomers were prepared on a 50 to 100 µmol scale. Fmoc-Gly-WANG resin (loading 0.30 mmol/g, 0.05mmol) was placed in the reaction vessel, and pre-swelled with DMF for 1-2h. For amino acids at the C-terminus, all coupling steps were performed on the *Liberty Blue* system (CEM). Fmoc-α-Xaa-OH (5 eq relative to the resin loading), DIC (5 eq) and Oxyma (5 eq) were dissolved in DMF. The mixture was added into the reaction vessel through *Liberty Blue* system and was irradiated under microwave (170W, 75°C, 15sec) + (35W, 90°C,110sec) for all Fmoc-α-Xaa-OH. All the coupling steps were performed once except for the Arg (two times). The Fmoc-Val-Thr(psiMe-MePro)-OH (4 eq) was introduced with HATU (3,8 eq)/DIEA (0,8 eq) at 75°C/25 W for 20 minutes and the Fmoc-Asp(tBu)-(Dmb)Gly-OH (4 eq) with HATU (3,8 eq) and DIEA (8 equiv.) at 75°C/25 W for 5 minutes twice. The resin was then filtered off and washed with DMF. Fmoc removal was carried out with 20% of piperidine in DMF under microwave irradiation (155W, 75°C, 15 sec) + (35W, 90°C, 50 sec). Regarding the azide part, the same reaction vessel was installed inside the microwave reactor CEM *Discover Bio*. The first urea residue (Fmoc-Gln(Trt)-OSu) **M11** was introduced in Fmoc-strategy with DIEA (7eq) at r.t. for 2h twice. Fmoc removal was performed as described before. Otherwise, the azide protected succinimidyl carbamate activated monomer (1.5 eq relative to the resin loading) were dissolved into DMF (1.5 mL) and were added to the reaction vessel (CEM), followed by DIEA (3eq). The reaction vessel was then irradiated under microwave (50°C, 50W, 15min). A double coupling was performed systematically. The resin was filtered off and successively washed with DMF (3 × and two times with a mixture of 1,4-dioxane:H_2O (7:3 v/v) in order to perform the azide reduction (3 × 3mL). Azide reduction was performed under microwave irradiation (50°C, 50W, 15 min) in the mixture 1,4- dioxane:H_2O (1.5mL) with 1M PMe_3 solution in THF (10 equiv. relative to the resin loading) as reducing agent. This reduction step was systematically performed twice. The resin was then filtered off and washed with DMF (4 × 3 mL). These coupling and reduction steps were monitored with the chloranil test[9]. To couple the remaining amino acids, the reaction vessel was let inside the microwave reactor (CEM Discover Bio). Fmoc-α-Xaa-OH (5 eq relative to the resin loading), DIC (5 eq) and Oxyma (5 eq), were dissolved in DMF. The resulting mixture was manually added to the reaction vessel. The vessel was then irradiated (75°C,25 W, 5 min) once for coupling. The resin was filtered off and washed with DMF (3 × 3 mL). Fmoc removal was carried out with 20% of piperidine in DMF (2mL) under microwave irradiation (75°C, 155W, 15 sec) + (90°C, 35W, 50 sec). After final washing with CH_2Cl_2, the resin was swelled in a mixture TFA/TIS/H_2O (95:2.5:2.5, v/v/v) and let to react for 4hrs under slight shaking. Then, the resin was filtered off, washed with TFA (2 x 2mL) and CH_2Cl_2 (2 x 2mL) and the filtrate was evaporated under reduced pressure. After precipitation in cold Et_2O, the crude oligomer was analyzed on RP-HPLC and lyophilized. The crude oligomer was purified by preparative RP-HPLC using the appropriate gradient to a final purity ≥ 95% and lyophilized.

p3: ASTEEK WARLARRIAGAGGVTLDGFG

The compound **p3** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa, following the general procedure D without the introduction of urea residue. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 10%. ESI-MS (*Mw* 2690.0): m/z 1345.6 $[M+2H]^{2+}$, 897.5 $[M+3H]^{3+}$, 673.5 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 4.71 min.

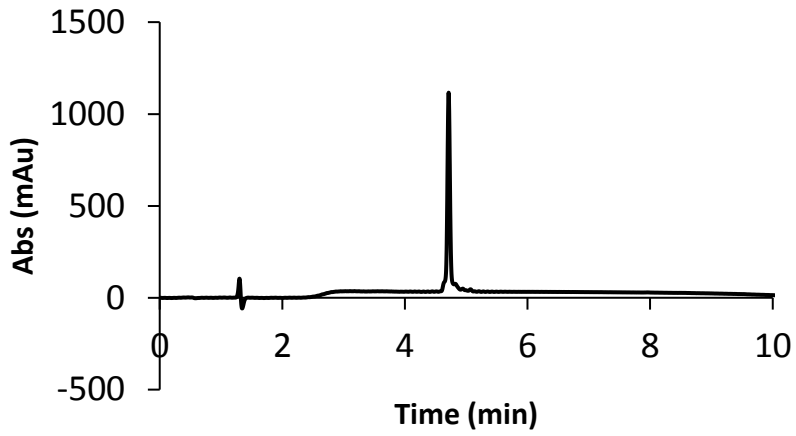


Figure 130: RP-HPLC chromatogram of p3 (10-100% ACN in 10 min, 50°C)

p11: ASTEEK WARLARRI-Ava-DAP^{u}-VTLDGFG

The compound **p11** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa, following the general procedure D with manual introduction of **M21** (5 eq) with DIEA (7eq) at r.t. 15 min x 3 and **M22** (5 eq) with BOP (5 eq) and DIEA (7 eq) at r.t. 20 min x 2. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 9. ESI-MS (*Mw* 2689.1): m/z 672.3 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 5.65 min.

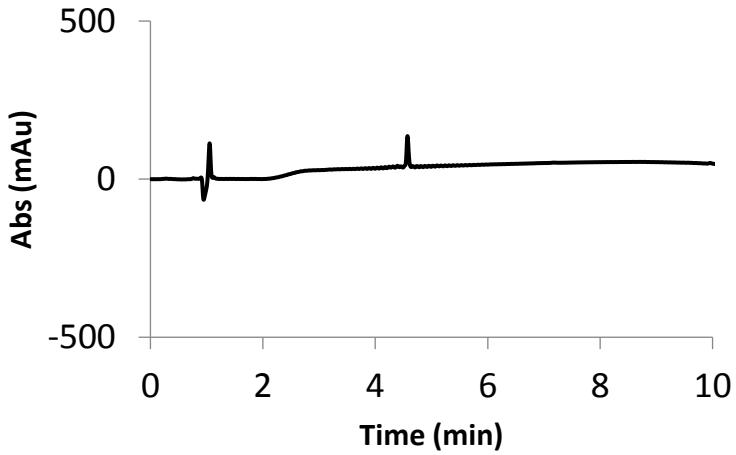


Figure 131: RP-HPLC chromatogram of p11 (10-100% ACN in 10 min, 50°C)

p12: ASTEEK WARLARRI-Glu-γAbu-VTLDGFG

The compound **p12** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa and **M23**, following the general procedure D. Fmoc-GABA-OH (5 eq) and **M23** (5eq) were coupled manually with BOP(6eq)/DIEA (6eq). The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 19%. ESI-MS (*Mw* 2389.7): m/z 1195.5 $[M+2H]^{2+}$, 797.3 $[M+3H]^{3+}$, 598.2 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 3.35 min.

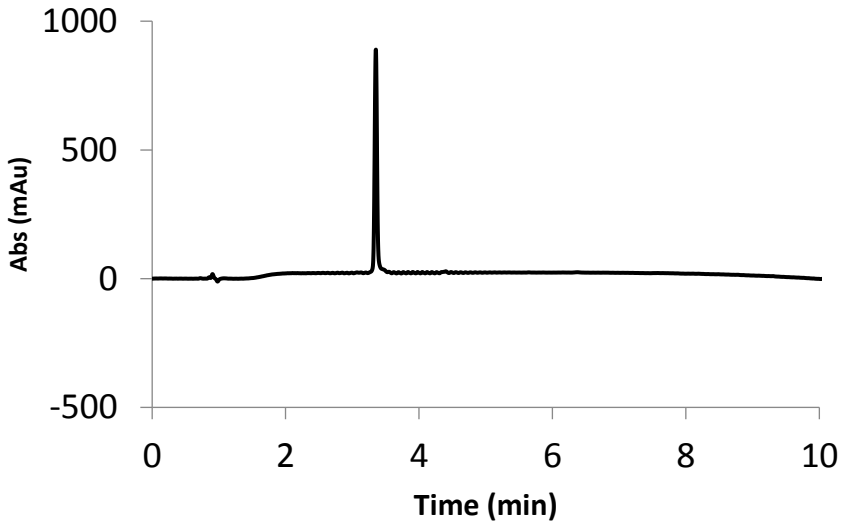


Figure 132: RP-HPLC chromatogram of p12 (10-100% ACN in 10 min, 50°C)

p13: ASTEEK WARLARRI-Asp-γAbu-VTLDGFG

The compound **p13** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa and **M24**, following the general procedure D. Fmoc-GABA-OH (5 eq) and **M24** (5eq) were coupled manually with BOP(6eq)/DIEA (6eq).The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 14%. ESI-MS (*Mw* 2375.7): m/z 1188.6 $[M+2H]^{2+}$, 792.8 $[M+3H]^{3+}$, 594.8 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 3.75 min.

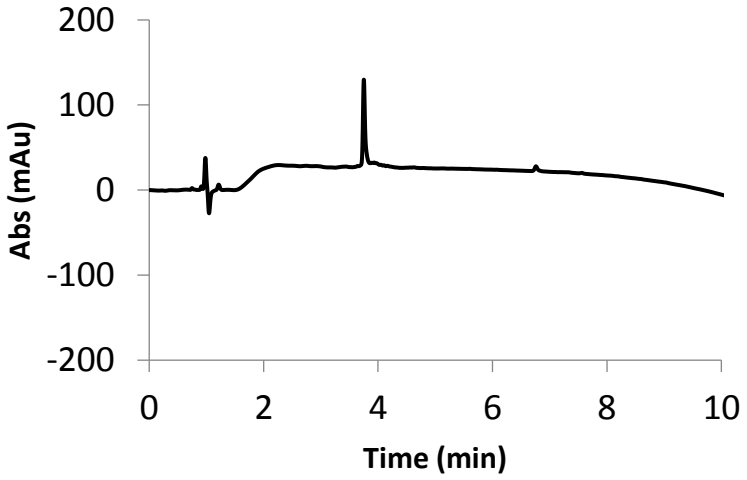


Figure 133: RP-HPLC chromatogram of p13 (10-100% ACN in 10 min, 50°C)

ch19: EK-Nal^u-Arg^u-Leu^u-Gln^u-RIAGAGGVTLDGFG

The compound **ch19** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa and monomers **M3**, **M6**, **M11**, and **M16**, following the general procedure D with introduction of the first urea residue in Fmoc-strategy. The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 60% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 6%. ESI-MS (*Mw* 2299.7): m/z 767.8 $[M+3H]^{3+}$; gradient 20-100% B, 5 min: t_R= 5.41 min.

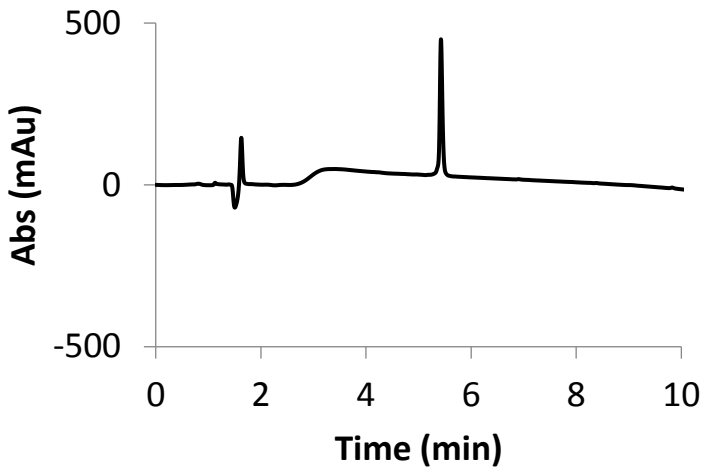


Figure 134: RP-HPLC chromatogram of ch19 (10-100% ACN in 10 min, 50°C)

ch20: EK-Nal^{u}-Arg^{u}-Leu^{u}-Gln^{u}-RI-Glu(-γAbu-LDGFG)

The compound **ch20** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa including Fmoc-GABA-OH (*Iris BIOTECH GMBH*) and monomers **M3**, **M6**, **M11**, **M16** and **M23**, following the general procedure D with introduction of the first urea residue in Fmoc-strategy. Fmoc-GABA-OH (5 eq) and **M23** (5eq) were coupled manually with BOP(6eq)/DIEA (6eq). The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 13%. ESI-MS (*Mw* 2000.3): m/z 1001.1 $[M+2H]^{2+}$, 667.7 $[M+3H]^{3+}$, 501.1 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 4.13 min.

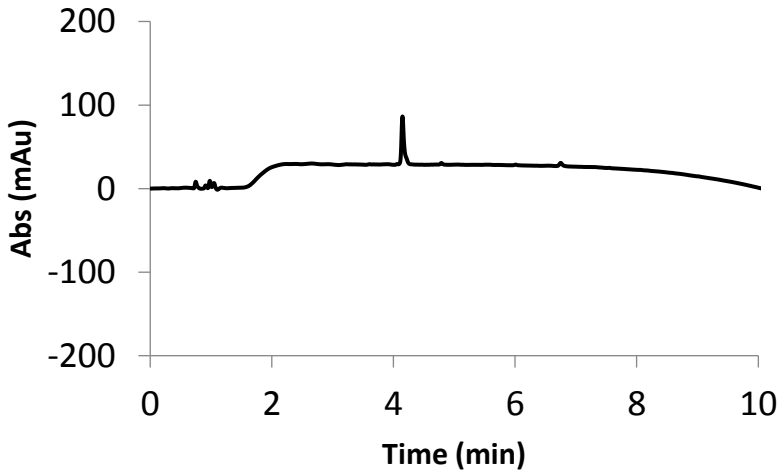


Figure 135: RP-HPLC chromatogram of ch20 (10-100% ACN in 10 min, 50°C)

ch21: EK-Nal^{u}-Arg^{u}-Leu^{u}-Gln^{u}-RI-Asp(-γAbu-LDGFG)

The compound **ch21** was synthesized on 50 µmol scale with commercially available Fmoc-α-Xaa including Fmoc-GABA-OH (*Iris BIOTECH GMBH*) and monomers **M3**, **M6**, **M11**, **M16** and **M24**, following the general procedure D with introduction of the first urea residue in Fmoc-strategy. Fmoc-GABA-OH (5 eq) and **M24** (5eq) were coupled manually with BOP(6eq)/DIEA (6eq). The desired product was purified by preparative RP-HPLC: H_2O (0.1% TFA), ACN (0.1% TFA); a gradient 20 to 50% of B in 20 min with a flow of 20 mL/min was used. The product was recovered with an overall yield of 13%. ESI-MS (*Mw* 1986.3): m/z 994.1 $[M+2H]^{2+}$, 663.1 $[M+3H]^{3+}$, 497.5 $[M+4H]^{4+}$; gradient 20-100% B, 5 min: t_R= 4.16 min.

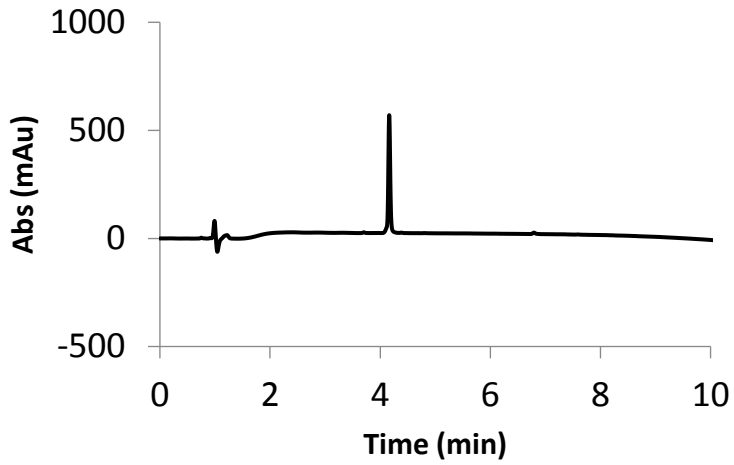


Figure 136: RP-HPLC chromatogram of ch21 (10-100% ACN in 10 min, 50°C)

D. Affinity experiments performed at CEA Saclay

Protein production and purification

Recombinant $(His)_6$ – GST – Tev site – hAsf1a 1-156 fusion proteins were produced from the pETM30 plasmid. Unlabelled and ^{15}N uniformly labeled proteins were obtained as described in by Mousson *et al.*[10] Briefly, soluble $(His)_6$-tagged GST fusion proteins were first purified on GSH agarose (Sigma). After a Tev site cleavage, the (His)6-tagged TEV protease used and the $(His)_6$-tagged GST were trapped in a Ni-NTA agarose column (Qiagen). The flow through fraction containing Asf1 proteins was finally purified by anion exchange chromatography using a Resource Q 6 mL column (GE Healthcare). Asf1 proteins were diafiltrated against a 50 mM Tris-HCl pH 7.4 storage buffer and concentrated using an Amicon device (Millipore) and an YM10 regenerated cellulose membrane (Millipore). The storage buffer was further replaced by NMR buffer (below) using the same system. Protein and peptide concentrations were determined by UV spectrophotometry (NanoDrop, Thermo Scientific) for molecules containing tryptophan residue (absorption at 280 nm) or Naphthalene containing residue (absorption at 277 nm). For peptides

missing any aromatic residues, the concentration was determined by mass spectrometry (S. Dubois, CEA-Saclay).

Isothermal Titration Calorimetry (ITC)

All IsoThermal Calorimetry (ITC) experiments were performed in a VP-ITC titration calorimeter (Malvern). All experiments were carried out at 20°C, in a 50 mM Tris-HCl pH7.4 buffer with Asf1 proteins in the sample cell and peptides/foldamers in the syringe. Protein and ligand concentrations varied between 5 µM – 15 µM and 50 µM - 300 µM, respectively, with a protein/molecule ratio ranging from 1:10 to 1:20 depending on binding affinities. Protein and molecule samples were prepared in the same buffer and extensively degassed (ThermoVac, Malvern) before loading on instrument. After equilibrating the cell at 20°C, the rotating syringe (310 rpm) injected during 12 or 20 sec ligand solution aliquots of 6 or 10 µL into Asf1 protein solution at intervals of 280 sec until saturation was observed. Raw ITC data were processed with the Origin 7.0 Software (OriginLab, Malvern) using the One-Set of sites fitting model. The sigmoidal curves obtained allowed the determination of the following parameters: stoichiometry (*N*), association constant (K_a) (K_D was deduced as $K_D = 1/K_a$) and the change in enthalpy (ΔH). The standard molar Gibbs energy change (ΔG) and standard molar entropic contribution ($T\Delta S$) were then calculated using the relationships $\Delta G = -RTln(KA) = T\Delta s - \Delta H$, where *R* is the gas constant (8.314472 $J.K^{-1}mol^{-1}$) and *T* is the temperature in Kelvin (293.15 K). All ITC experiments were performed in duplicate.

Nuclear magnetic resonance (NMR) binding experiments

NMR experiments were carried out in a Bruker DRX-600 spectrometer equipped with a cryoprobe. Interaction experiments and resonance assignments were adapted from (Mousson *et al.*[10]. All NMR experiments were performed at a protein concentration of 50 µM in NMR buffer (10 mM Tris-HCl pH 7.4, 0.1 mM EDTA, 0.1 mM DSS, 0.1 mM NaN_3, protease inhibitors (Roche), 10% D_2O). Heteronuclear single quantum correlation (HSQC) spectra of the uniformly ^{15}N labeled Asf1 proteins were recorded at 298°K. For titration experiments, inhibitory molecule aliquots at a concentration of 13 µM were added successively and HSQC spectra were recorded after each addition until saturation was observed. Proton chemical shifts (in ppm) were referenced relative to internal DSS and ^{15}N reference was set indirectly relative to DSS using frequency ratios[11]. NMR data were processed using Topspin (Bruker) and analyzed using Sparky (T.D. Goddard and D.G. Kneller, UCSF). Chemical shift variation upon binding was calculated as $\Delta\delta = [\left(\delta_{HN}^{f} - \delta_{HN}^{b}\right)^2 + (0.17 \times \left(\delta_{N}^{f} - \delta_{N}^{b}\right)^2]^{1/2}$, where *f* and *b* refer to free and bound forms respectively. The scaling factors normalize the magnitude of the ^1H and ^{15}N chemical shift changes (in ppm)[12]. Assignments of HSQC spectra of Asf1 bound to inhibitory foldamers were obtained by following progressive variations of chemical shifts upon titration.

Crystallization, data collection and structure determination

Asf1-**ch5** were mixed and the complex was concentrated to 8.6 mg/mL in a buffer 50 mM Tris-HCl pH7.4. Crystals of the complex were grown by sitting drop vapor diffusion at 20 °C against reservoir solution containing 100 mM Citrate pH 4.2, 300 mM LiSO4, 26% PEG3350. Crystals were soaked in a 100 mM Citrate pH 4.2, 300 mM $LiSO_4$, 26% PEG3350, 20% Glycerol cryo-protectant solution before being flash-frozen in liquid nitrogen. Diffraction data were collected on the PROXIMA-1 beamline at the synchrotron SOLEIL (Saint Aubin, France) at a temperature of 100 K with X-ray wavelength of 0.97857Å. Diffraction images recorded with PILATUS 6 M detector were processed using the XDS package[13]. Best data sets were obtained from two crystals of Asf1-**ch5** belonging to space group P1 which diffracted up to 1.8Å resolution. The two data sets were merged and scaled and the structure of Asf1-ch5 complex was determined by molecular replacement using MOLREP[14] with the human Asf1a structure (chain A of PDB entry 2I32) as model probe. Best solution contained two complexes per asymmetric unit. Model building was performed with Coot[15] and structure refinement was achieved with BUSTER version 2.10 (Cambridge, United Kingdom). Final refinement statistics are presented in Table 25. Structure representations presented in all figures were drawn with PYMOL software (Schrödinger).

Asf1-ch5 complex	
Data collection	
Space group	*P1*
Cell dimensions	
a. *b*. *c* (Å)	36.73. 51.28. 56.08
a. b. g (°)	112.38. 104.87. 101.190
Resolution (Å)	48.29-1.8 (1.91-1.8)
R_{merge}	8.4(66.6)
I / s*I*	9.41(2.14)
Completeness (%)	96.1(70.5)
Redundancy	3.9(2.6)
Refinement	
Resolution (Å)	1.8
No. unique reflections	31049
R_{work} / R_{free} (%)	17.64/20.72
No. atoms	
Protein	2576
Ligand/ion	369
Water	180
B-factors	
Protein	35.2
Ligand	47.1
Ion/small molecules	87.8
Water	44.3
R.m.s. deviations	
Bond lengths (Å)	0.010
Bond angles (°)	1.10

*Number of xtals for the structure: 2.
wavelength= 0.978570
*Values in parentheses are for highest-resolution shell.

Table 25: Refinement statistics of ch5-Asf1

Bibliography

(1) Douat-Casassus, C.; Pulka, K.; Claudon, P.; Guichard, G. Microwave-Enhanced Solid-Phase Synthesis of N , N′ -Linked Aliphatic Oligoureas and Related Hybrids. *Org. Lett.* **2012**, *14* (12), 3130–3133.

(2) Zhang, Z.; Fan, E. Solid-Phase and Solution-Phase Syntheses of Oligomeric Guanidines Bearing Peptide Side Chains. *J. Org. Chem.* **2005**, *70* (22), 8801–8810.

(3) Collie, G. W.; Pulka-Ziach, K.; Lombardo, C. M.; Fremaux, J.; Rosu, F.; Decossas, M.; Mauran, L.; Lambert, O.; Gabelica, V.; Mackereth, C. D.; et al. Shaping Quaternary Assemblies of Water-Soluble Non-Peptide Helical Foldamers by Sequence Manipulation. *Nat. Chem.* **2015**, *7* (11), 871–878.

(4) Rodriguez, M.; Llinares, M.; Doulut, S.; Heitz, A.; Martinez, J. A Facile Synthesis of Chiral N-Protected β-Amino Alcohols. *Tetrahedron Lett.* **1991**, *32* (7), 923–926.

(5) Sureshbabu, V. V.; Naik, S. A.; Hemantha, H. P.; Narendra, N.; Das, U.; Guru Row, T. N. N -Urethane-Protected Amino Alkyl Isothiocyanates: Synthesis, Isolation, Characterization, and Application to the Synthesis of Thioureidopeptides. *J. Org. Chem.* **2009**, *74* (15), 5260–5266.

(6) Sharma, K. S.; Durand, G.; Giusti, F.; Olivier, B.; Fabiano, A.-S.; Bazzacco, P.; Dahmane, T.; Ebel, C.; Popot, J.-L.; Pucci, B. Glucose-Based Amphiphilic Telomers Designed to Keep Membrane Proteins Soluble in Aqueous Solutions: Synthesis and Physicochemical Characterization. *Langmuir* **2008**, *24* (23), 13581–13590.

(7) Chakrabartty, A.; Baldwin, R. L. Stability of Alpha-Helices. *Adv. Protein Chem.* **1995**, *46*, 141–176.

(8) Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I. Color Test for Detection of Free Terminal Amino Groups in the Solid-Phase Synthesis of Peptides. *Anal. Biochem.* **1970**, *34* (2), 595–598.

(9) Vojkovsky, T. Detection of Secondary Amines on Solid Phase. *Pept. Res.* **1995**, *8* (4), 236–237.

(10) Mousson, F.; Lautrette, A.; Thuret, J.-Y.; Agez, M.; Courbeyrette, R.; Amigues, B.; Becker, E.; Neumann, J.-M.; Guerois, R.; Mann, C.; et al. Structural Basis for the Interaction of Asf1 with Histone H3 and Its Functional Implications. *Proc. Natl. Acad. Sci.* **2005**, *102* (17), 5975–5980.

(11) Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D. 1H, 13C and 15N Random Coil NMR Chemical Shifts of the Common Amino Acids. I. Investigations of Nearest-Neighbor Effects. *J. Biomol. NMR* **1995**, *5* (1), 67–81.

(12) Farmer, G.; Colgan, J.; Nakatani, Y.; Manley, J. L.; Prives, C. Functional Interaction between P53, the TATA-Binding Protein (TBP), AndTBP-Associated Factors in Vivo. *Mol. Cell. Biol.* **1996**, *16* (8), 4295–4304.

(13) Kabsch, W. Integration, Scaling, Space-Group Assignment and Post-Refinement. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66* (Pt 2), 133–144.

(14) Vagin, A.; Teplyakov, A. Molecular Replacement with MOLREP. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66* (Pt 1), 22–25.

(15) Emsley, P.; Cowtan, K. Coot: Model-Building Tools for Molecular Graphics. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60* (Pt 12 Pt 1), 2126–2132.