



**UNIVERSITÉ DE STRASBOURG**

**ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES – ED222**

**Institut de Science et d'Ingénierie Supramoléculaires – UMR 7006**

**THÈSE** présentée par :

**Régis BOEHRINGER**

soutenue le : **30 janvier 2018**

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

Discipline/ Spécialité : Chimie

**Synthèse chimique de protéines pour  
l'étude structurale et fonctionnelle de  
fibres amyloïdes**

**THÈSE dirigée par :**

**Monsieur TORBEEV Vladimir**

Maître de conférences, HDR, université de Strasbourg

**RAPPORTEURS :**

**Monsieur BODE Jeffrey**

Professeur, ETH Zurich

**Monsieur BOTURYN Didier**

Directeur de recherche, CNRS, université de Grenoble 1

---

**AUTRES MEMBRES DU JURY :**

**Monsieur TORBEEV Vladimir**

Maître de conférences, HDR, université de Strasbourg

**Monsieur KIEFFER Bruno**

Professeur, université de Strasbourg



## Acknowledgements

Je voudrais tout particulièrement et avant tout remercier mon superviseur Dr. Vladimir Torbeev de m'avoir fait l'honneur de m'accueillir au sein de son équipe. Je suis heureux d'avoir pu travailler à ses côtés. Nos discussions, nos échanges et nos aventures scientifiques ont été très enrichissants d'un point de vue personnel et professionnel. Son soutien, ses encouragements, sa confiance et ses précieux conseils m'ont permis d'avancer grandement dans mes projets. Merci de m'avoir permis d'apprendre tant de choses et de m'avoir aidé à repousser mes limites.

Je remercie Dr. Patrick Schultz, Dr. Jésus Raya, Dr. Alicia Schirer, Dr. Petra Hellwig, Corinne Crucifix, Dr. Bruno Vincent et Dr. Jean-Louis Schmitt pour leur aide, leurs conseils et leur collaboration.

Je tiens également à remercier mon collègue de projet, Dr. Jérémy Ruiz, pour les recherches que nous avons menées ensemble, pour les bons moments, les discussions et ses bons conseils. Je remercie également Dr. Boris Schmidtgal pour nos échanges toujours très passionnés, pour son expertise et sa bienveillance. Je remercie Dr. Robrecht Vergauwe et Dr. Abhishek Baral pour leur gentillesse et pour nos discussions scientifiques et moins scientifiques. Un grand merci à Annia Bertrand et Dr. Corinne Baehr d'avoir été là pour moi quand j'en avais besoin.

Un remerciement spécial pour mes amis thésards, Elise Naudin, Aromal Asokan et Valentin Bauer. Nous avons vraiment passé de très bons moments ensemble. Nos aventures, nos délires, nos péripéties scientifiques ont énormément comptés pour moi. Merci pour l'ambiance de travail que l'on a pu créer ensemble, pour votre aide, votre soutien et votre amitié.

Merci aux autres membres du laboratoire, passés et présents, et particulièrement à Claire Meyer et Marcel Grogg qui m'ont beaucoup aidé dans mes recherches.

Je tiens à remercier ma famille, mes parents et ma sœur Julie, Stéphane, Elina et Alicia. Vous êtes tout pour moi et je vous dois tout.

Merci à Sébastien, Adriano, Hervé, Ryan et Salen, vous êtes ma deuxième famille. Merci d'avoir toujours été toujours là pour moi.

## **Table of contents**

Table of tables.....	7
Table of figures.....	8
Abbreviations.....	10
Résumé de thèse.....	15
<b>1. Introduction .....</b>	<b>26</b>
<b>1.1 Amyloid related diseases.....</b>	<b>28</b>
1.1.1 Mechanisms of amyloid formation.....	29
1.1.2 Prion diseases .....	31
1.1.3 Functional amyloids .....	32
<b>1.2 The structural properties of amyloid fibrils.....</b>	<b>33</b>
<b>1.3 Structural polymorphism in amyloid fibrils .....</b>	<b>35</b>
<b>1.4 Inhibition of amyloid formation.....</b>	<b>37</b>
<b>1.5 Protein synthesis .....</b>	<b>38</b>
1.5.1 Solution-phase peptide synthesis.....	38
1.5.2 Recombinant protein synthesis.....	39
1.5.3 Solid-phase peptide synthesis.....	41
1.5.4 The chemical ligation of peptides.....	43
<b>1.6 Thesis outline .....</b>	<b>49</b>
<b>2. Chiral recognition in amyloid fiber growth .....</b>	<b>50</b>
<b>2.1 Introduction .....</b>	<b>50</b>
<b>2.2 Chemical synthesis and properties of covalently-linked constructs based on the [20-41]- fragment of <math>\beta</math>2-microglobulin .....</b>	<b>53</b>
<b>2.3 Chiral recognition in the growth of amyloid fibers from the constructs based on the [20- 41]-fragment of <math>\beta</math>2-microglobulin.....</b>	<b>59</b>
<b>2.4 Chemical synthesis and properties of full-length <math>\beta</math>2-microglobulin enantiomers .....</b>	<b>62</b>
<b>2.5 Chiral recognition during amyloid growth of human <math>\beta</math>2-microglobulin enantiomers.....</b>	<b>64</b>
<b>2.6 Discussion.....</b>	<b>68</b>
<b>2.7 Conclusions .....</b>	<b>72</b>
<b>3. Self-assembly of distinct amyloid structures.....</b>	<b>73</b>

<b>3.1</b>	<b>Introduction .....</b>	<b>73</b>
<b>3.2</b>	<b>Objectives .....</b>	<b>75</b>
<b>3.3</b>	<b>The covalent approach .....</b>	<b>76</b>
3.3.1	Synthesis of the covalent trimer of [20-41] $\beta$ 2m .....	76
3.3.2	Study and comparison of the amyloidogenicity of the different building blocks .....	77
3.3.3	Kinetic studies of amyloid growth .....	80
3.3.4	Characterization of [20-41] $\beta$ 2m covalent trimer amyloids by cryo-EM .....	82
3.3.5	Characterization of [20-41] $\beta$ 2m covalent trimer amyloids by solid state NMR .....	82
3.3.6	Characterization of [20-41] $\beta$ 2m covalent trimer amyloids by FTIR .....	83
<b>3.4</b>	<b>Supramolecular approach for controlling amyloid polymorphism .....</b>	<b>85</b>
3.4.1	Synthesis of the different [20-41] $\beta$ 2m-conjugates .....	85
3.4.2	Study of [20-41] $\beta$ 2m-conjugates with bulky hydrophobic groups .....	85
3.4.3	Study of a library of [20-41] $\beta$ 2m variants .....	88
3.4.4	Study of [20-41] $\beta$ 2m-(Ellman's-thiol) conjugate .....	90
<b>3.5</b>	<b>Discussion .....</b>	<b>93</b>
<b>3.6</b>	<b>Conclusions .....</b>	<b>95</b>
<b>4.</b>	<b>Total chemical synthesis and properties of covalently tethered oligomer mimicking amyloid structure .....</b>	<b>96</b>
<b>4.1</b>	<b>Introduction .....</b>	<b>96</b>
<b>4.2</b>	<b>Design of the structure of covalently-tethered oligomer .....</b>	<b>97</b>
<b>4.3</b>	<b>Strategy used for the <i>N</i>-methylation of the external trimers .....</b>	<b>99</b>
<b>4.4</b>	<b>Design of the synthetic route .....</b>	<b>101</b>
<b>4.5</b>	<b>Design and synthesis of the central linker .....</b>	<b>103</b>
<b>4.6</b>	<b>Synthesis of the covalently-tethered oligomers .....</b>	<b>104</b>
<b>4.7</b>	<b>Description of the different oligomer constructs .....</b>	<b>108</b>
<b>4.8</b>	<b>Study of oligomerization by size-exclusion chromatography .....</b>	<b>114</b>
<b>4.9</b>	<b>Study and comparison of the amyloidogenicity of the different “covalent trimers” .....</b>	<b>117</b>
<b>4.10</b>	<b>Study and comparison of the amyloidogenicity of the covalently-tethered constructs .....</b>	<b>118</b>
<b>4.11</b>	<b>Study of the <sup>15</sup>N-labeled “trimer of trimers” by solution NMR .....</b>	<b>120</b>
<b>4.12</b>	<b>Discussion .....</b>	<b>124</b>

4.13	Conclusions .....	125
5.	<b>Towards molecular recognition and selective inhibition of distinct amyloid polymorphs .....</b>	<b>126</b>
5.1	<b>Introduction .....</b>	<b>126</b>
5.2	<b>Design of the <i>N</i>-methylated peptidic constructs .....</b>	<b>129</b>
5.2.1	Description of the different <i>N</i> -methylated variants .....	129
5.3	<b>Synthesis of the <i>N</i>-methylated peptides .....</b>	<b>131</b>
5.3.1	General procedure used for the synthesis of <i>N</i> -methyl peptides by Fmoc-SPPS .....	131
5.3.2	General procedure for the synthesis of covalent dimers .....	134
5.3.3	General procedure for the synthesis of covalent trimers .....	135
5.4	<b>Characterization of the <i>N</i>-methylated constructs.....</b>	<b>136</b>
5.4.1	Study of the amyloidogenic properties of the constructs.....	136
5.4.2	Study of the secondary structure by circular dichroism .....	137
5.4.3	Study of the stability of the apparent $\beta$ -stranded structures .....	139
5.4.4	Solution NMR measurements.....	140
5.4.5	Study of the oligomerization propensity by size-exclusion chromatography and TEM. ..	141
5.5	<b>Study of the inhibition of amyloid growth .....</b>	<b>142</b>
5.6	<b>Discussion.....</b>	<b>145</b>
5.7	<b>Conclusions .....</b>	<b>147</b>
6.	<b>Significance and future perspectives .....</b>	<b>148</b>
7.	<b>References .....</b>	<b>150</b>
7.1	<b>Materials and methods.....</b>	<b>166</b>
7.2	<b>Chapter 2.....</b>	<b>173</b>
7.3	<b>Chapter 3.....</b>	<b>180</b>
7.4	<b>Chapter 4.....</b>	<b>194</b>
7.5	<b>Chapter 5.....</b>	<b>228</b>
8.	<b><i>Curriculum vitae</i>.....</b>	<b>242</b>

## **Table of tables**

Table 1: List of some amyloid related diseases and their related proteins or peptides.....	28
Table 2: List of some functional amyloids and the related biological functions.....	32
Table 3: Description of amyloid growth experiments for variants of [20-41] $\beta$ 2m peptide.....	88
Table 4: TEM images for amyloids of [20-41] $\beta$ 2m variants.....	89
Table 5: List of all the designed <i>N</i> -methylated variants of [20-41] $\beta$ 2m..	130
Table 6: Final yields of the <i>N</i> -methylated variants of [20-41] $\beta$ 2m .....	133
Table 7: Final yields of the <i>N</i> -methylated variants of [20-41] $\beta$ 2m homodimers. ....	134
Table 8: Final yields of the <i>N</i> -methylated variants of [20-41] $\beta$ 2m homotrimers.....	135

## **Table of figures**

Figure 1: Study of amyloid growth kinetics. ....	29
Figure 2: Representation of the different aggregation stages in amyloid formation. ....	31
Figure 3: Representation of the different types of cross $\beta$ -models. ....	34
Figure 4: The cross- $\beta$ -sheets revealed by X-ray diffraction ....	35
Figure 5: Principle of protein production by recombinant synthesis. ....	40
Figure 6: Principle of solid-phase peptide synthesis. ....	42
Figure 7: Non-native thioester bond formation ....	44
Figure 8: The native chemical ligation ....	45
Figure 9: Different methods to prepare peptide $\alpha$ -thioesters ....	46
Figure 10: The traceless Staudinger ligation. ....	47
Figure 11: The $\alpha$ -ketoacid-hydroxylamine (KAHA) ligation. ....	48
Figure 12: The [20-41] $\beta$ 2m fragment (K3 peptide) ....	54
Figure 13: Chemically synthesized disulfide-linked dimers of the [20-41] $\beta$ 2m fragment. ....	55
Figure 14: Synthesis of mixed <i>L/D</i> -[20-41] $\beta$ 2m dimers. ....	56
Figure 15: Amyloid fibers of <i>L/L</i> -[20-41] $\beta$ 2m ....	57
Figure 16: Properties of amyloids grown from racemic mixtures. ....	58
Figure 17: “Seeding” of amyloid growth with preformed amyloids. ....	59
Figure 18: Amyloid growth of <i>L/L</i> -[20-41] $\beta$ 2m ( <i>L/L</i> -K3) is inhibited by the opposite enantiomer. ....	60
Figure 19: Chemically synthesized <i>L</i> - and <i>D</i> -enantiomers of $\beta$ 2-microglobulin ....	63
Figure 20: “Seeding” of amyloid growth with preformed amyloids is enantioselective for $\beta$ 2m. ....	65
Figure 21: Amyloid growth of <i>L</i> - $\beta$ 2m is inhibited by the opposite <i>D</i> - $\beta$ 2m enantiomer. ....	66
Figure 22: Properties of amyloid fibrils obtained from racemic mixtures of <i>L</i> - and <i>D</i> - $\beta$ 2m ....	67
Figure 23: Depictions of parallel and antiparallel $\beta$ -sheets of homochiral and heterochiral composition ..	69
Figure 24: Structural models for amyloid fibrils based on numerous constraints ....	74
Figure 25: Covalent tethering and supramolecular strategies. ....	75
Figure 26: Synthesis of covalent [20-41] $\beta$ 2m trimer ....	77
Figure 27: TEM images of amyloids. ....	78
Figure 28: Fluorescence measurements of several dyes bound to amyloid fibers. ....	79
Figure 29: Kinetics of growth. ....	81
Figure 30: Properties of [20-41] $\beta$ 2m covalent trimer amyloids ....	82
Figure 31: Properties of [20-41] $\beta$ 2m covalent trimer amyloids ....	84
Figure 32: Synthesis of the different conjugates of [20-41] $\beta$ 2m. ....	86
Figure 33: Modulation of the morphology and structural properties of amyloids ....	87
Figure 34: Properties of [20-41] $\beta$ 2m-(Ellman’s thiol) conjugate. ....	91
Figure 35: Cryo-EM analysis of amyloids obtained for the [20-41] $\beta$ 2m-(Ellman’s thiol) conjugate. ....	92
Figure 36: Molecular model of covalently-tethered oligomer. ....	98
Figure 37: Molecular models of the <i>N</i> -methylated [20-41] $\beta$ 2m segments. ....	101
Figure 38: Retrosynthetic analysis of the covalently-tethered oligomer ....	102
Figure 39: Chemical structures of the central linkers. ....	103
Figure 40: Synthesis of the central linker used in the synthesis of non-symmetric covalent trimers. ....	104
Figure 41: Strategy employed for the on-resin conjugation of the central linker. ....	105
Figure 42: Strategy employed for the synthesis of non-symmetric covalent trimers. ....	106

Figure 43: Structures of different covalent trimers. ....	107
Figure 44: Successive native chemical ligations to form the covalently-tethered oligomer. ....	109
Figure 45: Characterization of TofT-1.. ....	110
Figure 46: Characterization of TofT-2.. ....	111
Figure 47: Characterization of TofT-3. ....	112
Figure 48: Molecular model of TofT-2. ....	113
Figure 49: SEC chromatographic analysis of the covalently-tethered oligomers. ....	114
Figure 50: <sup>1</sup> H NMR of TofT-1 .....	115
Figure 51: <sup>1</sup> H NMR of TofT-2 .....	116
Figure 52: Kinetics of growth determined by ThT fluorescence using a plate reader.....	117
Figure 53: TEM images for aggregates of <i>N</i> -methylated trimer-1-A, amyloids of <i>N</i> -methylated and trimer-3-A and amyloids of central trimer-2-A. ....	118
Figure 54: Fluorescence measurements of several dyes interacting with TofT-1 and TofT-2.....	119
Figure 55: TEM image of TofT-1 fibrils.....	120
Figure 56: <sup>1</sup> H NMR of TofT-3. ....	121
Figure 57: A) <sup>1</sup> H- <sup>15</sup> N HSQC NMR spectrum of TofT-3 and 2D projection of the <sup>15</sup> N-edited NOESY spectrum. ....	122
Figure 58: <sup>1</sup> H- <sup>13</sup> C HSQC NMR spectrum of TofT-3.....	123
Figure 59: Molecular model of the [20-41]β2m protofibril .....	129
Figure 60: Study of kinetics of growth determined by ANS fluorescence using a plate reader.....	136
Figure 61: Study of the kinetics of growth determined by ThT fluorescence on a plate reader.....	137
Figure 62: CD spectra of the different <i>N</i> -methylated constructs.....	138
Figure 63: Melting curves of different <i>N</i> -methylated constructs .....	139
Figure 64: <sup>1</sup> H-TOCSY , <sup>1</sup> H-NOESY spectra of Inh-1-trimer and <sup>1</sup> H-TOCSY, <sup>1</sup> H-NOESY spectra of Inh-2-trimer. ....	140
Figure 65: Size-exclusion chromatograms of Inh-1-trimer and Inh-2-trimer alone and the 1:1 mixture of both trimers.....	141
Figure 66: TEM images of Inh-1-trimer, Inh-2-trimer and 1:1 mixture of both trimers.....	142
Figure 67: Kinetics of [20-41]β2m amyloid growth determined by ThT fluorescence using a plate reader in the presence of Inh-8-trimer, Inh-8-dimer or Inh-8 as a monomer. ....	145

# Abbreviations

<b>Å</b>	Angström
<b>A<math>\beta</math></b>	amyloid beta
<b>AD</b>	Alzheimer's disease
<b>AF</b>	amyloid fibrils
<b>AFM</b>	atomic force microscopy
<b>Aib</b>	$\alpha$ -aminoisobutyric acid
<b>alloc</b>	allyloxycarbonyl
<b>ANS</b>	1-anilinonaphthalene-8-sulfonic acid
<b>Ar</b>	aryl
<b>ATR</b>	attenuated total reflectance
<b>a.u.</b>	arbitrary unit
<b><math>\beta</math>2m</b>	human $\beta$ 2-microglobulin
<b>Bis-ANS</b>	4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid
<b>Boc</b>	<i>tert</i> -butyloxycarbonyl
<b>Bn</b>	benzyl
<b>°C</b>	degree Celsius
<b>Cbz</b>	carboxybenzyl
<b>CD</b>	circular dichroism
<b>chAla</b>	cyclohexylalanine
<b>COSY</b>	correlation spectroscopy
<b>CP-MAS</b>	cross-polarization magic angle spinning
<b>cryo-EM</b>	cryo-electron microscopy
<b>csp</b>	cold shock protein

<b>CW</b>	continuous wave
<b><math>\delta</math></b>	chemical shift
<b>d</b>	doublet
<b>DCM</b>	dichloromethane
<b>DIC</b>	<i>N,N'</i> -diisopropylcarbodiimide
<b>DIPEA</b>	<i>N,N</i> -diisopropylethylamine
<b>DNA</b>	deoxyribonucleic acid
<b>DMF</b>	<i>N,N</i> -dimethylformamide
<b>DMSO</b>	dimethylsulfoxide
<b>DP-MAS</b>	direct polarization magic angle spinning
<b>DTT</b>	<i>DL</i> -dithiothreitol
<b>EDTA</b>	ethylenediaminetetraacetic acid
<b>equiv</b>	equivalent
<b>EPR</b>	electron paramagnetic resonance
<b>ESI</b>	electrospray ionization
<b>FALS</b>	familial amyotrophic lateral sclerosis
<b>FAP</b>	familial amyloid polyneuropathy
<b>Fmoc</b>	9-fluorenylmethyloxycarbonyl
<b>FRET</b>	Förster resonance energy transfer
<b>FTIR</b>	Fourier transform infrared spectroscopy
<b>Gn.HCl</b>	guanidine hydrochloride
<b>HETCOR</b>	heteronuclear correlation
<b>h</b>	hour
<b>HATU</b>	<i>O</i> -(7-azabenzotriazole-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate
<b>HBTU</b>	<i>O</i> -(benzotriazole-1-yl)- <i>N,N,N',N'</i> -tetramethyluronium hexafluorophosphate

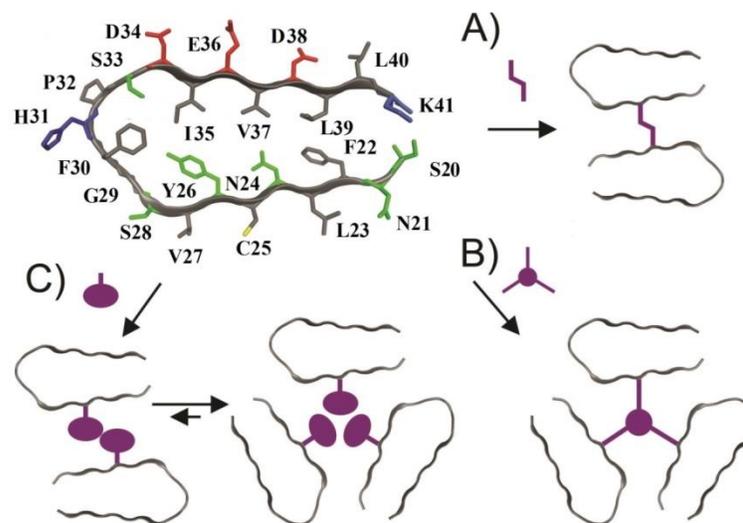
<b>hCT</b>	human calcitonin
<b>HIV-1</b>	human immunodeficiency virus-1
<b>HSQC</b>	heteronuclear single quantum coherence
<b>Hz</b>	hertz
<b>IAPP</b>	islet amyloid polypeptide
<b>IL</b>	interleukines
<b>K3</b>	[20-41]-fragment of human $\beta$ 2-microglobulin
<b>KAHA</b>	$\alpha$ -ketoacid-hydroxylamine
<b>LC-MS</b>	liquid chromatography-mass spectrometry
<b>LS</b>	long straight
<b>M</b>	mole per liter
<b><i>M</i></b>	molecular mass
<b><i>m/z</i></b>	mass-to-charge ratio
<b>MAS</b>	magic angle spinning
<b>Me</b>	methyl
<b>MesNa</b>	sodium 2-mercaptoethane sulfonate
<b>min</b>	minute
<b>MPAA</b>	4-mercaptophenylacetic acid
<b>MRE</b>	mean residue ellipticity
<b>NCL</b>	native chemical ligation
<b>NMP</b>	<i>N</i> -methyl-2-pyrrolidone
<b>NMR</b>	nuclear magnetic resonance
<b>NOESY</b>	nuclear Overhauser effect spectroscopy
<b>Orn</b>	ornithine
<b><i>p</i></b>	para

<b>Pbf</b>	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
<b>Pd/C</b>	palladium on carbon
<b>PDB</b>	protein data bank
<b>pfPhe</b>	perfluorophenylalanine
<b>pH</b>	potential of hydrogen
<b>PMD</b>	protein misfolding disease
<b>ppm</b>	parts per million
<b>Prp</b>	prion protein
<b>PTM</b>	post-translational modification
<b>Pup</b>	prokaryotic ubiquitin-like protein
<b>RNA</b>	ribonucleic acid
<b>rpm</b>	revolution per minute
<b>(RP)-HPLC</b>	reverse-phase high-performance liquid chromatography
<b>r.t.</b>	room temperature
<b>s</b>	singlet
<b>SEC</b>	size-exclusion chromatography
<b>S<sub>N</sub>Ar</b>	nucleophilic aromatic substitution
<b>SOD-1</b>	superoxide dismutase-1
<b>SPPS</b>	solid-phase peptide synthesis
<b>ssNMR</b>	solid-state nuclear magnetic resonance
<b><sup>t</sup>Bu</b>	<i>tert</i> -butyl
<b>t</b>	triplet
<b>TCEP</b>	tris(2-carboxyethyl)phosphine
<b>TEM</b>	transmission electron microscopy
<b>TFA</b>	trifluoroacetic acid

<b>TFE</b>	trifluoroethanol
<b>TfOH</b>	trifluoromethanesulfonic acid
<b>TIPS</b>	triisopropylsilane
<b>ThT</b>	thioflavin T
<b>Thz</b>	thiazolidine-4-carboxylic acid
<b>TOCSY</b>	total correlation spectroscopy
<b>TOF</b>	time of flight
<b>Tris</b>	tri(hydroxymethyl)aminomethane
<b>Trt</b>	trityl
<b>TSE</b>	transmissible spongiform encephalopathies
<b>TTR</b>	transthyretin
<b>UV</b>	ultraviolet irradiation
<b>Vis</b>	visible irradiation
<b>v/v</b>	volume per volume percentage concentration
<b>w/v</b>	weigh per volume percentage concentration
<b>WL</b>	worm like

## Résumé de thèse

Les recherches que je mène actuellement au laboratoire se concentrent sur l'étude des peptides impliqués dans les processus d'amyloïdogénèse. Les fibres amyloïdes sont souvent à l'origine de nombreuses maladies dégénératives telles que la maladie d'Alzheimer ou la maladie de Parkinson<sup>1,2</sup>. La formation de ces plaques insolubles est due à une agrégation anormale de protéines. Les études structurales et biologiques des amyloïdes (empilements de feuillets bêta le long de l'axe de la fibre) sont hautement complexes du fait de leur organisation sous forme de superstructures unidirectionnelles composées d'une infinité d'unités peptidiques ou protéiques, mais aussi à cause de leur hétérogénéité conformationnelle et polymorphique<sup>1,2</sup>.

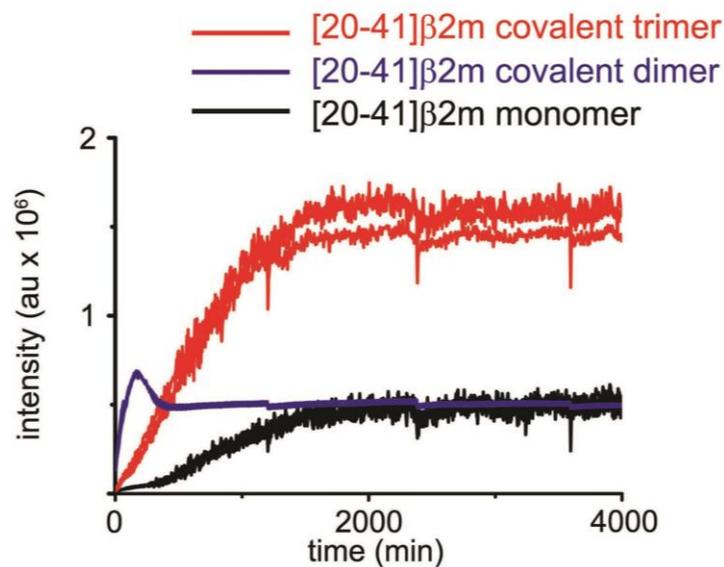


**Figure 1** : Contrôle du polymorphisme de [20-41] $\beta$ 2m selon une approche covalente (A et B) et selon une approche supramoléculaire (C).

On retrouve cette diversité au niveau des plaques de la bêta-amyloïde ( $A\beta$ ), peptide impliqué dans la maladie d'Alzheimer<sup>1,2</sup>. On observe notamment différents nombres de sous-unités peptidiques par coupe transversale des fibres, possédant des symétries rotationnelles d'ordre deux ou trois<sup>3</sup>. Inspirés par le polymorphisme de  $A\beta$ , nous avons mis au point des méthodes permettant d'obtenir des fibres amyloïdes qui s'auto-assemblent de manière dimérique ou trimérique. Pour

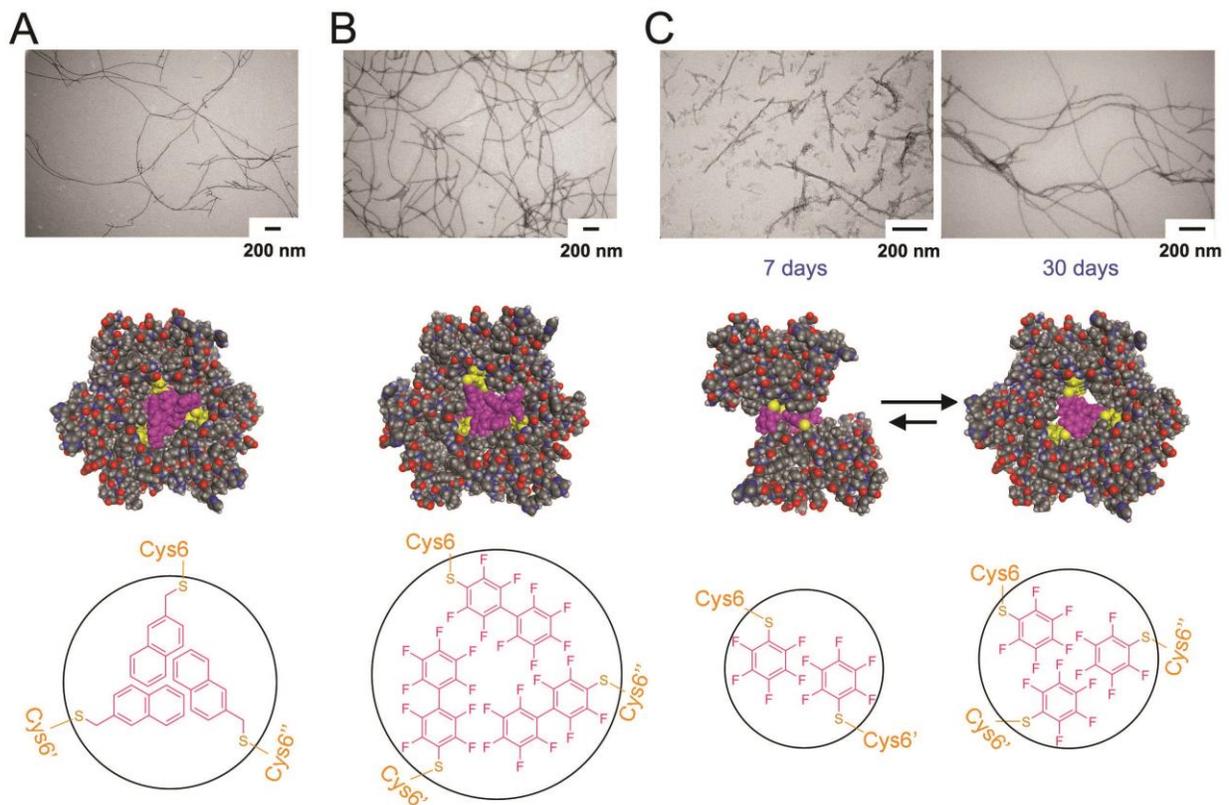
cela, nous nous sommes intéressés au segment 20-41 de la protéine  $\beta 2$ -microglobuline humaine ([20-41] $\beta 2m$ ) en tant que peptide modèle<sup>4</sup>. Il possède la même structure<sup>5</sup> en arc- $\beta$  que le peptide A $\beta$ <sup>6</sup>. La première approche (**Figure 1-A** et **Figure 1-B**) de ce concept consiste à utiliser un lien covalent entre les peptides, par pont disulfure pour la formation du dimère covalent ou par l'introduction d'un linker abiotique central pour la formation du trimère covalent.

Grâce à cette méthode, nous avons pu préparer au laboratoire des fibres amyloïdes homogènes et bien définies avec un haut contrôle de la stœchiométrie. Ceci permettra de faciliter l'étude structurale de tels édifices, la mise au point de tests immunologiques et la recherche d'inhibiteurs spécifiques de certains polymorphes. La cinétique de la croissance des fibres a été contrôlée par fluorescence (**Figure 2**) en utilisant différentes sondes fluorogéniques et par mesure de la turbidité par diffusion de la lumière. On a pu observer que le trimère covalent, le dimère covalent ainsi que le monomère du segment [20-41] $\beta 2m$  possèdent des cinétiques différentes au niveau de la croissance des fibres amyloïdes.



**Figure 2** : Etude cinétique de la croissance des fibres amyloïdes déterminée par mesure de la fluorescence en présence de thioflavine T.

La seconde approche (**Figure 1-C**) consiste à introduire des acides aminés modifiés possédant des chaînes latérales composées de groupements encombrants en position 6 du segment peptidique (cystéine 25). Des groupements hydrophobes de type naphtalène ou perfluoroaryle ont été conjugués au segment [20-41] $\beta$ 2m via la chaîne latérale de la cystéine. Cette dernière approche permet d'obtenir des auto-assemblages par complémentarité stérique favorisant la formation de fibres possédant une symétrie rotationnelle d'ordre 3 (**Figure 3**).



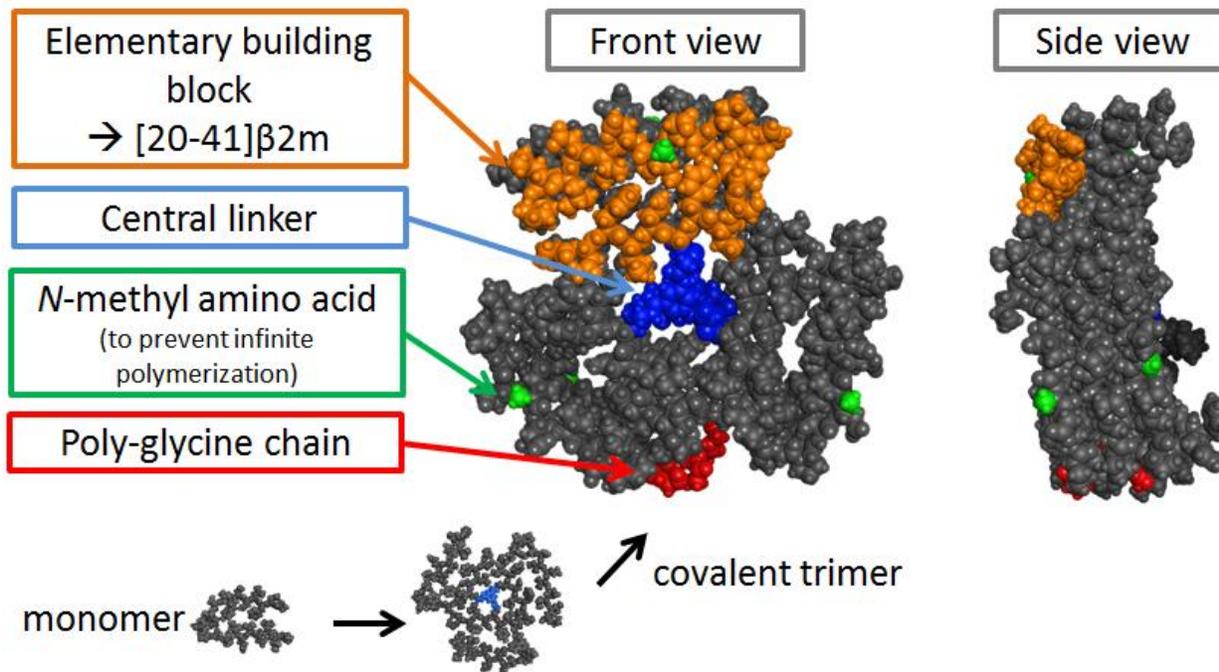
**Figure 3** : Le groupement attaché au niveau de la chaîne latérale en position 6 module la morphologie et les propriétés structurales des fibres amyloïdes. Au-dessus : images TEM ; en milieu : les modélisations moléculaires correspondantes ; en-dessous : les structures chimiques des groupements conjugués à la cystéine 6 (résidu 25 dans le segment [20-41] $\beta$ 2m).

Une dizaine d'autres analogues du segment [20-41] $\beta$ 2m ont été synthétisés afin d'observer l'effet des substitutions en acide aminé sur l'auto-assemblage en fibres amyloïdes. Nous avons remplacé la cystéine 25 par différents acides aminés à chaîne latérale hydrophobe (alanine, isoleucine, cyclohexylalanine), aromatique (tyrosine, phénylalanine, pentafluorophénylalanine,

tryptophane), polaire (glutamine) et chargée (lysine, acide aspartique, arginine). Tous les analogues aromatiques ou à chaîne latérale hydrophobe encombrée forment des fibres amyloïdes bien définies. De même, l'analogue Cys25Gln forme des fibres amyloïdes. En effet, les résidus glutamines sont souvent présents dans la séquence des peptides amyloïdogéniques. Cependant, aucune fibre n'a été détectée lorsque la cystéine a été substituée par une alanine ou par des acides aminés chargés. Ces résultats montrent qu'il est nécessaire que l'espace entre les  $\beta$ -arcs doive être occupé afin de stabiliser l'auto-assemblage et que la tolérance pour des acides aminés chargés au sein du cœur de la structure amyloïde est faible.

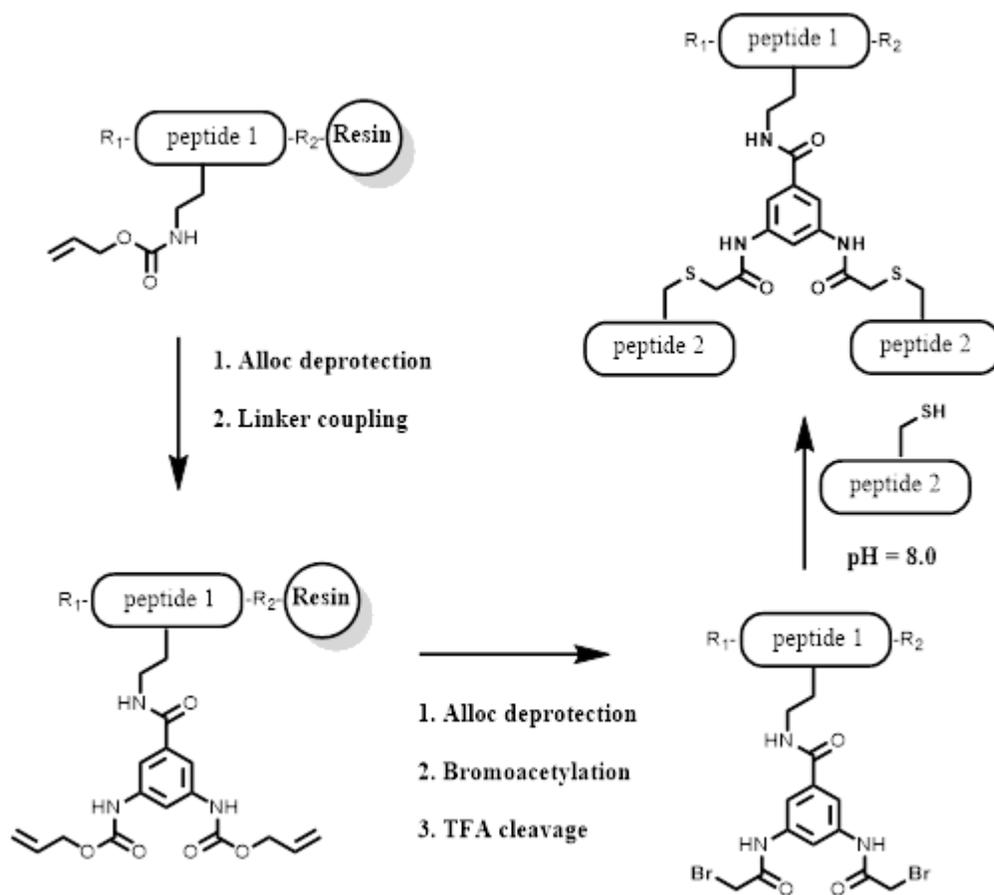
Afin d'en apprendre davantage sur les interactions peptide-peptide, les aspects structuraux, les distances entre les unités peptidiques et sur l'assemblage supramoléculaire au sein d'une fibre amyloïde, des oligomères composés de trois trimères covalents ont été synthétisés (**Figure 4**). Ces oligomères peuvent être assimilés à des coupes transversales d'une fibre amyloïde trimérique composée de neuf sous-unités peptidiques analogues au segment [20-41] $\beta$ 2m. Nous nous attendons à ce que ces édifices s'auto-assemblent et adoptent une structure compacte, distincte et bien définie contrairement aux précurseurs sauvages de fibres amyloïdes qui sont des oligomères métastables et poly-dispersés. Ces derniers sont par conséquent difficiles à étudier et sont d'autant plus soupçonnés d'être hautement cytotoxiques<sup>7</sup>.

Concernant leurs synthèses, trois monomères peptidiques [20-41] $\beta$ 2m sont liés de façon covalente par le biais d'un linker abiotique central (représenté en bleu) afin de composer un trimère covalent. Trois trimères covalents sont ainsi synthétisés puis liés entre eux par ligature chimique native. Des espaceurs de type chaînes poly-glycine (représentés en rouge) se trouvent entre chaque trimère. À l'extrémité de l'édifice (au niveau des trimères covalents extérieurs) ont été incorporés des acides aminés *N*-méthylés au cœur des feuilletts  $\beta$  des segments peptidiques. Ces *N*-méthylations permettent de prévenir la polymérisation de cet édifice peptidique en empêchant un autre oligomère d'interagir<sup>8</sup>. En effet, les groupements méthyles entraînent une gêne stérique et empêchent la formation de liaisons hydrogènes. Tout ceci donne lieu à un trimère covalent de trimères covalents. Pour le premier oligomère synthétisé, nous avons introduit une *N*-méthylation par peptide et deux *N*-méthylations pour le second oligomère synthétisé.



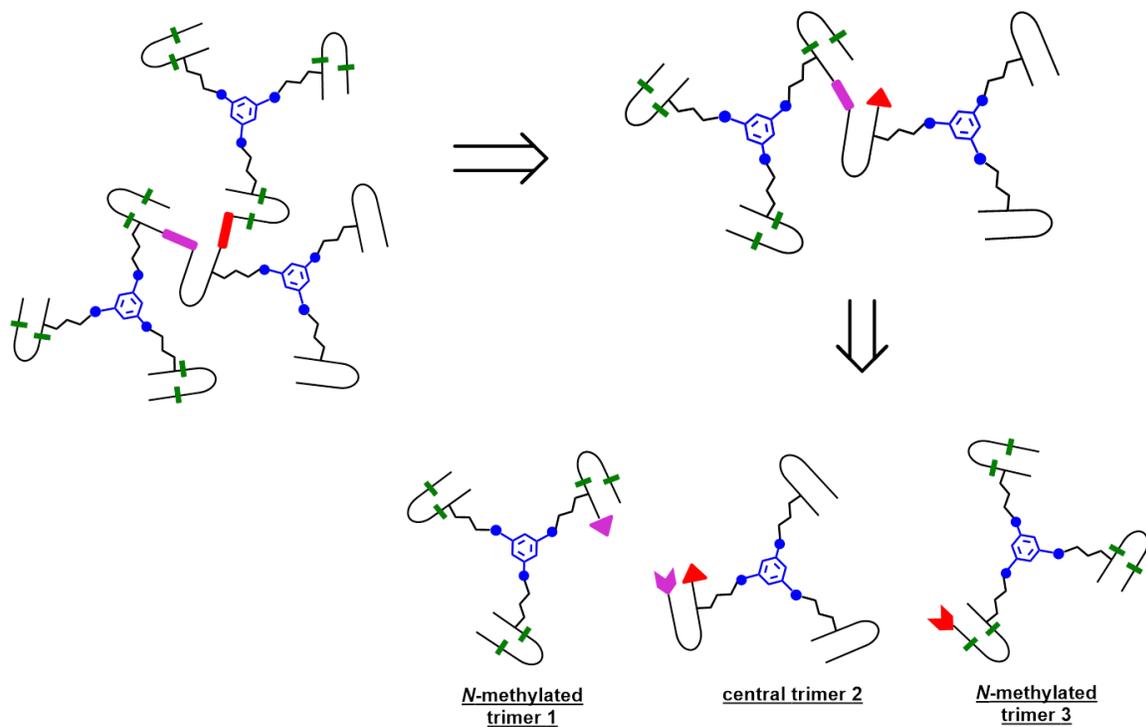
**Figure 4** : Modélisation de l'oligomère (vue de face et vue de côté).

La **Figure 5** décrit la formation du peptide conjugué au linker central. Le peptide est synthétisé sur support solide. La cystéine 25 a été remplacée par une ornithine protégée par une fonction alloc au niveau de la chaîne latérale. Une fois l'élongation du peptide terminée, le groupement protecteur alloc de l'ornithine est retiré de façon orthogonale. Le linker central, protégé au niveau des amines aromatiques par des groupements alloc, est couplé à la chaîne latérale de l'ornithine. Les amines aromatiques du linker sont ensuite déprotégées puis bromoacétylées.



**Figure 5** : Stratégie employée pour la synthèse des trimères covalents. Le groupe R1 peut représenter l'extrémité N-terminal native ou un espaceur de type tri-glycine + cystéine selon les cas. Le groupement R2 peut représenter une fonction acide carboxylique ou une fonction hydrazide nécessaire pour la ligature chimique native.

Enfin, le peptide est coupé de la résine, extrait puis purifié (**Figure 5**). Le peptide portant le linker synthétisé précédemment est incubé en présence du segment [20-41] $\beta$ 2m comportant une cystéine (**Figure 5**) dans un tampon phosphate à pH 8 en condition dénaturante afin d'obtenir un trimère covalent. Les différents trimères ont été synthétisés suivant cette méthode mise au point au laboratoire. Les trois trimères sont ensuite liés entre eux par ligature chimique native<sup>9</sup> afin d'obtenir un trimère covalent de trimères covalents (**Figure 6**).



**Figure 6** : Analyse rétro-synthétique de l'oligomère. L'oligomère final est obtenu par une succession de ligatures chimiques natives des différents trimères covalents. Les segment  $[20-41]\beta 2m$  sont représentés par les arcs noirs. Les *N*-méthylations sont représentées en vert et sont localisées au niveau des trimères externes. Le linker central est représenté en bleu. Au niveau du produit final, les chaînes de type polyglycine sont représentées en rouge et en magenta.

Les propriétés structurales de ces édifices peptidiques ont été étudiées par résonance magnétique nucléaire. Leurs états d'agrégation ont été étudiés par chromatographie d'exclusion stérique et microscopie électronique. Nous avons conclu que l'oligomère comportant une méthylation par peptide ne formait pas de structure unique, précipitait au fil du temps et réagissait très nettement à des sondes fluorescente couramment utilisées pour la détection de fibres amyloïdes. Néanmoins, l'oligomère comportant deux méthylations par peptide forme une structure unique, stable et ne forme pas d'agrégats. Ces résultats sont à confirmer par d'autres analyses structurales.

Dans ce même registre, nous avons voulu étudier plus en profondeur les propriétés structurales et fonctionnelles de peptides *N*-méthylés et de leurs propriétés inhibitrices de la propagation des fibres amyloïdes pour les trois types d'assemblage (monomère, dimère et trimère). Pour cela, des

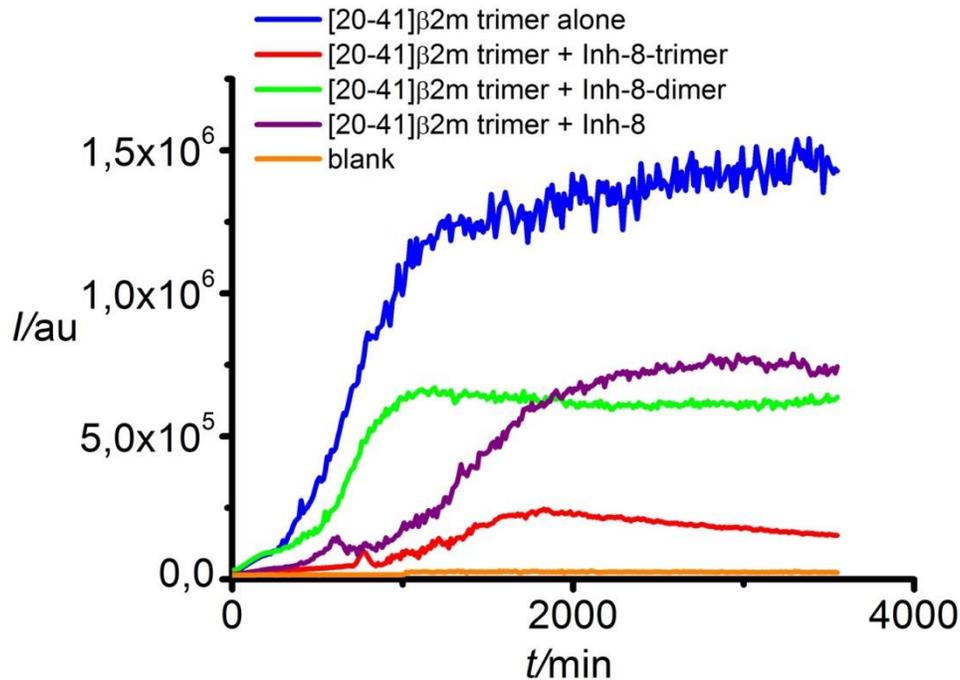
analogues  $N^\alpha$ -méthylés du segment [20-41] $\beta$ 2m ont été synthétisés. En effet, de nombreuses études ont montré l'efficacité de peptides  $N^\alpha$ -méthylés en tant qu'inhibiteur<sup>8</sup>. Néanmoins, une fibre amyloïde possède une structure polaire ne présentant pas la même interface à chaque extrémité (interface A et interface B). Il faut donc pouvoir inhiber la propagation des deux côté de la fibre, si elle présente une élongation bidirectionnelle.

**Tableau 1** : Liste des différents analogues  $N$ -méthylés synthétisés. En rouge sont représentés les acides aminés  $N$ -méthylés.

Nom du peptide (monomères)	Nombre de $N$ -méthylation	Séquence	Interface $N$ -méthylée
Inh-1	4	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVD <b>L</b> LK	B
Inh-2	4	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVD <b>L</b> LK	A
Inh-3	2	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVDLLK	B
Inh-4	2	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVD <b>L</b> LK	B
Inh-5	2	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVD <b>L</b> LK	A
Inh-6	2	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVDLLK	A
Inh-7	2	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVDLLK	B
Inh-8	2	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVDLLK	B
Inh-9	2	SNFLN <b>C</b> YVSGFHPSD <b>I</b> EVDLLK	A and B

C'est pourquoi plusieurs inhibiteurs (**Tableau 1**) pouvant reconnaître une des deux interfaces (A ou B) ont été synthétisés pour chaque architecture peptidique (monomérique, dimérique et trimérique). De plus, nous avons fait varier le nombre et la position des acides aminés  $N$ -méthylés au sein du fragment [20-41] $\beta$ 2m. Nous observons que ces analogues  $N$ -méthylés ne forment pas de fibres amyloïdes ni d'autres agrégats. La mesure du dichroïsme circulaire indique la présence de feuilletts  $\beta$  distordus. Ceci est dû à la présence d'acides aminés  $N$ -méthylés. Certains variants ont été également caractérisés par RMN, microscopie électronique et chromatographie d'exclusion stérique. Les tests d'inhibition sont encore à approfondir afin d'évaluer leur pouvoir inhibiteur sur la formation des fibres amyloïdes. Cependant, nous pouvons déjà affirmer que le

peptide *N*-méthylé « Inh-8 » inhibe efficacement la croissance des fibres amyloïdes (**Figure 7**) quelle que soit la structure (monomère, dimère ou trimère).



**Figure 7:** Etude cinétique de la croissance des fibres amyloïdes déterminée par fluorescence (thioflavine T) en présence d'Inh-8-trimère, Inh-8-dimère et Inh-8 (monomère)

L'inhibition de la formation de fibres amyloïdes par des peptides *D* est également un point sur lequel nous nous sommes penchés. En effet, les peptides *D* possèdent une grande stabilité vis-à-vis des enzymes protéolytiques, provoquent moins de réactions immunitaires et sont, par conséquent très intéressants pour des visées thérapeutiques. Cependant, la compréhension des mécanismes intervenant lors de la reconnaissance chirale des peptides *D* vis-à-vis des amyloïdes *L* n'est pas complète. Nous avons pour cela mené des expériences sur la croissance des amyloïdes du dimère covalent du fragment [20-41]β2m sur des échantillons comportant séparément l'énantiomère *L* et *D* et sur le mélange racémique des deux énantiomères. Les expériences d'EPR et de spectroscopie infrarouge sur des peptides marqués au <sup>13</sup>C ont révélés que pour le mélange racémique, chaque énantiomère précipitait sous la forme de fibres homochirales intactes.

Nous avons également étudié ce phénomène dans le cas de la protéine entière, la  $\beta$ 2-microglobuline humaine. Nous en avons conclu que le mélange racémique conduisait à la formation de structures fibrillaires mixtes mais constituées de domaines énantiopures. Les fibres ainsi formées ne possédaient donc pas de structures en hélice généralement observées pour les fibres amyloïdes. Concernant le fragment [20-41] $\beta$ 2m *L* et la protéine entière *L*, nous avons observé que la présence de leurs énantiomères *D* inhibait la formation des fibres amyloïdes *L*. Ceci indique que les peptides ou les protéines d'une certaine chiralité peuvent interagir avec des fibres amyloïdes composées de peptides ou protéines de chiralité inverse. Nous nous dirigeons donc vers l'étude de l'inhibition de la formation de fibres amyloïdes par des peptides *D*, *N*-méthylés afin de combiner les deux types de modification.

Au cours de ces différents travaux de thèse en collaboration avec différents laboratoires d'analyses structurales, nous avons développé plusieurs outils de synthèse tant pour la formation de différents polymorphes de fibres amyloïdes que pour la formation d'espèces oligomériques de tailles conséquentes qui sont un challenge du point de vue synthétique et méthodologique mais aussi pour leur caractérisation. Ces différentes avancées permettront de mieux comprendre les mécanismes de formation de fibres amyloïdes et de préparer des échantillons homogènes pour les analyses structurales et biologiques, notamment pour l'étude de la cytotoxicité des différents variants et pour le screening de possibles inhibiteurs pour chaque type de structure. L'étude de modifications chimiques telles que la *N*-méthylation ou les polypeptides *D* est également un enjeu important pour l'élucidation des interactions vis-à-vis des structures amyloïdogéniques monomériques, oligomériques de transition ou des fibres amyloïdes. Ces recherches ont finalement pour but de mettre au point des agents thérapeutiques ou des outils de diagnostic basés sur des structures peptidiques simples ou hybrides conjuguées à des sondes ayant des propriétés de fluorescence, électroniques ou photoniques.

1. Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 (2014).
2. Greenwald, J. & Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **18**, 1244–1260 (2010).

3. Paravastu, A. K., Leapman, R. D., Yau, W.-M. & Tycko, R. Molecular structural basis for polymorphism in Alzheimer's  $\beta$ -amyloid fibrils. *Proc. Natl. Acad. Sci.* **105**, 18349–18354 (2008).
4. Kozhukh, G. V. *et al.* Investigation of a peptide responsible for amyloid fibril formation of  $\beta$ 2-microglobulin by *Achromobacter* protease I. *J. Biol. Chem.* **277**, 1310–1315 (2002).
5. Iwata, K *et al.* 3D structure of amyloid protofilaments of  $\beta$ 2-microglobulin fragment probed by solid-state NMR. *Proc. Natl. Acad. Sci.* **103**, 18119–18124 (2006).
6. Tycko, R. Solid-state NMR studies of amyloid fibril structure. *Annu. Rev. Phys. Chem.* **62**, 279–299 (2011).
7. Walsh, D. M. *et al.* Naturally secreted oligomers of amyloid  $\beta$  protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* **416**, 535–539 (2002).
8. Sciarretta, K. L., Gordon, D. J. & Meredith, S. C. Peptide-based inhibitors of amyloid assembly. In *Methods in Enzymology* **413**, 273–312 (Academic Press, 2006).
9. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).

# 1. Introduction

Proteins play highly important roles in cellular and extracellular functions. Together with nucleic acids, sugars and lipids, they belong to the four major classes of biological macromolecules that are the building blocks of all living organisms. Proteins are bio-polymers that are composed of twenty canonical amino acids common to all living species and possess unique structural properties, functions, complexity and diversity. The function of a protein is intimately correlated to its three-dimensional structure. The fold of polypeptide chains is defined by the amino acid sequence, resulting in diverse biological roles that range from catalysis (enzymes), cell signaling and regulation of gene expression to structural functions. In some cases, upon distinct environmental stress conditions, the protein does not adopt its native conformation and loses its properties and functions. In most cases, misfolded proteins are degraded and recycled by the cell. But some can accumulate and aggregate in the cell, leading to cell malfunctioning.

Amyloid fibrils (AF) are insoluble aggregates resulting from the self-assembly of infinite copies of a  $\beta$ -sheet-rich peptide or protein. Such aggregates are associated with many neurodegenerative diseases such as Alzheimer's disease, Huntington's disease, Parkinson's disease or bovine spongiform encephalopathy but also with type II diabetes or dialysis related amyloidosis<sup>1,2</sup>. The main distinctive feature of these disorders is the formation of insoluble plaques formed by amyloid deposition leading to cell death and disruption of functions of tissues and organs. However, the mechanisms of amyloid formation remain poorly understood. Misfolding of the involved protein leads to critical conformational changes. The soluble monomeric protein switches into highly insoluble aggregates by aberrant recruitment of the same protein. Amyloidogenic peptides or proteins self-assemble into long fibrils characterized by their cross- $\beta$ -sheet structures<sup>3</sup>. Two or more  $\beta$ -sheets are stacked repetitively against each other with  $\beta$ -strands running nearly perpendicularly to the fiber axis. Multiple copies of these protein sub-units interact resulting in highly ordered fibrils<sup>4</sup>.

Different techniques are used to reveal the presence of AF in body tissues. AF can easily be identified by light microscopy using cross polarizers. One of their key signatures is their ability to be stained with the Congo Red dye and to display green-yellow birefringence<sup>5</sup>. Some other

extrinsic fluorescent dyes are also used to detect the presence of amyloid fibers and prefibrillar oligomers such as thioflavin T (ThT) (**Figure 1**), 1-anilinonaphthalene-8-sulfonic acid (ANS), 4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid (bis-ANS) or Nile Red <sup>6,7</sup>. AF can also be identified by X-ray diffraction analysis by showing characteristic cross- $\beta$  diffraction patterns<sup>8</sup>. Other biophysical methods that are used to characterize amyloid fibrils include solid-state nuclear magnetic resonance (NMR) spectroscopy<sup>9</sup>, electron microscopy<sup>10</sup>, atomic force microscopy (AFM) or electron paramagnetic resonance (EPR) spectroscopy<sup>11</sup>. Such experimental techniques can provide detailed information about the structural properties, the heterogeneity, the length and width of the fibrils. But amyloids remain difficult to study. Because of the mono-directional fibrillary structure, they are poor candidates for high resolution three dimensional (3D) X-ray crystallographic analysis due to difficulties in obtaining single crystals. Moreover, they often display polymorphism leading to high structural diversity, which complicates their structural and biological studies. In fact, each polymorph may present different biophysical properties and biological activities<sup>12</sup>. There is a need for a better understanding of structural properties and mechanism of fiber formation at a molecular level, which may contribute to finding new treatments against amyloid deposition. Especially, the early stages of fiber-formation are not well understood.

Total chemical synthesis of peptides and proteins is a powerful tool, which can also help the studying the amyloid structural properties. By synthesizing wild type sequences or chemical variants of disease-associated peptides or proteins, detailed insights can be obtained about the processes involved in the early stages of formation of oligomeric species and the factors that trigger the polymerization into fibrils. To synthesize proteins, various methods have been developed. One of the most important techniques is solid-phase peptide synthesis (SPPS), which allows rapid preparation of peptides with high purity, so that sequences of pronounced hydrophobic character can be obtained readily in large quantities. A further advantage is the accessibility of peptides possessing post-translational modifications (PTM), unnatural moieties, isotope labels or fluorescent probes. It can be particularly interesting for studying protein-protein interactions and molecular mechanisms of amyloidosis. Furthermore, it will also help in finding new types of peptide-based inhibitors of amyloid fiber formation.

## 1.1 Amyloid related diseases

The interest in amyloid fibrils has substantially increased over the past decades in many scientific fields such as chemistry, medicine, pharmacology, physics and biology. The study of these protein aggregates has become increasingly essential and fundamental. Amyloid fibrils are related to more than 40 human diseases (**Table 1**)<sup>1,12,13</sup>. They have become a significant health problem and have thus a growing degree of awareness worldwide. Since life expectancy is increasing, more and more people are diagnosed with neurodegenerative disorders, which are closely related to ageing, such as Alzheimer's disease or Parkinson's disease. Many other different organs than brain can be affected by amyloid deposition such as the liver, the spleen and the peripheral nervous system.

The characteristic signature of these protein misfolding disorders (PMDs) is the accumulation of amyloid fibrils into plaques leading to malfunction of the organs. The involved proteins do not share sequence similarity, the number of residues can differ greatly and they possess very different native structures ranging from intrinsically disordered to all- $\beta$  or all- $\alpha$  structures. But they all can self-assemble into cross- $\beta$  structured fibrils.

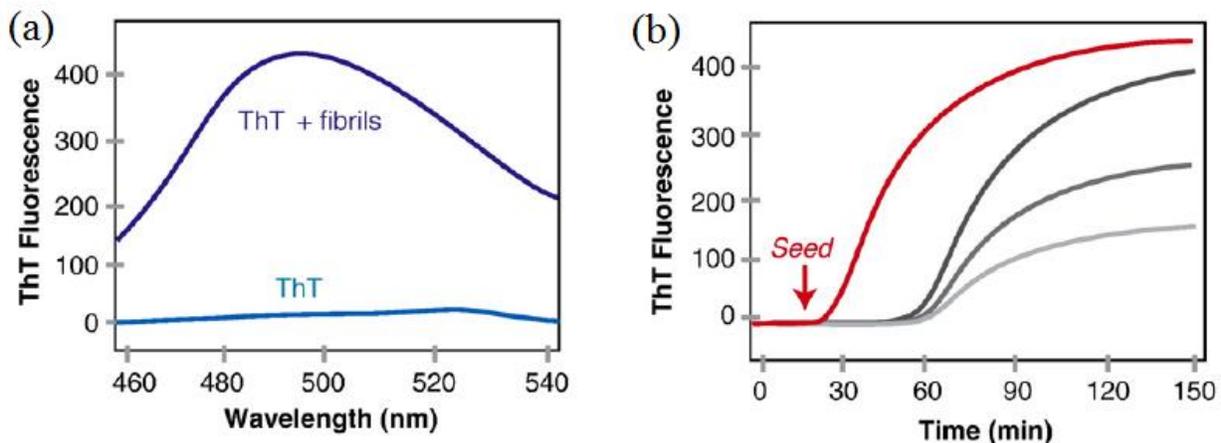
**Table 1:** List of some amyloid related diseases and their related proteins or peptides

Diseases	Related proteins or peptides
Alzheimer's disease	$\beta$ -amyloid peptide
Huntington's disease	huntingtin
Parkinson's disease	$\alpha$ -synuclein
Medullary thyroid cancer	calcitonin
Type II diabetes	amylin
Hemodialysis-related amyloidosis	$\beta$ 2-microglobulin
Familial amyloid polyneuropathy	transthyretin
Atherosclerosis	apolipoprotein AI
Bovine spongiform encephalopathy	prion protein

But the plaques might not necessarily be the most toxic species involved in amyloidosis. More and more studies are pointing out the implication of transient intermediate oligomers or protofilaments as the most cytotoxic species<sup>14</sup>.

### 1.1.1 Mechanisms of amyloid formation

The mechanisms involved in amyloid fibril formation are rather complex and are not fully understood. Nevertheless, it is agreed that the amyloids are forming through a nucleated growth mechanism involving a stochastic nucleus formation followed by a rapid exponential growth phase. Three phases were clearly identified by *in vitro* ThT fluorescence measurement of amyloid growth kinetics<sup>15</sup>. The experiment in **Figure 1-b** shows that there is a lag phase in the beginning with no ThT fluorescence followed subsequently by a rapid increase in the fluorescence intensity suggesting a rapid polymerization. The third stage is characterized by steady state propagation, when finally the fluorescence intensity reaches a plateau (equilibrium stage). Pre-seeding induces a substantial acceleration of fibril formation via templating effect when rate-limiting step of the formation of nucleus is omitted (**Figure 1-b**).



**Figure 1:** (a) Characteristic increase of the ThT fluorescence upon addition of amyloid fibrils. (b) Schematic representation of amyloid growth kinetics monitored by ThT fluorescence and the influence of seeding on the fibrillation process<sup>15</sup>. Figure is taken from *Biochim. Biophys. Acta* 1804, 1405-1412 (2010).

The most distinct property of amyloid fibrils is their high content of  $\beta$ -sheets. As previously mentioned, amyloids are long straight fibrils of infinitely stacked  $\beta$ -sheets with  $\beta$ -strands running nearly perpendicular to the fibril axis. Amyloidogenic peptides or proteins must then expose their backbone amides and carbonyl groups in order to interact through hydrogen bonding with the backbone of other partners.

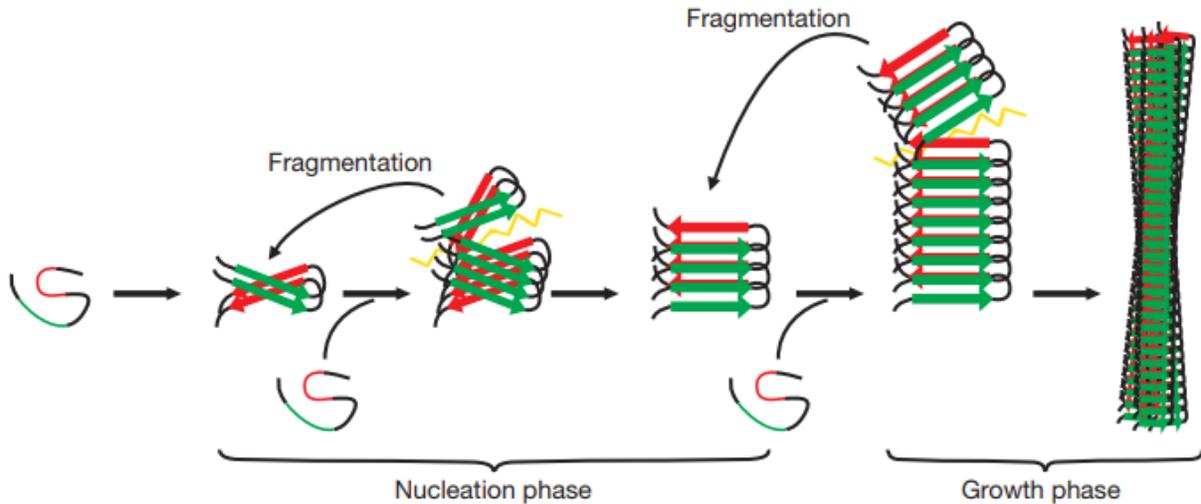
Several experimental conditions can trigger a transformation of a protein or a peptide into a non-native amyloidogenic isoform with exposed backbone amide groups:

- a local or complete denaturation or alteration of the secondary structure of a native peptide or protein can result in a partially or fully misfolded state;
- proteolytic cleavage of a functional folded protein can yield amyloidogenic peptides. For instance, the proteolysis of the amyloid- $\beta$  precursor protein by different type of secretases leads to the formation of various A $\beta$  peptides related to Alzheimer's disease with different chain lengths and distinct amyloidogenicities and cytotoxicities<sup>16</sup>.

The amyloidogenic propensity is not sufficient for the successful formation of fibrils since nucleation is a stochastic process. The local concentration of the involved protein must be high enough to counteract the entropy loss during the formation of ordered fibrils.

Nuclei are formed by several copies of a protein as amorphous species via reversible conformational isomerizations. They are disordered and present no prevalent cross- $\beta$  sheet structures<sup>13</sup>. Once the aggregation-prone state is reached the propagation process occurs. Recruitment of more monomeric or oligomeric species by template effects leads to the formation of larger aggregates called protofilaments.

Protofilaments contain  $\beta$ -sheet structures and bind Congo Red and Thioflavin T dyes, suggesting more ordered oligomers with regular structural motifs. Many different shapes were observed. They can adopt spherical or annular to more linear and curly structures. These are still quite metastable species and of particular interest because they exhibit high cytotoxicity levels<sup>17</sup>. More and more studies point out that these oligomers have a crucial role in the neurodegeneration processes and might be the most toxic species. Their identifications and structural characterizations are of high importance and might help for the discovery of new therapeutic agents preventing amyloid deposition.



**Figure 2:** Representation of the different aggregation stages in amyloid formation<sup>18</sup>. *Figure is taken from Nature 501, 45-51 (2013).*

The final state is the formation of highly ordered fibrils resulting from the infinite polymerization and interaction between multiple protofilaments. They are no longer soluble and start to precipitate. They can also fragment into smaller fibrillary aggregates and become new additional seeds and template the formation of new long fibrils, therefore, amplifying the propagation of the disease (**Figure 2**).

### 1.1.2 Prion diseases

Prion diseases, also known as transmissible spongiform encephalopathies (TSE),<sup>19</sup> are fatal neurodegenerative disorders such as Creutzfeldt-Jakob disease in human or scrapie in sheep or goat. They are caused by infectious agents named prion proteins. Contrary to viruses or bacteria, the infection does not rely on nucleic acids as the genetic support but seems to be a protein-based process. The normal protein ( $\text{PrP}^{\text{C}}$ ) is a glycoprotein anchored on the surface of cells. The tertiary structure of the prion can be altered and the abnormal change of conformation renders the  $\text{PrP}^{\text{C}}$  isoform resistant to proteases and highly insoluble. Moreover, the misfolded prion ( $\text{PrP}^{\text{Sc}}$ ) is able to change the conformation of native  $\text{PrP}^{\text{C}}$  into  $\text{PrP}^{\text{Sc}}$  and, therefore, self-propagate the

information and spread the infection. The  $\alpha$ -helical structure of PrP<sup>C</sup> is converted into  $\beta$ -sheet rich PrP<sup>Sc</sup> able to form amyloid fibrils.

### 1.1.3 Functional amyloids

In contrast with disease-related protein amyloids, many different amyloid strains have been found to be harmless and have a beneficial role. They are called functional amyloids<sup>20</sup> (**Table 2**). It seems that upon evolution, some species (bacteria, fungi, insects and even mammals) have taken advantages of such protein assemblies for a wide variety of biological functions.

**Table 2:** List of some functional amyloids and the related biological functions.

Involved protein	Biological function
curli	<i>Escherichia coli</i> : Inert surfaces colonization, biofilm formation, adhesion
Ure2p	<i>Saccharomyces cerevisiae</i> : Nitrogen uptake promotion, regulatory functions
Sup35	<i>Saccharomyces cerevisiae</i> : Translation termination, better read-through of stop codons
PmeI17	<i>Homo sapiens</i> : Involvement in skin pigmentation
HET-s	<i>Podospora anserina</i> : Complex cell death phenomenon programming

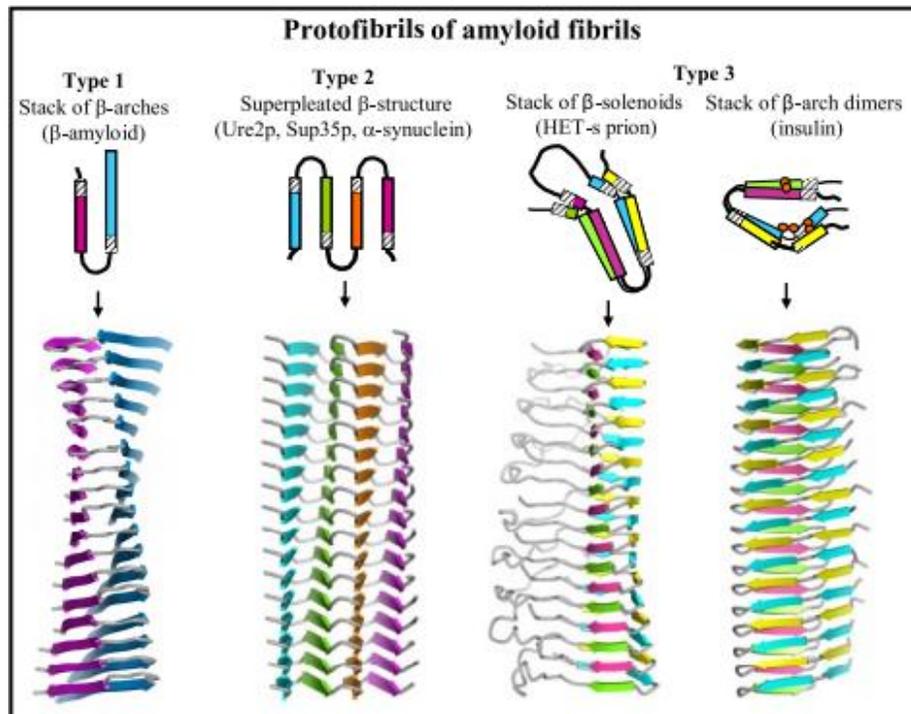
Lower organisms such as bacteria or fungi use the particular mechanical properties of amyloids. For instance, *Escherichia coli* produces an amyloidogenic protein called curli<sup>21</sup> that is forming amyloid fibrils in the extracellular matrix and has an important role in the formation of biofilms and in the colonization of inert surfaces.

It was found that some protein or peptide hormones could be stored in secretory granules in an amyloid-like conformation<sup>22</sup>. Upon the appropriate biochemical stimuli, granules are secreted and monomeric hormones dissociate from the amyloid aggregate. Pmel17 is another interesting example of non-pathological amyloid found in mammals<sup>23</sup>. Pmel17 amyloids are involved in the synthesis of melanin, which plays an important role in neutralizing pathogens or protection of UV radiation. More and more functional amyloids are discovered and the list will certainly increase over the years, with the discovery of new functions and properties.

## **1.2 The structural properties of amyloid fibrils**

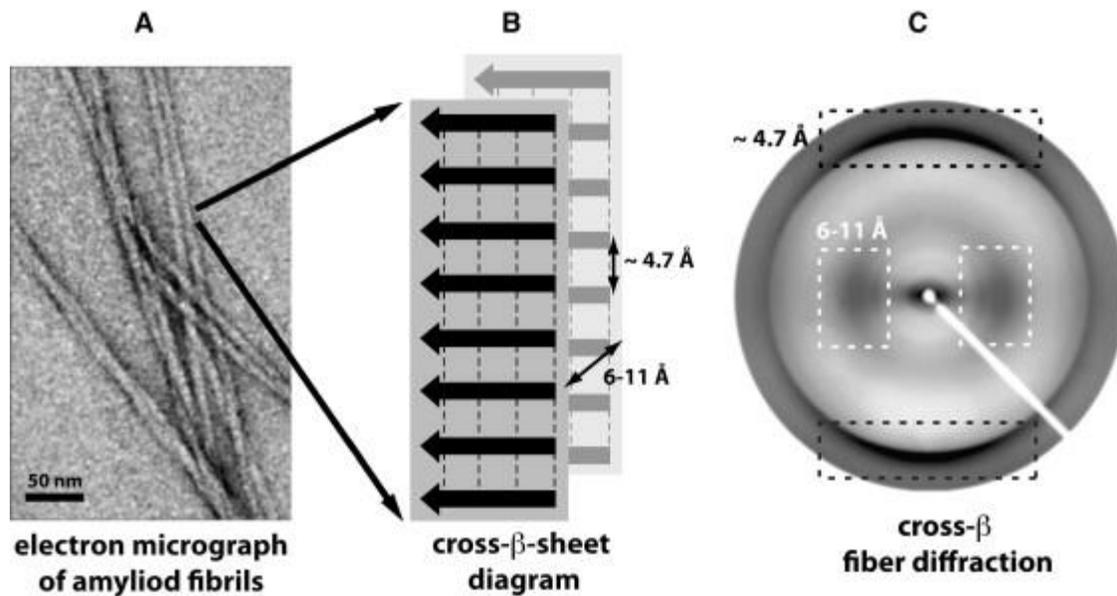
Regarding the sequence and the length of the involved peptide or protein, amyloid fibrils can adopt very different structures<sup>24</sup>. One aspect that all AF have in common is the presence of recurring motifs with similar cross- $\beta$  structures. And they can be classified according to the nature of  $\beta$ -arches stacking along the axis. A  $\beta$ -arch motif adopts a strand-loop-strand conformation and has to be differentiated from  $\beta$ -hairpin motifs, which are more commonly present in globular proteins.

$\beta$ -Hairpins interact through the amides of the peptide backbone, where the  $\beta$ -strands are connected by a loop called  $\beta$ -turn, whereas  $\beta$ -arches are composed of two or more  $\beta$ -strands separated by a  $\beta$ -arc. In the latter case, the  $\beta$ -strands interact through their side chains via hydrophobic interactions creating a tight interdigitating interface called steric zipper. The presence of glutamine or asparagine residues in the  $\beta$ -arches increases the stability of the structure due to additional hydrogen bonding via the amide groups on their side-chains, creating an additional H-bonding network along the fiber axis<sup>25</sup>. It is also the case for oppositely charged side chains at the outer surface of the fiber. In most cases, the peptidic units stack in a parallel fashion, where the  $\beta$ -stands are in-register (all the residues involved in the cross- $\beta$ -structure are facing the same identical residues along the fiber).



**Figure 3:** Representation of the different types of cross  $\beta$ -models classified by Andrey Kajava<sup>24</sup>. Top: axial view of the repetitive units, where the colored rectangles represent the  $\beta$ -strands. Bottom: lateral view of the related protofilaments. Orange circles in schematic representation of insulin highlight cysteine residues forming disulfide bridges. *Figure is taken from FASEB J. 24, 1311-1319 (2010).*

Andrey Kajava classified the “ $\beta$ -arcades” into three distinct types of assembly (**Figure 3**). The first type corresponds to a simple stack of  $\beta$ -arches, which forms a double layer of parallel  $\beta$ -sheets. This topology is adopted by A $\beta$  peptide<sup>26</sup>, the [20-41]-fragment of  $\beta$ 2-microglobulin<sup>27</sup> or by amylin<sup>25</sup>. The second type is called super-pleated  $\beta$ -structure that is seen for prion proteins like Ure2p and Sup35p or for  $\alpha$ -synuclein<sup>28,29</sup>. In this case, several  $\beta$ -stands (more than two) are forming the cross- $\beta$  motif generating an array of parallel  $\beta$ -sheets within the fibril. The third type includes the fungal HET-s prion protein, where the unit coils and occupies two levels within the fiber<sup>30</sup>. For insulin, it is a stack of  $\beta$ -arch dimers connected by disulfide bridges, where two different polypeptide chains alternate along the fibril<sup>25</sup>.



**Figure 4:** The cross- $\beta$ -sheets can be revealed by X-ray diffraction. (A) TEM image of negatively stained amyloid fibers. (B) Schematic diagram of the cross- $\beta$ -sheet. Hydrogen bonds are represented by dashed lines. (C) Characteristic X-ray diffraction pattern of amyloid fiber<sup>2</sup>. *Figure is taken from Structure 18, 1244-1260 (2010).*

Amyloids reveal a meridional reflection around 4.7-4.8 Å corresponding to the space between two  $\beta$ -stands along the fiber axis and an equatorial reflection within a range from 6 to 11 Å depending on the sequence and corresponding to the distance between  $\beta$ -sheets (**Figure 4**).

Another feature of amyloid fibrils is the presence of a helical twist along the fiber axis<sup>31</sup>. It can be observed by atomic force microscopy or electron microscopy. Amyloidogenic polypeptide chains made of *L*-amino acids are expected to self-assemble into fibers with a left-handed twist and form a chiral super-structure. Therefore, mirror-image polypeptides (*D*-peptides) are expected to fold into right-handed helical fibers.

### 1.3 Structural polymorphism in amyloid fibrils

As a result of evolution, functional amyloids are produced as highly homogeneous fibers that are required in order to elicit their biological functions<sup>2</sup>. This is not the case for disease related

amyloids. As they appear due to stress events (protein misfolding, abnormally high local concentration), their formation can lead to different structural states. We can postulate that the involved proteins or peptides were not “meant” to form fibers, and, therefore, their primary and secondary structure did not evolve for that purpose. This is why upon misfolding the non-native structure may possess different local energy minima that can give rise to various amyloid structures<sup>2</sup>.

Structural polymorphism can manifest itself in different ways. In amyloid fibrils, most of the units self-assemble to form parallel  $\beta$ -sheets. But it is not rare to observe the same polypeptide adopting different types of packing. Co-existence of units packed in parallel or anti-parallel fashion within the same fiber is possible<sup>32</sup>. The same residues are usually involved but they are not facing their identical counterparts. For instance, X-ray diffraction analyses and solid-state NMR revealed that self-assembly of [20-29]-fragment of amylin can lead to two types of packing polymorphs, one composed of only parallel  $\beta$ -sheets and the other one having anti-parallel strands within the  $\beta$ -sheet layers<sup>33</sup>. It was also reported that human calcitonin (hCT) can form two stable and ordered fibril structures<sup>34</sup> composed of either a parallel or an antiparallel arrangement. In the latter case, the anti-parallel conformation provides the highest content of  $\beta$ -sheets leading to a more stable architecture.

It was found that the same polypeptide can lead to different segmental polymorphs. In each isomorph the  $\beta$ -strand will differ by its sequence or length, leading to different structures and geometries. An example is A $\beta$  peptide, which can adopt remarkably different states<sup>35,36</sup>.

Polymorphism can manifest itself at the amino-acid side-chain level and was observed in crystals of the peptide LVEALYL. The glutamic acid residue can adopt two most favored configurations having different hydrogen bonding networks. Both configurations result in fibrils with different surface features and solvent interactions<sup>2</sup>.

Hierarchical assembly of amyloidogenic peptides can offer an even higher degree of complexity. It has been shown that a various amount of peptide subunits can form a single fibril, leading to different sizes, geometry and morphologies. A $\beta$ (1-40) was found to form fibrils with either twofold or threefold symmetries with two or three peptides per cross section of the fibril, respectively, according to solid-state NMR experiments and mass-per-length measurements by

electron microscopy<sup>37</sup>. The 11-residue fragment of the protein transthyretin (TTR(105-115)) can self-assemble into twofold symmetry fibers with either four, six or eight protofilaments per cross section<sup>9</sup>. Moreover, once the fibrils are formed, they can interact to form higher order aggregates resulting in intertwined polymeric fibers with different thicknesses.

## 1.4 Inhibition of amyloid formation

In protein misfolding diseases, the inhibition of amyloid formation has been proposed as a possible therapeutic strategy<sup>38</sup>. Indeed, recent clinical trials with anti-amyloid antibodies developed for treatment of Alzheimer's disease (AD) showed that administering the antibodies to early-stage AD patients helped to clear amyloid deposits and to slow down cognitive decline<sup>39</sup>, providing experimental support for the notion that amyloids formed by the A $\beta$  peptide are a causative factor in AD. Nevertheless, the repertoire of antibodies currently approved for advanced clinical trials is rather limited and many candidate antibodies in the trials have serious drawbacks, such as brain swelling.

Peptide-based inhibitors are a potential alternative to antibodies as a means to curtail amyloid propagation. Rational design of inhibitors based on the sequence of the amyloidogenic peptides allows for better targeting of specific regions and inhibition rates. The peptide-based inhibitors are developed as  $\beta$ -sheet breakers. Their complementary sequence will allow binding to the target peptide or protofilament and prevent additional monomers to be added. Peptides with modified backbones are particularly interesting, such as amide-to-ester substitution<sup>40</sup>, which destabilizes peptide assembly by removal of intramolecular hydrogen bonding. A particularly interesting backbone modification is *N* <sup>$\alpha$</sup> -methylation. *N*-methylated peptide are known to be efficient inhibitors against amyloid propagation<sup>41</sup>. *N*-methyl groups eliminate hydrogen-bonding interaction and create steric hindrance that is perturbing the binding of another partner. In the same manner, modification of the  $\alpha$ -carbon can also perturb the interactions and fibril growth by suppression of hydrogen bond interaction and creation of steric hindrance. Moreover, it creates severe conformational constraints<sup>42</sup> unfavorable to  $\beta$ -strand conformation. Peptides containing *D*-amino acids can also be designed as potent inhibitors. Mixtures of poly-*L*-lysine and poly-*D*-

lysine form fibrils more readily compared to enantiomerically pure solutions<sup>43</sup>. This suggests that *D*-peptides can possess a strong affinity for *L*- counterparts and, therefore, be modified to acquire strong inhibiting properties of *L*-amyloid formation. Recently, it was also demonstrated for A $\beta$ (1-42) peptide that mixing of both enantiomers accelerates the formation of fibrils that were identified as non-toxic<sup>44</sup>. So, in this case, it is not about inhibiting amyloid propagation but consuming all of the pro-toxic A $\beta$ (1-42) peptide to form non-toxic racemic fibrils instead.

## 1.5 Protein synthesis

### 1.5.1 Solution-phase peptide synthesis

In the beginning of the twentieth century, little was known about the composition and the structure of proteins. At that time, Emil Fischer wanted to elucidate enzyme's functioning by chemical synthesis. He became famous for the synthesis of carbohydrates and purines and was honored with a Nobel Prize in 1902 for his work. Beyond that, his contribution to the understanding of protein structure and polypeptide synthesis was remarkable. Emil Fischer synthesized the first dipeptide (glycyl-glycine) in 1901, which he obtained by opening of diketopiperazine (piperazine-2,5-dione)<sup>45</sup>. Some years later in 1907 he synthesized an octadecapeptide in solution composed of leucine and glycine residues (Leu-Gly<sub>3</sub>-Leu-Gly<sub>3</sub>-Leu-Gly<sub>9</sub>) by the use of acid chlorides for the coupling reactions<sup>46</sup>. It was a stepwise synthesis with successive peptide segment couplings in solution.

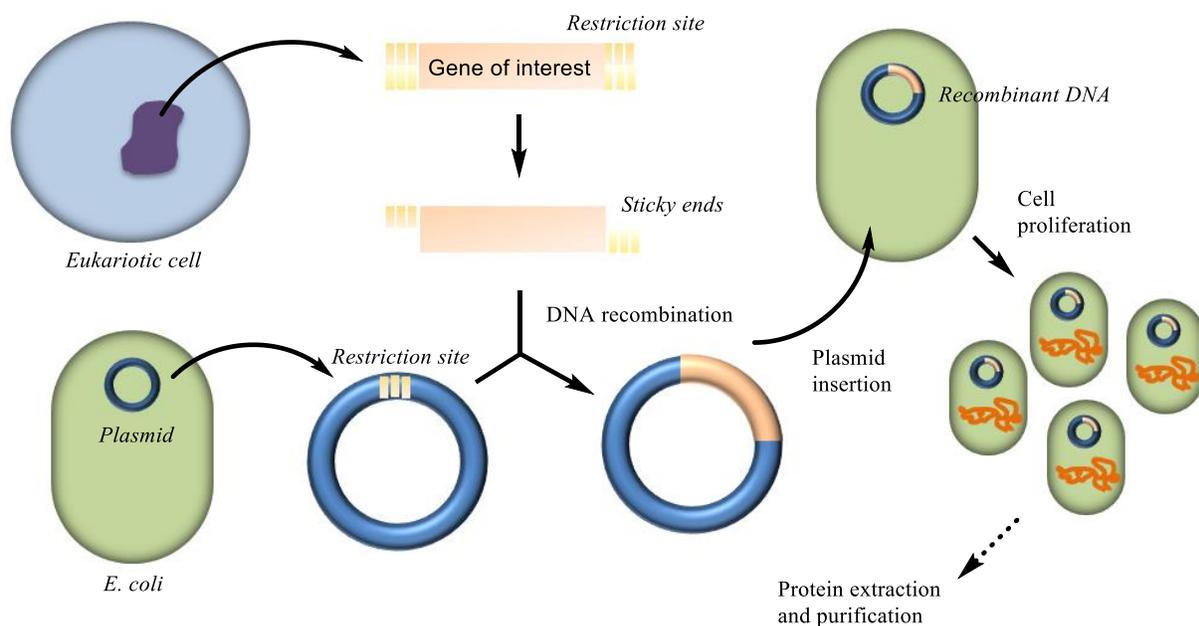
In 1954, Vincent du Vigneaud and co-workers synthesized for the first time biologically active oxytocin, a nine amino acid long peptide hormone<sup>47</sup>. The N-terminal parts of the segments were protected with carboxybenzyl (Cbz) groups and the side chain of cysteine residues were protected by benzyl (Bn) groups to avoid undesirable side reactions. Vincent du Vigneaud received a Nobel Prize in chemistry in 1955 for this work.

In the seventies, Peter Sieber and co-workers synthesized biologically active human insulin by segment condensation<sup>48</sup>. The synthesis of the fully active form of the ribonuclease A (124

amino acids) was accomplished in 1981 by Yajima and Fujii<sup>49</sup> and the *Aequorea* green fluorescent protein (238 amino acids) in 1998 by Shumpei Sakakibara and colleagues by fully protected segment coupling in solution<sup>50</sup>. Solution-phase peptide synthesis is a powerful, robust and well-established method for the synthesis of medium-sized proteins. But this technique suffers from many limitations. The full protection of the peptidic fragments leads to serious solubility issues and only low concentrations can be achieved in organic solvents, which is not convenient for coupling reactions. As a result, the reactions are slow and not complete. Moreover, the chiral integrity can be lost due to highly activated reaction conditions. It is also very hard to purify and characterize the intermediates possessing high molecular weights. All the steps in the synthesis are quite long and the yields are typically low. This method of protein synthesis by segment condensation of fully protected peptides is too specialized and in academic laboratories alternative methods are commonly used. These include bio-engineering strategies via recombinant DNA technologies<sup>51</sup> introduced in 1973 by Herbert Boyer and Stanley Cohen and solid-phase peptide synthesis<sup>52</sup> in combination with chemical ligation techniques<sup>53</sup> (described in sections 1.5.2 – 1.5.4).

### 1.5.2 Recombinant protein synthesis

Each protein is coded by a distinct nucleotide sequence in the genome and is produced by ribosomes in the cells. The ribosome uses messenger RNA as a template for the synthesis. The messenger RNA is read, one codon (three nucleotides) after another. Each codon is related to a specific transfer RNA that will deliver the required amino acids into the ribosome machinery. One by one, the amino acids are joined together to form the polypeptide chain. At the end of the synthesis, the chain is released and allowed to serve its predefined biological purpose. Sequencing of eukaryotic genomes has revealed the presence of more than 20 000 open reading frames which possibly encode for the same number of proteins<sup>54</sup>. Moreover, PTM of proteins and messenger RNA splicing increase the diversity of the total protein population.



**Figure 5:** Principle of protein production by recombinant synthesis from plasmid insertion to protein extraction.

To study the details of protein structure and function, intra- and inter-molecular interactions, they can be expressed in *Escherichia coli* via recombinant DNA-based technology using plasmids and restriction enzymes<sup>55</sup> (**Figure 5**). It has also become a powerful tool to synthesize mutants in order to investigate the role of specific amino acids within the sequence and to see the effects on the folding or reactivity of the mutated protein. For instance, alanine scanning mutagenesis of a polypeptide sequence allows for elucidation of the contribution of each individual amino acid<sup>56</sup>.

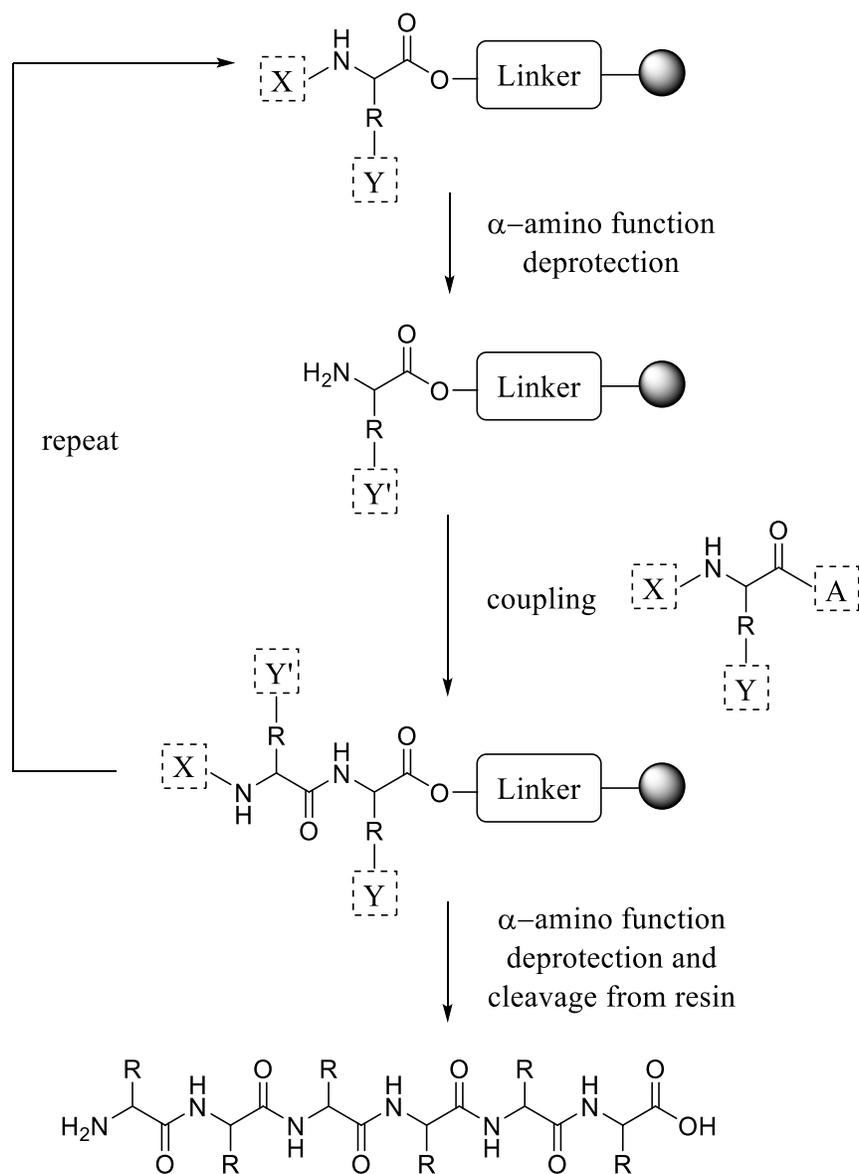
But this approach comes up against several limitations in terms of flexibility and diversity. Only the canonical amino acids can be promptly incorporated. It is true that many efforts have been done to expand the genetic repertoire<sup>57,58</sup>. By using engineered orthogonal transfer RNA-aminoacyl-transfer RNA synthetase pairs, dozens of unnatural amino acids with different side chain features can be incorporated into proteins via recombinant synthesis, increasing the scope for protein function and structure studies<sup>59</sup>. But this approach remains laborious and is not commonly used. Moreover, the production of post-translationally modified proteins at specific positions is challenging or even impossible.

Total chemical protein synthesis overcomes many limitations encountered by recombinant synthesis and allows rigorous atomic-level precision of the protein chemical structure. Various kinds of side chains are conceivable with different architectures or moieties. Moreover, non-peptidic scaffolds can be inserted. The peptide backbone can be modified and modulated; unnatural, mirror-image and “exotic” amino-acids can easily be introduced. Stable isotope labeling at specific positions can be obtained. Preparing conjugated proteins or peptides by introducing different bio-orthogonalities during solid-phase peptide synthesis (SPPS) is a prevailing advantage. It permits the easy insertion of fluorophores, linkers, ligands and the functionalization of specific amino acid side chains. It also allows the introduction at specific positions of many PTMs such as methylation, phosphorylation, acetylation, glycosylation or lipidation. All of these valuable tools offer many possibilities to study protein-protein interactions and molecular mechanisms.

### 1.5.3 Solid-phase peptide synthesis

To overcome all the laborious synthesis and purification steps associated with peptide synthesis in solution, in 1963 Robert Bruce Merrifield introduced an innovative technique that highly simplified the chemical synthesis of peptides called the solid-phase peptide synthesis<sup>52</sup> (SPPS) (**Figure 6**). This technique relies on the use of a cross-linked polymeric support<sup>60</sup> to which a  $N^\alpha$ -protected C-terminal amino acid of a peptidic sequence is covalently attached through its C-terminus. The  $N^\alpha$ -protecting group is then removed with the appropriate deprotecting agent and the support is washed and filtered. The solid support remains insoluble in the solvent commonly used for peptide synthesis but the growing polypeptide chain is microscopically solubilized. This double facet allows for deprotection and coupling reactions on the support and rapid purification by simple washings and filtrations. Consequently, a large excess of the next pre-activated amino acid can be introduced to drive the reaction to completion. The reaction is therefore faster and almost quantitative, yielding a dipeptide in high purity. Deprotection and coupling steps are repeated until the target sequence is obtained. In the final step, the peptide is detached from the support and the side-chain protecting groups are removed to afford the crude peptide product. Robert Bruce Merrifield received the Nobel Prize in chemistry in 1984 for his

breakthrough for the development of methodology for chemical synthesis on a solid matrix. Solid-phase peptide synthesis revolutionized and facilitated the synthesis of peptides. The simple approach and the standardization have made this technique universal and widely used around the world<sup>61</sup>.



**Figure 6:** Principle of solid-phase peptide synthesis.

The two most used strategies in SPPS are the Boc/Bn and Fmoc/<sup>t</sup>Bu approaches. Boc-SPPS relies on the use of *N*<sup>α</sup>-Boc (*tert*-butyloxycarbonyl) protected amino acid building blocks. It is

simply removed by trifluoroacetic acid (TFA) treatment. The cleavage from the resin and global side-chain deprotection are realized using strong acids. Typically, liquid hydrogen fluoride is employed. Some alternatives to this hazardous chemical have been investigated like trifluoromethanesulfonic acid (triflic acid, TfOH) to allow this technique to be safer<sup>62</sup>. An alternative method, Fmoc-SPPS, has been developed to avoid highly hazardous conditions. In this case, *N*<sup>α</sup>-Fmoc (9-fluorenylmethoxycarbonyl) protected building blocks are coupled and the amino group is deprotected with 20% piperidine in *N,N*-dimethylformamide (DMF). The cleavage and global deprotection is performed with 95% TFA. These conditions are easier to set up and furthermore enable automation of the procedure by peptide synthesizers.

But only small to medium sized peptides can be readily prepared by this method. In fact, this technique is limited to approximately fifty amino acid long sequences. Above this limit, coupling reactions strongly tend to be incomplete, leading to an increase in the number of by-products, which complicates purification of target polypeptide. This limitation was overcome later by the introduction of peptide ligation techniques.

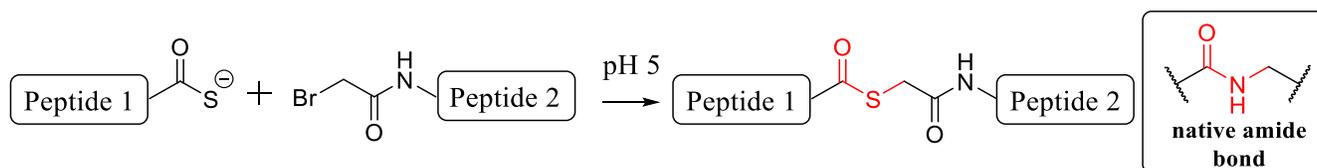
#### 1.5.4 The chemical ligation of peptides

The concept of peptide ligation relies on the use of unprotected peptidic segments to synthesize longer sequences and, therefore, accomplish total chemical synthesis of proteins in a convergent manner, thus overcoming peptide size limitations in SPPS.

In the first step, small to medium-sized fragments are synthesized by SPPS and purified by preparative reversed phase high performance liquid chromatography. Unprotected peptides present many advantages compared to those used in segment condensation. A large quantity of the starting material for the peptide ligation can readily be obtained in high purity. They can be more easily handled, because solubility issues are overcome and a highly concentrated solution can be prepared using a mixture of water and acetonitrile or buffers containing chaotropes such as 6 M guanidine hydrochloride or 8 M urea.

The ligation reaction must be highly chemoselective, taking into account that peptides contain many chemical functionalities with distinct reactivities (phenols, amines, alcohols, heteroaromatic rings, thiols or thioethers). Many strategies have been introduced to ligate peptides together using bio-orthogonal chemistries in aqueous media.

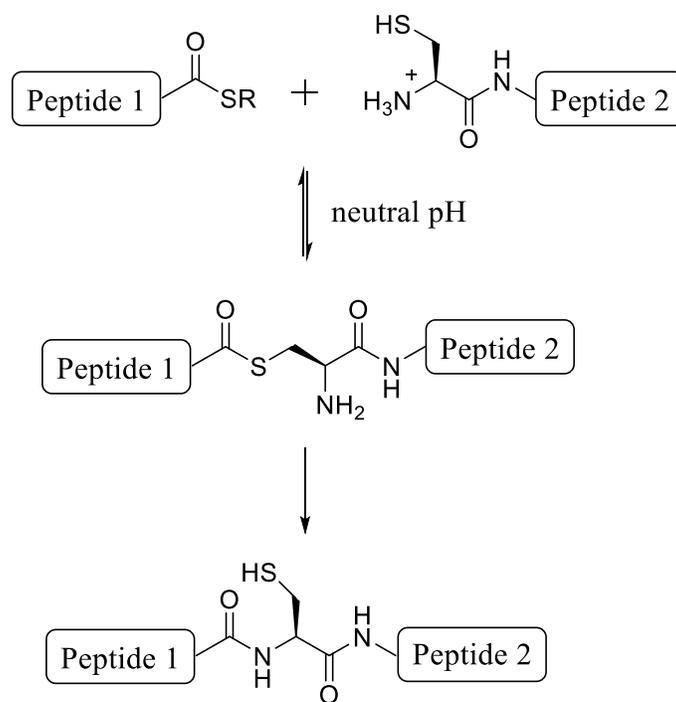
The first chemoselective ligation of unprotected peptides was accomplished by Martina Schnölzer and Stephen Kent in 1992<sup>63</sup> for the synthesis of a human immunodeficiency virus-1 (HIV-1) protease analogue composed of 99 amino acids. They described the formation of a non-native bond at the ligation site between two peptide segments (**Figure 7**). The ligation occurred between a C-terminal peptide- $\alpha$ -carboxy thioacid and an N-terminal bromoacetylated peptide forming a thioester bond at a glycine-glycine junction. In this case, the peptide bond surrogate (or pseudo-peptide) possesses structural similarities with a native amide bond. They prepared a backbone-engineered protein with unique properties allowing for structure-activity studies. It resulted in the formation of a fully active HIV-1 protease analogue.



**Figure 7:** Non-native thioester bond formation, nearly isosteric to native amide bond for the production of protein analogues.

Two years later, in 1994, Kent and co-workers published an unprecedented chemical synthesis of 72 amino-acid long human interleukin 8 (IL-8) by chemical ligation<sup>53</sup> between two unprotected peptides yielding a native amide bond at the ligation site. This was a significant breakthrough in the field of chemical protein synthesis since it is an easily applicable, robust and efficient method to produce proteins with medium molecular weights. Convergent synthesis with several peptide fragments can be realized giving access to protein with more than 200 residues. Native chemical ligation allowed many research teams around the world to synthesize protein analogues with high purity and atomic-level control and, currently, it is the most frequently used method to prepare synthetic proteins.

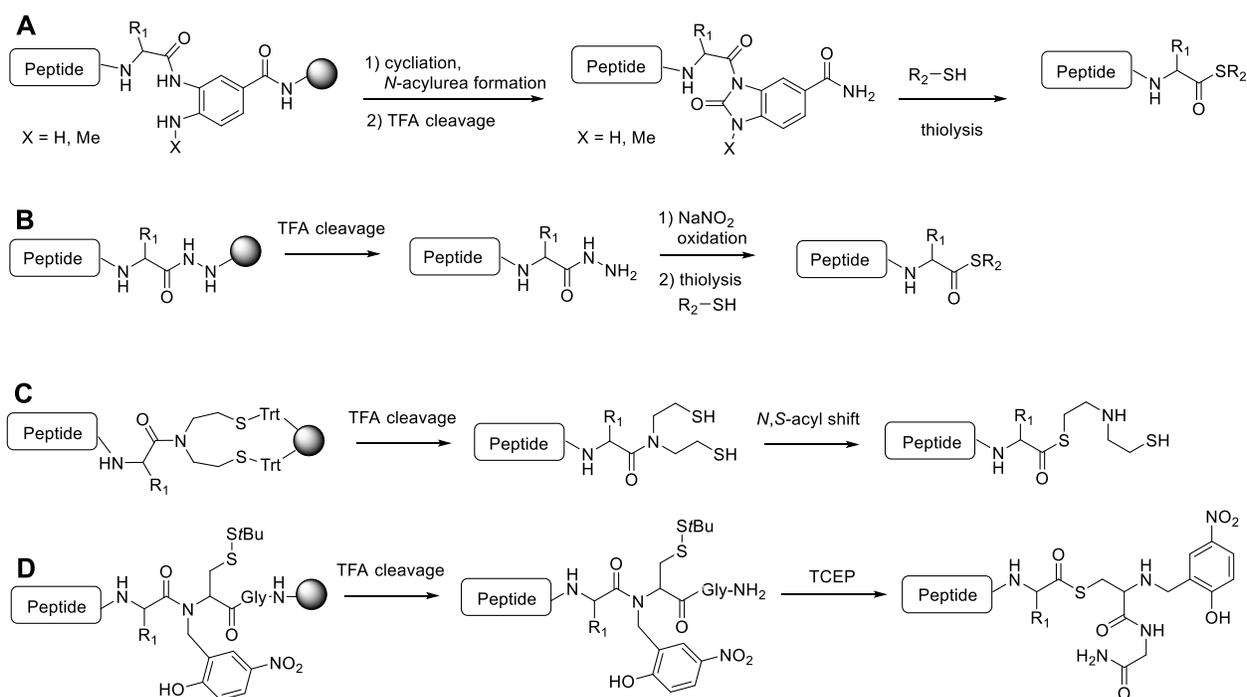
The technique relies on the distinct reactivity of thiols towards thioesters in aqueous media. The ligation occurs between a C-terminal  $\alpha$ -thioester of one peptide and an N-terminal cysteine of another peptide (**Figure 8**). The thiol moiety of the N-terminal cysteine will selectively react with the thioester through a transthioesterification reaction producing a thioester-linked intermediate. The thiol-thioester exchange is reversible but the intermediate rapidly undergoes *S*-to-*N* acyl shift with the free  $\alpha$ -amine of the cysteine residue to form an irreversible native amide bond. The rate of the ligation depends on the nature of the amino acid bearing the thioester moiety. Bulky residues will slow down the rearrangement step. It is also dependent on the nature of the thiol leaving group during the transthioesterification. This reaction is often catalyzed by the addition of aryl thiols<sup>64,65</sup>. Thiol-thioester exchange with aromatic thiols produces highly reactive thioesters that accelerate the ligation reaction and increase the yield at neutral pH.



**Figure 8:** The native chemical ligation, a reversible and rate-determining transthioesterification followed by an irreversible intramolecular *S*-to-*N* acyl shift.

In 2001, Philip Dawson and Liang Yan widened the scope of possibilities by combining native chemical ligation with an additional desulfurization step of the thiol moiety of the cysteine that is involved in the ligation reaction<sup>66</sup>. In this way, the cysteine is converted into an alanine residue. This extension allows the synthesis of proteins without cysteine residues especially because cysteines are rather rare. Later on, synthesis of amino acid building blocks with different thiolated side chains<sup>67-70</sup> and auxiliaries<sup>71,72</sup> extended the concept to many other residues.

In order to perform native chemical ligation, preparation of peptide  $\alpha$ -thioesters is required, which can be readily obtained by Boc-SPPS after cleavage from the solid support<sup>53,73,74</sup>. This is not the case with Fmoc-SPPS due to repetitive usage of nucleophilic amines during Fmoc removal steps which are not compatible with thioester linkages. Different strategies were established in order to circumvent this problem. Peptides with C-terminal modifications called crypto-thioesters were designed in order to be synthesized on resins and be converted afterwards into thioesters. Some methods are presented in **Figure 9**.

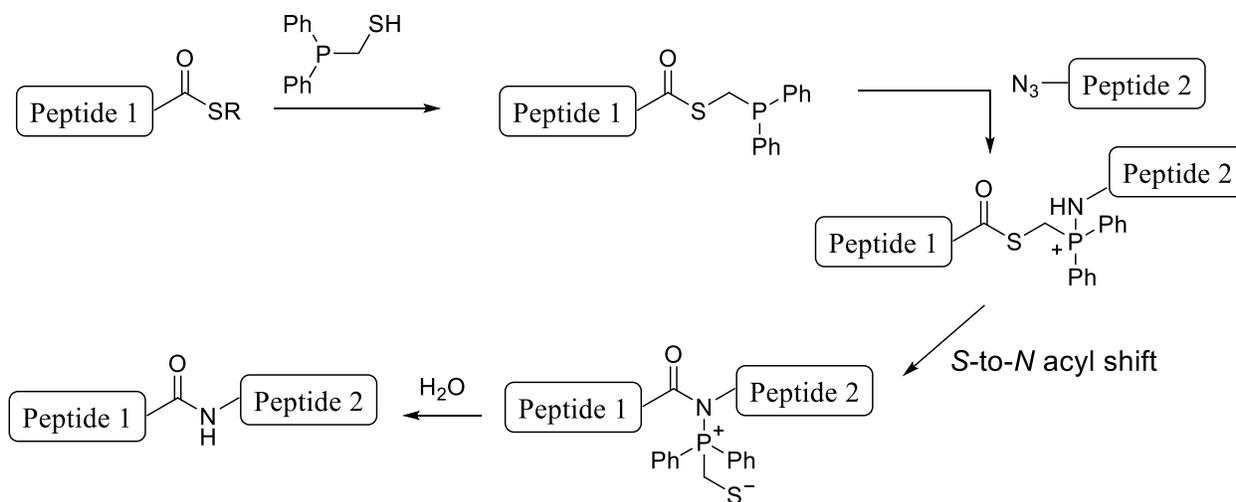


**Figure 9:** Different methods to prepare peptide  $\alpha$ -thioesters. A) Dawson's acylurea strategy<sup>75,76</sup>, B) Liu's peptide hydrazides strategy<sup>77,78</sup>, C) Melnyk's SEA (bis(2-sulfanylethyl)amino) strategy<sup>79,80</sup> and D) Aucagne's *N*-(2-hydroxybenzyl)cysteine strategy<sup>81</sup>.

In 2000, the traceless Staudinger ligation of peptides was developed by Ronald Raines, which is based on the Staudinger reduction reaction and on the chemoselectivity of phosphines towards azides<sup>82-84</sup>. The principle of this peptide ligation technique is the reaction between a C-terminal phosphinothioester peptide with an N-terminal azidopeptide (**Figure 10**).

Mechanistically, the phosphine reacts with the azide moiety to form a phosphazide intermediate, which rearranges *via* an iminophosphorane and liberates dinitrogen. The iminophosphorane undergoes acylation through *S*-to-*N* acyl transfer to yield the native amide bond after hydrolysis.

This technique overcomes the necessity to have a cysteine residue or derivatives at the ligation site and can be potentially applied with any kind of amino acid. The stereochemistry is retained during the reaction and there is no need for special conditions.

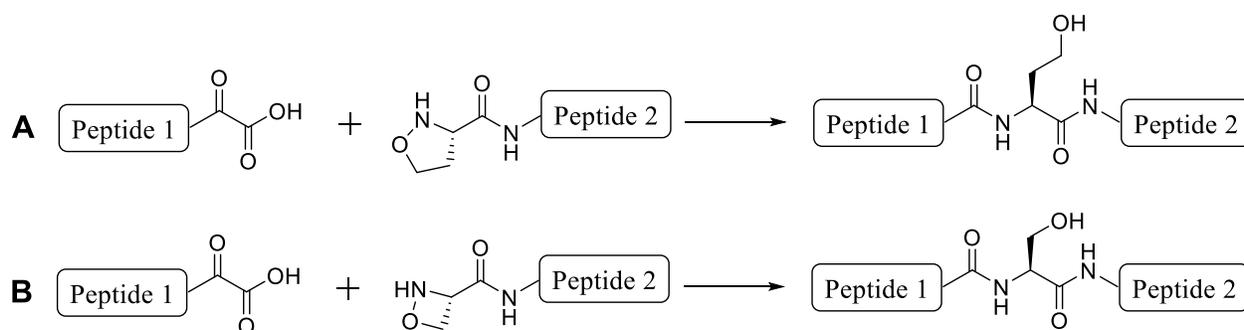


**Figure 10:** The traceless Staudinger ligation.

In 2012, Jeffrey Bode and co-workers reported the chemical synthesis of the prokaryotic-ubiquitin-like protein (Pup, 63 amino acids) and the cold shock protein A (cspA, 66 amino acids) by chemoselective  $\alpha$ -ketoacid-hydroxylamine ligation between a cyclic N-terminal alkoxyamine (5-membered ring 5-oxaproline) peptide and a C-terminal  $\alpha$ -ketoacid peptide<sup>85</sup> (**Figure 11-A**). The postulated mechanism is rather complex. The condensation of the alkoxyamine on the ketoacid produces a hemiaminal, which upon water elimination gives a nitron intermediate. The

nitrene undergoes cyclization by attack of carboxylate to form an  $\alpha$ -lactone that upon rearrangement and decarboxylation yields the native amide bond<sup>86</sup>. In those two cases, the use of 5-oxaproline leads to the formation of homoserine, a non-canonical residue at the ligation site. This method was also used to chemically synthesize a 184-amino acid long nitrophorin 4 protein, which upon folding was able to insert ferric heme and bind nitric oxide<sup>87</sup>, proving the robustness of the technique.

In 2015, the synthesis of a four-membered ring alkoxyamine, the Fmoc-oxazetidione, allowed rapid serine-forming ligations (**Figure 11-B**) and enabled the synthesis of a native calcium-binding protein (S100A4, 101 amino acids)<sup>88</sup>.



**Figure 11:** The  $\alpha$ -ketoacid-hydroxylamine (KAHA) ligation using the 5-membered ring 5-oxaproline (A) or 4-membered ring oxazetidione (B) leading to the formation of homoserine or native serine, respectively, at the ligation site.

## 1.6 Thesis outline

Chemical protein synthesis is a powerful tool that allows obtaining chemically modified polypeptides such as isotopically-labeled, mirror-image or *N*-methylated or various proteins. In the following chapters, I report on the use of solid-phase peptide synthesis, peptide ligations and total chemical synthesis of proteins to produce very distinct and innovative peptide assemblies to dissect structural properties and functions of amyloids. To do so, we used many advanced analytical methods such as solid-state NMR, solution NMR, ATR-FTIR spectroscopy, TEM, cryo-EM, circular dichroism or fluorescence spectroscopy. In the second chapter, I report on the synthesis of mirror-image polypeptides in order to study the chiral recognition in amyloid fibers and the use of *D*-peptides for the inhibition of amyloid growth. In the third chapter, I present a novel approach to produce distinct amyloid polymorphs by the use of covalent tethering or by supramolecular assemblies. In the fourth chapter, I report a highly innovative method for the synthesis of covalently-tethered oligomers that have structural properties similar to amyloid fibrils. Finally, in the fifth chapter, I demonstrate the use of *N*-methylated peptides for the selective recognition and inhibition of specific amyloid polymorphs.

## **2. Chiral recognition in amyloid fiber growth**

Based on the published manuscript: Vladimir Torbeev, Marcel Grogg, Jérémy Ruiz, Régis Boehringer, Alicia Schirer, Petra Hellwig, Gunnar Jeschke and Donald Hilvert, “Chiral recognition in amyloid fiber growth”. *Journal of Peptide Science* **22**, 290-304 (2016).

### **2.1 Introduction**

In protein misfolding diseases, the inhibition of amyloid formation has been proposed as a possible therapeutic strategy. Interestingly, the complete inversion of peptide backbone stereochemistry, i.e. conversion of *L*-peptide inhibitors into *D*-peptides, can afford commensurate or even enhanced inhibitory efficacy<sup>89</sup>. A significant pharmacological advantage of *D*-peptides is their stability against proteases<sup>90</sup> and reduced immunogenicity upon repeated administration<sup>91</sup>. *D*-peptide inhibitors with sequences unrelated to the target peptide or protein have been designed computationally against tau protein fibers<sup>92</sup> and identified by “mirror-image” phage display against A $\beta$ (1-42) amyloids<sup>93</sup>. The high selectivity of *D*-peptides for A $\beta$ (1-42) was exploited in the development of novel diagnostic agents for Alzheimer’s disease<sup>94</sup>. Despite their attractiveness as potential inhibitors, it is not entirely clear how *D*-peptides interact with *L*-amyloid surfaces and whether there are any general rules that might be used in structure-guided design of novel therapeutics or diagnostics for protein misfolding diseases. Experiments on racemic mixtures of enantiomeric amyloidogenic peptides provide some insight into the nature of these interactions and show that the *D*-enantiomer of a particular peptide may not always act as an inhibitor of amyloid growth.

One of the earliest studies on co-aggregation of enantiomeric polypeptides was reported for a mixture of poly(*L*-lysine) and poly(*D*-lysine)<sup>95</sup>. While solutions of the individual enantiomers remained clear above pH 11, a precipitate forms readily upon mixing the two solutions. Recent re-investigation of this finding confirmed that the precipitate is composed of amyloid-like fibers containing racemic, antiparallel  $\beta$ -sheets<sup>96</sup>. The explanation for such an outcome is that racemic

cross- $\beta$  architecture with “knobs-into-holes” packing better accommodates the side chains of lysine residues and results in a more stable aggregate. A racemic mixture of poly(*L*-glutamic acid) and poly(*D*-glutamic acid) similarly produces amyloid-like fibers at pH 4 and 65°C. However, in this case, the racemic precipitate was found to be less structured and less stable than fibers obtained from the individual enantiomers<sup>97</sup>. Mixed racemic  $\beta$ -sheet structures were also observed for racemates of other homo-oligomers, such as poly-valine and poly-leucine, albeit not in fibrous form<sup>98</sup>.

Co-polymerization of *D*- and *L*-peptides into composite amyloid-like fibers has been reported for de novo designed 20-residue  $\beta$ -hairpin peptides. These copolymers form hydrogels that have enhanced mechanical rigidity when compared to hydrogels formed by individual enantiomers<sup>99</sup>. In another study, designed 8-residue amphipathic *L*- and *D*-peptides co-assembled into racemic amyloids<sup>100</sup>. Moreover, it was shown that mixed racemic amyloids are thermodynamically more stable than fibers composed entirely of a single enantiomer. Molecular modelling of racemic and enantiopure antiparallel  $\beta$ -sheets for these peptides showed that the racemic structure has the advantage that the  $\beta$ -strands are perfectly aligned, whereas in enantiomerically pure  $\beta$ -sheets neighboring  $\beta$ -strands are one residue out-of-register, leaving hydrogen bond donors/acceptors and side-chains of terminal amino acids unpaired.

Opposite behavior was observed for two *L*- and *D*-enantiomers of a synthetic hybrid peptide, in which hydrophobic tetra-leucine fragments were attached to both the N- and C-termini of a pH-responsive (Lys)<sub>8</sub> fragment<sup>101</sup>. At pH 9, the pure enantiomers readily formed amyloid-like fibers with opposite helical twist, whereas the racemic mixture produced only spherical aggregates. Nevertheless, by using FRET dyes as labels, it was possible to show that the two enantiomers interact and form hybrid heterochiral complexes. However, it may well be that although a heterochiral  $\beta$ -sheet-like interface forms, the more complex heterochiral fiber superstructure is not as stable as in the corresponding enantiomers. Alternatively, for this particular hybrid peptide the heterochiral  $\beta$ -sheet layer may not be as stable as its homochiral counterpart.

Adding preformed amyloids (‘seeds’) as a template to induce fiber propagation from solutions of non-polymerized amyloidogenic peptides was shown to be specific with respect to amino acid sequence and polypeptide conformation<sup>102</sup>. Such complementarity in molecular

recognition underlies biological mechanisms for transmission of molecular information in prions. Analogously, molecular interactions between chiral surfaces of amyloid seeds formed by mirror image enantiomers of a polypeptide and the corresponding peptides in solution are expected to be complementary as well and exhibit high levels of enantioselectivity. An illustrative example of biomolecular stereoselectivity is the chemically synthesized *D*-enantiomer of HIV-1 protease, which catalyzes efficient hydrolysis of *D*-peptide substrates but is inactive against the corresponding *L*-enantiomers<sup>103</sup>.

Strikingly, for amyloidogenic *D*-polyglutamine and *L*-polyglutamine peptides seeding was found to be non-enantioselective<sup>104</sup>. Instead, mirror-image amyloids had similar seeding efficacy as amyloids derived from the cognate enantiomer. It was demonstrated that both *D*-polyglutamine and *L*-polyglutamine peptides are taken up by mammalian cells and exhibit comparable levels of toxicity. Furthermore, both could sequester ribosomally produced polyglutamine polypeptides. Finally, it was shown that equimolar mixtures of *L*- and *D*-peptides form amyloid aggregates faster than the individual enantiomers. A definitive conclusion on the chiral composition of these fibers was not reached, however.

The situation is completely reversed for A $\beta$ (1-40) peptides, where amyloid growth was shown to be highly stereoselective. Preformed fibers ('seeds') of either all-*L* or all-*D* enantiomers were only found to be good templates for deposition of cognate enantiomers from solutions of non-polymerized A $\beta$ (1-40). Moreover, *D*-A $\beta$ (1-40) enantiomer did not inhibit deposition of *L*-A $\beta$ (1-40) onto *L*-A $\beta$ (1-40) fibers. Analogously, in experiments with the A $\beta$ (25-35) fragment, monomers of *D*-A $\beta$ (25-35) did not extend seeds of *L*-A $\beta$ (25-35) seeds and vice versa<sup>105,106</sup>.

A strong preference for homochiral self-association between two  $\beta$ -stands was confirmed for a model  $\beta$ -sheet peptide, which was engineered in such a way that only one  $\beta$ -strand edge was available for dimerization and infinite  $\beta$ -sheet assembly was impaired by an unnatural building block<sup>107</sup>. In a low polarity solvent such as chloroform, predominantly homochiral dimers were observed: such high enantioselectivity was presumably possible because of favorable stacking of hydrophobic side chains, which is easier to realize if the dimer interface has the geometry of a homochiral  $\beta$ -sheet.

Overall, from the examples summarized above, it is evident that our understanding of chiral recognition in amyloid growth is still incomplete. To explain this phenomenon, it is essential to consider not only molecular interactions between two adjacent  $\beta$ -strands or a single  $\beta$ -sheet layer, but also to include fiber superstructure in the analysis. Given the increasing availability of structural information on polypeptide organization in amyloids, it is tempting to speculate that chiral specificity is dependent on the length of the polypeptide and the conformation it adopts within fibers, whether it is a stack of  $\beta$ -sheets formed by rather short peptides or a more complex organization exemplified by super-pleated  $\beta$ -structures and  $\beta$ -helices ( $\beta$ -solenoids).

In this chapter, data on amyloid formation for two distinct polypeptides of different length and structural organization are provided:

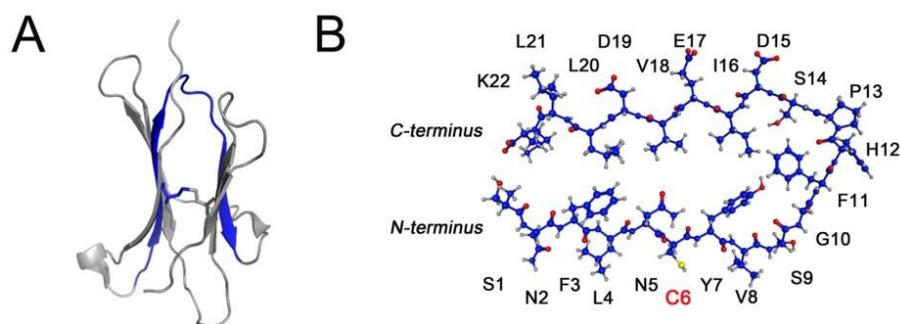
- a 44-residue covalently-linked dimer derived from the [20-41]-fragment of human  $\beta$ 2-microglobulin;
- the full-length 99-residue human  $\beta$ 2-microglobulin ( $\beta$ 2m).

We analyzed the chiral specificity of amyloid growth, as well as chiral cross-seeding and cross-inhibition effects for these systems.

## **2.2 Chemical synthesis and properties of covalently-linked constructs based on the [20-41]-fragment of $\beta$ 2-microglobulin**

The 99-residue protein  $\beta$ 2-microglobulin ( $\beta$ 2m) is the light chain of the type I major histocompatibility complex and is implicated in dialysis-related amyloidosis. Upon proteolysis of  $\beta$ 2m (**Figure 12-A**) by a lysine-specific protease from *Achromobacter lyticus* and subsequent reduction of disulfide bridges, the 22-residue [20-41]fragment (called K3 peptide) was isolated and shown to form amyloid fibers spontaneously. It was hypothesized that [20-41] $\beta$ 2m may represent the minimal sequence needed for  $\beta$ 2m amyloid formation. Indeed, a certain level of structural homology between amyloids formed by K3 and full-length  $\beta$ 2m exists since [20-41] $\beta$ 2m amyloid seeds accelerate amyloid growth of  $\beta$ 2m and vice versa<sup>108</sup>.

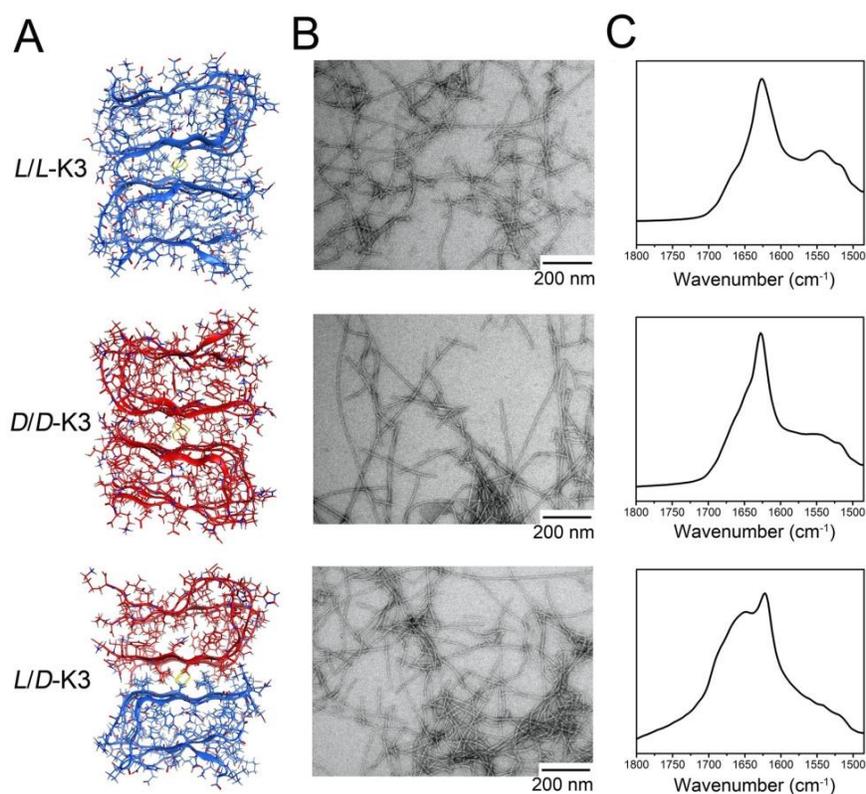
The structure of [20-41] $\beta$ 2m within amyloid protofilaments was solved by solid-state NMR (ssNMR). The peptide assembles as an infinite stack of staggered  $\beta$ -arches along the long fiber axis<sup>27</sup> (**Figure 12-B**). Both  $\beta$ -sheets in these cross- $\beta$  structures are parallel and in-register. Importantly, one exterior face of the protofilament consists mostly of hydrophilic residues, whereas the other is more hydrophobic. The latter also contains a single cysteine residue in the center of each strand which can be used for covalent linking to form the dimeric construct. For the published structural study, samples were prepared in 20% (v/v) 2,2,2-trifluoroethanol (TFE) containing 10 mM HCl in order to generate protofilaments of homogeneous thickness (~1.5 nm on average), which are required for well-resolved and interpretable ssNMR spectra. The presence of the organic co-solvent (TFE) presumably helps to stabilize the amphiphilic [20-41] $\beta$ 2m conformation against lateral aggregation. Indeed, under different solvent conditions, such as aqueous buffers at acidic or close to neutral pH, fibers were obtained with heterogeneous thickness distributions (up to 9 nm) and variable morphology, including bundles of multiple proto-fibers. Such heterogeneity was attributed to (partial) oxidation of the cysteine in [20-41] $\beta$ 2m to give covalently tethered disulfide dimers. When a [20-41] $\beta$ 2m homodimer was obtained in pure form, it was found to possess enhanced aggregation rates both at acidic and neutral pHs<sup>109</sup>.



**Figure 12:** The [20-41] $\beta$ 2m fragment (K3 peptide). (A) Corresponding fragment (in blue) in the context of natively folded full-length human  $\beta$ 2m. (B) The  $\beta$ -arch loop conformation of the K3 peptide elucidated by solid-state NMR in K3 amyloid protofilaments (PDB ID 2E8D). The axis of the amyloid fiber is perpendicular to the image. *Figure is taken from J. Pept. Sci.* 22, 290-304 (2016).

In a separate study, the enantiomeric *L*- and *D*-forms of [20-41] $\beta$ 2m were synthesized. Both produced amyloids with mirror-image circular dichroism spectra. Moreover, the *L/D*-mixture also formed amyloids morphologically identical to fibers obtained from individual enantiomers.

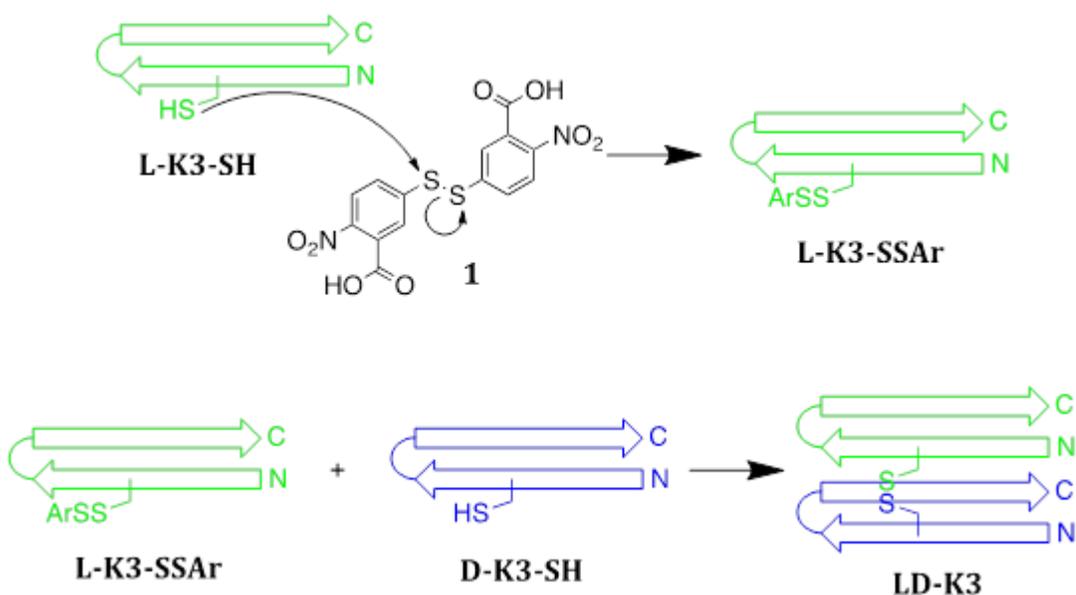
Cross-seeding was shown to be enantioselective but the chiral composition of fibers obtained from the racemic mixture was not unambiguously resolved<sup>110</sup>. In this study, we build on these observations and extend research on chiral recognition in amyloid growth to dimeric disulfide-linked K3 constructs of homochiral (*L/L*-K3 and *D/D*-K3) as well as heterochiral (*L/D*-K3) composition. The amphiphilic nature of the monomeric K3 protofilament suggests that two such protofilaments are predisposed to bind together along their more hydrophobic external surface.



**Figure 13:** Chemically synthesized disulfide-linked dimers of the [20-41] $\beta$ 2m fragment of homochiral and heterochiral composition. (A) Molecular models of disulfide-linked [20-41] $\beta$ 2m homodimers (*L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m) synthesized from either all-*L* (in blue) or all-*D* (in red) amino acids, respectively, and a heterodimer (*L/D*-[20-41] $\beta$ 2m), in which an all-*L* [20-41] $\beta$ 2m -peptide (in blue) is covalently linked to an all-*D* [20-41] $\beta$ 2m -peptide (in red). Disulfide bonds linking  $\beta$ -arches are shown in yellow. According to molecular modelling the *trans* orientation of  $\beta$ -arches (as depicted) in covalent dimers is advantageous in comparison to *cis* based on better steric, H-bonding and electrostatic interactions at the dimer interfaces. (B) TEM images of fibers obtained from the respective constructs taken after 2 days of growth. (C) ATR-FT-IR spectra of the corresponding amyloids. *Figure is taken from J. Pept. Sci. 22, 290-304 (2016).*

As a consequence, the disulfide bond formed by two cysteine residues in the dimeric constructs should covalently cement such assembly. Molecular models of the corresponding dimers are shown in **Figure 13-A**.

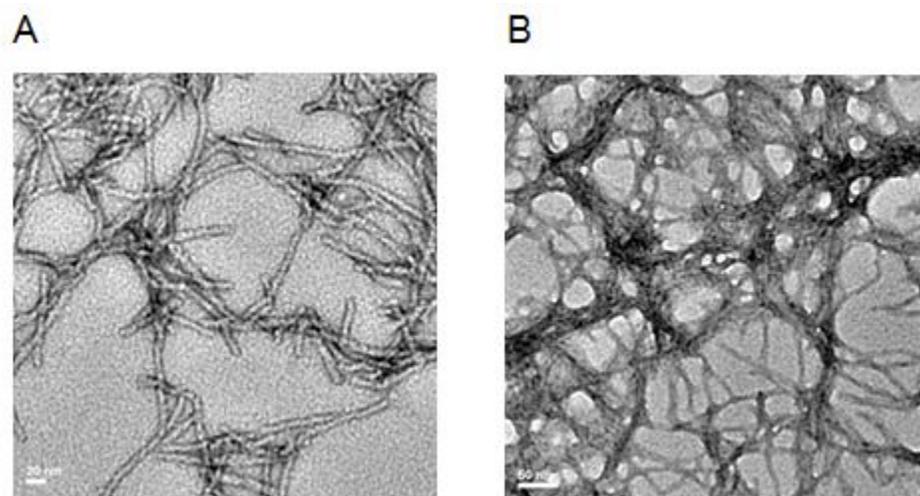
Enantiomeric homodimers were synthesized by air-oxidation of the corresponding *L*- and *D*-[20-41] $\beta$ 2m peptides in 7 M guanidine hydrochloride at pH 9. For the synthesis of the *L/D*-[20-41] $\beta$ 2m heterodimer we used a directed disulfide-bond formation strategy, first synthesizing an activated mixed disulfide with Ellman's reagent, *L*-[20-41] $\beta$ 2m-Cys(S-S-Ar), and then displacing 2-nitro-5-thiobenzoate by the cysteine group of the *D*-[20-41] $\beta$ 2m peptide (**Figure 14**).



**Figure 14:** Synthesis of mixed *L/D*-[20-41] $\beta$ 2m dimers. (A) Activation of *L*-[20-41] $\beta$ 2m (L-K3-SH) with Ellman's reagent (1); (B) reaction between activated *L*-[20-41] $\beta$ 2m (L-K3-SSAr) and the free thiol of *D*-[20-41] $\beta$ 2m (D-K3-SH). *Figure is taken from J. Pept. Sci.* 22, 290-304 (2016).

Amyloid fibers were grown for all the synthesized constructs at pH 7.5 with and without TFE (**Figure 15**). Within a few days it was possible to obtain fibers for both homodimers and the heterodimer. The fibers were rather short, of uniform thickness, with flat surfaces and without noticeable helical twist. This morphology did not change upon longer incubation (up to 25 days). The fiber thickness measured in negative stain TEM images was identical for all constructs

within the experimental error (*L/L*-[20-41] $\beta$ 2m,  $5.2 \pm 0.8$  nm; *D/D*-[20-41] $\beta$ 2m,  $6.2 \pm 0.9$  nm; *L/D*-[20-41] $\beta$ 2m,  $5.3 \pm 0.6$  nm). Qualitatively similar data were obtained by atomic force microscopy (AFM) (*L/L*-[20-41] $\beta$ 2m,  $4.0 \pm 1.3$  nm; *D/D*-[20-41] $\beta$ 2m,  $3.7 \pm 1.0$  nm; *L/D*-[20-41] $\beta$ 2m,  $3.3 \pm 1.3$  nm).

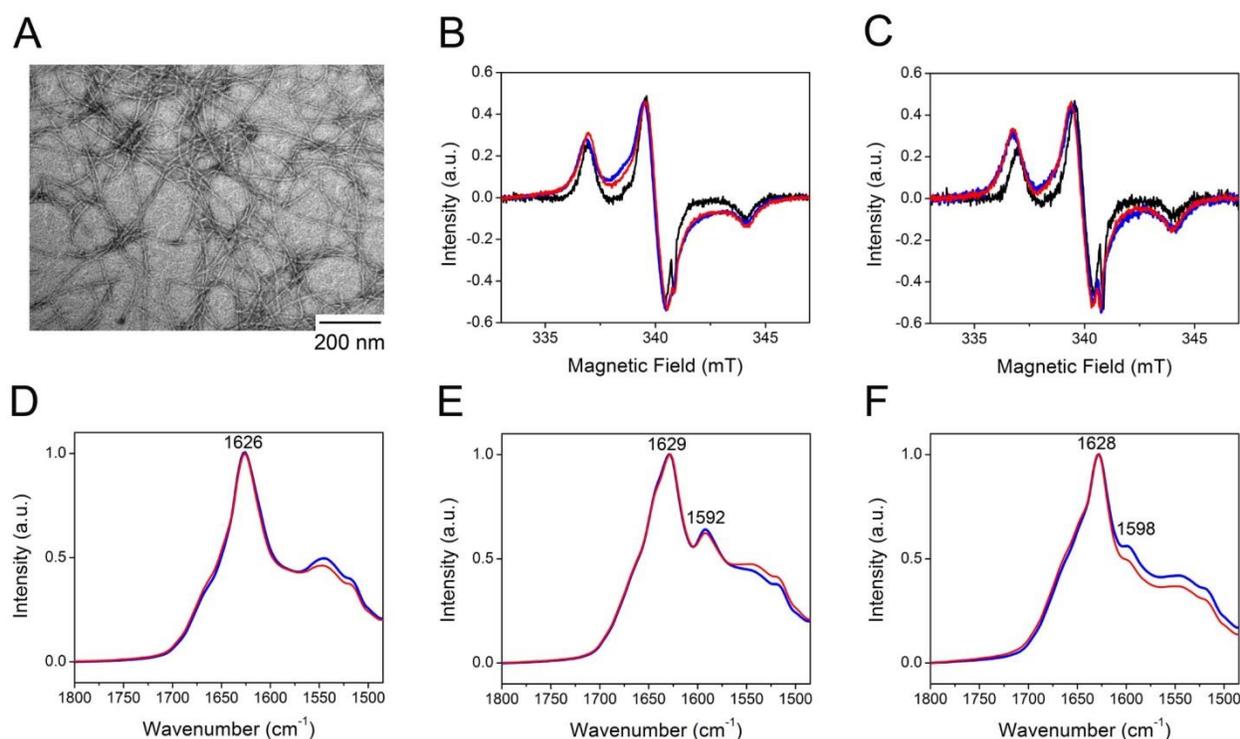


**Figure 15:** Amyloid fibers of *L/L*-[20-41] $\beta$ 2m obtained in the presence of trifluoroethanol (TFE). (A) 10% (v/v) TFE. (B) 20% (v/v) TFE. *Figure is taken from J. Pept. Sci. 22, 290-304 (2016).*

Kinetics of amyloid growth was monitored by thioflavin T (ThT) fluorescence in a continuous assay. Homodimers formed fibrils much faster than [20-41] $\beta$ 2m monomers, without a lag phase, as reported previously<sup>109</sup>. In contrast, the *L/D*-[20-41] $\beta$ 2m heterodimer formed fibrils with slower kinetics and with a significant lag period. This result suggests that formation of a heterochiral  $\beta$ -sheet interface slows down fiber elongation.

Attenuated total reflection Fourier transform IR (ATR-FT-IR) spectra were recorded for all three constructs (**Figure 16**). The amide I band shows characteristic frequencies ( $1626\text{ cm}^{-1}$ ) for a parallel  $\beta$ -sheet in the case of the *L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m homodimer constructs, whereas the corresponding band for the *L/D*-[20-41] $\beta$ 2m heterodimer is slightly red-shifted ( $1621\text{ cm}^{-1}$ ). Secondary structure deconvolution of the amide I and amide II bands for the *L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m dimers shows similar contributions from  $\beta$ -strand (50-60%), turn

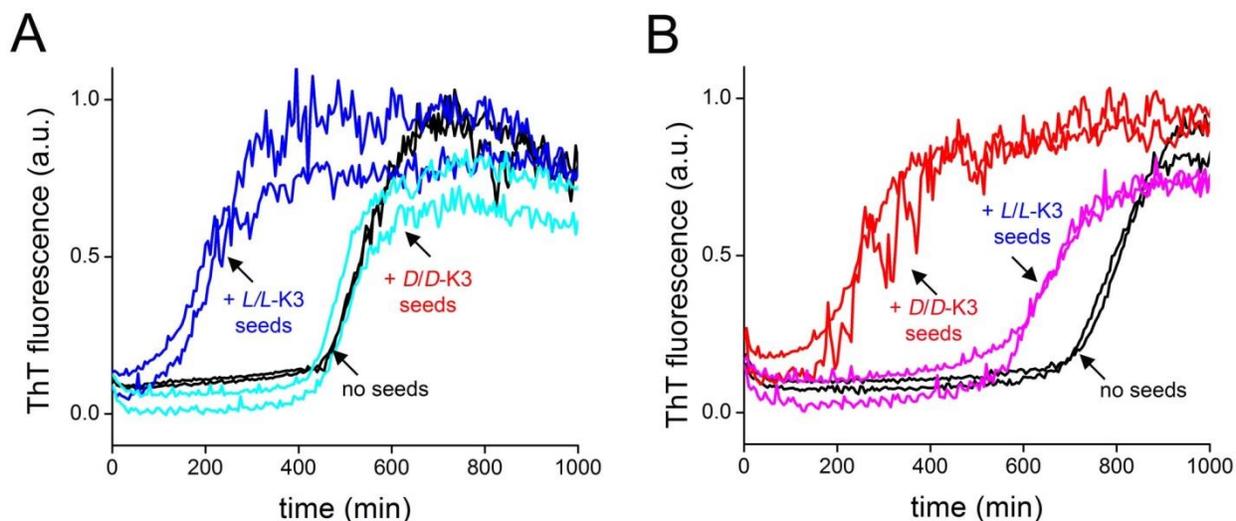
(25-30%) and random coil (20-25%) structures. Inspection of the amide I and the amide II bands for *L/D*-[20-41] $\beta$ 2m revealed an additional frequency at 1691  $\text{cm}^{-1}$  and contributions from two  $\beta$ -sheet structures under the amide I band (at 1635 and 1621  $\text{cm}^{-1}$ ), possibly originating from a mixture of parallel and antiparallel  $\beta$ -sheets. Deconvolution of the spectrum suggests that the *L/D*-[20-41] $\beta$ 2m heterodimer sample contains 44%  $\beta$ -strand (parallel and antiparallel), 33% turns, and 23% random coil.



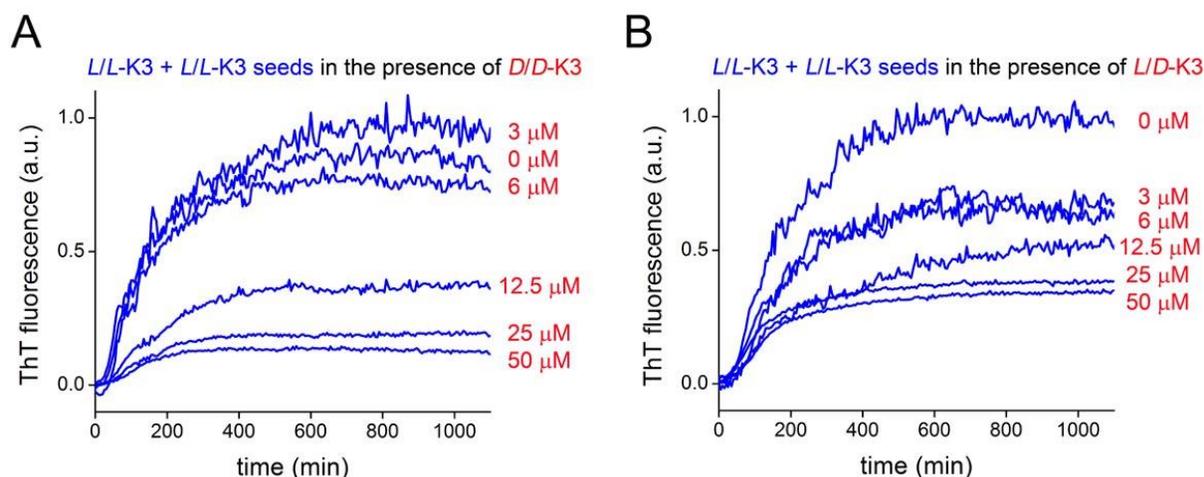
**Figure 16:** Properties of amyloids grown from racemic (*L/L*-[20-41] $\beta$ 2m + *D/D*-[20-41] $\beta$ 2m) mixtures. (A) TEM image of amyloid fibers grown from a 1:1 mixture of *L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m (total peptide conc. 100  $\mu\text{M}$ ). (B) CW-EPR spectrum of amyloids obtained from a quasi-racemic mixture of N-terminally nitroxide-labeled *L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m (in red) overlaid on spectra for the corresponding nitroxide-labeled *L/L*-[20-41] $\beta$ 2m in an amyloid state (in blue) and unpolymerized (in black). (C) The same as in (B) but for C-terminally nitroxide-labeled *L/L*-[20-41] $\beta$ 2m. (D) ATR-FT-IR spectra for amyloids of *L/L*-[20-41] $\beta$ 2m (in blue) and the (*L/L*-[20-41] $\beta$ 2m + *D/D*-[20-41] $\beta$ 2m) racemate (in red). (E) ATR-FT-IR spectra for [ $^{13}\text{C}$ -carbonyl]Phe3-labeled *L/L*-[20-41] $\beta$ 2m alone (in blue) and as a 1:1 mixture with *D/D*-[20-41] $\beta$ 2m (in red). (F) ATR-FT-IR spectra for [ $^{13}\text{C}$ -carbonyl]Val8-labeled *L/L*-[20-41] $\beta$ 2m alone (in blue) and as a 1:1 mixture with *D/D*-[20-41] $\beta$ 2m (in red). *Figure is taken from J. Pept. Sci.* 22, 290-304 (2016).

### 2.3 Chiral recognition in the growth of amyloid fibers from the constructs based on the [20-41]-fragment of $\beta$ 2-microglobulin

We first performed ‘seeding’ experiments in which preformed amyloids of either *L/L*-[20-41] $\beta$ 2m or *D/D*-[20-41] $\beta$ 2m homodimers were added to freshly prepared solutions of the corresponding enantiomers at pH 7.5. Since the kinetics of amyloid growth induced by the *L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m constructs did not display a lag phase, we needed to find different conditions to evaluate whether seeding is enantioselective. Adding 20% (v/v) of TFE as a co-solvent resulted in much slower kinetics and a significant lag period. Importantly, although the thioflavin T response for fibers grown with TFE was lower than without TFE, the morphology of the resulting fibers was identical. As shown in **Figure 17**, adding cognate seeds triggered almost immediate amyloid growth, whereas a considerable lag phase was observed in the case of reciprocal seeds. This result demonstrates that seeding is enantioselective.



**Figure 17:** “Seeding” of amyloid growth with preformed amyloids is enantioselective for *L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m. Solutions of *L/L*-[20-41] $\beta$ 2m or *D/D*-[20-41] $\beta$ 2m (conc. 50  $\mu$ M) were incubated with or without “seeds” (~5% of total protein concentration) of either *L/L*-[20-41] $\beta$ 2m or *D/D*-[20-41] $\beta$ 2m amyloids in 50 mM Tris·HCl, 100 mM NaCl, 0.05% (w/v) NaN<sub>3</sub>, pH 7.5 in the presence of 20% (v/v) TFE. Kinetics of amyloid growth was monitored by binding of fluorescent thioflavin T (ThT) dye. Upon addition of homologous seeds, amyloids propagated without delay, whereas for non-cognate seeds a “lag” period was observed similar to control experiments when no seeds were added. *Figure is taken from J. Pept. Sci.* 22, 290-304 (2016).



**Figure 18:** Amyloid growth of *L/L*-[20-41] $\beta$ 2m (*L/L*-K3) is inhibited by the opposite enantiomer. (A) Inhibition of *L/L*-K3 amyloid growth by the enantiomeric *D/D*-[20-41] $\beta$ 2m (*D/D*-K3) peptide. (B) Inhibition of *L/L*-K3 amyloid growth by the *L/D*-[20-41] $\beta$ 2m (*L/D*-K3) heterodimer. Inhibition assays were run at a constant concentration of *L/L*-K3 (50  $\mu$ M) in the presence of cognate seeds and variable concentrations of the relevant inhibitor in 50 mM Tris-HCl, 100 mM NaCl, 0.05% (w/v) NaN<sub>3</sub> at pH 7.5. Figure is taken from *J. Pept. Sci.* 22, 290-304 (2016).

We also examined the effect of the enantiomerically pure *D/D*-[20-41] $\beta$ 2m (*D/D*-K3) and heterochiral *L/D*-[20-41] $\beta$ 2m (*L/D*-K3) dimers on *L/L*-[20-41] $\beta$ 2m (*L/L*-K3) fibrillation. Different concentrations of *D/D*-[20-41] $\beta$ 2m or *L/D*-[20-41] $\beta$ 2m were added to solutions containing identical concentrations of *L/L*-[20-41] $\beta$ 2m, and amyloid growth was triggered by addition of preformed *L/L*-[20-41] $\beta$ 2m seeds. As depicted in **Figure 18-A**, *D/D*-[20-41] $\beta$ 2m inhibited amyloid growth by *L/L*-[20-41] $\beta$ 2m (the concentration that affords 50% inhibition is around 9  $\mu$ M). Heterochiral *L/D*-[20-41] $\beta$ 2m also inhibited amyloid growth of *L/L*-[20-41] $\beta$ 2m, but to a lesser extent (the concentration corresponding to 50% inhibition is ~14  $\mu$ M) (**Figure 18-B**). These observations indicate that both *D/D*-[20-41] $\beta$ 2m and *L/D*-[20-41] $\beta$ 2m bind to the surface of growing *L/L*-[20-41] $\beta$ 2m amyloids and prevent fiber extension. Even if only half of the molecule is composed of *D*-amino acids, a significant inhibitory effect is observed.

Although *L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m cross-inhibit amyloid growth of the respective enantiomeric peptide, a 1:1 racemic mixture of *L/L*-[20-41] $\beta$ 2m and *D/D*-[20-41] $\beta$ 2m afforded amyloids (**Figure 16-A**). The resulting fibers were morphologically identical to fibrils obtained from the pure enantiomers. They were  $5.4 \pm 0.8$  nm thick in negative stain TEM images,

similar to amyloids obtained for the individual enantiomers and for the heterodimeric *L/D*-[20-41] $\beta$ 2m construct.

To determine whether these amyloids are truly racemic or a mechanical mixture (conglomerate) of enantiomorphous fibers, we synthesized two *L/L*-[20-41] $\beta$ 2m constructs equipped with a nitroxide spin-label (*L/L*-[20-41] $\beta$ 2m-NO). A nitroxide-containing 2,2,5,5-tetramethyl-3-pyrroline-1-oxyl-3-carboxylate moiety was conjugated to  $\beta$ -amino group of  $\beta$ -amino-alanine, which was introduced either at the N- or C-termini of the *L/L*-[20-41] $\beta$ 2m dimer. Both labeled constructs produced fibers alone and as 1:1 mixtures with the unlabeled *D/D*-K3 peptide that were indistinguishable from the corresponding non-labeled dimer by TEM. Continuous-wave EPR (CW-EPR) spectra were recorded as fingerprints for both samples. As shown in **Figure 16-B** and **Figure 16-C**, the CW-EPR spectra for the non-polymerized *L/L*-[20-41] $\beta$ 2m-NO monomers (in black) have rather narrow linewidths. In contrast, the superposed spectra obtained from *L/L*-[20-41] $\beta$ 2m-NO amyloids and the corresponding *L/L*-[20-41] $\beta$ 2m-NO amyloids polymerized in the presence of *D/D*-[20-41] $\beta$ 2m are much broader, indicating strong dipole-dipole coupling due to close proximity of the nitroxide labels in the amyloid fibers. The fact that the fingerprint spectra are essentially identical for pure *L/L*-[20-41] $\beta$ 2m-NO amyloids and amyloids obtained from (*L/L*-[20-41] $\beta$ 2m-NO + *D/D*-[20-41] $\beta$ 2m) mixtures indicates that the *D/D*-[20-41] $\beta$ 2m dimer does not co-polymerize with the *L/L*-[20-41] $\beta$ 2m-NO nitroxide spin-labeled enantiomers.

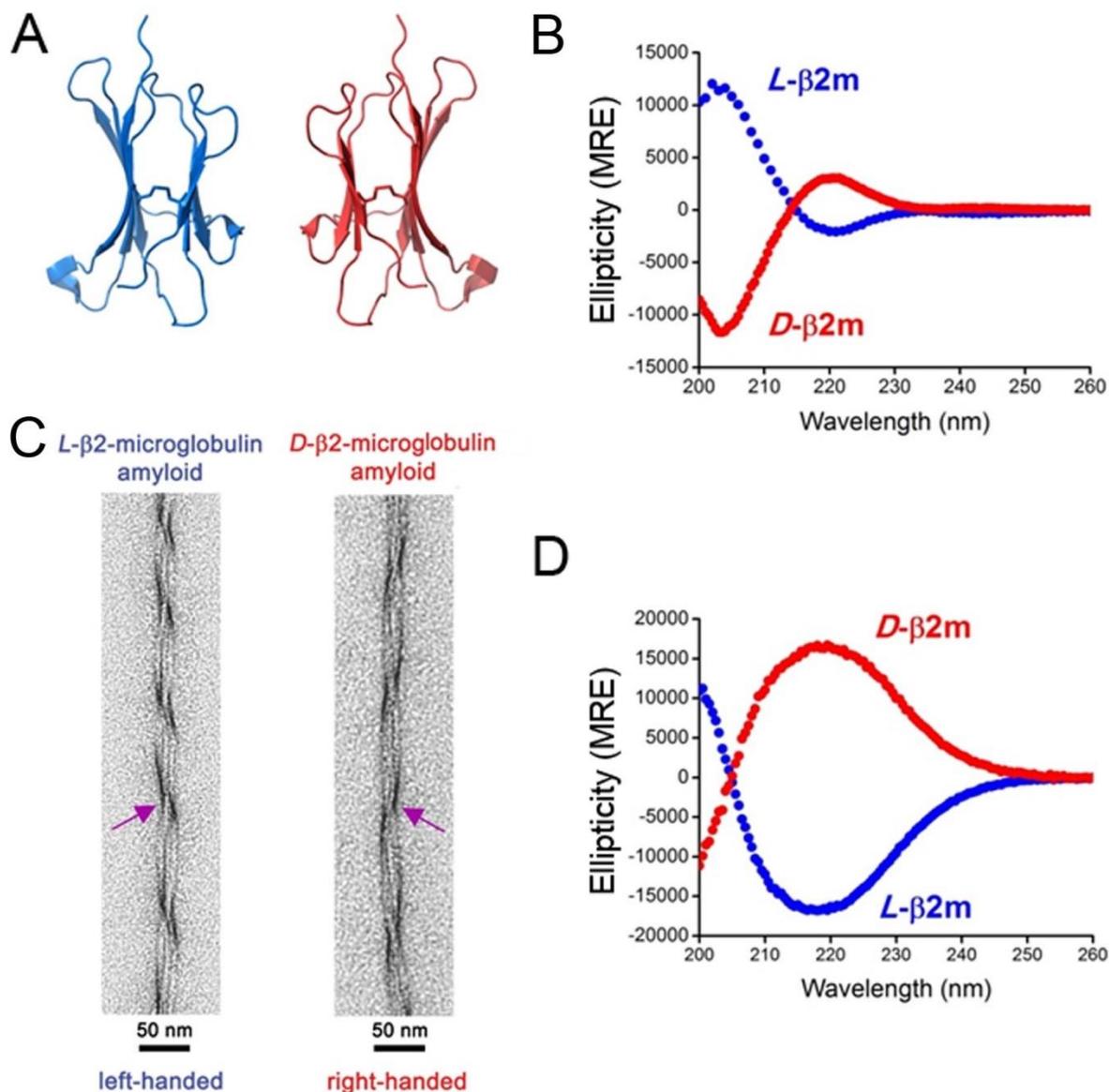
Furthermore, two other *L/L*-[20-41] $\beta$ 2m constructs were synthesized, in which the carbonyl groups of either Phe23 or Val28 were selectively  $^{13}\text{C}$ -labeled. The ATR-FT-IR spectra of amyloids grown from these samples showed well separated amide I' bands with shifts of  $37\text{ cm}^{-1}$  and  $30\text{ cm}^{-1}$  relative to the main amide I band ( $^{13}\text{C}$ Phe3-*L/L*-[20-41] $\beta$ 2m,  $1592\text{ cm}^{-1}$ ;  $^{13}\text{C}$ Val8-*L/L*-[20-41] $\beta$ 2m,  $1598\text{ cm}^{-1}$ ), respectively (**Figures 16-E** and **Figure 16-F**). The magnitude of these shifts, which originate from dipole-dipole interactions between  $^{13}\text{C}$ -labeled carbonyl groups, indicates that the labels are located close to each other (within  $\sim 5\text{ \AA}$ )<sup>111,112</sup> and, therefore, unambiguously confirms an in-register, parallel mode of assembly of  $\beta$ -strands. As seen for the CW-EPR data, the ATR-FT-IR spectra of fibers grown from 1:1 mixtures of the  $^{13}\text{C}$ -labeled *L/L*-[20-41] $\beta$ 2m constructs and non-labeled *D/D*-[20-41] $\beta$ 2m were nearly identical to the reference spectra of individually labeled *L/L*-[20-41] $\beta$ 2m amyloids. The combined CW-EPR and FT-IR

results provide compelling evidence for spontaneous Pasteur-like resolution of the peptides into conglomerates of left- and right-handed amyloids.

## **2.4 Chemical synthesis and properties of full-length $\beta$ 2-microglobulin enantiomers**

Although [20-41] $\beta$ 2m peptide may constitute the minimal sequence needed for  $\beta$ 2m amyloid formation, the sequence of the full-length protein clearly modulates the kinetics of amyloid formation as well as the structure and properties of the resulting fibers. Thus, the wild-type protein forms long straight (LS) amyloids at pH 2.5 and worm-like (WL) fibrils at pH 4 in high salt buffers (0.4 M NaCl)<sup>113</sup>. While the organization of  $\beta$ 2m amyloids is not known at atomic resolution, recent ssNMR studies of the fibers established that the polypeptide is arranged as parallel in-register  $\beta$ -sheets<sup>114</sup>. To probe how chiral recognition phenomena influence fibrillation of wild-type  $\beta$ 2m, we prepared both *L*- and *D*-forms of the full-length protein (**Figure 19**).

Chemical synthesis was carried out according to our previously published protocol using native chemical ligation of three peptide fragments<sup>115</sup>. The folded enantiomers of  $\beta$ 2m had reciprocal CD spectra as expected (**Figure 19-B**) and readily formed amyloids under acidic conditions. The long and straight fibers of wild-type  $\beta$ 2m are composed of several protofilaments supercoiled into left-handed bundles. We found that enantiomeric *D*- $\beta$ 2m produced mirror-image right-handed, super-helically twisted amyloids as depicted in the negatively stained TEM images in **Figure 19-C**. When suspended in solution, the *L*- $\beta$ 2m and *D*- $\beta$ 2m amyloids also had reciprocal CD spectra with a strongly pronounced minimum or maximum at 218 nm, respectively (**Figure 19-D**), indicating mostly  $\beta$ -sheet composition.

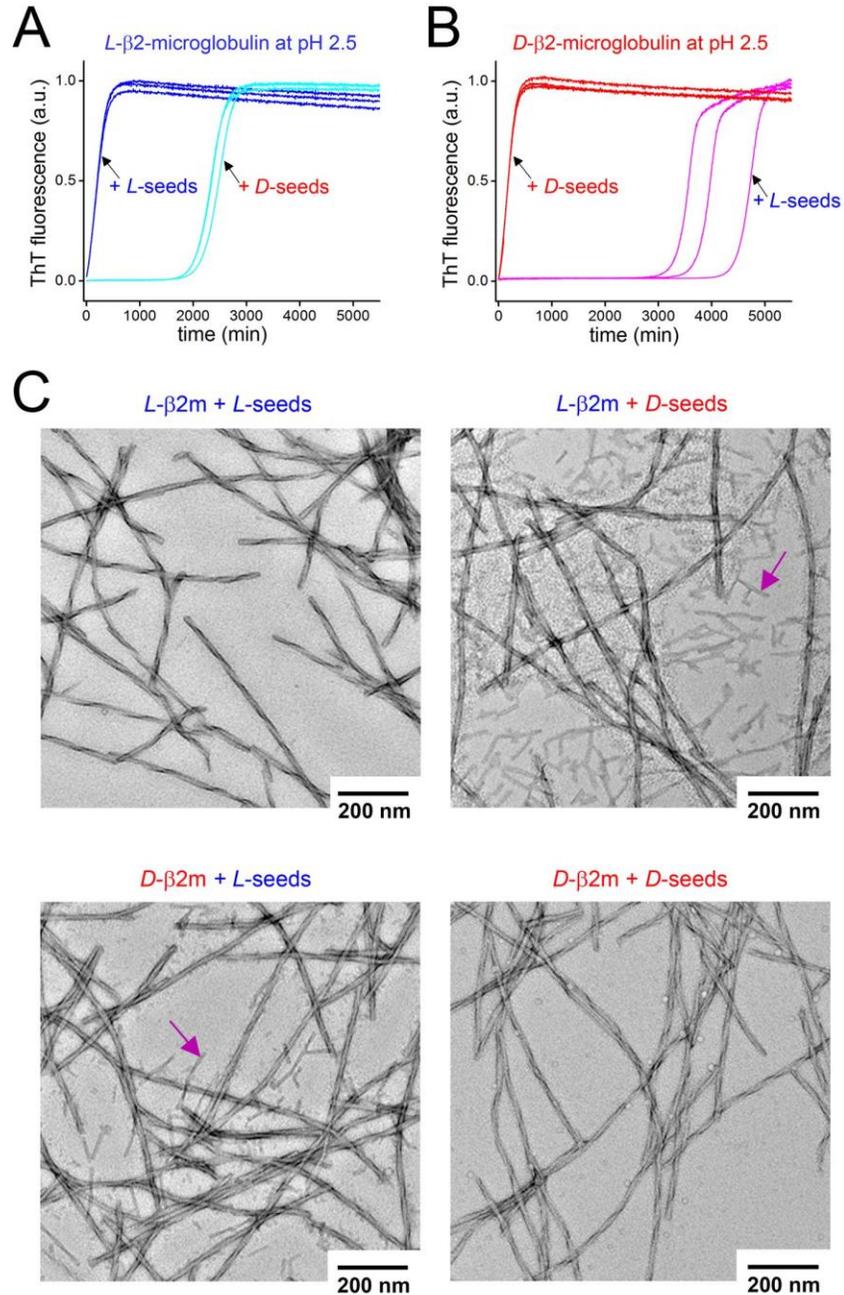


**Figure 19:** Chemically synthesized  $L$ - and  $D$ -enantiomers of  $\beta$ 2-microglobulin ( $L$ - $\beta$ 2m and  $D$ - $\beta$ 2m) and their properties. (A) Three-dimensional representation of folded  $L$ - $\beta$ 2m and  $D$ - $\beta$ 2m. Coordinates of the  $D$ -enantiomer were obtained by inversion of reported coordinates (PDB ID 2YXF). (B) Mirror-image circular dichroism (CD) spectra of folded  $L$ - $\beta$ 2m and  $D$ - $\beta$ 2m at pH 7.5. (C) Negatively stained transmission electron microscopy (TEM) images of long, straight amyloid fibrils grown from the  $L$ - $\beta$ 2m and  $D$ - $\beta$ 2m samples at pH 2.5, respectively. Super-helical structures composed of several protofilaments clearly indicate mirror-image left- and right-handed morphologies for these enantiomorphous fibrils. (D) Corresponding mirror-image CD spectra of the  $L$ - $\beta$ 2m and  $D$ - $\beta$ 2m amyloids obtained at pH 2.5. *Figure is taken from J. Pept. Sci. 22, 290-304 (2016).*

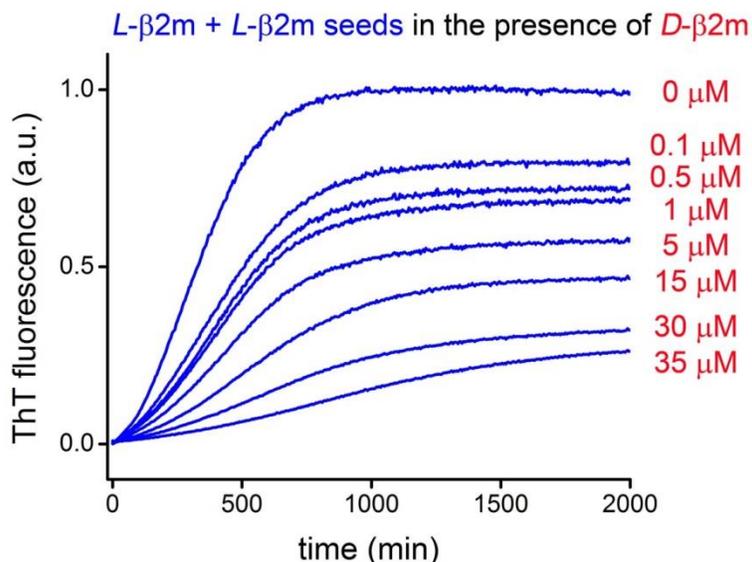
## 2.5 Chiral recognition during amyloid growth of human $\beta$ 2-microglobulin enantiomers

As in the case of the [20-41] $\beta$ 2m constructs, templated growth of *L*- and *D*- $\beta$ 2m amyloids in the presence of seeds was found to be highly enantioselective. Seeds, prepared by sonication of solutions containing *L*- or *D*- $\beta$ 2m amyloids, were added to the samples so that the concentration of seeds corresponded to ~5% of the total protein concentration in freshly prepared solutions. **Figure 20-A** and **Figure 20-B** show growth curves monitored by ThT fluorescence. The *L*- $\beta$ 2m and *D*- $\beta$ 2m proteins began to form amyloids immediately upon seeding with cognate seeds but a lag time was evident when reciprocal seeds were added. The solutions were analyzed by TEM upon completion of the fibrillation process (**Figure 20-C**). The negatively stained TEM images clearly show that *L*- and *D*- $\beta$ 2m amyloids obtained with cognate seeds are left- and right-handed long straight fibers, respectively. In contrast, in experiments with seeds from amyloids of reciprocal chirality, we observed mixtures of long straight fibers and numerous smaller fragments (which we attribute to non-cognate seeds generated by sonication). Unlike cognate seeds, these stereochemically ‘wrong’ seeds were not incorporated into the final amyloids.

The *D*- $\beta$ 2m protein was found to inhibit the growth of long straight *L*- $\beta$ 2m amyloids when added at increasing concentrations (**Figure 21**) with a mid-point inhibition concentration of ~5  $\mu$ M. Again, the samples were inspected by TEM upon completion of each experiment. Along with typical long straight (LS) helical fibers, we observed the appearance of worm-like (WL) filaments, especially at higher concentrations of *D*- $\beta$ 2m. When equimolar amounts of *L*- and *D*- $\beta$ 2m were used to grow amyloids in aqueous buffer at pH 2.5 (without additives or with 15% or 30% (v/v) TFE), WL filaments were the only fibers obtained (**Figure 22-A** and **Figure 22-C**). These structures resemble the WL amyloid polymorph obtained in experiments with enantiopure *L*- $\beta$ 2m at high salt concentrations (0.4 M NaCl). Such WL amyloids were also studied by ssNMR and it was found that they are less structured (or homogeneous) than the LS amyloid polymorph. Nevertheless, it was concluded that the general structural features of  $\beta$ 2m in WL and LS amyloids are similar<sup>116</sup>.



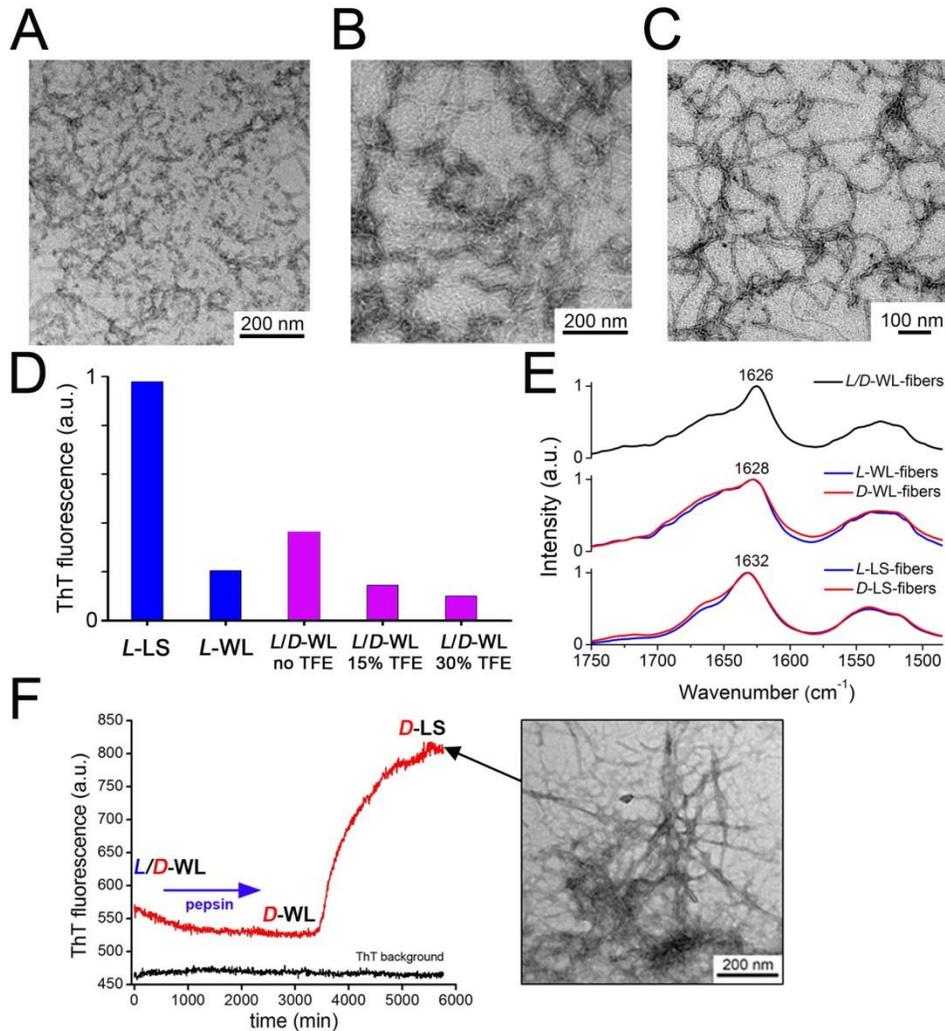
**Figure 20:** “Seeding” of amyloid growth with preformed amyloids is enantioselective for β2m. Kinetics of amyloid growth was monitored for *L*-β2m (A) and *D*-β2m (B) by binding of fluorescent thioflavin T (ThT) dye. Upon addition of homologous seeds, amyloids propagated without delay, whereas for non-cognate seeds a ‘lag’ period was observed. (C) TEM images of end-points of the fibrillation experiments depicted in (A) and (B). Clear long, straight (LS) amyloids were observed in the case of homologous seeding, whereas non-cognate seeds were not incorporated into mature amyloid fibrils (the fragmented fibrils are indicated by purple arrows). *Figure is taken from J. Pept. Sci.* 22, 290-304 (2016).



**Figure 21:** Amyloid growth of *L*-β2m is inhibited by the opposite *D*-β2m enantiomer. Inhibitions assay was run at a constant concentration of *L*-β2m (*c* 35 μM) in the presence of cognate seeds and variable concentrations of the *D*-β2m in 100 mM NaCl, 0.05% (w/v) NaN<sub>3</sub> at pH 2.5. *Figure is taken from J. Pept. Sci.* 22, 290-304 (2016).

Inspection of ThT fluorescence intensities at 430 nm upon addition of identical quantities of either LS, WL or amyloids prepared from (*L*-β2m + *D*-β2m) mixtures showed that WL amyloids generally have the same fluorescence intensity irrespective of chiral composition (**Figure 22-D**). ATR-FT-IR spectra were also measured for LS, WL and *L,D*-WL fibers (**Figure 22-E**). The amide I spectrum for LS and WL fibers showed that both aggregates are dominated by parallel β-sheet arrangements (amide I maximum at ~1630 cm<sup>-1</sup>). The *L,D*-WL amyloids share some common properties with enantiopure WL amyloids with the maximum corresponding to a slightly red-shifted amide I band (1626 cm<sup>-1</sup>). A sample of *L,D*-WL fibers was subjected to proteolytic degradation by pepsin, which has an optimal enzymatic activity at acidic pH and can only cleave all-*L* polypeptides. The starting ThT fluorescence intensity initially decreased by approximately 50% but, after a certain time, increased again to the level normally seen for LS fibrils. Indeed, when this sample was analyzed by TEM, we detected the presence of LS amyloids (**Figure 22-F**). It is likely that pepsin removes *L*-β2m peptides that cap the ends of *D*-β2m amyloids or that are adsorbed on the lateral surfaces of *D*-β2m amyloids, thus allowing rearrangement of remaining *D*-β2m worm-like fragments and oligomerization as long straight fibers. Overall, the results suggest that *L,D*-WL amyloids consist of enantiopure fragments,

which coexist in mixed fibrillar structures. The presence of *L*- and *D*-segments in the same fiber may explain why these fibers do not form more complex helically-twisted hierarchical fibrillar bundles.



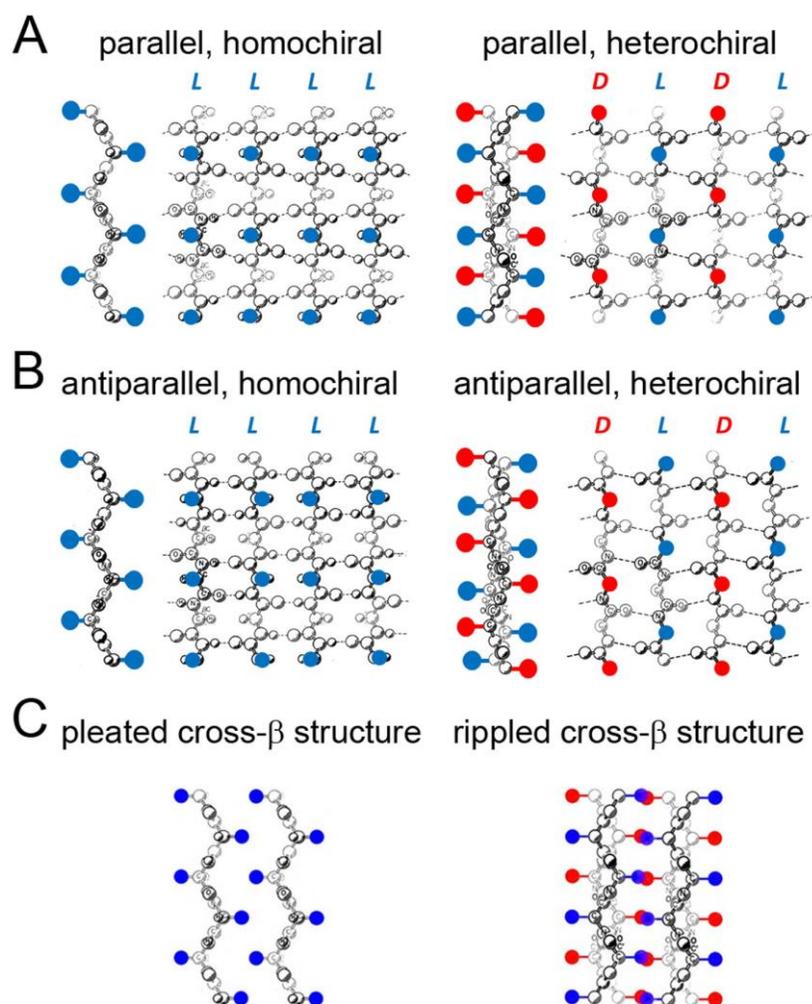
**Figure 22:** Properties of amyloid fibrils obtained from racemic mixtures of *L*- and *D*- $\beta$ 2m. Worm-like fibrous aggregates of *L/D*- $\beta$ 2m were obtained from (A) 100 mM NaCl, pH 2.5, (B) 100 mM NaCl, pH 2.5, 15% (v/v) TFE; (C) 100 mM NaCl, pH 2.5, 30% (v/v) TFE. (D) ThT fluorescence levels in the presence of the same quantities of long straight (*L*-LS) *L*- $\beta$ 2m fibrils, worm-like (*L*-WL) *L*- $\beta$ 2m, and fibrils obtained from racemic mixture of *L*- and *D*- $\beta$ 2m as depicted in (A), (B) and (C), respectively. (E) ATR-FT-IR spectra of *L,D*-worm like (*L,D*-WL) fibrils obtained from an *L,D*- $\beta$ 2m mixture, spectra for worm-like (*L*-WL and *D*-WL) fibrils obtained for individual enantiomers from high-salt (0.4 M NaCl) buffer, and the corresponding spectra for long-straight (*L*-LS and *D*-LS) fibrils obtained under low-salt (0.1 M NaCl) conditions. (F) Proteolytic degradation of the *L*-component in worm-like fibrils obtained from *L,D*- $\beta$ 2m racemic mixtures (*L,D*-WL) by pepsin at pH 2.5. Initially, the ThT fluorescence decreases by half presumably due to degradation of the *L*- $\beta$ 2m component, which prevents hierarchical self-assembly as racemic worm-like fibrils. The remaining *D*- $\beta$ 2m WL fragments can subsequently rearrange and co-assemble to produce long, straight *D*- $\beta$ 2m fibrils (*D*-LS) as confirmed by TEM images (inset). *Figure is taken from J. Pept. Sci.* 22, 290-304 (2016).

## 2.6 Discussion

Linus Pauling and Robert Corey were the first to provide a detailed structural description of both parallel and antiparallel  $\beta$ -sheets (so-called *pleated* sheets)<sup>117</sup>. In their models, these structures are entirely planar with amino acid residues adopting a conformation with Ramachandran ( $\phi, \psi$ )-angles of approximately  $-135^\circ$  and  $135^\circ$ , respectively. In a contemporaneous report, they also considered the peculiar stereochemical situation arising when  $\beta$ -sheets are formed from a racemic mixture of alternating isotactic (homochiral) polypeptide chains composed entirely of either *L*-amino acids or *D*-amino acids. They called the resulting structure a *rippled* sheet. The major difference between pleated and rippled sheets, both parallel and antiparallel, is the distribution of side chains on the surface of the  $\beta$ -sheet (**Figure 23-A** and **Figure 23-B**). In-register pleated sheets have the side chains of neighboring  $\beta$ -strands aligned parallel to each other forming bulges; the pleats are parallel as well. In contrast, in racemic rippled sheets, the side chains adopt an alternating chessboard pattern to produce a corrugated surface. Further, Pauling and Corey noted that enantiopure pleated sheets and racemic rippled fragments may coexist within a single structure generated from a mixture of all-*L* and all-*D* peptides, given that the hydrogen-bonding patterns in all configurations are nearly the same if the  $\beta$ -sheet is flat<sup>118</sup>.

The situation is more complicated in the case of amyloids, where at least two  $\beta$ -sheets interact via  $\beta$ -sheet surfaces and not through their edges. The differences in side-chain arrangement between pleated and rippled sheets result in two possible distinct interfaces (**Figure 23-C**). In the case of a cross- $\beta$  structure formed by pleated sheets, the interface is formed by bulges of one layer embedded into grooves of the second layer (**Figure 23-C**, left). For the rippled sheet cross- $\beta$  architecture, packing at the interface between two layers essentially follows the ‘knobs-into-holes’ principle.

Based on these models, it is not obvious *a priori*, which of the interfaces is most advantageous (or more stable). Given the high diversity of amino acid side-chain functionalities the answer probably depends on multiple parameters, including steric (or shape) complementarity, electrostatic charge, and hydrogen-bonding properties of the two layers.



**Figure 23:** Depictions of parallel and antiparallel  $\beta$ -sheets of homochiral and heterochiral composition, as well as the corresponding cross- $\beta$  structures. The original drawings of Pauling and Corey were modified by highlighting the  $\beta$ -carbon atoms of *L*-amino acids in blue and of *D*-amino acids in red. (A) A parallel pleated  $\beta$ -sheet that is composed of all-*L*  $\beta$ -strands (left) and a parallel rippled  $\beta$ -sheet that is composed of alternating all-*L* and all-*D*  $\beta$ -strands (right). (B) An antiparallel pleated  $\beta$ -sheet that is composed of all-*L*  $\beta$ -strands (left) and an antiparallel, rippled  $\beta$ -sheet that is composed of alternating all-*L* and all-*D*  $\beta$ -strands. Note the differences in distribution of side chains: side chains form parallel bulges in pleated in-register  $\beta$ -sheets but are aligned diagonally in rippled  $\beta$ -sheets. The alignment may change if the  $\beta$ -sheets are out-of-register. An important property of a pleated  $\beta$ -sheet, that is not evident in this figure, is a pronounced helical twist with predominantly right-handed chirality (if viewed along the peptide chain direction). In the racemic rippled structure, the *L*- and *D*-strands may still twist at the level of individual  $\beta$ -strands. However, the expectation is that the sign of the twist for enantiomeric fragments will be opposite. As a consequence, it is likely that when combined in an alternating manner the entire rippled  $\beta$ -sheet will lack a twist. (C) Cross- $\beta$  structure formed by pleated (left) and rippled sheets (right). *Figure is taken from J. Pept. Sci.* 22, 290-304 (2016).

The alignment of two pleated  $\beta$ -sheets may tolerate some offset along the backbone H-bonding direction when two layers form a dimer interface, whereas for two rippled  $\beta$ -sheets the two layers must fit precisely to form a bilayer. In the latter case, two possibilities must be considered: when every *D*-strand in the first layer lies above a *D*-strand of another layer or, alternatively, above an *L*-strand. Intuitively, the side chains are more evenly distributed on the surface of a rippled sheet, which may result in a better steric match when a ‘knobs-into-holes’ interface is formed by especially sterically demanding  $\beta$ -branched residues.

Overall, there are three distinct possibilities for racemic *L/D*-mixtures of amyloidogenic peptides:

- Racemic fibers are more stable and form more readily than their enantiomeric counterparts, and the individual enantiomers do not cross-inhibit each other’s growth;
- Racemic peptide or protein mixtures produce mechanical mixtures (conglomerates) of enantiomorphous fibrils (although some heterochiral interfaces may be present) and in this case enantiomers cross-inhibit amyloid growth;
- Fibers are not formed from racemic mixtures at all, because one enantiomer acts as an inhibitor of fibril growth of the opposite enantiomer.

Such significantly different outcomes do not depend on the relative stabilities of rippled versus pleated  $\beta$ -sheet interfaces *per se*. Mixing of *L*- and *D*-enantiomers to produce an amyloid entropically favors formation of racemic interfaces and there are no fundamental preferences against heterochiral peptide assembly, at least at the level of short sequence fragments: if the arrangement of the polypeptide backbone allows for efficient organization into racemic amyloids, then fibers appear. In contrast, heterochiral rippled interfaces may be very stable, but if the conformation of the complex (*D/L* dimer or oligomer) is not compatible with the cross- $\beta$  architecture of amyloids, then self-assembly stops at the level of short oligomers. It is clear that more complex  $\beta$ -sheet-related conformations adopted by longer sequences (e.g., super-pleated  $\beta$ -structures or  $\beta$ -solenoids) may have distinct self-assembly preferences and the results will strongly depend on the stereochemical arrangement of the misfolded *L*- and *D*-polypeptides at the propagation interface.

In the present study, we investigated how peptide chirality influences amyloid growth in two systems: 44-residue disulfide-bridged dimers of the [20-41]-fragment of  $\beta$ 2m and the 99-residue full length protein. The fibrillar cross-section of the engineered [20-41] $\beta$ 2m dimers corresponded to two  $\beta$ -arches stitched together by a single disulfide bond. Although detailed structural information on the polypeptide conformation in fibers of full-length  $\beta$ 2m is not available, it is known that polypeptides in  $\beta$ 2m amyloids are arranged in a parallel in-register configuration<sup>114,116</sup>. For both polypeptides, we were able to obtain amyloid fibers from racemic mixtures. For [20-41] $\beta$ 2m, the fibrils were morphologically identical to amyloids obtained separately from the individual enantiomers and, in all cases, their dimensions corresponded to the cross-section of a single covalently tethered dimer. CW-EPR measurements on nitroxide-labeled constructs and FT-IR measurements with <sup>13</sup>C-carbonyl labeled analogues showed that the composition of fibrils obtained from the racemic mixtures was enantiopure, i.e. they self-sorted analogous to classical Pasteur resolution observed in re-crystallization experiments with racemates of small molecules<sup>119</sup>.

For full-length  $\beta$ 2m, fibers obtained from *D/L*-mixtures were morphologically distinct from those obtained from the homochiral samples. Instead of the helical supercoiled fibers composed of several protofilaments observed for *L*- $\beta$ 2m and *D*- $\beta$ 2m, thinner amyloids corresponding to single protofilaments were observed. FT-IR measurements and proteolysis experiments suggest that these amyloids are largely enantiopure. Nevertheless, some heterochiral interfaces are also present since the individual fibers are likely constructed from alternating enantiopure fragments of variable length. This hypothesis is supported by experiments with non-racemic mixtures of *L*- $\beta$ 2m and *D*- $\beta$ 2m. For example, a 1:4 (*L:D*) mixture resulted in fibers of similar thickness as the racemate but much longer, moreover, long, straight amyloids concomitantly precipitated as well.

For both [20-41] $\beta$ 2m and full-length  $\beta$ 2m, the individual enantiomers cross-inhibited seeded amyloid growth of their mirror-image counterpart. This observation indicates that enantiomeric peptides can interact with the surface of amyloids of opposite chirality. Importantly, the homochiral *D/D*-[20-41] $\beta$ 2m dimer was found to be a better inhibitor than the heterochiral *L/D*-[20-41] $\beta$ 2m, suggesting that binding and inhibition via a heterochiral interface is more efficient than with a peptide that has a 50% homologous composition.

Most studies on amyloid inhibitors have focused on short peptides (20 residues or smaller) that can modulate or block only a portion of the growing ends of amyloid filaments. In this work, we showed that a mirror-image 99-residue *D*-protein is also capable of inhibiting the growth of long straight *L*- $\beta$ 2m amyloids (mid-point inhibition concentration  $\sim 5 \mu\text{M}$ ); instead worm-like fibers with distinct morphology precipitate. This result is noteworthy because it was not obvious *a priori* whether a misfolded *D*-protein of such length would adopt a conformation that supports a stable heterochiral protein-protein interface. For comparison, a 12-residue *D*-peptide inhibitor identified by four rounds of selection in “mirror-image” phage display experiments bound  $\text{A}\beta(1-42)$  with a dissociation constant ( $K_D$ ) of either  $0.4 \mu\text{M}$  or  $4 \mu\text{M}$  depending on the algorithm used to fit the binding data<sup>93</sup>. The fact that a nearly 100-residue long *D*-protein without any optimization or high-throughput screening has inhibitory efficiency comparable to a short peptide selected from a large combinatorial library highlights the potential of long *D*-polypeptides as candidates for inhibition of amyloid propagation.

## 2.7 Conclusions

Mirror-image proteins are valuable tools for studying protein-protein interactions and understanding the regulation of biological processes, as well as for discovering and designing modulating reagents and inhibitors. The transformative experimental technique that made these advances possible is total synthesis of proteins via native chemical ligation, which is the only method that currently allows for an efficient preparation of *D*-proteins. Our study of amyloid assembly from mixtures of *L*- and *D*-proteins complements previous work on heterochiral protein-protein interactions and provides guidelines for future applications of synthetic protein science in addressing the challenging problem of protein misfolding and aggregation.

### **3. Self-assembly of distinct amyloid structures**

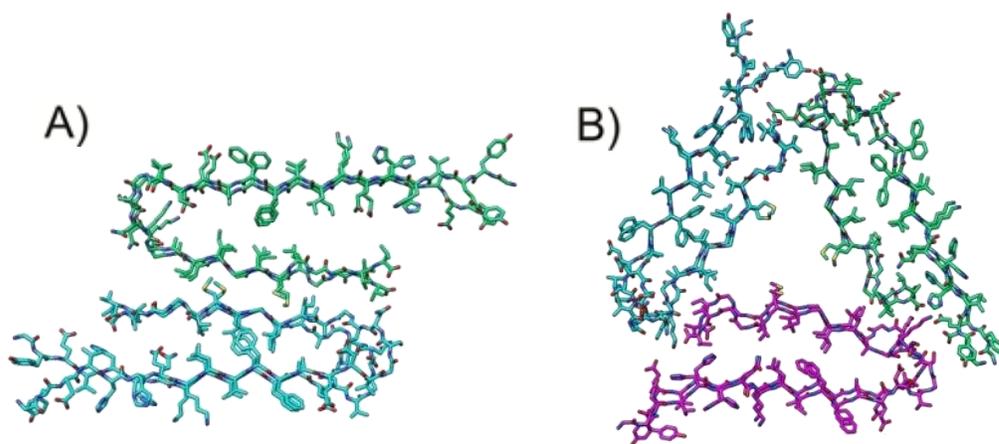
Based on the published manuscript: Jérémy Ruiz\*, Régis Boehringer\*, Marcel Grogg, Jésus Raya, Alicia Schirer, Corinne Crucifix, Petra Hellwig, Patrick Schultz and Vladimir Torbeev. Covalent tethering and residues with bulky hydrophobic side chains enable self-assembly of distinct amyloid structures. *ChemBioChem* **17**, 2274-2285 (2016) [\* = co-first authors].

#### **3.1 Introduction**

Systematic studies of amyloids are convoluted with several obstacles such as molecular-level polymorphism, conformational heterogeneity, and complex hierarchical organization<sup>9,120,121</sup>. The reported work on Alzheimer's disease-related 40- and 42-residue amyloid- $\beta$  peptides (A $\beta$ (1-40) and A $\beta$ (1-42)) illustrates the structural diversity that is typical of amyloids. With the help of solid-state NMR (ssNMR) it was shown that A $\beta$ (1-40) peptide adopts a  $\beta$ -arch conformation ( $\beta$ -strand-loop- $\beta$ -strand)<sup>122</sup>, where two  $\beta$ -strands are linked through a loop sequence and interact at their side chains (not by H-bonding of the backbone amides as in  $\beta$ -hairpins). The relative orientations of  $\beta$ -strands in  $\beta$ -arches and the part of A $\beta$ (1-40) that corresponds to a loop vary for different samples, according to a number of reports. Within the fibers, the  $\beta$ -arches stack in a parallel fashion<sup>37</sup>, although, an antiparallel assembly was also detected for an Iowa D23N A $\beta$ (1-40) mutant<sup>32</sup>. It was also found that identical  $\beta$ -arches can self-assemble into fibers that contain different numbers of peptide subunits per fiber cross-section, that is, assemblies with two- and three-fold symmetry<sup>37</sup>. In a study of one particular A $\beta$ (1-40) amyloid polymorph, it was concluded that the basic building block of A $\beta$ (1-40) fibers is an asymmetric dimer, based on two sets of chemical shifts in the ssNMR spectra, despite the uniform fiber morphology characterized by transmission electron microscopy<sup>123</sup>. More recently, a completely novel twofold symmetric peptide architecture (distinct from the  $\beta$ -arch) was determined for Osaka E22 $\Delta$  A $\beta$ (1-40) deletion mutant amyloids<sup>124</sup>.

The structures of (1-42)-isoform of A $\beta$  have also been studied. Early work suggested a peptide  $\beta$ -arch conformation similar to that of A $\beta$ (1-40)<sup>26,125</sup>, whereas more recent work has

shown that the two additional C-terminal residues result in a distinct S-shape conformation<sup>126,127</sup>. One or multiple peptide subunits per fiber cross-section have been reported for A $\beta$ (1-42) amyloids<sup>126,127</sup>. Studies of A $\beta$ (1-40) and A $\beta$ (1-42) amyloids by cryo-electron microscopy (cryo-EM)<sup>10,128</sup> revealed molecular models of polypeptide arrangement different to the findings by ssNMR. However, it should be noted that non-identical conditions were used for sample preparation (e.g., pH, peptide concentration, incubation conditions); this could influence self-assembly and result in polymorphism<sup>129</sup>.

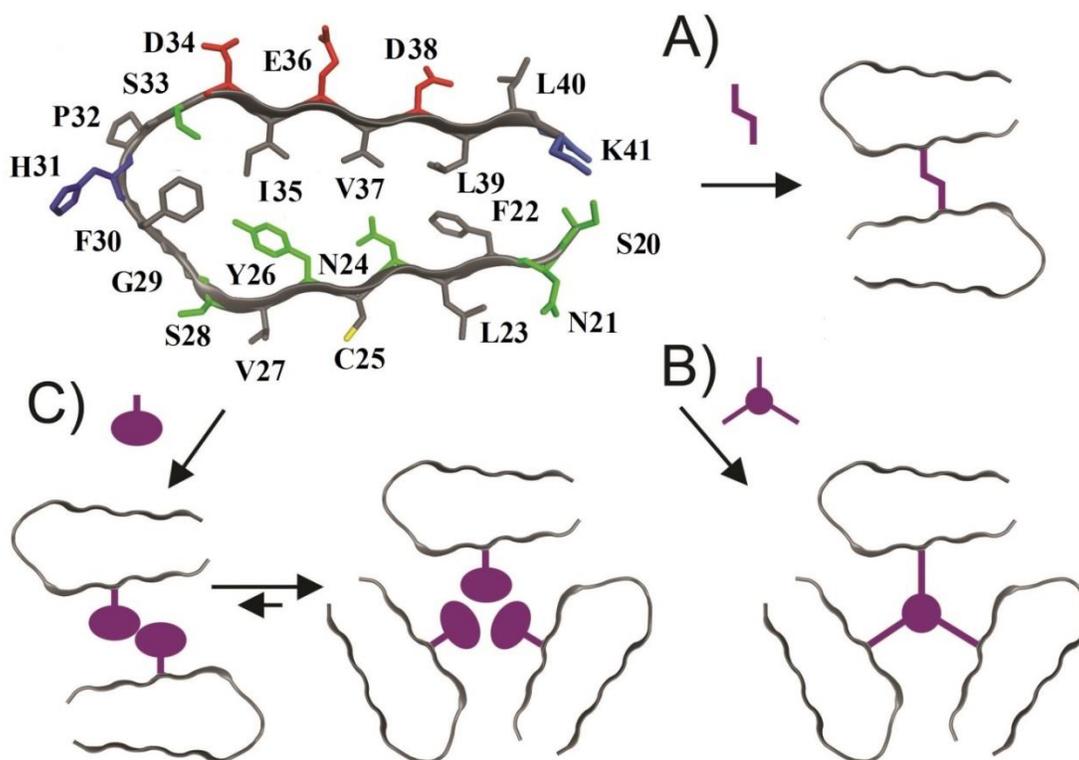


**Figure 24:** Structural models for amyloid fibrils based on numerous constraints from solid state NMR and electron microscopy<sup>37</sup>. A) 40-residue  $\beta$ -amyloid fibril with two-fold symmetry, PDB ID: 2LMN. B) 40-residue  $\beta$ -amyloid fibril with three-fold symmetry, PDB ID: 2LMP. Figure is taken from *ChemBioChem* 17, 2274-2285 (2016).

Designing amyloids is an approach that might help to resolve the complexity of amyloid structures and the self-assembly mechanisms. Strategies have to be developed for the precise control of a particular  $\beta$ -arch (or  $\beta$ -hairpin) conformation, relative orientation and register of the corresponding  $\beta$ -sheets, as well as stoichiometry and symmetry when multiple  $\beta$ -arches or  $\beta$ -hairpins constitute the cross-section of an amyloid. Such an approach will help in the preparation of homogeneous samples for high-accuracy structural studies and, moreover, the elucidation of biological activities of various amyloid polymorphs.

## 3.2 Objectives

In this study, the objective was to control the stoichiometry of the amyloid core, more specifically, trimeric versus dimeric assembly (**Figure 24**). The first approach involved a covalent linker to tether the required number of amyloid-forming peptides thereby facilitating precipitation of an amyloid polymorph with the desired stoichiometry at the fiber cross-section. In the second approach, a bulky hydrophobic group was attached at a specific site in the  $\beta$ -arch to disfavor dimeric association (by steric clashes), in order to direct self-assembly towards a trimeric structure in which the void within the fiber core is able to accommodate these bulky groups (**Figure 25**).



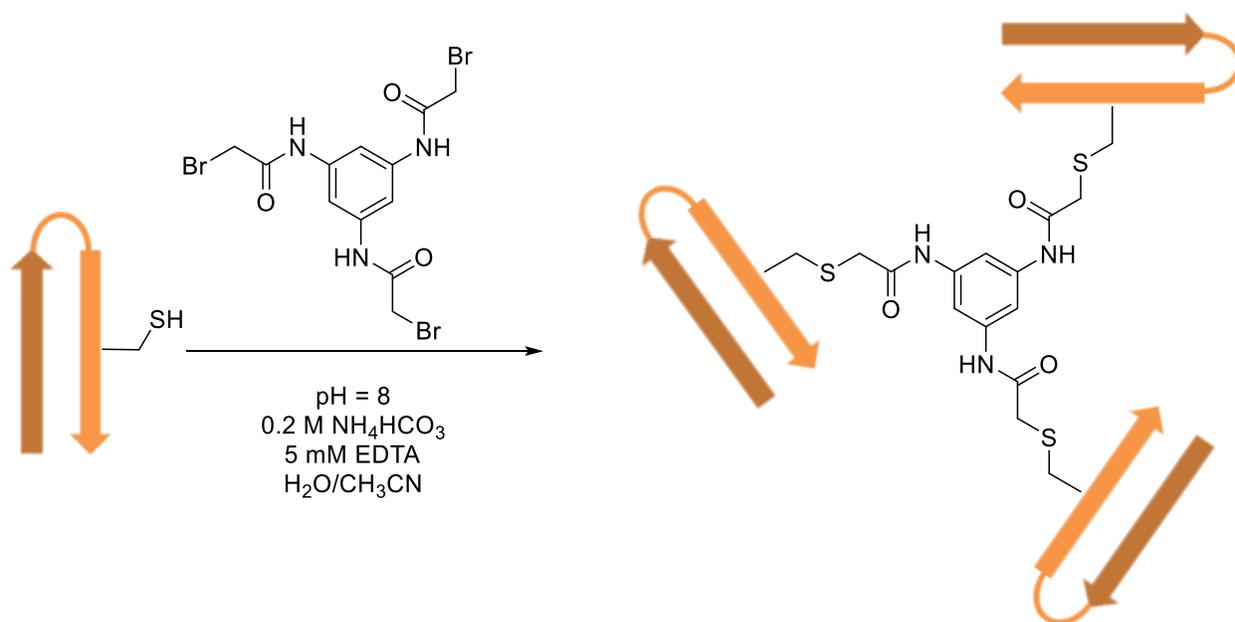
**Figure 25:** Covalent tethering and supramolecular strategies to control the self-assembly of [20–41]-peptide fragment of human  $\beta$ 2-microglobulin into amyloids with distinct structures. Note that fiber cross-sections are depicted (i.e., the fiber axis is perpendicular to the plane of the figure). Amyloids are formed by stacking multiple copies of the peptide in  $\beta$ -arch conformation. Residue side chains are color coded based on their properties: positively charged (blue), negatively charged (red), polar (green), and hydrophobic (gray). Cys25 (orange) was used for peptide modification. A), B) Two or three  $\beta$ -arches are covalently tethered to form the corresponding two- or threefold symmetric structures. C) Attachment of a bulky group to the peptide prevents dimeric association, and favors trimeric assembly. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*

We chose the 22-residue amyloidogenic peptide that corresponds to the [20-41]-fragment of human  $\beta$ 2-microglobulin ( $\beta$ 2m) to validate this concept. The structure of this peptide within amyloid protofilaments, as solved by ssNMR<sup>27</sup>, is a hydrogen-bonded assembly of in-register parallel  $\beta$ -arches (PDB ID: 2E8D), and resembles the structures characterized by a U-shaped  $\beta$ -strand-loop- $\beta$ -strand motif determined for A $\beta$ (1-40) peptides. The important feature of [20-41] $\beta$ 2m protofilaments is the amphiphilic character of their external surfaces. One side is composed mainly of hydrophilic charged residues, whereas the other is decorated with non-charged polar and hydrophobic side chains, with the cysteine 25 in the middle of the corresponding  $\beta$ -sheet. Previous studies showed that the dimeric construct of [20-41] $\beta$ 2m obtained by oxidation of Cys25 into disulfide bonds ("covalent dimer") produces amyloids more readily than the parent [20-41] $\beta$ 2m monomer (as mentioned in Chapter 2 and in the literature<sup>109</sup>). Furthermore, using isotope-edited FTIR spectroscopy we demonstrated that the [20-41] $\beta$ 2m covalent dimer amyloids are composed of parallel in-register  $\beta$ -sheets similar to what was observed for the [20-41] $\beta$ 2m monomer fibers<sup>27</sup>.

### 3.3 The covalent approach

#### 3.3.1 Synthesis of the covalent trimer of [20-41] $\beta$ 2m

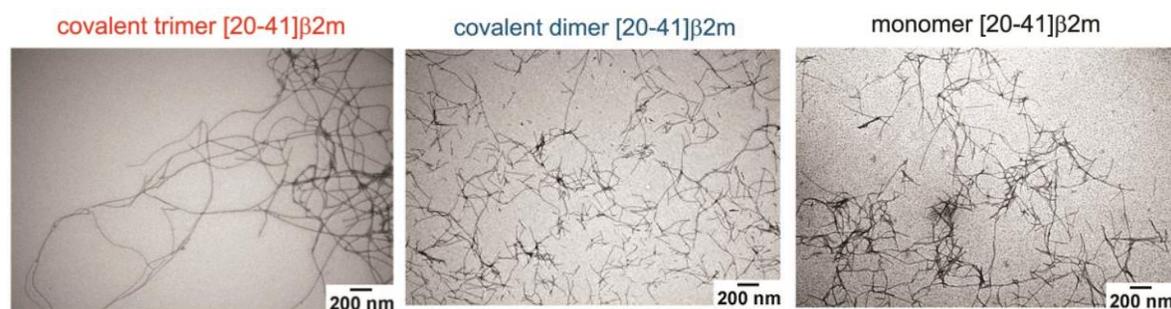
A "covalent trimer" of [20-41] $\beta$ 2m by tethering the peptide was synthesized by reaction of the thiol group of Cys25 with a bromoacetyl derivative of the trivalent *N,N,N'*-(benzene-1,3,5-triyl)tris(acetamide) linker (**Figure 26**). The reaction is done in a mixture of water and acetonitrile at pH 8 and 30°C. Four equivalents of [20-41] $\beta$ 2m peptide are introduced compared to the trivalent linker in order to drive the reaction to completion avoiding the formation of side-products. The reaction conditions were optimized in order to circumvent incomplete tethering (formation of species with only one or two reacted peptides) and hydrolysis of the bromide moieties that can occur if the kinetics of the conjugation reaction is too slow.



**Figure 26:** Synthesis of covalent [20-41] $\beta$ 2m trimer. The peptidic chains of [20-41] $\beta$ 2m are represented in a  $\beta$ -arch conformation in orange with the thiol group of Cys25 displayed on one of the  $\beta$ -stands. At pH 8, the thiols selectively react with the linker by bromide displacement (nucleophilic substitution) to form the trimeric construct.

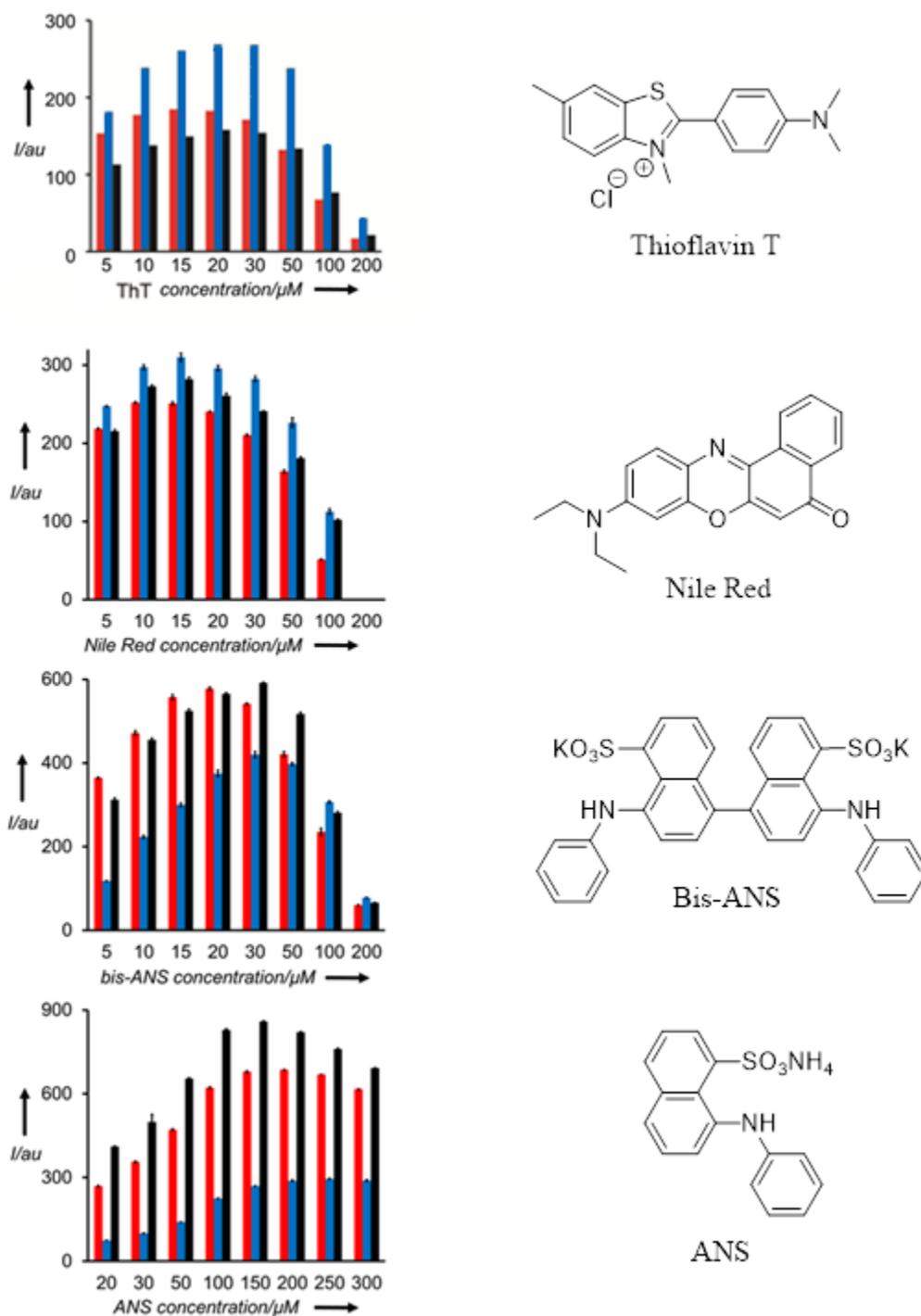
### 3.3.2 Study and comparison of the amyloidogenicity of the different building blocks

This trimeric construct was found to be amyloidogenic, but with morphology distinct from those of monomers or covalent dimers (**Figure 27**). Long fibers (up to several microns) of uniform thickness were observed without noticeable helical twist, in contrast to the much shorter amyloids detected for covalent dimers or the fibers of poly-disperse morphology observed for monomer<sup>109</sup>.



**Figure 27:** TEM images of amyloids prepared from covalently tethered trimeric or dimeric constructs and the parent [20–41]β<sub>2</sub>m. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*

The binding of specific dyes was used to reveal tinctorial differences between amyloid fibers formed by covalent trimers, covalent dimers, and monomer (**Figure 28**). Thioflavin T (ThT, 4-(3,6-dimethyl-1,3-benzothiazol-3-ium-2-yl)-*N,N*-dimethylaniline), Nile Red (9-diethylamino-5-benzo[α]phenoxazinone), bis-ANS (4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic acid), and ANS (8-anilino-1-naphthalenesulfonic acid) at a range of concentrations were added to suspensions of amyloids. For all dyes, the three amyloid samples produced different fluorescence responses, which may indicate distinct environments and/or number of binding sites for these dyes. Data for covalent trimer are depicted in red, for covalent dimer in blue and for monomer in black. Intensities for fluorescence maxima for ThT ( $\lambda_{\text{ex}} = 442 \text{ nm}$ ,  $\lambda_{\text{em}} = 480 \text{ nm}$ ), bis-ANS ( $\lambda_{\text{ex}} = 400 \text{ nm}$ ,  $\lambda_{\text{em}} = 475 \text{ nm}$  for covalent trimer and covalent dimer,  $\lambda_{\text{em}} = 491 \text{ nm}$  for monomer), ANS ( $\lambda_{\text{ex}} = 400 \text{ nm}$ ,  $\lambda_{\text{em}} = 475 \text{ nm}$ ) and Nile Red ( $\lambda_{\text{ex}} = 554 \text{ nm}$ ,  $\lambda_{\text{em}} = 630 \text{ nm}$ ) are plotted. Total concentration of peptide in each measurement was kept constant and equal to that used in the amyloid growth for each construct ([covalent trimer] = 33.3  $\mu\text{M}$ ; [covalent dimer] = 50  $\mu\text{M}$ ; [monomer] = 100  $\mu\text{M}$ ), whereas concentration of dyes was varied. At high concentrations of dyes, the intensities of fluorescence response are diminished presumably due to fluorescence quenching when more dye molecules are bound to amyloids. Each construct does not show the same fluorescence intensity for each dye, which might indicate distinct environments and/or number of binding sites for these dyes.



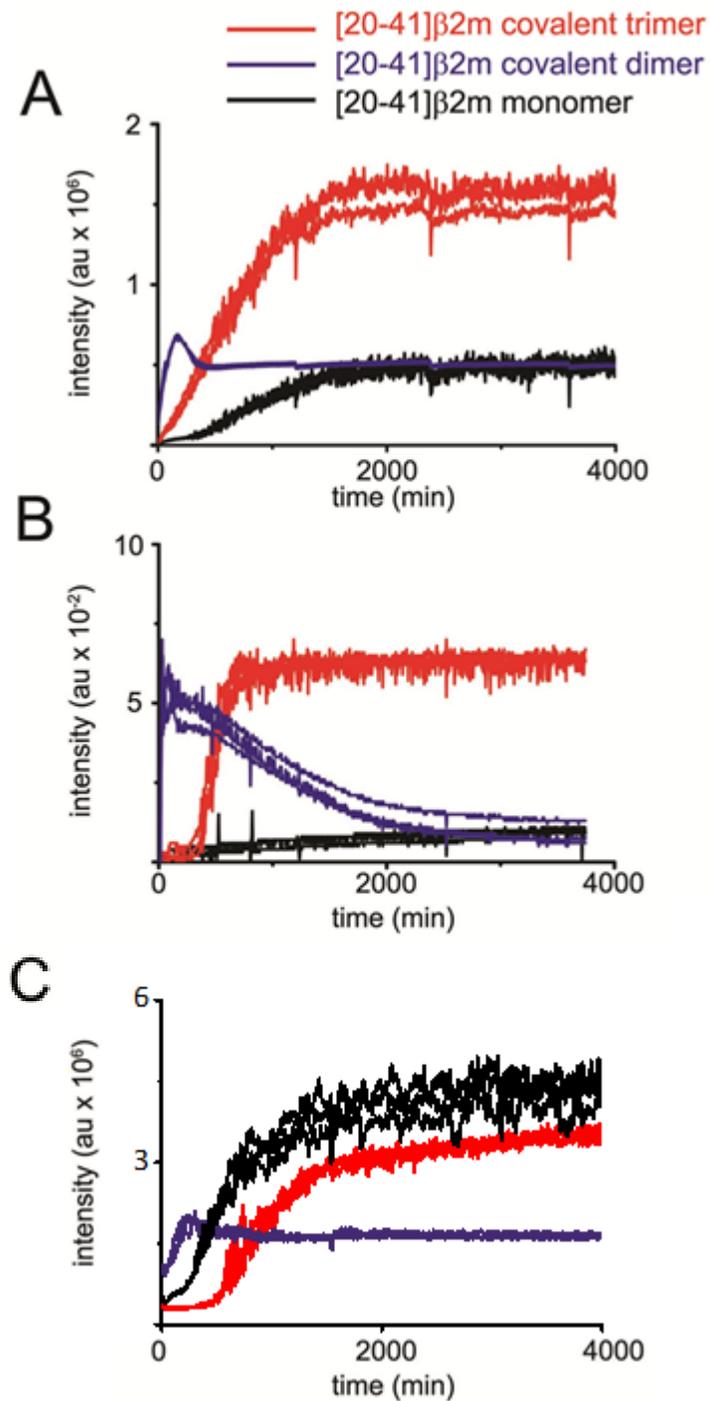
**Figure 28:** Fluorescence measurements of several dyes bound to amyloid fibers that were obtained from [20-41] $\beta$ 2m covalent trimer, covalent dimer and monomer. On the left, bar charts for each constructs (intensity of fluorescence versus dye concentration) are represented and on the left the corresponding chemical structures of the dyes are shown. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*

### 3.3.3 Kinetic studies of amyloid growth

Fluorescence of ThT or bis-ANS was also used to investigate the kinetics of fiber growth in a continuous format by using a plate reader; in addition, turbidity measurements were performed. Data obtained with fluorescent dyes and turbidity showed distinct growth kinetics for these three peptides. It should be noted that the order of fluorescence intensity obtained by bulk fluorescence measurements for ThT was different to that obtained by the plate reader (**Figure 28** and **Figure 29-A**) due to use of optical filters with bandpass parameters that were not at the excitation/emission maxima of ThT. This was not the case for bis-ANS, where values obtained in bulk measurements (1 cm × 0.2 cm cuvette) and the final fluorescence intensities recorded on the plate reader (96-well plates) had the same order for the three peptides (compare **Figure 28** and **Figure 29-C**), because the excitation/emission wavelengths of the filter set were similar to those used in the spectrophotometer.

Kinetic studies by time-resolved fluorometry can be affected by the presence of extrinsic dyes in the buffer. Therefore, without addition of exogenous compounds, kinetics of amyloid growth was also studied by turbidity measurements (**Figure 29-B**). The turbidity measurements depend only on light scattering and were realized at 635 nm. No absorbance from the sample was detected at this wavelength, so the measured intensity only correlated to light scattering. The monitoring was done using a plate reader and showed kinetics qualitatively similar to those observed by fluorescence measurements (covalent dimer > covalent trimer ≈ monomer).

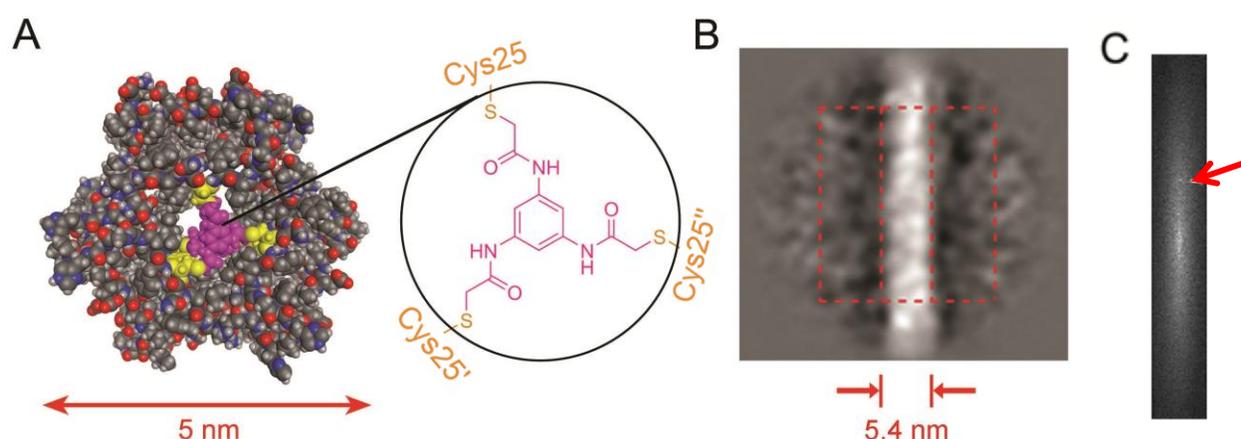
For both techniques, the initial increase then decrease of signal intensity for covalent dimers in A and B are attributed to the tendency of the amyloids to precipitate from the optical detection volume. This was confirmed by bulk measurements in a cuvette by using a spectrophotometer.



**Figure 29:** Covalent trimer, covalent dimer and [20–41]β2m monomer have different kinetics of amyloid growth. A) Kinetics of growth determined by ThT fluorescence using a plate reader. B) Kinetics monitored by turbidity measurements ( $\lambda=635$  nm) in a plate-reader. C) Kinetics of growth determined by Bis-ANS fluorescence in a plate reader. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*

### 3.3.4 Characterization of [20-41] $\beta$ 2m covalent trimer amyloids by cryo-EM

In order to accurately measure the thickness of the covalent trimer amyloids we used cryogenic transmission electron microscopy (cryo-EM) to visualize the frozen hydrated fibers. A total of 4050 images of short fiber patches were obtained from over 200 randomly chosen fibers. The average of the aligned images showed a mean diameter of 5.4 nm, which is in a good agreement with the dimensions estimated by molecular modeling (**Figure 30-A** and **Figure 30-B**). The Fourier transforms of the cryo-EM images showed a characteristic diffraction peak at 4.9 Å, which corresponds to the inter- $\beta$ -strand distance typical of amyloids (**Figure 30-C**).



**Figure 30:** Properties of [20–41] $\beta$ 2m covalent trimer amyloids. A) Molecular model of amyloid with a triangular core (cross-section view, long axis perpendicular to the image). Detail: structure of the central linker. B) Cryo-EM to determine the thickness of amyloids (the average of 4050 aligned images is depicted). The mean thickness (5.4 nm) is in a good agreement with the molecular model. C) Fourier transform of amyloid fibrils obtained from covalent trimer of [20-41] $\beta$ 2m, the red arrow indicates a diffraction peak corresponding to 4.9 Å periodic distance. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*

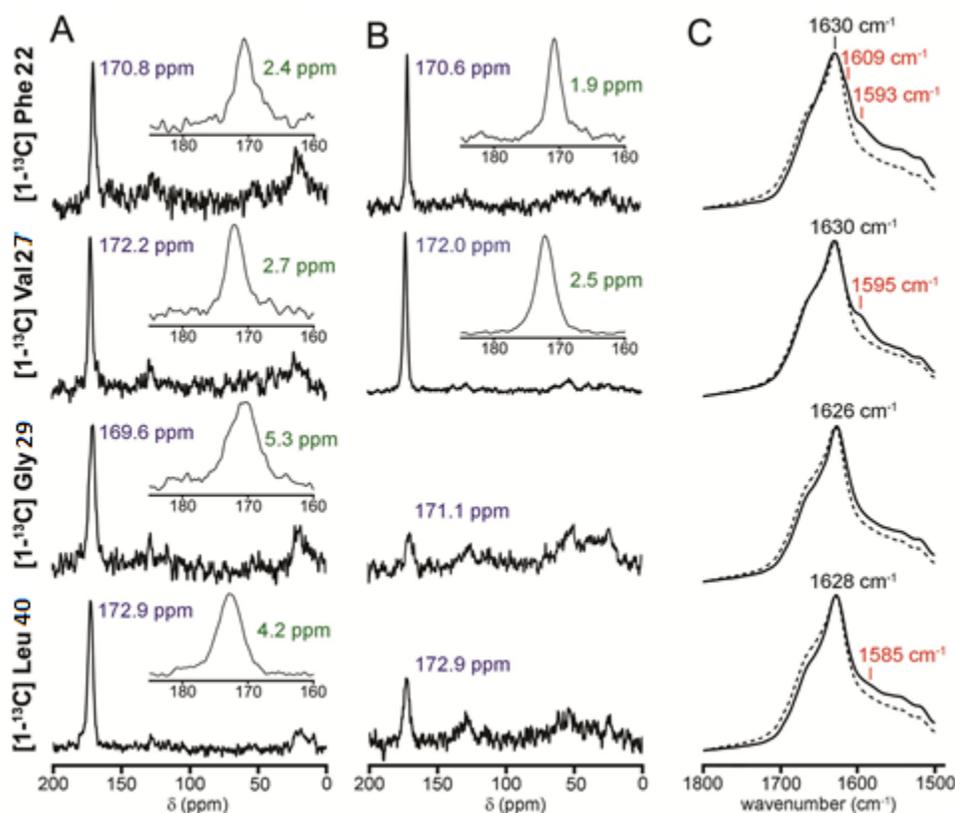
### 3.3.5 Characterization of [20-41] $\beta$ 2m covalent trimer amyloids by solid state NMR

To gain a more detailed insight into the molecular structures of fibers, four variants of [20-41] $\beta$ 2m covalent trimer were synthesized with  $^{13}\text{C}$ -labels introduced individually at the carbonyls of Phe22, Val27, Gly29, and Leu40, and the corresponding amyloids were studied by ssNMR. Direct polarization magic-angle spinning (DP-MAS)  $^{13}\text{C}$  ssNMR spectra are depicted in **Figure**

**31-A** and show well-defined peaks for  $^{13}\text{C}$ -labeled carbonyls in each sample. Cross-polarization magic-angle spinning (CP-MAS)  $^{13}\text{C}$  ssNMR spectra were also recorded for all four samples (**Figure 31-B**). Whereas symmetric single peaks were obtained for  $[1-^{13}\text{C}]\text{Phe22}$  and  $[1-^{13}\text{C}]\text{Val27}$  samples, only very weak signals were observed for covalent trimer amyloids with  $[1-^{13}\text{C}]\text{Gly29}$  and  $[1-^{13}\text{C}]\text{Leu40}$ . Weaker signal intensity in CP-MAS might have resulted from less efficient  $^1\text{H}$ -to- $^{13}\text{C}$  polarization transfer, due to enhanced conformational mobility on the  $\mu\text{s}$ -ms timescale<sup>130</sup>. Both Phe22 and Val27 are within the inner  $\beta$ -sheet connected by a central linker and are likely to be rigidified, whereas Gly29 and Leu40 are in the loop and solvent-exposed  $\beta$ -sheet, respectively, and might indeed undergo enhanced conformational dynamics. The peak line-widths (**Figure 31-A** and **Figure 31-B**) are in agreement with this interpretation: sharper for  $[1-^{13}\text{C}]\text{Phe22}$  and  $[1-^{13}\text{C}]\text{Val27}$ , and broader for  $[1-^{13}\text{C}]\text{Gly29}$  and  $[1-^{13}\text{C}]\text{Leu40}$ . Two-dimensional  $^{13}\text{C},^1\text{H}$  heteronuclear correlation (HETCOR) experiments were performed in order to improve the spectral resolution. For  $[1-^{13}\text{C}]\text{Val27}$ , we observed a single  $^{13}\text{C}$  chemical shift, whereas in the  $[1-^{13}\text{C}]\text{Phe22}$  sample we observed cross-peaks with slightly different  $^{13}\text{C}$  chemical shifts ( $\Delta\delta < 0.3$  ppm). Nevertheless, these data confirm high structural homogeneity of the samples, especially in the amyloid core, and a similar conformation on average for the three  $\beta$ -arches in the covalent trimer.

### 3.3.6 Characterization of [20-41] $\beta$ 2m covalent trimer amyloids by FTIR

In accordance with the ssNMR findings, differences for the conformational properties of Phe22, Val27, Gly29, and Leu40 were also noticeable in the ATR-FTIR data (**Figure 31-C**). Secondary structure deconvolution of the amide I and amide II bands for the non-labeled covalent trimer identified contributions from  $\beta$ -strands (45%), turns (30%), and random coils (25%). A well-defined  $^{13}\text{C}$ -edited amide I' band shifted to  $1593\text{ cm}^{-1}$  was detected for the amyloid core  $[1-^{13}\text{C}]\text{Val27}$  in the corresponding  $^{13}\text{C}$ -labeled construct. Such distinctive frequencies correspond to dipole-dipole interactions between  $^{13}\text{C}$ -labeled carbonyl groups in close proximity<sup>131</sup>.



**Figure 31:** Properties of [20–41] $\beta$ 2m covalent trimer amyloids. A) DP-MAS, and B) CP-MAS  $^{13}\text{C}$ -ssNMR spectra of amyloids obtained from covalent trimers labeled with [1- $^{13}\text{C}$ ]Phe22, [1- $^{13}\text{C}$ ]Val27, [1- $^{13}\text{C}$ ]Gly29, or [1- $^{13}\text{C}$ ]Leu40. The  $^{13}\text{C}$  chemical shifts are in blue, and line-widths (full width at half maximum) are in green. C) ATR-FT-IR spectra shown in absorbance mode of the corresponding  $^{13}\text{C}$ -labeled construct (solid line). The spectrum of amyloid of non-labeled covalent trimer is overlaid (dashed lines, amide I maximum is at 1628  $\text{cm}^{-1}$ ). The spectral bands emerging from  $^{13}\text{C}$  substitution are in red. Maxima were determined by spectral deconvolution and inspection of the corresponding second derivative FT-IR spectra. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*

The fact that Val27 is not in the middle of the corresponding  $\beta$ -strand but towards one of its ends indicates an in-register parallel arrangement of  $\beta$ -sheets (as opposed to antiparallel  $\beta$ -sheets, where strong transition dipole coupling are expected only for in-register residues in the middle of  $\beta$ -strands)<sup>112</sup>. The C-terminal [1- $^{13}\text{C}$ ]Leu40 also produced  $^{13}\text{C}$ -edited amide I' signal shifted from the main amide I band to 1585  $\text{cm}^{-1}$ . For the [1- $^{13}\text{C}$ ]Phe22-labeled sample, in addition to the amide I' peak analogously shifted to 1595  $\text{cm}^{-1}$ , we observed a second signal at 1609  $\text{cm}^{-1}$ . Together with the HETCOR ssNMR data, this suggests a locally populated alternative conformation for Phe22.

### 3.4 Supramolecular approach for controlling amyloid polymorphism

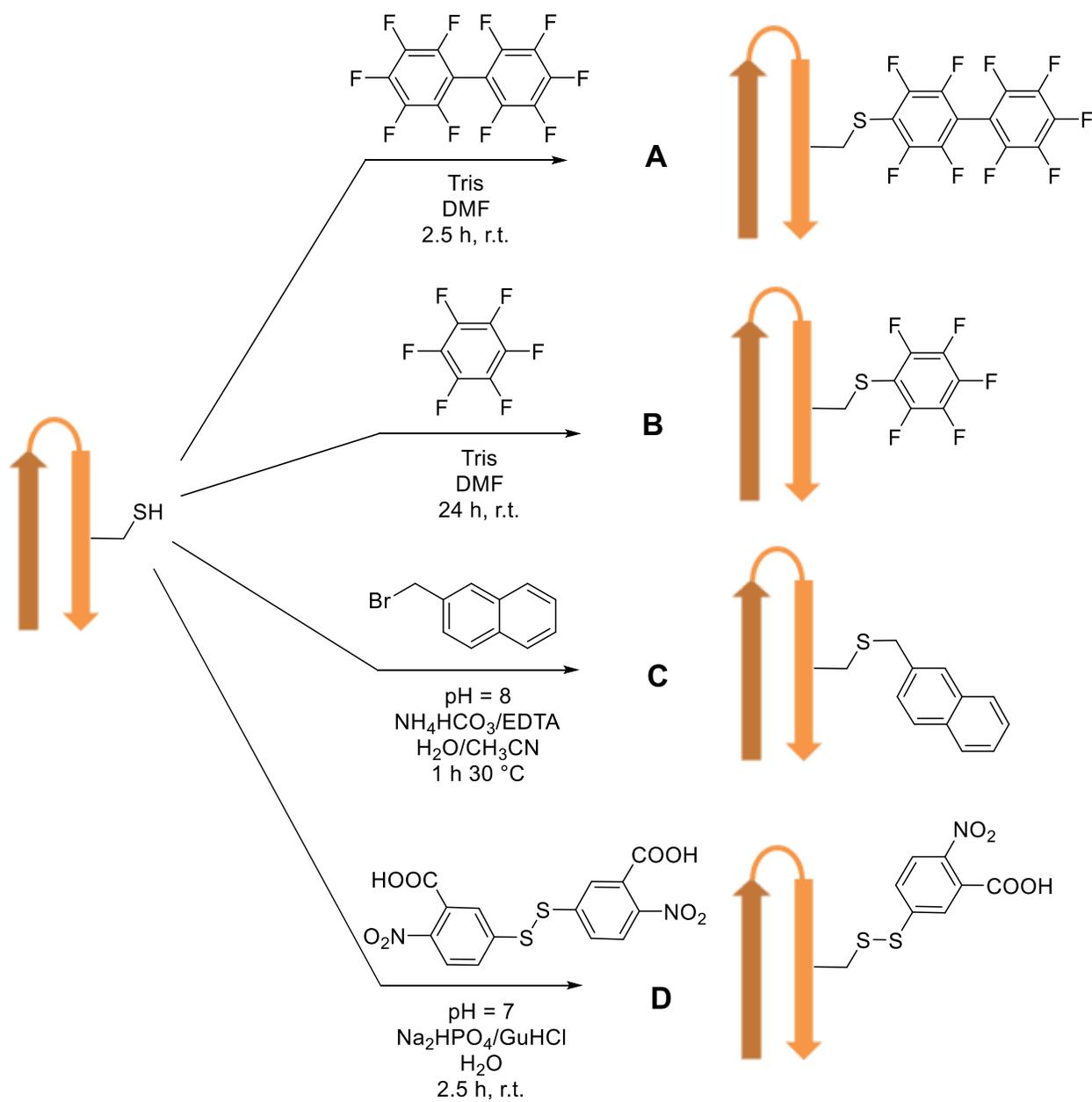
Controlling the stoichiometry of [20-41] $\beta$ 2m amyloid self-assembly is also possible by noncovalent interactions, by modifying the properties of the hydrophobic side of the  $\beta$ -arch. For this purpose, a single cysteine residue was ideally positioned to introduce various modifications and produce different types of [20-41] $\beta$ 2m conjugates (**Figure 25** and **Figure 32**).

#### 3.4.1 Synthesis of the different [20-41] $\beta$ 2m-conjugates

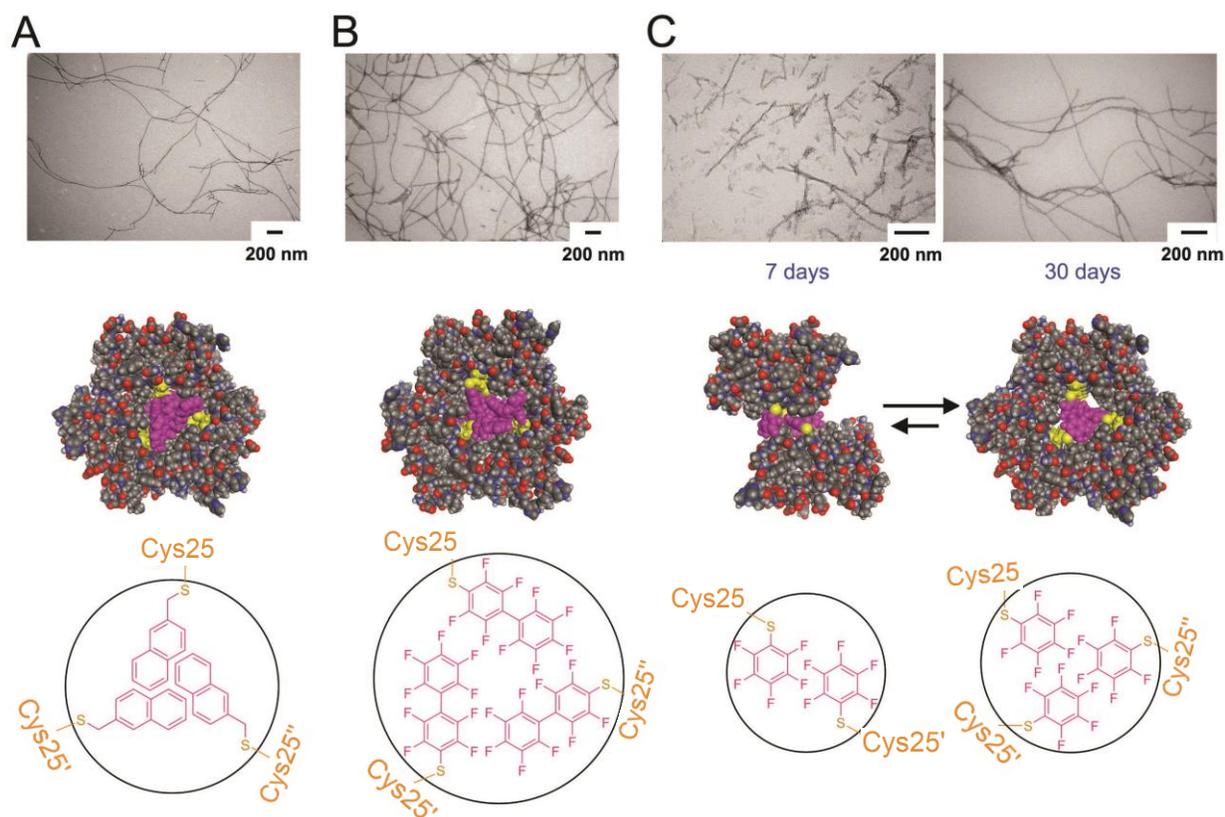
The groups that were tested were perfluoro-[1,1'-biphenyl]-4-yl, pentafluorophenyl naphthalen-2-ylmethyl and Ellman's thiol moieties. The naphthalen-2-ylmethyl group was attached by nucleophilic substitution by using 2-(bromomethyl)naphthalene at pH 8 in a mixture of water and acetonitrile. Perfluoroaryl moieties were attached by using  $S_NAr$  (nucleophilic aromatic substitution) chemistry in *N,N*-dimethylformamide in the presence of Tris base (tris(hydroxymethyl)aminomethane). The Ellman's thiol conjugate was obtained by reacting [20-41] $\beta$ 2m with the Ellman's reagent at pH 7 in aqueous buffer leading to the formation of the desired mixed disulfide.

#### 3.4.2 Study of [20-41] $\beta$ 2m-conjugates with bulky hydrophobic groups

All these constructs were able to form amyloids (**Figure 33**, TEM images). Strikingly, in the case of naphthalen-2-ylmethyl and perfluoro-[1,1'-biphenyl]-4-yl, the fiber morphologies were similar to that of covalent trimer, that is, long regular fibers were observed (**Figure 33-A**). Molecular modeling of the fibers with a triangular core showed that the attached hydrophobic moieties fit well into the space that is formed if three copies of amyloidogenic peptide constitute the cross-section of the fiber (**Figure 33**, molecular modeling).



**Figure 32:** Synthesis of the different conjugates of [20-41]β2m. The peptidic chains of [20-41]β2m are represented in a β-arch conformation in orange with the thiol group of Cys25 pointing out of one of the β-stands. A) Synthesis of [20-41]β2m-(perfluoro-[1,1'-biphenyl]-4-yl) conjugate. B) Synthesis of [20-41]β2m-pentafluorophenyl conjugate. C) Synthesis of [20-41]β2m-(naphthalene-2-ylmethyl) conjugate. D) Synthesis of [20-41]β2m-(Ellman's thiol) conjugate.



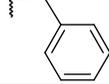
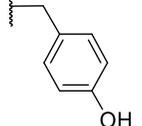
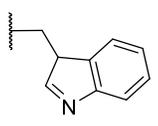
**Figure 33:** The attached group at Cys25 modulates the morphology and structural properties of amyloids. A, B) [20–41]β2m conjugated to naphthalen-2-ylmethyl or perfluoro-[1,1'-biphenyl]-4-yl produced long regular amyloids that are morphologically similar to covalent trimer fibers. C) The pentafluorophenyl moiety initially resulted in metastable helically twisted shorter fibers; upon prolonged incubation (30 days) these transformed into long amyloids of regular thickness. Top: TEM images; middle: corresponding molecular models; bottom: chemical structures of moieties conjugated to Cys25. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*

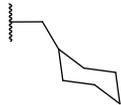
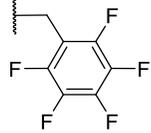
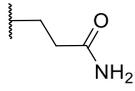
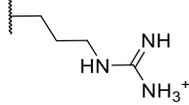
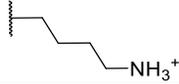
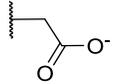
The situation was different for the pentafluorophenyl adduct: initially short fibers were detected, the majority of which were composed of two filaments supercoiled into left-handed helices (**Figure 33-C**). However, upon extended incubation (30 days) these structures transformed into longer amyloid fibers that again were morphologically similar to the amyloids of covalent trimers. This observation suggests that the smaller size of the pentafluorophenyl substituent enables dimeric assemblies at first (tending to associate as helically twisted fibers), whereas longer incubation allows these metastable structures to rearrange into more stable, longer amyloids with a triangular core (**Figure 33-C**).

### 3.4.3 Study of a library of [20-41] $\beta$ 2m variants

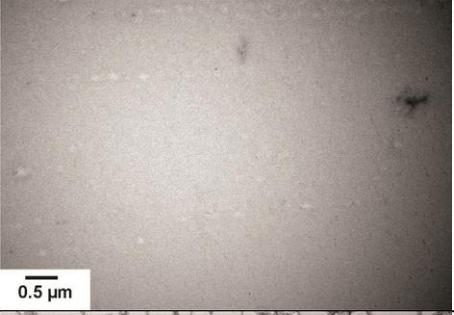
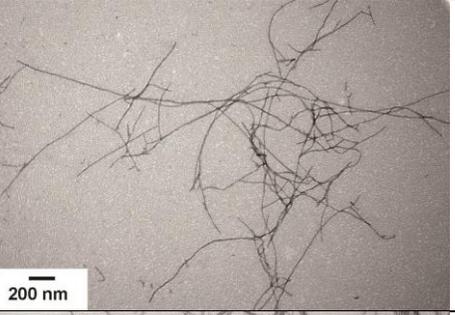
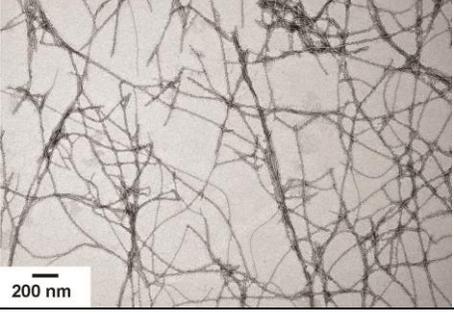
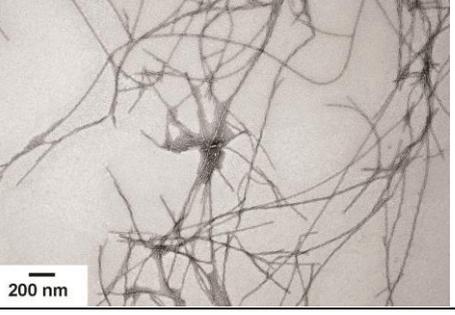
In order to rigorously probe the influence of the substituted structures and their properties for self-assembly of [20-41] $\beta$ 2m amyloid, we synthesized a library of variants where Cys25 was substituted with various amino acids: Ala, Ile, Phe, pfPhe (pentafluorophenylalanine), chAla (cyclohexylalanine), Tyr, Trp, Gln, Asp, Lys, and Arg. Overall, all substituents with bulky hydrophobic groups yielded well-defined amyloid fibers, albeit with differences in morphology (**Table 3** and **Table 4**). Amyloids were also observed in the case of the Cys25Gln substitution; indeed, glutamine residues are frequently found in amyloids as their side chains can form H-bonds that structurally reinforce the  $\beta$ -sheet. In contrast, no fibers were detected for variants where amino acids with polar charged groups (acidic or basic) were inserted, or for Cys25Ala. These results underlie the necessity to fill the space between  $\beta$ -arches to stabilize the assembly as well as the low tolerance for charged side chains in the amyloid core.

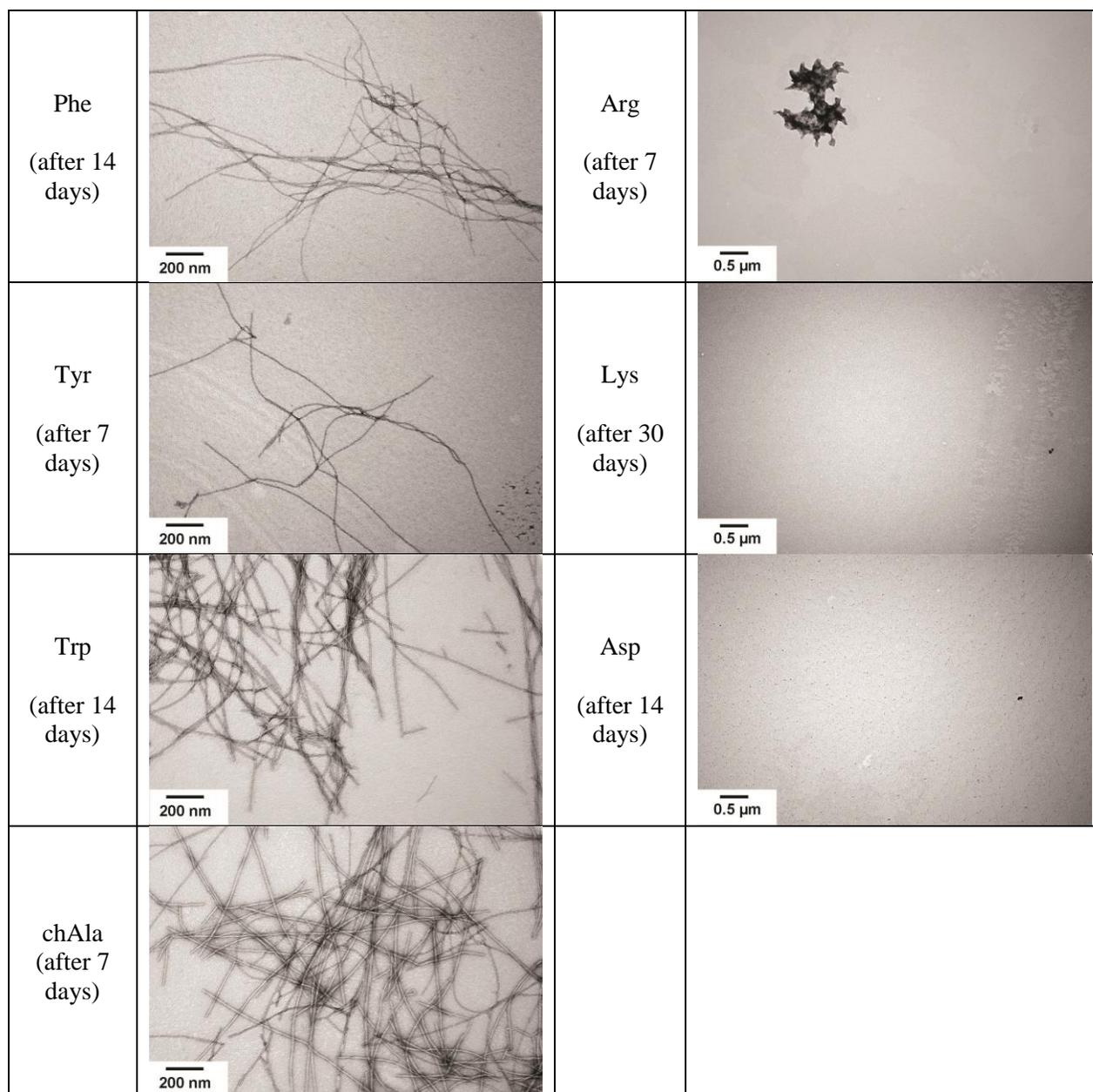
**Table 3:** Description of amyloid growth experiments for variants of [20-41] $\beta$ 2m peptide where Cys25 was substituted with Xaa = Ala, Ile, Phe, Tyr, Trp, chAla (cyclohexylalanine), pfPhe (pentafluorophenylalanine), Gln, Asp, Lys and Arg. Concentration of peptide is 100  $\mu$ M in 50 mM sodium phosphate, 100 mM NaCl, 0.05% NaN<sub>3</sub>, pH 7.5. Table is taken from *ChemBioChem* 17, 2274-2285 (2016).

Analogue	Structure of residue Xaa6 side chain	Description
Ala		No fibers by TEM (inspected after 3, 7, 14 and 30 days of incubation)
Ile		Long fibers of regular thickness by TEM (inspected after 3, 7, 14 and 30 days of incubation)
Phe		Long fibers of regular thickness by TEM (inspected after 3, 7, 14 and 30 days of incubation)
Tyr		Short fibers of regular thickness by TEM (detected after 3 days of incubation); long fibers of regular thickness by TEM (detected after 7, 14 and 30 days of incubation)
Trp		Short fibers of different thickness by TEM (after 3 days of incubation); Short and long fibers composed of several protofilaments (helically twisted: left-handed) after 7, 14 and 30 days of incubation

chAla		Long fibers of different thickness (detected after 3, 7, 14 and 30 days of incubation)
pfPhe		Long fibers of different thickness (detected after 3, 7, 14 and 30 days of incubation)
Gln		After 3 days: only few fibers are visible. After 7 days: long fibers of irregular thickness appear. These fibers are composed of several protofilaments and are helically twisted (left-handed) (inspected after 14 and 30 days)
Arg		No fibers by TEM (inspected after 3, 7, 14 and 30 days of incubation)
Lys		No fibers by TEM (inspected after 3, 7, 14 and 30 days of incubation)
Asp		No fibers by TEM (inspected after 3, 7, 14 and 30 days of incubation)

**Table 4:** TEM images for amyloids of [20-41] $\beta$ 2m with Xaa25 = Ala, Ile, Phe, Tyr, Trp, chAla (cyclohexylalanine), pfPhe (perfluorophenylalanine), Gln, Asp, Lys and Arg (see Table 3-1). Time of incubation corresponding to each image is indicated in parentheses. *Table is taken from ChemBioChem 17, 2274-2285 (2016).*

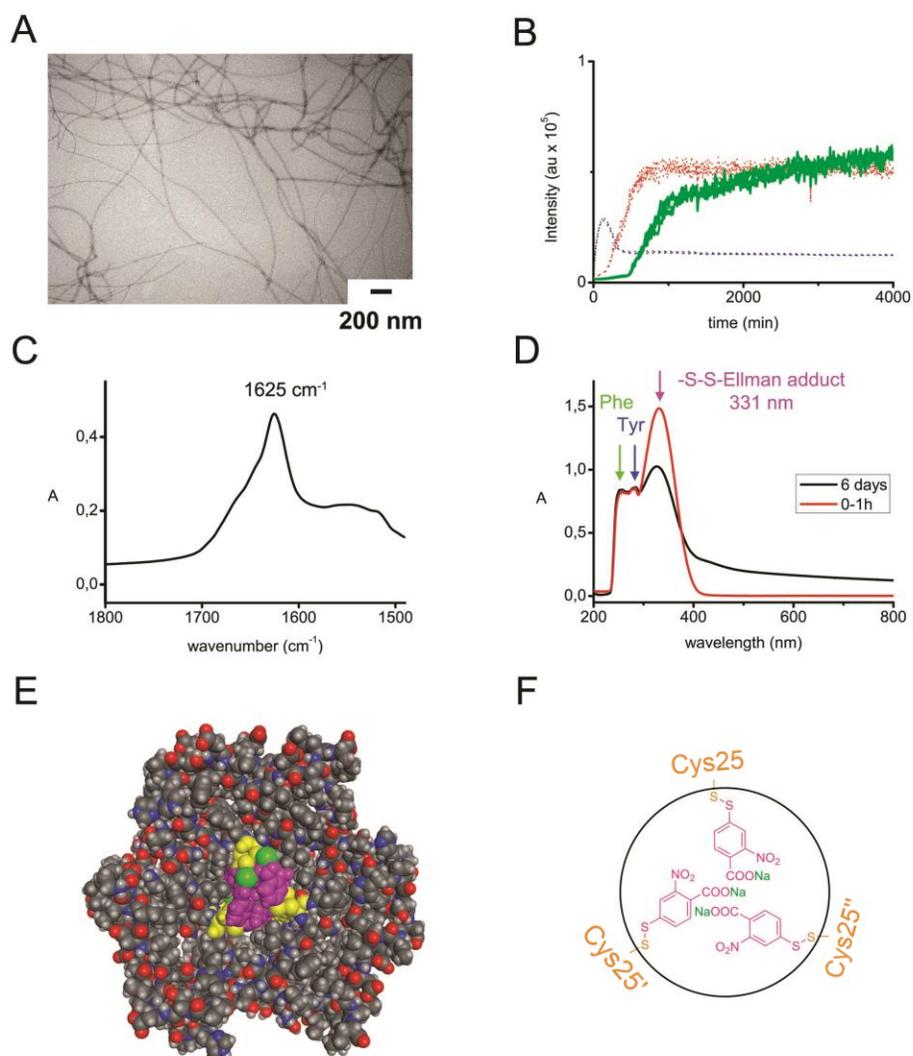
Analogue	TEM image	Analogue	TEM image
Ala (after 30 days)		pfPhe (after 3 days)	
Ile (after 14 days)		Gln (after 7 days)	



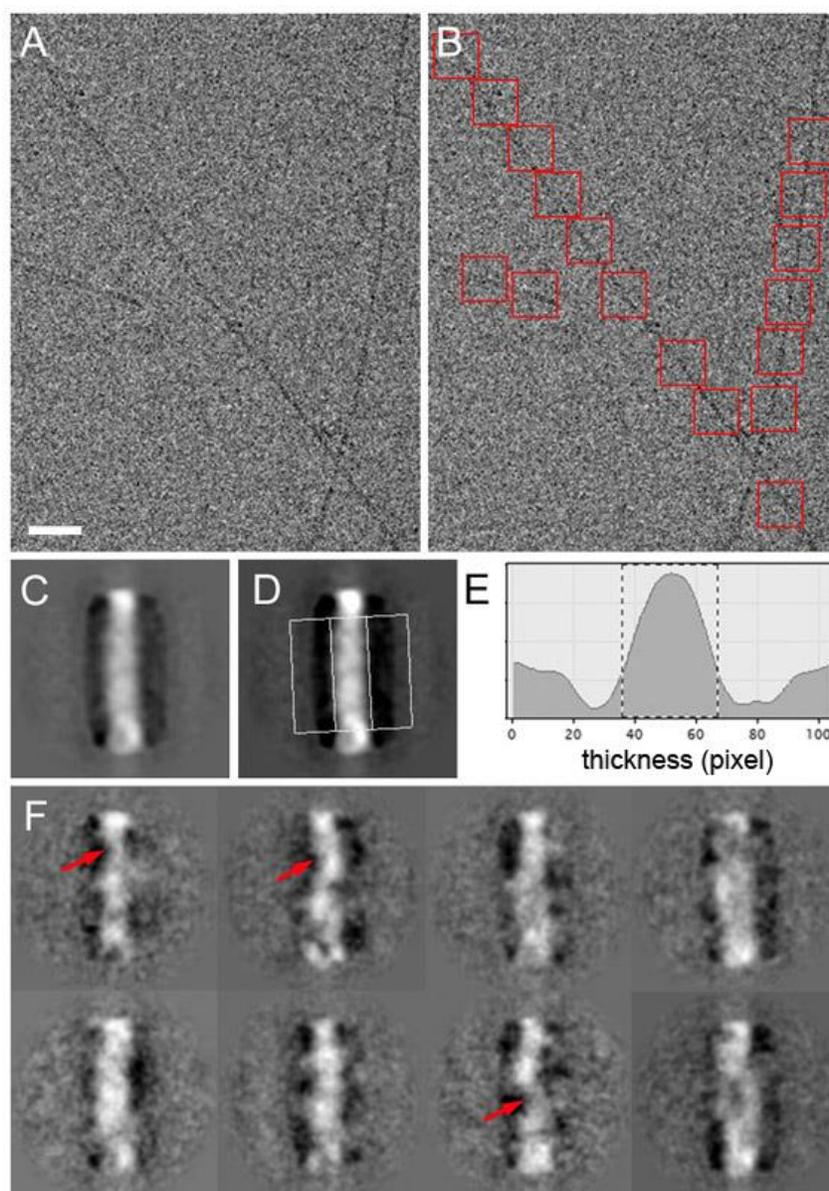
#### 3.4.4 Study of [20-41] $\beta$ 2m-(Ellman's-thiol) conjugate

An additional construct in this study was [20-41] $\beta$ 2m-(Ellman's thiol), where the Ellman's thiol group was conjugated by disulfide exchange with Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid)dithiol). Despite the fact that the Ellman thiol moiety contains a carboxylic group, long regular amyloids were observed at pH 7.5, at which the carboxylic group is

negatively charged (**Figure 34**). Moreover, the growth kinetics resembled that for the covalent trimer (**Figure 34-B**), with similar ATR-FT-IR spectra of the fibrils (**Figure 34-C**).



**Figure 34:** Properties of [20-41] $\beta$ 2m-(Ellman's thiol) conjugate (A) TEM image of the negatively stained long fibers after 20 days of incubation. The long regular amyloids are morphologically similar to covalent trimer fibers (for cryo-EM analysis, see Figure 3-12). (B) Kinetics of growth for amyloids for [20-41] $\beta$ 2m-(Ellman's thiol) conjugate (in green solid line) and covalent trimer (in red dot), covalent dimer (in blue dot) (provided for comparison) monitored by ThT dye fluorescence on a plate-reader. (C) ATR-FT-IR spectrum shown in absorbance mode of amide I band region. The maximum was determined by spectral deconvolution and inspection of the corresponding second derivative spectrum. Secondary structure deconvolution of the amide I band identifies contributions from  $\beta$ -strand (43%), turn (29%) and random coil (28%) structures. (D) UV-Vis absorption spectrum shown in absorbance mode at 0-1 h and after 6 days of incubation. (E) Molecular modeling of the cross-section of amyloids with a triangular core. The long fiber axis is perpendicular to the plane of the figure. (F) Chemical structure of the central core is depicted as a circular inset. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*



**Figure 35:** Cryo-EM analysis of amyloids obtained for the [20-41] $\beta$ 2m-(Ellman's thiol) conjugate. (A) A representative image of fibers. A total of 427 frames of 4096 $\times$ 4096 pixels were recorded with a pixel size of 0.178 nm. Out of these images 43 (10%) were selected for analysis. (B) Red squares illustrate the selection of sub-images to determine the average profile along the fiber. 3179 sub-images of 192 x 192 pixels in size and centered on visible filaments were identified and iteratively aligned so that the filaments are depicted vertically. (C) The average image of the aligned patches. (D) The boxing made to determine the average profile along the fiber. (E) An average diameter corresponds to 5.8 nm. (F) A set of representative image classes, which reveal structural heterogeneity. Thinner filaments can be observed with a diameter of half the average diameter (about 2.5-3.0 nm) indicated by red arrows. This observation suggests that there are deviations from covalent trimer architecture presumably due to non-covalent nature of interactions between subunits (as opposite to covalent tethering in covalent trimer amyloids). The scale bar (shown in A) corresponds to 50 nm in A and B, and 8.5 nm in C, D, and F. *Figure is taken from ChemBioChem 17, 2274-2285 (2016).*

Furthermore, cryo-EM and image analysis determined an average diameter of 5.8 nm (**Figure 35**), which is also close to the value for the covalent trimer fibrils (**Figure 30-B**). Overall, these observations suggest the presence of the trimeric amyloid core with negatively charged carboxylates of Ellman's thiol balanced by counter-ions (e.g., Na<sup>+</sup>); however, proving this conjecture requires further study. Alternatively, more substantial conformational changes of the peptide building block might accompany self-assembly of the corresponding amyloids.

### 3.5 Discussion

Designing amyloids has a strong potential to aid the elucidation of the role of primary sequence and spatial conformational requirements in self-assembly mechanisms. Pioneering work in this field includes the determination of sequence requirements in short peptides for facilitation of the precipitation of fibers<sup>132</sup>, and the design and demonstration of amyloidogenic properties in peptide constructs binary patterned with alternating (hydrophilic-hydrophobic)<sub>n</sub> amino acids<sup>133</sup>. Elegant studies have shown that properly placed non-canonical  $\beta$ -turn-supporting residues (including *D*-amino acids) enhance amyloidogenicity<sup>134</sup>.

Furthermore, intra- and intermolecular covalent tethering of amyloidogenic peptides was shown to be efficient to alter the properties of amyloids. For example, the introduction of an intramolecular lactam bridge between the side chains of Asp23 and Lys28 of the A $\beta$ (1-40) peptide supported the  $\beta$ -arch conformation and resulted in a peptide variant with a 1000-times faster rate of amyloid formation<sup>135</sup>. Engineered intramolecular disulfides in A $\beta$ (1-42) (Leu17Cys/Leu34Cys or Ala21Cys/Ala30Cys) stabilized the  $\beta$ -hairpin peptide conformation and led to the precipitation of pre-fibrillar oligomers, thus providing an insight into the early stages of amyloid growth<sup>136-138</sup>. In the case of the Ala21Cys/Ala30Cys disulfide-linked construct, ssNMR analysis resulted in a structural model for hexameric assembly. Moreover, these species exhibit higher neural toxicity than the parent A $\beta$ 42. In addition to intramolecular covalent tethering, various intermolecular dimeric disulfide-linked A $\beta$  constructs have been produced; these differed

in fiber formation kinetics and exhibited biological effects previously undetected for wild-type A $\beta$ <sup>139</sup>.

Our work complements previous studies and suggests that covalent tethering is a promising technique for systematic studies of amyloid structure and function. We showed that with the help of an abiotic covalent linker we can direct self-assembly of [20-41] $\beta$ 2m peptide into amyloids with a triangular core. Various experimental approaches unambiguously confirmed such mode of self-assembly. Moreover, well-defined trimeric stoichiometry at the fiber cross-section results in the distinct amyloid growth kinetics and fiber morphology. Importantly, covalent tethering did not alter the parallel  $\beta$ -sheet interactions previously observed for [20-41] $\beta$ 2m monomer amyloids<sup>27</sup>. The data also suggest that the  $\beta$ -arch conformation of [20-41] $\beta$ 2m peptide building block is likely to be retained in the covalent trimer. For example, if  $\beta$ -arch unfolded into linear  $\beta$ -strand this would have resulted in thicker fibers than were observed in our study; or if  $\beta$ -arch isomerized into  $\beta$ -hairpin, then the  $\beta$ -sheets would become antiparallel. These results suggest that it will be possible to engineer even more complex covalently-tethered structures and potentially to reproduce the peptide conformations observed in, for example, the known structures of A $\beta$  amyloids.

In addition, we demonstrated that a noncovalent (supramolecular) approach can be used to modulate self-assembly of amyloidogenic peptides. Depending on the properties of the moieties attached to [20-41] $\beta$ 2m, we obtained a full spectrum of self-assembly behavior, from homogeneous long fibers to polymorphic amyloids to non-amyloidogenic peptides. Although a detailed understanding of the mechanisms governing amyloid growth for each of these constructs requires further study, empirical dependencies can be deduced from our experiments. Specifically, the formation of a well-structured amyloid is dictated by the volume, polarity, and charge of the attached group.

Our data are in an agreement with previous studies of different amyloids systems (such as A $\beta$ , islet amyloid polypeptide, and designed amyloids), where it was shown that hydrophobic contacts of aliphatic and aromatic side chains and  $\pi$ - $\pi$  stacking of aromatic residues modulate the growth kinetics and stability of amyloids<sup>140,141</sup>. The  $\pi$ - $\pi$  stacking is critical for the stability of amyloid structures, especially those formed from short peptides<sup>142</sup>.

The novelty in our work is the attempt to explore hydrophobic interactions and  $\pi$ - $\pi$  stacking to induce precipitation of amyloid fibers with a triangular core. The similarity of morphologies of long fibers produced by various modified [20-41] $\beta$ 2m constructs and the reference covalently tethered trimeric construct indicate that self-assembly of trimeric fibers indeed takes place. Such principles could be exploited for controlling stoichiometry of amyloid self-assembly in other amyloidogenic peptide systems.

### **3.6 Conclusions**

In summary, we reported chemical strategies that enable control of the formation of distinct amyloid polymorphs, such as (i) covalent tethering to define the structure of the molecular unit for self-assembly, and (ii) conjugation of various substituents to a hydrophobic part of the  $\beta$ -arch in order to modulate the stoichiometry of the fiber core by supramolecular interactions. The structural properties of the covalent trimer [20-41] $\beta$ 2m amyloids were characterized by cryo-EM, ssNMR, and FTIR. Moreover, we found that the individual peptide building block in the covalent trimer likely retains the  $\beta$ -arch conformation, similarly to [20-41] $\beta$ 2m covalent dimer and monomer amyloids. Such findings highlight the potential of amyloidogenic peptides with  $\beta$ -arch ( $\beta$ -strand-loop- $\beta$ -strand) conformation to serve as modular elements for a hierarchical design of distinct amyloid structures.

The possible future applications of these approaches are the preparation of homogenous amyloid samples for structural studies (e.g., ssNMR experiments with engineered A $\beta$  peptides), immunization experiments (to identify polymorph-specific antibodies), and screening of small-molecules to selectively inhibit fiber growth of a particular amyloid polymorph. Furthermore, the high diversity of the attached substituents tolerated by [20-41] $\beta$ 2m peptide while still retaining the ability to produce fibrillar structures suggests that this (and other amyloidogenic peptides engineered by similar ways) might serve as a scaffold for the controlled placement of various functional groups with unique electronic and photonic properties, in order to develop novel biomaterials.

## **4. Total chemical synthesis and properties of covalently tethered oligomer mimicking amyloid structure**

### **4.1 Introduction**

Amyloids are the hallmark of protein misfolding diseases in which soluble monomeric proteins or peptides undergo conformational isomerization to form various oligomers and subsequently insoluble fibrils. More and more structural information is available nowadays for amyloid fibrils with several high-resolution structures recently reported based on solid-state NMR<sup>124,126,127,143,144</sup> and cryo-EM studies<sup>145,146</sup>. Despite the recent success in structural elucidation of amyloid fibrils much less is known for metastable transient oligomers which may structurally deviate or resemble the final amyloids. Their characterization is difficult due to their high compositional and structural dynamics; they can adopt very different structures and present a lot of heterogeneity. Now it seems clear that the intermediate proto-fibrils or proto-filaments and small oligomers emerging at the early stages of amyloid formation play a key role in the development of protein and peptide misfolding diseases<sup>147,14</sup>. Those small oligomers appear to be the most toxic species and, therefore, represent an important target for the development of new therapeutic agents against neurodegenerative disorders such as Alzheimer's disease<sup>148</sup>.

One possible way to study properties of oligomers is by covalent cross-linking whereby the application of chemoselective residue-specific reagents enables to stabilize transient non-covalent assemblies. This method was used to prepare covalently stabilized oligomers of A $\beta$ (1-40) of different sizes (dimer and trimer) in pure state<sup>149</sup>. Cross-linking requires the presence of reactive aromatic moieties such as the phenol ring of tyrosine residues and the position of such residues within the oligomer are of high importance. For instance, the preparation of stable A $\beta$ (1-42) oligomers was only possible by changing the position of tyrosine residues for the cross-linking reaction<sup>150</sup>. Another study showed that cross-linking of A $\beta$  peptides produces oligomers limited in size, not larger than trimer for A $\beta$ (1-40) and tetramer for A $\beta$ (1-42)<sup>151</sup>. Moreover, the conformations of the oligomers were altered after the cross-linking reaction, leading to species

with less  $\beta$ -strand content than non-crosslinked oligomers<sup>151</sup>. Therefore, new approaches must be investigated.

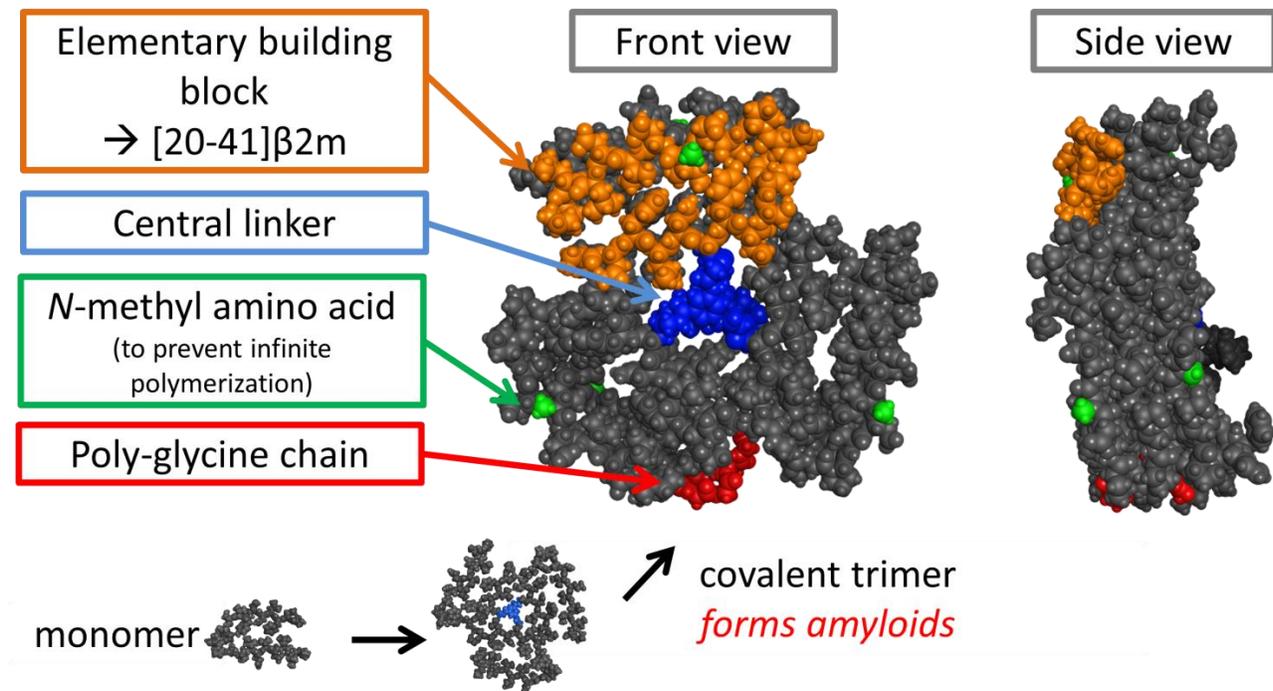
Total chemical synthesis of oligomers is another approach that can offer the possibility to prepare homogeneous samples, which can be characterized in details by structural methods and biological essays. Covalent tethering of several copies of amyloidogenic peptidic segments with atomic-level control will allow for the formation of unique and well-defined structures. In the literature, not much was so far reported about synthesis of such oligomers. A construct was synthesized, in which a cyclic peptide served to display three copies of the amyloidogenic [25-35]A $\beta$  fragment<sup>152</sup>, and it was found that this construct exhibited enhanced aggregation and cytotoxicity properties. A construct composed of four [16-37]A $\beta$  fragments attached to a cyclic peptide scaffold was synthesized and it was shown that the resulting protofibril was binding ThT and Congo Red and displayed  $\beta$ -sheet signature by circular dichroism<sup>153</sup>. A tetrameric A $\beta$  “mini-amyloid” was self-assembled on a polyol template<sup>154</sup>. The conformation of [32-40]A $\beta$  fragments led to a nearly parallel arrangement with a spacing of 5 Å but with an out-of-register orientation.

In the current study, we demonstrate the design principles of a covalently-tethered oligomer that mimics amyloid structure and elaborate method for its total chemical synthesis. Although, our design is inspired by existing structures of amyloid fibrils, the covalently-tethered oligomer may have features similar to those in supramolecular oligomers. In order to avoid formation of fibrillary structures and to render the oligomers soluble, several chemical modifications were introduced into the corresponding solvent exposed  $\beta$ -strands. Furthermore, the folding of the oligomer into distinct monomeric structures was confirmed by solution NMR measurements.

## **4.2 Design of the structure of covalently-tethered oligomer**

The design of covalently-tethered oligomer is based on the structural model of A $\beta$ (1-40) amyloid with 3-fold symmetric arrangement of amyloidogenic peptide subunits<sup>37</sup>. In this structure, three copies of A $\beta$ (1-40) adopt a  $\beta$ -arch ( $\beta$ -strand-turn- $\beta$ -strand) conformation in which

$\beta$ -strands interact via amino acid side chains and not through backbone hydrogen bonds. The amyloidogenic [21-40]-segment of human  $\beta$ 2-microglobulin ([21-40] $\beta$ 2m) adopts  $\beta$ -arch conformation similar to A $\beta$ (1-40)<sup>27</sup> but it represents a more tractable system to explore methodologies of chemical synthesis towards various multimeric constructs. In chapter 3, we showed that three copies of [21-40] $\beta$ 2m can be covalently joined with an appropriate trimeric linker and, as a result, can form morphologically distinct homogeneous amyloid fibrils. In this study, we engineered a covalently-tethered oligomer composed of three units of covalent trimer of [21-40] $\beta$ 2m, which comprises nine copies of a parent [21-40] $\beta$ 2m peptide in total (**Figure 36**).



**Figure 36:** Molecular model of covalently-tethered oligomer. The [20-41] $\beta$ 2m segment (PDB ID 2E8D), represented in orange is adopting a  $\beta$ -arch conformation within a protofibril. It is the elementary building block of the oligomer. Three copies of [20-41] $\beta$ 2m are covalently linked, to form a trimeric species (a “covalent trimer”), through an abiotic central linker represented in blue. Three copies of this covalent trimer are covalently attached together via poly-glycine chains (represented in red) to form the “covalent trimer of covalent trimers”. The external trimers are modified with *N*-methylated amino acid (represented in green) in order to prevent polymerization of the structure and to obtain a monomeric species.

The design of such complex peptide architecture was realized with the help of molecular modeling taking into account our knowledge about the structural properties of the [21-40] $\beta$ 2m covalent trimer. We envisaged three copies of covalent trimer [21-40] $\beta$ 2m to be stacked in parallel, in register orientation, such that the internal covalent trimer has the surroundings identical to those found in the amyloid fibril as it is “sandwiched” by two external covalent trimers. To prevent extensive fibrillation akin to amyloids, a subset of solvent exposed amides in the external trimers is modified by *N*-methylation with the purpose to disrupt hydrogen bonding of  $\beta$ -sheets necessary for amyloid propagation. Further modifications include conformationally flexible tetra-glycine chains used as covalent tethers between three covalent trimers to ensure their correct arrangement.

### **4.3 Strategy used for the *N*-methylation of the external trimers**

The covalent oligomer is composed of nine amyloidogenic [20-41] $\beta$ 2m fragments. Once the structure is folded, we predicted that it would mimic a cross section of an amyloid fiber with a three-fold symmetry. Molecular modelling of the oligomer is in agreement with our predictions (**Figure 36**). Each peptidic fragment would adopt a  $\beta$ -arch ( $\beta$ -strand-loop- $\beta$ -strand) conformation, and the folding of the entire structure would be driven by the formation of intramolecular  $\beta$ -sheets and the formation of a hydrophobic pocket at the center of the structure. Each covalent trimer would interact with the neighboring covalent trimers through hydrogen bonding with the peptidic backbone involved in the  $\beta$ -sheets. In this case, it is evident that such structure would interact with other partners through the same type of interactions and, therefore, will be able to form large amyloid aggregates with same morphology as previously observed for [20-41] $\beta$ 2m covalent trimer in chapter 3. Without any chemical modifications, nothing can prevent intramolecular stacking and formation of larger species that would ultimately lead to precipitation of the aggregates.

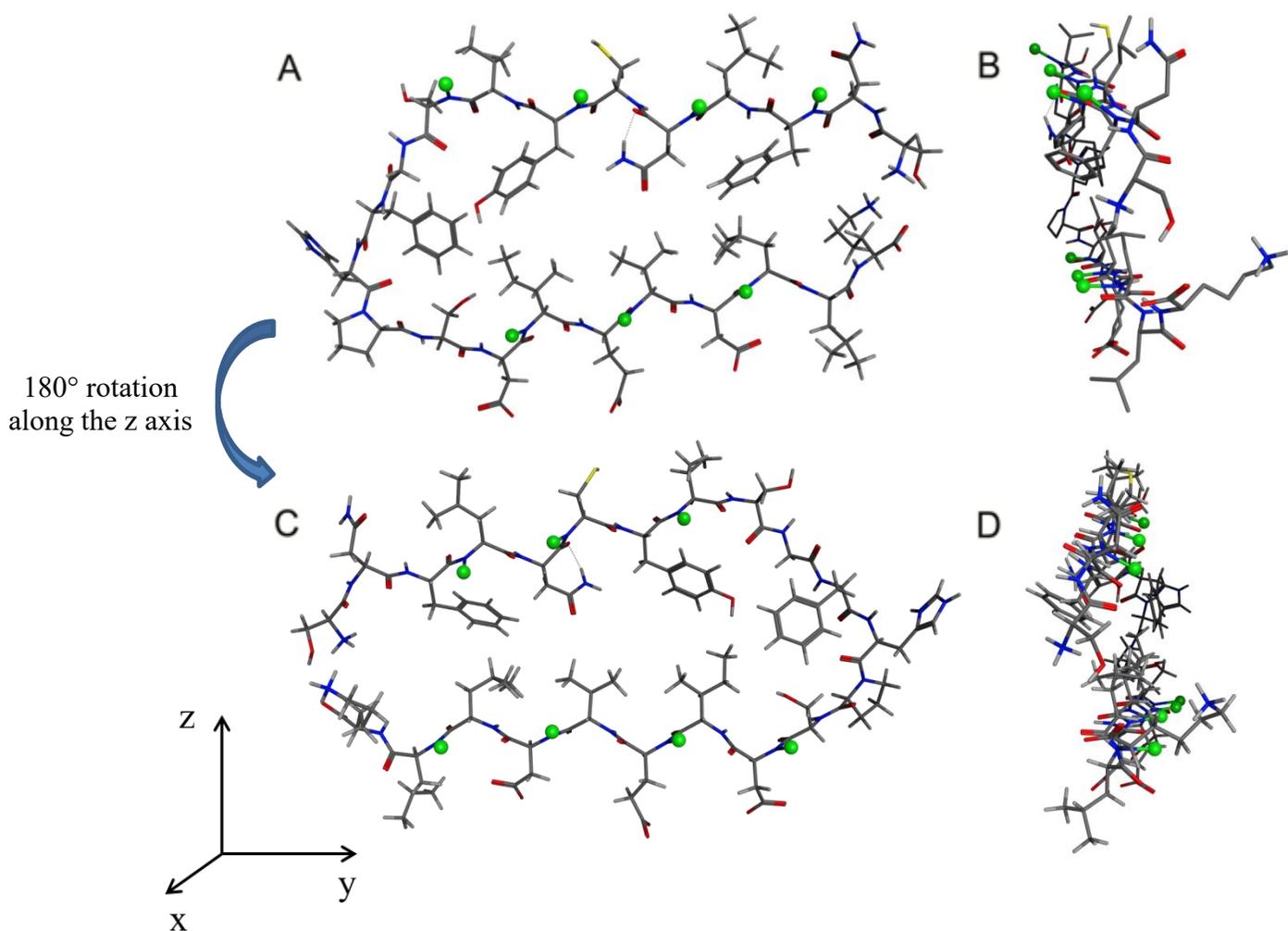
The external covalent trimers have one side of their structure with solvent-exposed backbone of  $\beta$ -sheets. To prevent stacking with other covalent oligomers, the interactions responsible for inter-molecular  $\beta$ -sheet formation must be suppressed or at least perturbed. One solution would

be to engineer the exposed backbone amide groups to repel other partners that would act like an “amyloid defense shield”, without disrupting the  $\beta$ -sheets inside the structure. We decided to replace a certain number of amino acid residues by *N*-methylated analogues at specific positions.

*N*-methyl amino acids are known to be  $\beta$ -sheet breakers<sup>155</sup> and are predicted to disrupt oligomer-oligomer interactions. According to the structure of [20-41] $\beta$ 2m protofilaments obtained by solid-state NMR<sup>27</sup>, we decided to place the methylations in the central parts of  $\beta$ -strands. It is important that the amino acids bearing the *N*-methylation are adopting a  $\beta$ -sheet conformation and that the methyl groups are pointing out of the folded structure. Replacing the hydrogen atom by a methyl group on the nitrogen atom of the amide will suppress one hydrogen bond interaction and create a steric hindrance that will perturb the hydrogen bond interactions of the residues near the modification<sup>156</sup>. Moreover, *N*-methylation of the peptide backbone induces rotational restrictions to the amide bond, influences the *cis-trans* equilibrium and can be used to modulate the conformation<sup>157</sup>. *N*-methylated peptides are less flexible and the residues next to the modified amino acid are conformationally affected and hence, restricted<sup>158</sup>.

It is important to mention that each side of the oligomer exposed to the solvent is different. Hence, the backbone of both external covalent trimers has to be modified with different methylation patterns. To better describe the strategy of [20-41] $\beta$ 2m methylation, each side of [20-41] $\beta$ 2m will be named according to the molecular models depicted in **Figure 37**. **Figure 37-A** represents one side of the [20-41] $\beta$ 2m fragment with a  $\beta$ -arch conformation that is called **Interface A**. In this case, the hydrogen atoms (highlighted in green) of the amide groups of residues Phe22, Asn24, Tyr26 and Ser28 in the upper  $\beta$ -stand and Ile35, Val37 and Leu20 in the lower  $\beta$ -stand are pointing towards the positive *x* direction. All of this amino acid positions are possible sites for *N*-methylation. One of the external “covalent trimer” has to be methylated on this side (Interface A) in order to have the methyl groups pointing out of the oligomer structure. **Figure 37-C** represents **Interface B** of the [20-41] $\beta$ 2m fragment: in this case the possible methylation sites are Leu23, Cys25 and Val27 in the upper  $\beta$ -stand and Asp34, Glu36, Asp38 and Leu21 in the lower  $\beta$ -stand and the hydrogen atoms are pointing towards the positive *x* direction. Therefore, the other external “covalent trimer” has to be methylated on this side (Interface B).

**Figure 37-B** and **Figure 37-D** represent the side views of Interface A and Interface B respectively with the location of every possible methylation sites and their orientations.

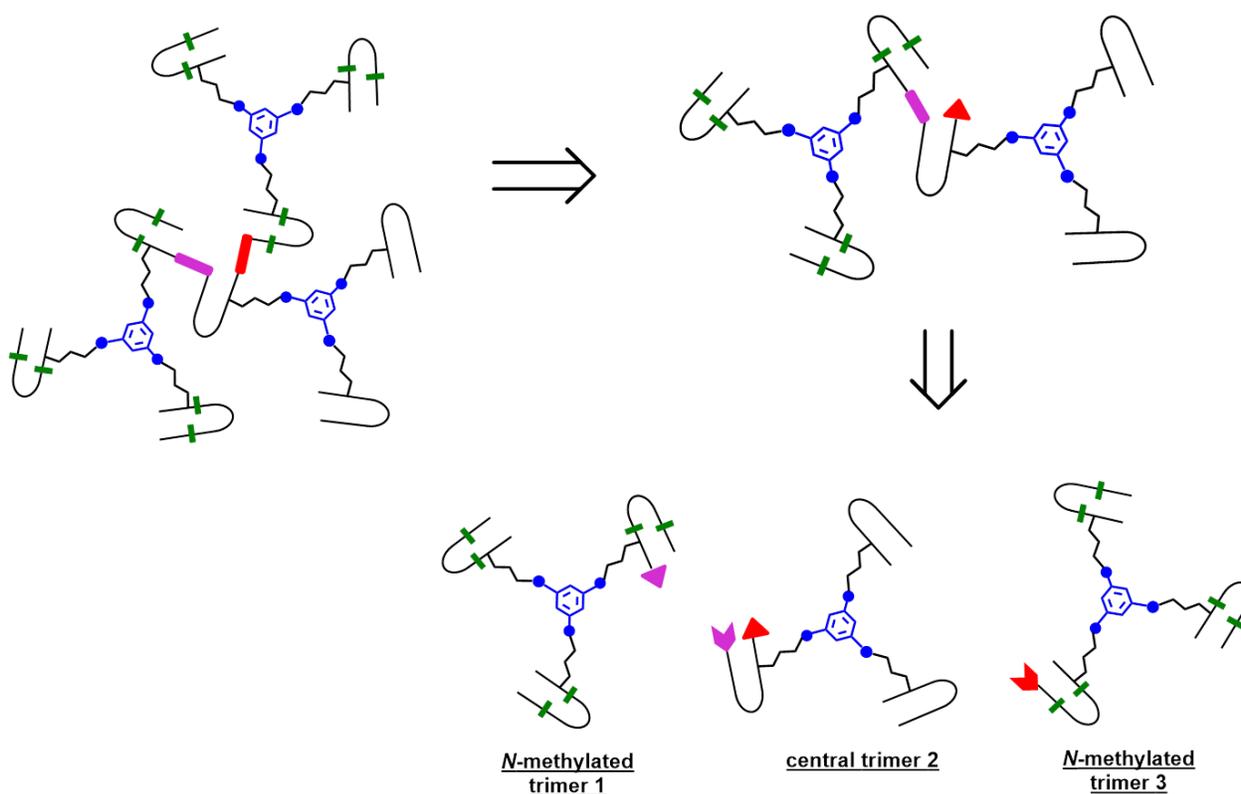


**Figure 37:** Molecular models of the [20-41] $\beta$ 2m segment based on the structure obtained by solid-state NMR in [20-41] $\beta$ 2m proto-fibrils (PDB ID 2E8D)<sup>27</sup>. The green spheres represent the possible sites for substitutions of the hydrogen atoms of the amide groups by methyl groups. A) Interface A of [20-41] $\beta$ 2m. B) Side view (from the right) of Interface A. C) Interface B of [20-41] $\beta$ 2m. D) Side view (from the left) of Interface B.

#### 4.4 Design of the synthetic route

Retrosynthetic analysis towards total synthesis of the designed peptide construct is depicted in **Figure 38**. The synthetic target represents peptide with a non-linear topology, resembling that

of comb-copolymer with the additional star AB<sub>2</sub> extensions. The tetra-glycine linkers, shown in magenta and in red, are connecting the corresponding covalent trimers.



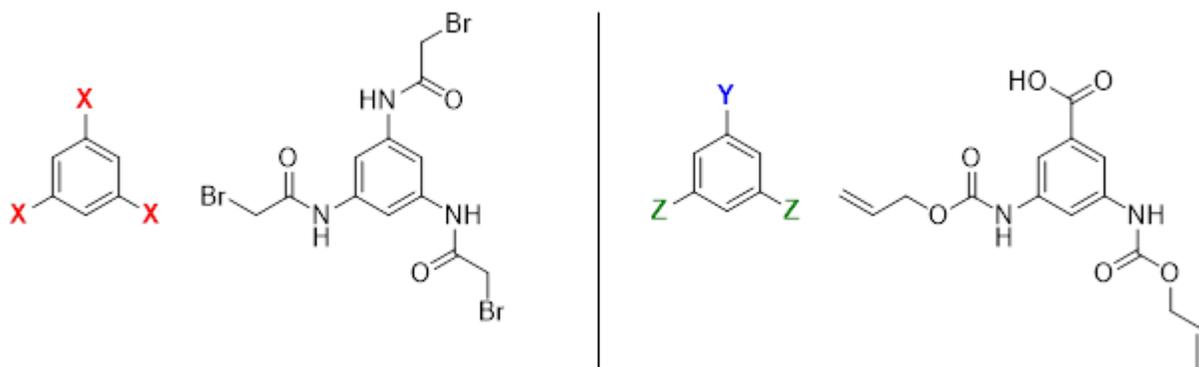
**Figure 38:** Retrosynthetic analysis of the covalently-tethered oligomer. The final construct is obtained by sequential native chemical ligation of the different covalent trimers. The parent 22-residue  $\beta$ -arch of [20-41] $\beta$ 2m is represented by black arches. *N*-methylations are represented in green and are localized on the external trimers and the central linkers are represented in blue. In the final construct, the poly-glycine chains are represented in red and in magenta.

The sequential tethering of the covalent trimers is to be carried out by native chemical ligation, therefore, one of the peptide subunits in each covalent trimers is modified with required functionalities such as a cysteine residue at N-terminus and  $\alpha$ hydrazide moiety at C-terminus. The cysteine residue is introduced in place of one of the glycine residue in a tetra-glycine linker (hence the final linker sequence is Cys-Gly-Gly-Gly). The  $\alpha$ hydrazides can efficiently be converted into  $\alpha$ thioesters<sup>77,78</sup>. As a result, three distinct non-symmetric covalent trimers need to

be prepared for the assembly of the desired covalently-tethered oligomer from six modified [21-40] $\beta$ 2m peptides.

#### 4.5 Design and synthesis of the central linker

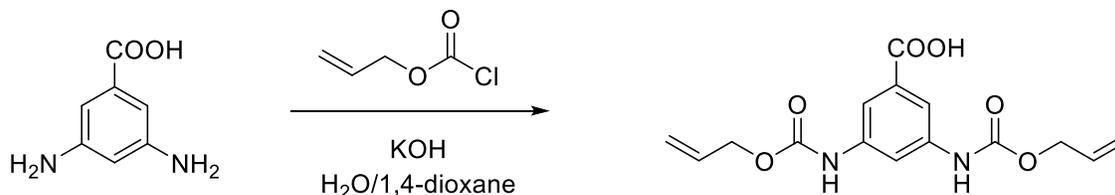
In our previous work on the covalent trimer [20-41] $\beta$ 2m amyloids, the trivalent *N,N,N'*-(benzene-1,3,5-triyl)tris(acetamide) linker (**Figure 39**, left side) was used to join three copies of [20-41] $\beta$ 2m peptide. The conjugation was based on a nucleophilic displacement of symmetrically substituted bromoacetamide derivative by a cysteine residue of the [20-41] $\beta$ 2m peptide. In this work, according to our design principles, the central linker needs to support the attachment of distinct peptide constructs that would enable introduction of the additional tetra-glycine linkers. Thus, the central core linker was designed to support synthesis of covalent trimers that are composed of two identical peptides and the one that is distinct. As a result, the structure of the new central linker 3,5-bis(((allyloxy)carbonyl)amino)benzoic acid contains two amines protected by Alloc ((allyloxy)carbonyl) groups and one reactive carboxylic acid (**Figure 39**, right side).



**Figure 39:** Chemical structures of the central linkers. Left: the linker used for the formation of symmetric trimers of [20-41] $\beta$ 2m peptide (“covalent trimer”). Right: the new central linker allowing the synthesis of covalent trimers with non-identical peptides.

The synthesis of the central linker was carried on a 6.5 mmol scale (**Figure 40**). One equivalent of 3,5-diaminobenzoic acid was mixed with three equivalents of potassium hydroxide in a mixture of water and 1,4-dioxane. A solution of 2.2 equivalents of allyl chloroformate in 1,4-

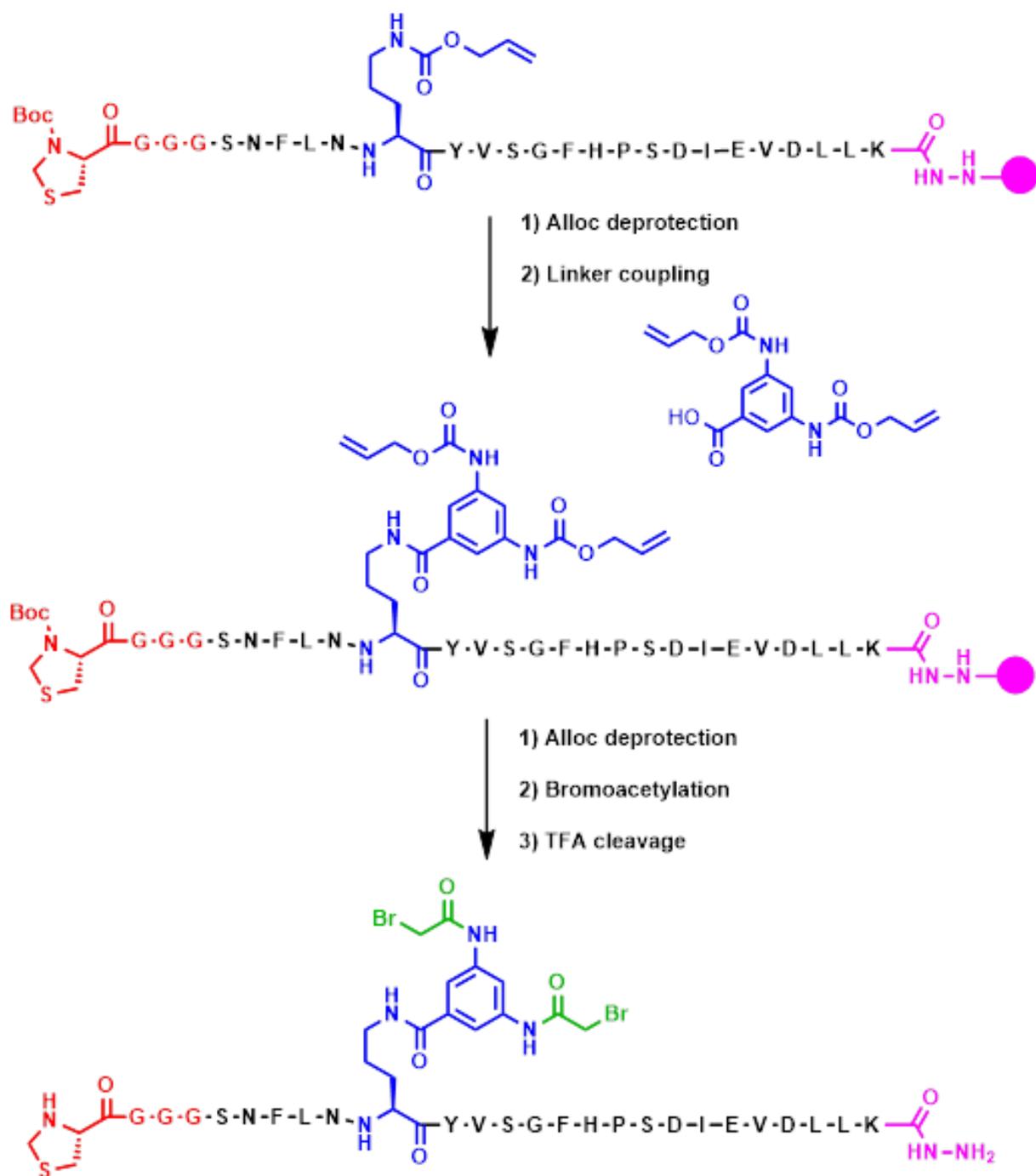
dioxane was added dropwise to the reaction mixture. It was let to react overnight. After purification by several liquid-liquid extractions, the final product was obtained with very high purity. It didn't require any other type of purification. The yield in the synthesis was 95%.



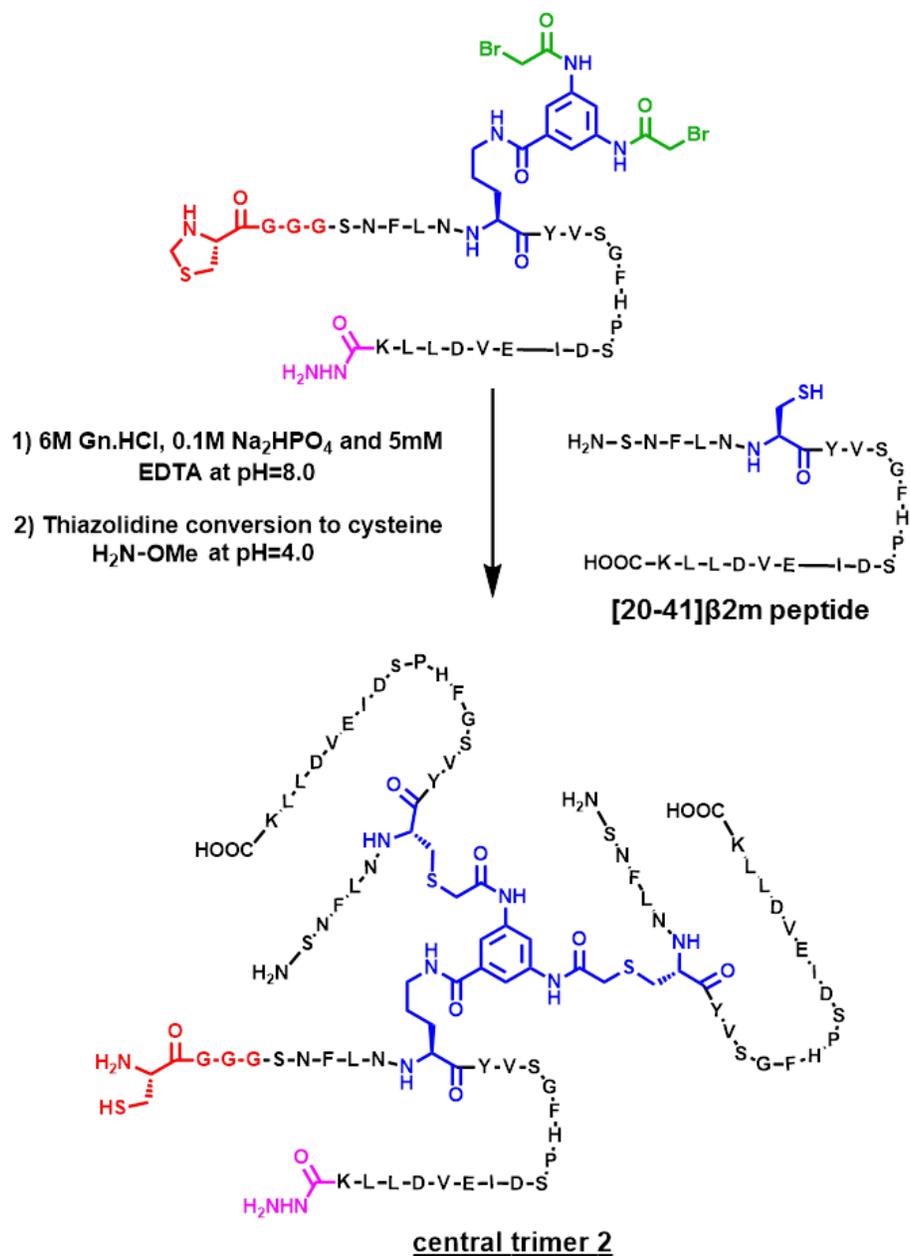
**Figure 40:** Synthesis of the central linker used in the synthesis of non-symmetric covalent trimers.

## 4.6 Synthesis of the covalently-tethered oligomers

The unprotected carboxylic acid was coupled to the side chain of ornithine introduced at position 25 through amidation. Ornithine was chosen because the length of the side chain is optimal and leads to the same number of atoms between the peptide backbone and the aromatic ring of the central linker as for the first generation linker used in the synthesis of covalent trimer [20-41] $\beta$ 2m. Corresponding peptide was synthesized with the side-chain of the ornithine protected by Alloc which is orthogonal to the side chain protecting groups of the other amino acid commonly used in Fmoc-SPPS. It was deprotected at the end of the chain assembly on the solid support and the linker was directly coupled (**Figure 41**). In the next step, Alloc-protecting groups were removed from the primary aromatic amino groups of the central linker and the amines were bromoacetylated. The final step is the cleavage of the modified peptide from the resin and purification by HPLC. The formation of the non-symmetric covalent trimers was then achieved by reaction with [20-41] $\beta$ 2m peptides having a cysteine residue at position 25 (**Figure 42**).

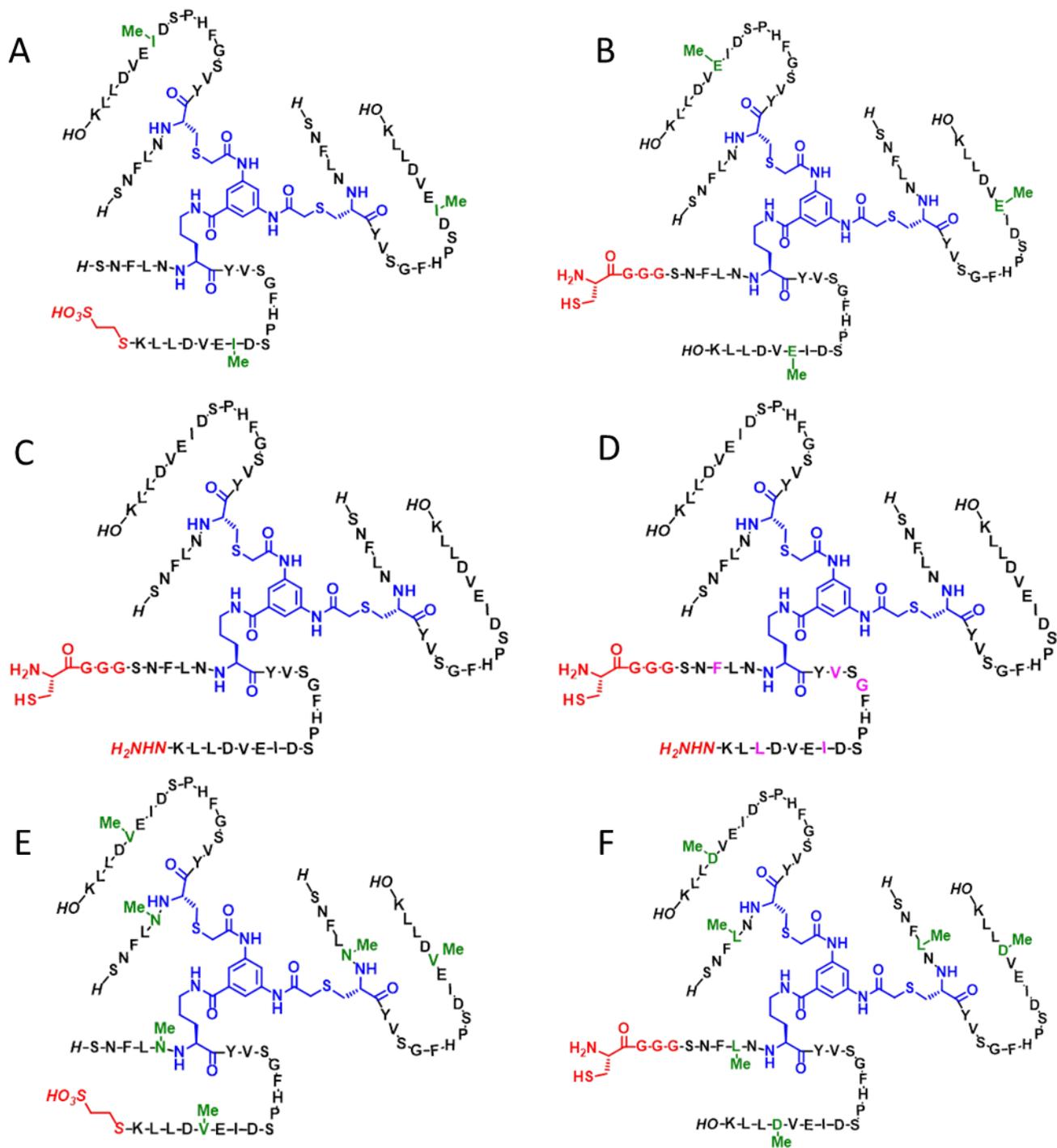


**Figure 41:** Strategy employed for the on-resin conjugation of the central linker via the side-chain of ornithine 25. This figure is illustrating the synthesis of the peptide-linker used for the production of the non-methylated central trimer of the oligomer. The same strategy is used for the production of the other covalent trimers. The circles in magenta represent the solid support.



**Figure 42:** Strategy employed for the synthesis of non-symmetric covalent trimers. The conjugation of two copies of distinct peptide unit was achieved by nucleophilic displacement of bromoacetamide derivatives by a cysteine residue of the [20-41]β<sub>2</sub>m peptide.

Several trimers were prepared according to this strategy. The *N*-methylated trimer-1-A (**Figure 43-A**) possesses one methyl group per peptide (pointing at the Interface A) at isoleucine 35 residues. The *N*-methylated trimer-3-A (**Figure 43-B**) possesses one methyl group per peptide (pointing at the Interface B) at glutamic acid 36 residues.



**Figure 43:** Structures of different covalent trimers. A) *N*-methylated trimer-1-A; B) *N*-methylated trimer-3-A; C) Central trimer-2-A; D) Central trimer-2-B, E) *N*-methylated trimer-1-B. F) *N*-methylated trimer-3-B. *N*-methyl amino-acids and  $^{15}\text{N}$ -isotopically labeled amino acids are represented in green and in magenta, respectively.

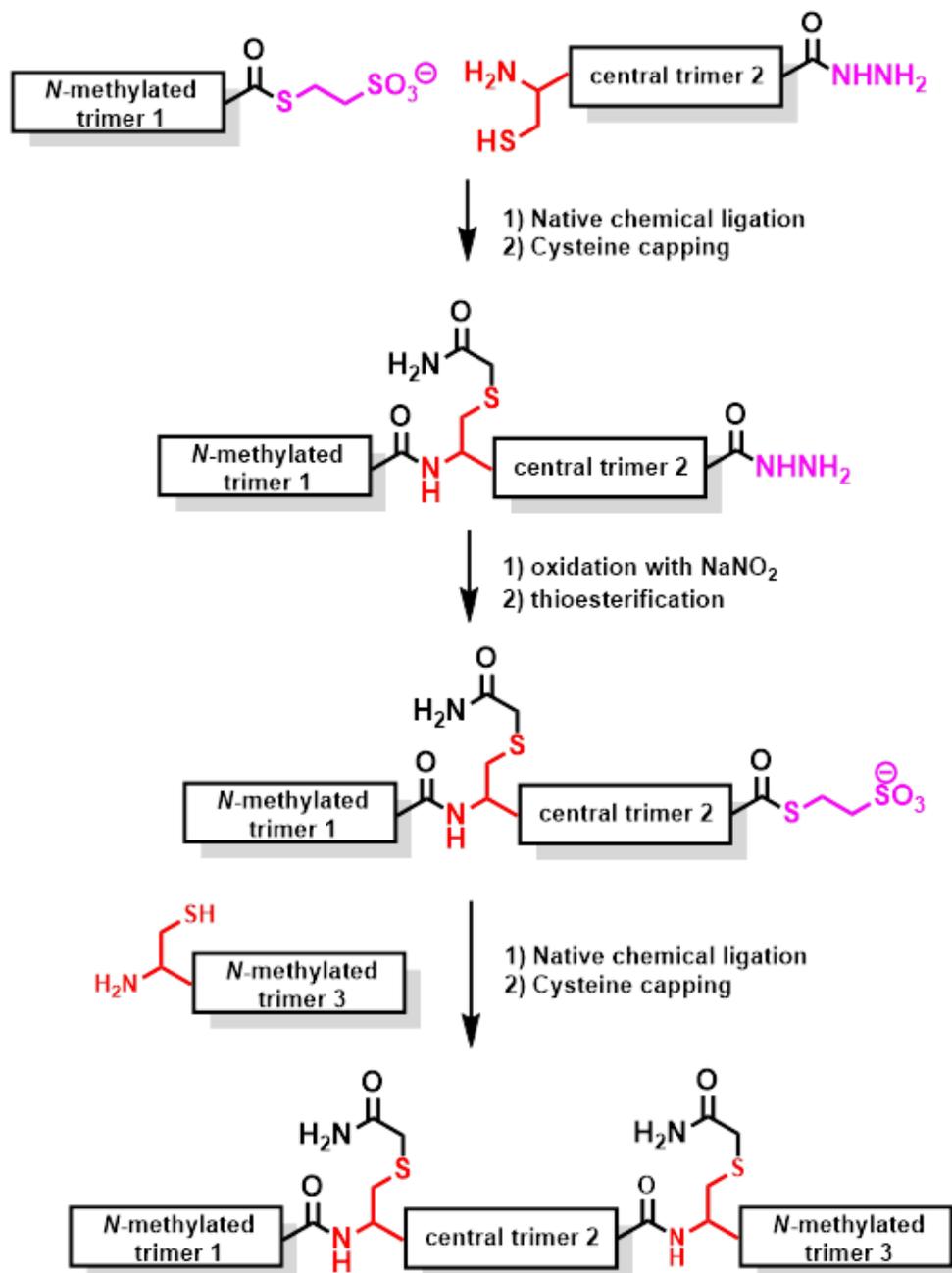
The central trimer-2-A and the central trimer-2-B possess the same chemical structure (**Figure 43-C** and **Figure 43-D**) except that the central trimer-2-B is <sup>15</sup>N-isotopically labeled at five positions (Phe22, Val27, Gly 29, Ile 35 and Leu39) within the same peptide. The *N*-methylated trimer-1-B (**Figure 43-E**) possesses two methyl groups per peptide (pointing at the Interface A) at asparagine 24 and valine 37 residues. The *N*-methylated trimer-3-B (**Figure 43-F**) possesses two methyl groups per peptide (pointing at the Interface B) at leucine 23 and aspartic acid 38 residues.

Native chemical ligation was used to tether the three non-symmetric covalent trimers in the required order (**Figure 44**). The *N*-methylated trimer 1 was obtained as <sup>α</sup>thioester and ligated to the central trimer; the Cys-residue at the ligation junction was alkylated with 2-bromoacetamide to avoid undesirable Cys-oxidation. After HPLC-purification, the <sup>α</sup>hydrazide group in the product of the first ligation was converted into <sup>α</sup>thioester and the second ligation was performed with the *N*-methylated trimer 3. After subsequent alkylation with 2-bromoacetamide and HPLC purification, the desired product was obtained with an overall yield of 13% (based on the quantity of *N*-methylated trimer 1 used for the first ligation reaction).

#### **4.7 Description of the different oligomer constructs**

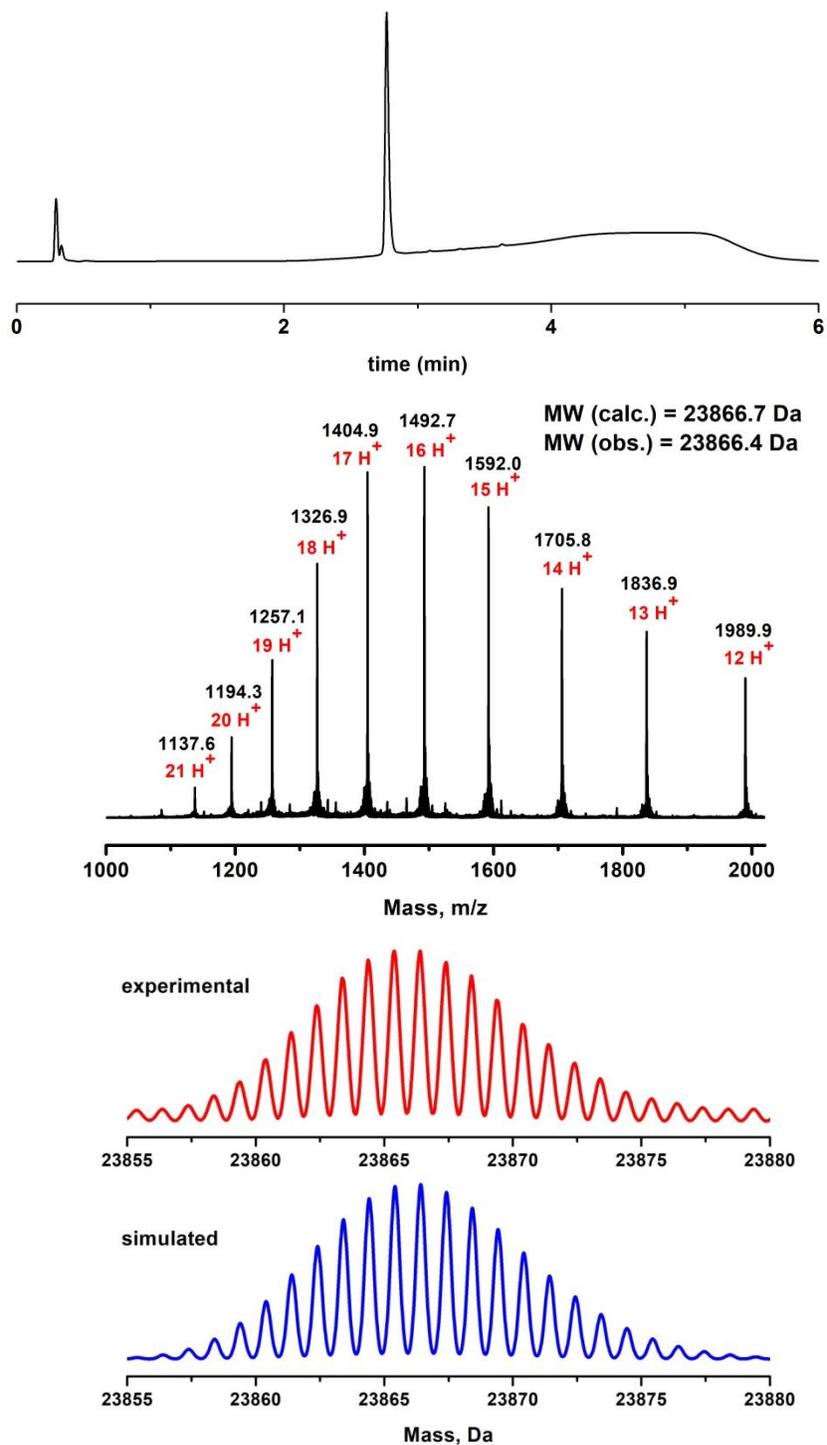
Several oligomers were prepared according to this strategy:

- First oligomer (called “trimer of trimers 1” or **TofT-1**) with the external trimers possessing one methyl group in β-arch. TofT-1 is the result of the ligation reactions between trimer-1-A, trimer-2-A and trimer-3-A;
- Second oligomer (called “trimer of trimers 2” or **TofT-2**) with the external trimers possessing two methyl groups in β-arch. TofT-2 is the result of the ligation reactions between trimer-1-B, trimer-2-A and trimer-3-B;
- Third oligomer with the external trimers possessing two methyl groups in β-arch and five <sup>15</sup>N-isotopically labelled amino acids in the central trimer (called “trimer of trimers 3” or **TofT-3**). TofT-3 is the result of the ligation reactions between trimer-1-B, trimer-2-B and trimer-3-B.

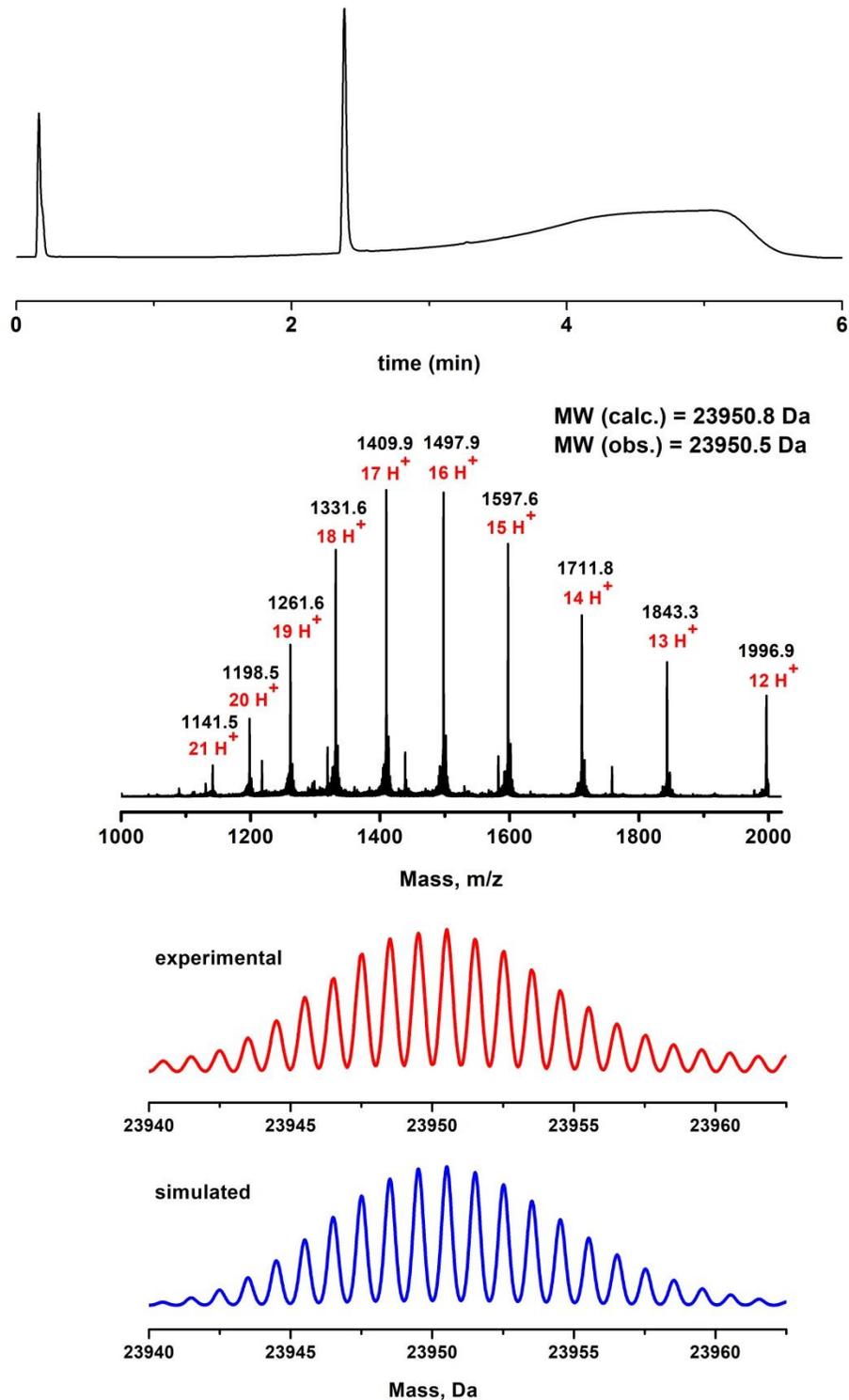


**Figure 44:** Successive native chemical ligations to form the covalently-tethered oligomer.

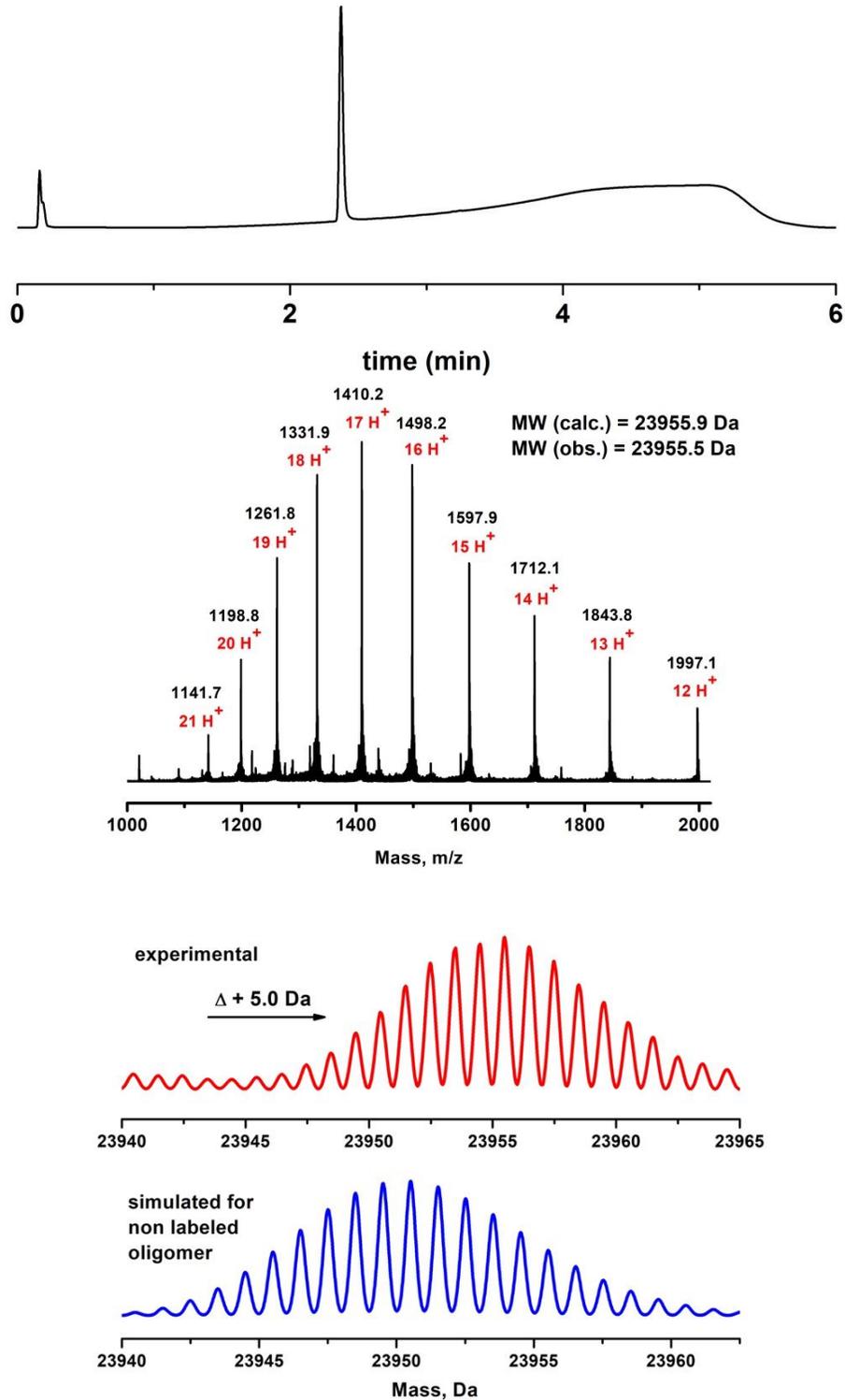
The three final constructs (TofT-1, TofT-2 and TofT-3) were analyzed by analytical HPLC and by electrospray ionization mass-spectrometry using Orbitrap mass-analyzer (ESI-MS-Orbitrap) (**Figures 45, 46 and 47**).



**Figure 45:** Characterization of ToFT-1. Top: Analytical HPLC ( $\lambda=220$  nm) trace; Middle: ESI-MS-Orbitrap mass spectrum; Bottom: Deconvolution of the experimental mass spectrum in red and theoretical simulation in blue.

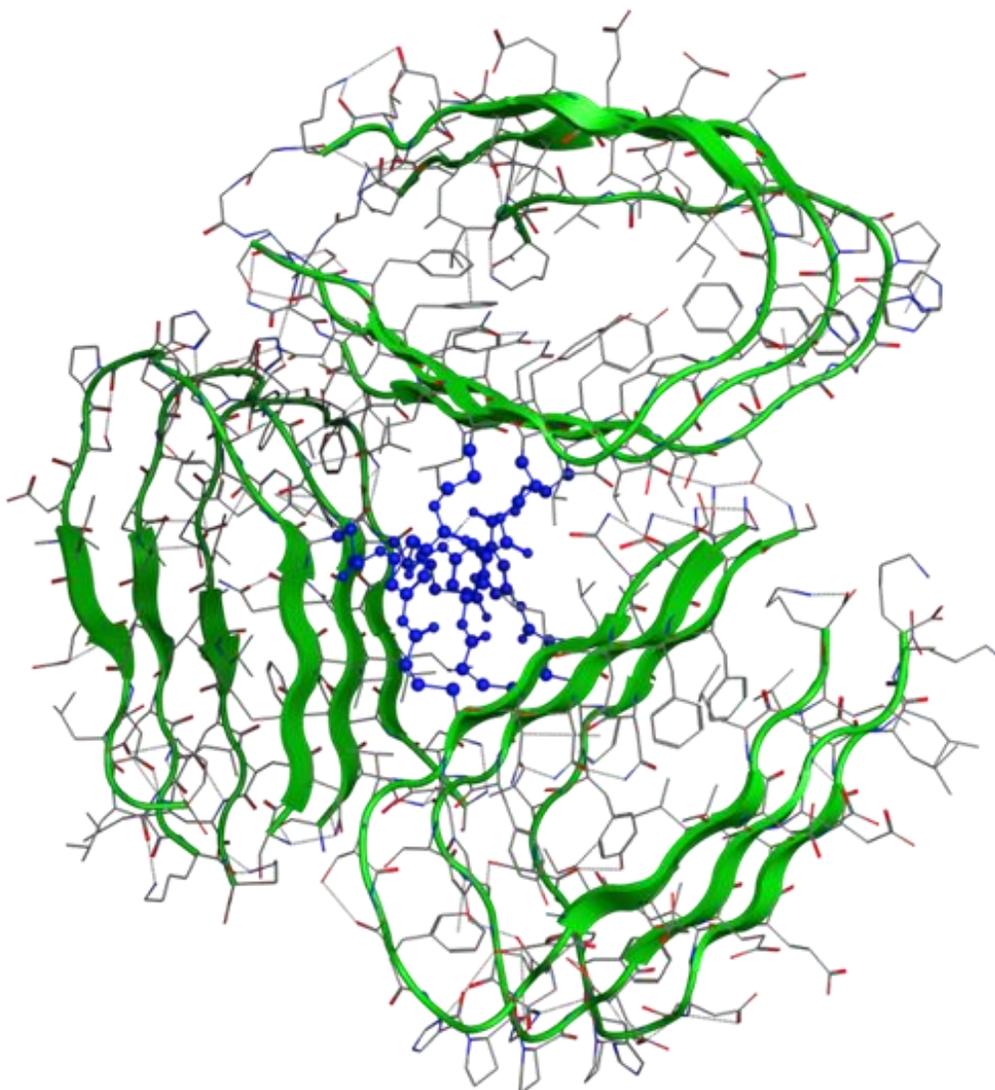


**Figure 46:** Characterization of ToFT-2. Top: Analytical HPLC ( $\lambda=220$  nm) trace; Middle: ESI-MS-Orbitrap mass spectrum; Bottom: Deconvolution of the experimental mass spectrum in red and theoretical simulation in blue.



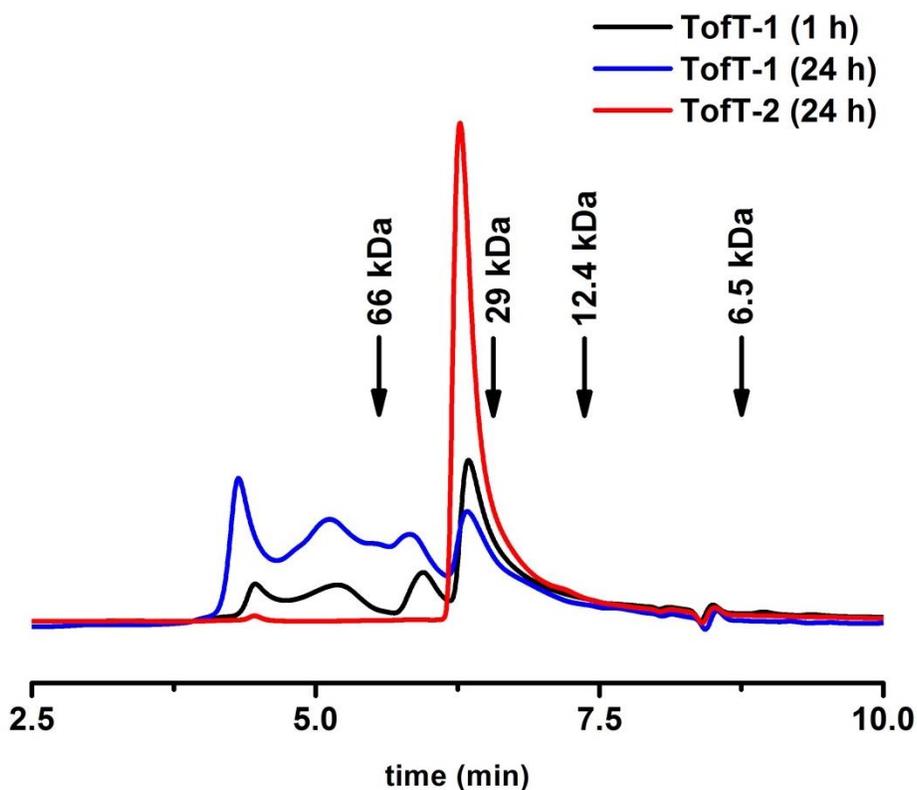
**Figure 47:** Characterization of ToFT-3. Top: Analytical HPLC ( $\lambda=220$  nm) trace; Middle: ESI-MS-Orbitrap mass spectrum; Bottom: Deconvolution of the experimental mass spectrum in red and theoretical simulation in blue.

In **Figure 48** a molecular model of **TofT-2** is shown. The modelling was performed using MOE software. The peptide backbones are represented in green and the central linkers in blue. Energies of the models were minimized by using the minimization force field AMBER10:EHT, with R-Field as a solvation mode. The model comprises nine copies of the [20-41] $\beta$ 2m fragment that remained in a  $\beta$ -arch conformation and possessed  $\beta$ -sheet secondary structures.



**Figure 48:** Molecular model of **TofT-2**.

## 4.8 Study of oligomerization by size-exclusion chromatography

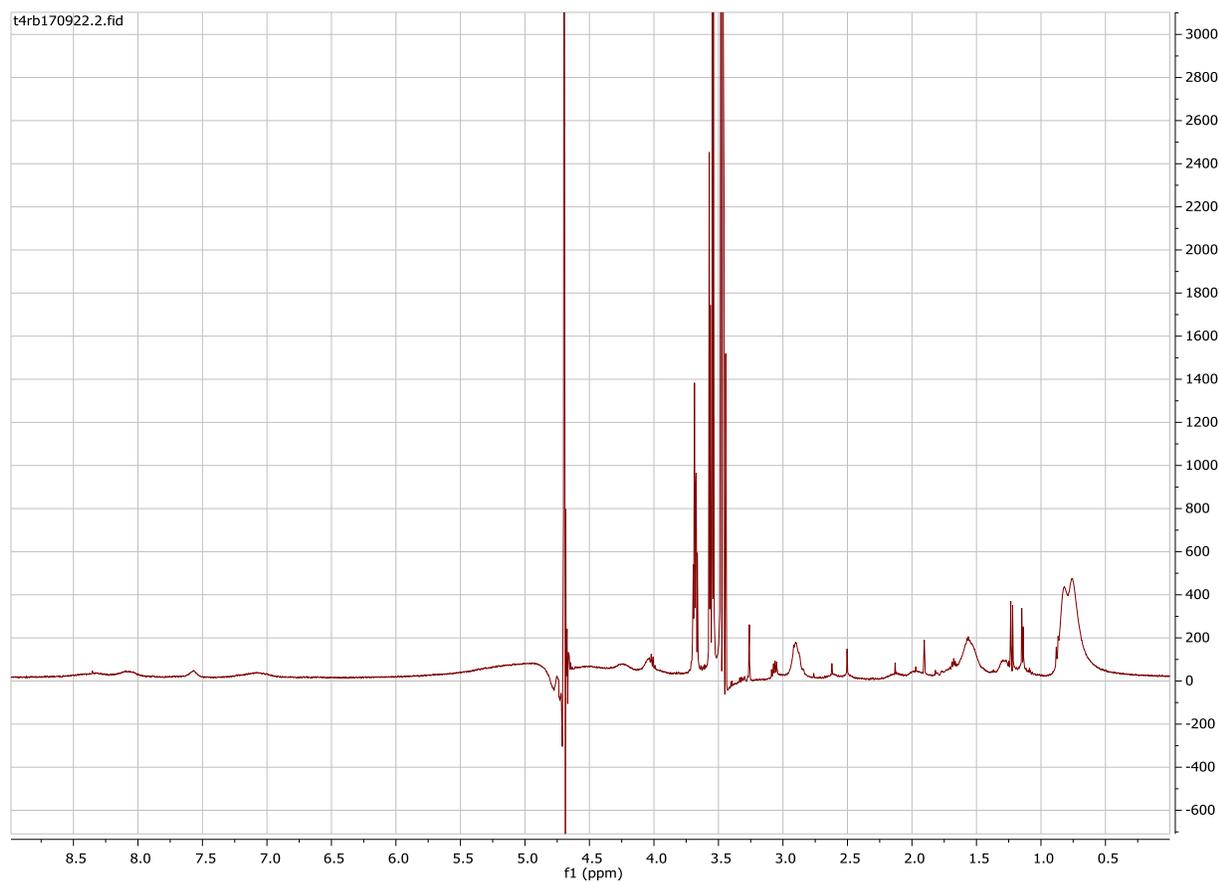


**Figure 49:** SEC chromatographic analysis of the covalently-tethered oligomers. Absorbance was monitored at 280 nm. The mobile phase was a buffer (pH 7.5) containing sodium phosphate (50 mM), sodium chloride (100 mM) and sodium azide (0.05%, w/v). Black arrows correspond to the retention times of the protein markers used for the calibration. The calculation of the molecular weights of the oligomers was based on the calibration curve.

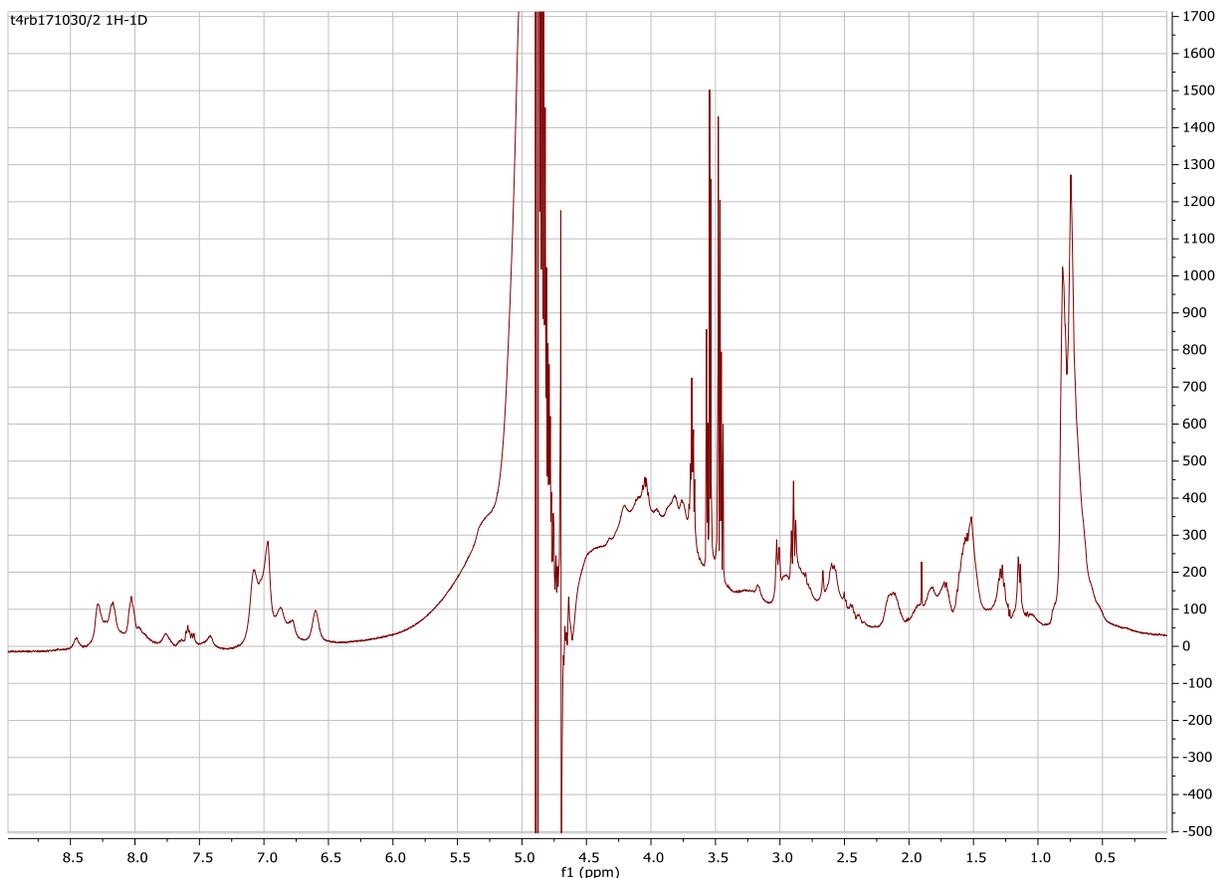
ToFT-1 (“trimer of trimer 1”), possessing one single methyl group in  $\beta$ -arch (six methyl groups in total) and (“trimer of trimer 2”), possessing two methyl groups in  $\beta$ -arch (twelve methyl groups in total) were analyzed by SEC in order to assess their aggregation properties. Samples with an oligomer concentration of 50  $\mu$ M in PBS buffer were prepared and injected after one hour and one day of incubation. The results clearly indicate that one methyl group in  $\beta$ -arch is not enough to prevent ToFT-1 from forming aggregates (**Figure 49**, black curve). The chromatogram indicates that the monomeric form of ToFT-1 is still present (peak around 6.6 min, calculated mass: ~31 kDa) but larger species are already present in the sample (eluting from 4.5 to 6.2 min

as broad peaks). The calculated mass of the monomeric form is higher than the theoretical, which can be due to the non-spherical shape of the oligomer tri-dimensional structure. After one day of incubation (**Figure 49**, blue curve), the percentage of larger species is higher, especially the species eluting around 4.4 min that corresponds to the void volume of the SEC column. This indicates that very large aggregates are formed and are not retained by the stationary phase.

Remarkably, TofT-2 elutes as a single peak after one hour and after one day (**Figure 49**, red curve) of incubation (peak around 6.6 min, calculated mass: ~31 kDa). Additional methyl group on each external peptide stopped the aggregation process. TofT-2 is highly homogeneous in size; no larger species are present in the sample.



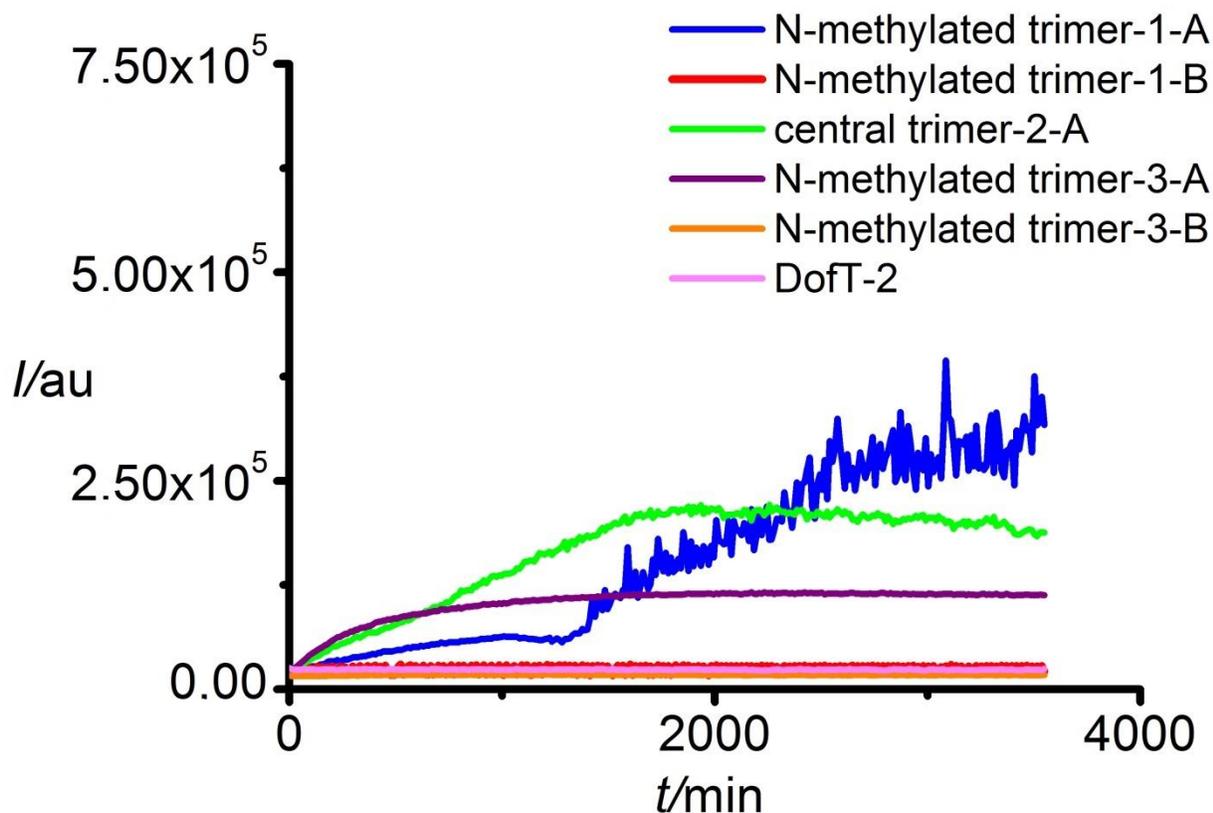
**Figure 50:** <sup>1</sup>H NMR of TofT-1 showing only broad peaks of peptide construct that indicates extensive aggregation.



**Figure 51:** <sup>1</sup>H NMR of ToFT-2 shows well populated peaks coming from peptide construct, which indicates smaller size and high solubility of the species.

ToFT-1 and ToFT-2 were also analyzed by <sup>1</sup>H NMR (**Figure 50** and **Figure 51**) at a concentration of 0.2 mM in buffer (pH 7.0) containing sodium phosphate (25 mM) and sodium azide (0.05%, w/v). No NMR signal of protein is present for ToFT-1 with only very broad and weak signals, which suggests that the major part of the protein may have aggregated and such species are not visible by solution NMR. However, ToFT-2 shows well detectable signals in the amide and aromatic region. The signals are still broad and might be due to the size of the construct. But most importantly, ToFT-2 is still in solution and these results are in accordance with the SEC experiments.

#### 4.9 Study and comparison of the amyloidogenicity of the different “covalent trimers”

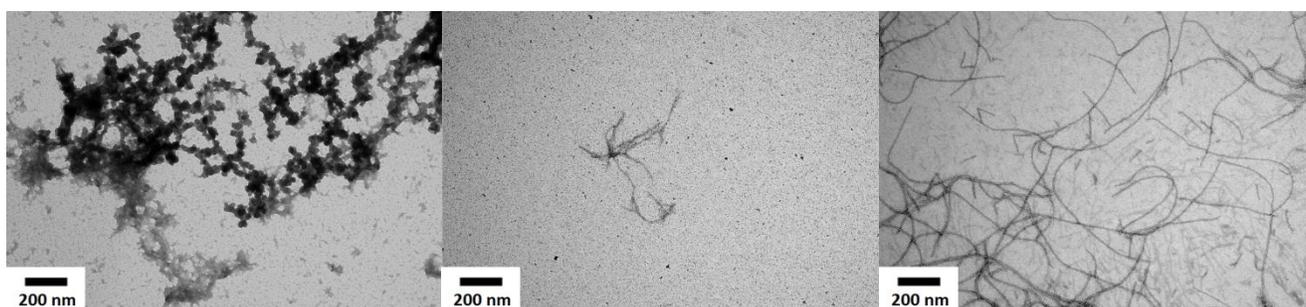


**Figure 52:** Kinetics of growth determined by ThT fluorescence using a plate reader.

Fluorescence of ThT was used to investigate the kinetics of fiber growth in a continuous format by using a plate reader (**Figure 52**). Data obtained showed distinct growth kinetics for the different constructs. As expected, the central trimer-2-A (non-methylated) shows an increase of the ThT fluorescence, which corresponds to unperturbed amyloid growth. The trimers with one methyl group in  $\beta$ -arch (*N*-methylated trimer-1-A and *N*-methylated trimer-3-A) also demonstrated an increase in ThT fluorescence. In accordance with the results obtained for ToFT-1, one methyl group is not sufficient to prevent the formation of aggregates or fibrils (alternatively, the location of this methylation within the peptide sequence is not adequate to

prevent interaction with other copies of the peptide). Trimers with two methyl groups in  $\beta$ -arch (*N*-methylated trimer-1-B and *N*-methylated trimer-3-B) do not show any increase of the ThT fluorescence. This is also the case for DofT-2 (“covalent dimer of covalent trimers”), which is the ligation product between *N*-methylated trimer-1-B and central trimer-2-A. The ligation between a non-methylated “covalent trimer” and a *N*-methylated “covalent trimer” (with two methyl groups in  $\beta$ -arch) seems to suppress aggregation process.

We correlated those data with TEM measurements of the central trimer-2-A, the *N*-methylated trimer-1-A and trimer-3-A (**Figure 53**), which showed an increase of the ThT fluorescence intensity. TEM experiments for the *N*-methylated trimer-1-A revealed not typical fibrils but instead very large globular shaped aggregates that seem to bind ThT to a similar extent as more typical amyloids. For *N*-methylated trimer-3-A, we observed more canonical amyloid fibrils. For central trimer-2-A, possessing no *N*-methyl group, very long amyloid fibrils were observed, morphologically similar to “covalent trimer” [20-41] $\beta$ 2m amyloids described in Chapter 3.

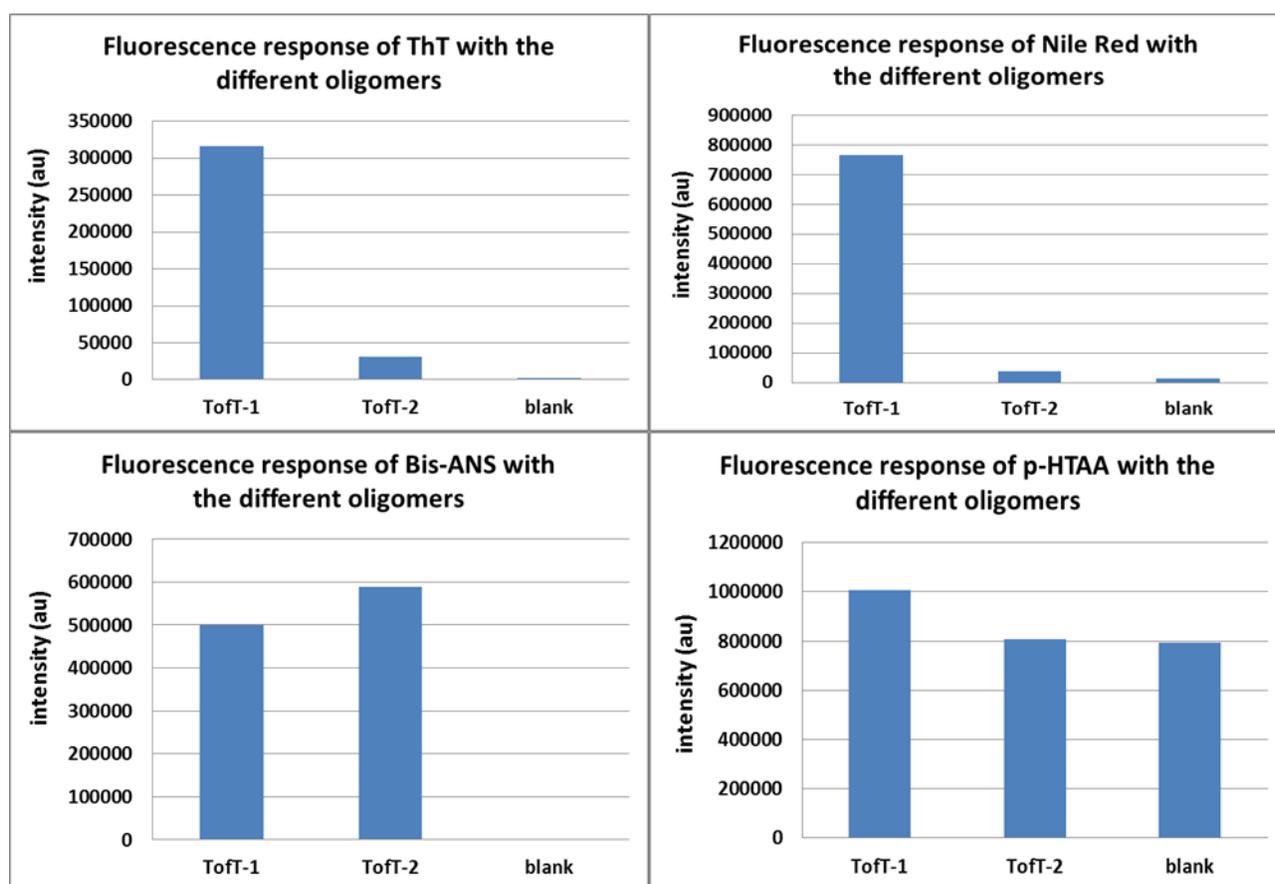


**Figure 53:** TEM images for aggregates of *N*-methylated trimer-1-A (left), amyloids of *N*-methylated and trimer-3-A (center) and amyloids of central trimer-2-A (right).

#### **4.10 Study and comparison of the amyloidogenicity of the covalently-tethered constructs**

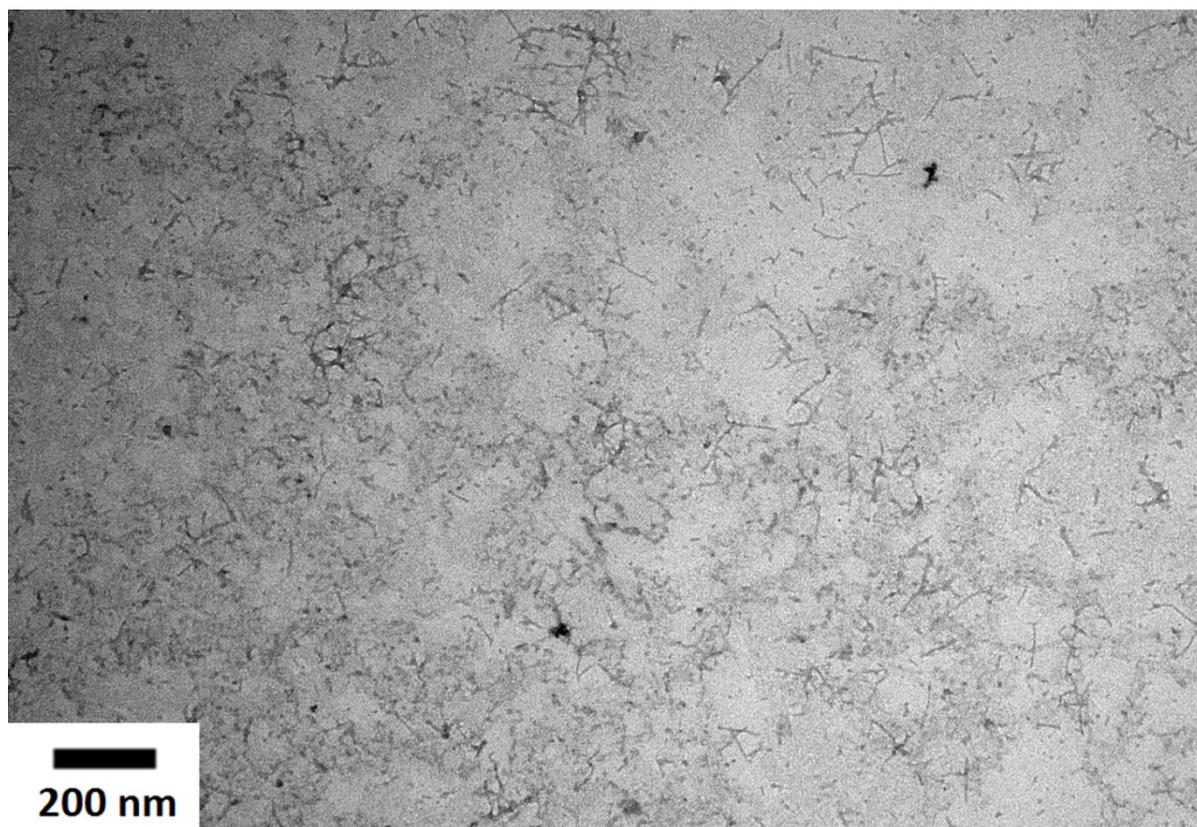
The final constructs (TofT-1 and TofT-2) were tested for their ability to bind fluorescent probes commonly used for detection of either amyloid fibrils and/or hydrophobic pockets in

proteins. We tested dyes such as ThT, Nile Red, bis-ANS and the sodium salt of pentamer hydrogen thiophene acetic acid (p-HTAA), which is used for the identification of amyloid aggregates but also for transient pre-fibrillar assemblies<sup>159</sup>. Dyes were added to solutions of covalently-tethered oligomers (30  $\mu\text{M}$ ) at concentrations of 90  $\mu\text{M}$  giving a fluorescent dye concentration ratio of 3:1 (dye:oligomer). Three measurements with each dye and oligomer were made and average values were calculated. TofT-1 gives a very strong ThT and Nile Red response compared to TofT-2 (**Figure 54**). However, with bis-ANS, both oligomers give a strong fluorescence response and it is even higher for TofT-2. Bis-ANS is indeed an efficient probe for oligomer and protofibril assemblies and is known to bind to hydrophobic patches in soluble proteins<sup>6</sup>. Therefore, bis-ANS may be revealing the expected hydrophobic core inside the structure of especially TofT-2. For p-HTAA, we detected higher response for TofT-1 than for TofT-2.



**Figure 54:** Fluorescence measurements of several dyes interacting with TofT-1 and TofT-2.

TEM measurements were done on both oligomers (TofT-1 and TofT-2). TofT-1 was found to form short, non-homogeneous amyloid fibers (**Figure 55**), which is consistent with the high response of ThT and Nile Red towards TofT-1 samples. On the other hand, no aggregates were observed for TofT-2.

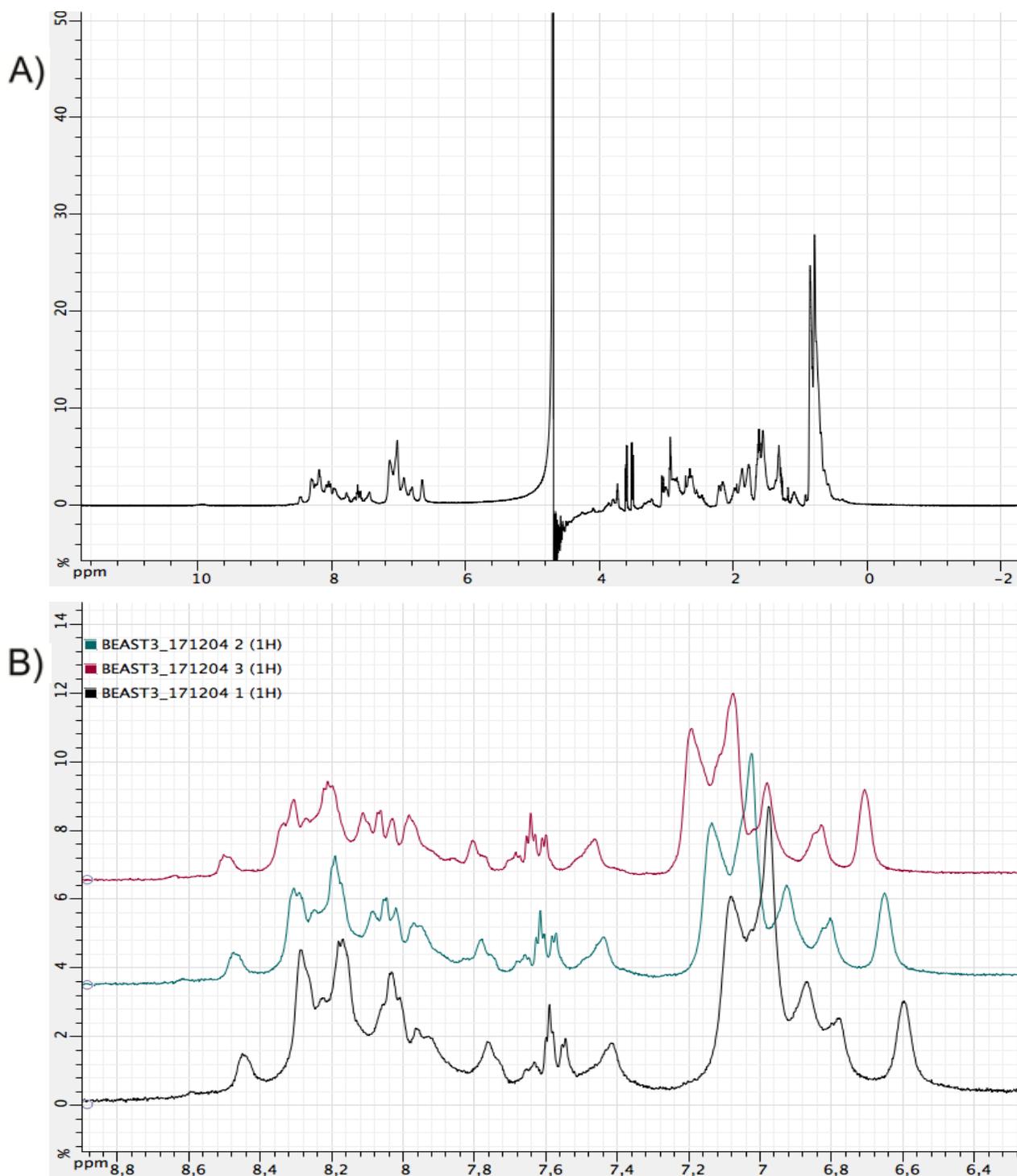


**Figure 55:** TEM image of TofT-1 fibrils.

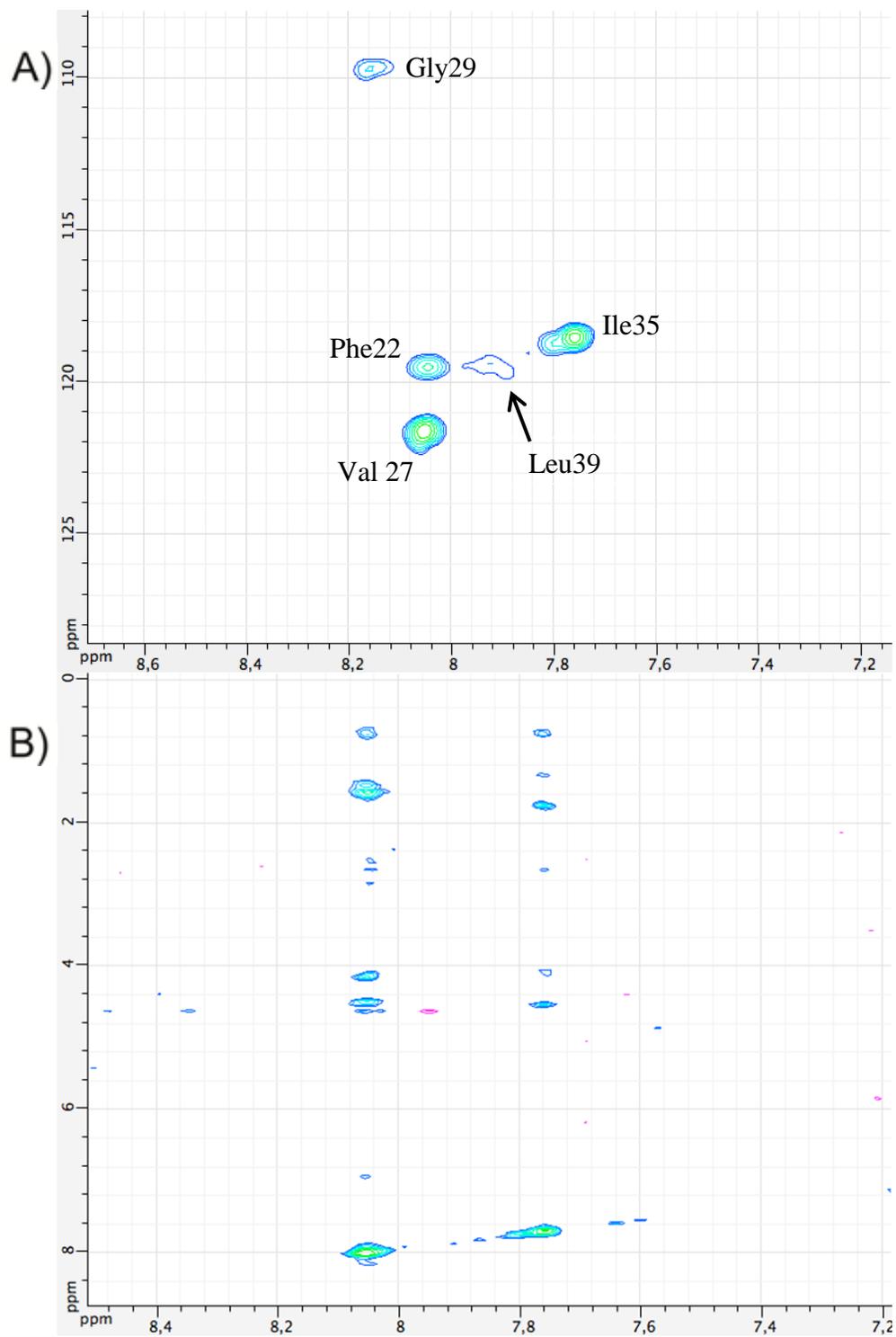
#### **4.11 Study of the $^{15}\text{N}$ -labeled “trimer of trimers” by solution NMR**

TofT-3 (which is an  $^{15}\text{N}$ -labelled version of TofT-2) was also analyzed by NMR at a concentration of 0.2 mM in buffer (pH 7.0) containing sodium phosphate (25 mM). The spectra were recorded on a Avance III 700 MHz spectrometer equipped with a TCI cryo-probe. The proton spectrum of TofT-3 displays interesting features (**Figure 56-A**), notably a well dispersed region of amide resonances as previously observed for TofT-2. Moreover, the temperature dependence of proton spectrum was recorded at 298 K, 303 K and 308 K (**Figure 56-B**). The

dispersion of amide resonances improves at 308 K as well as the line-width. Subsequent spectra were recorded at 308 K.

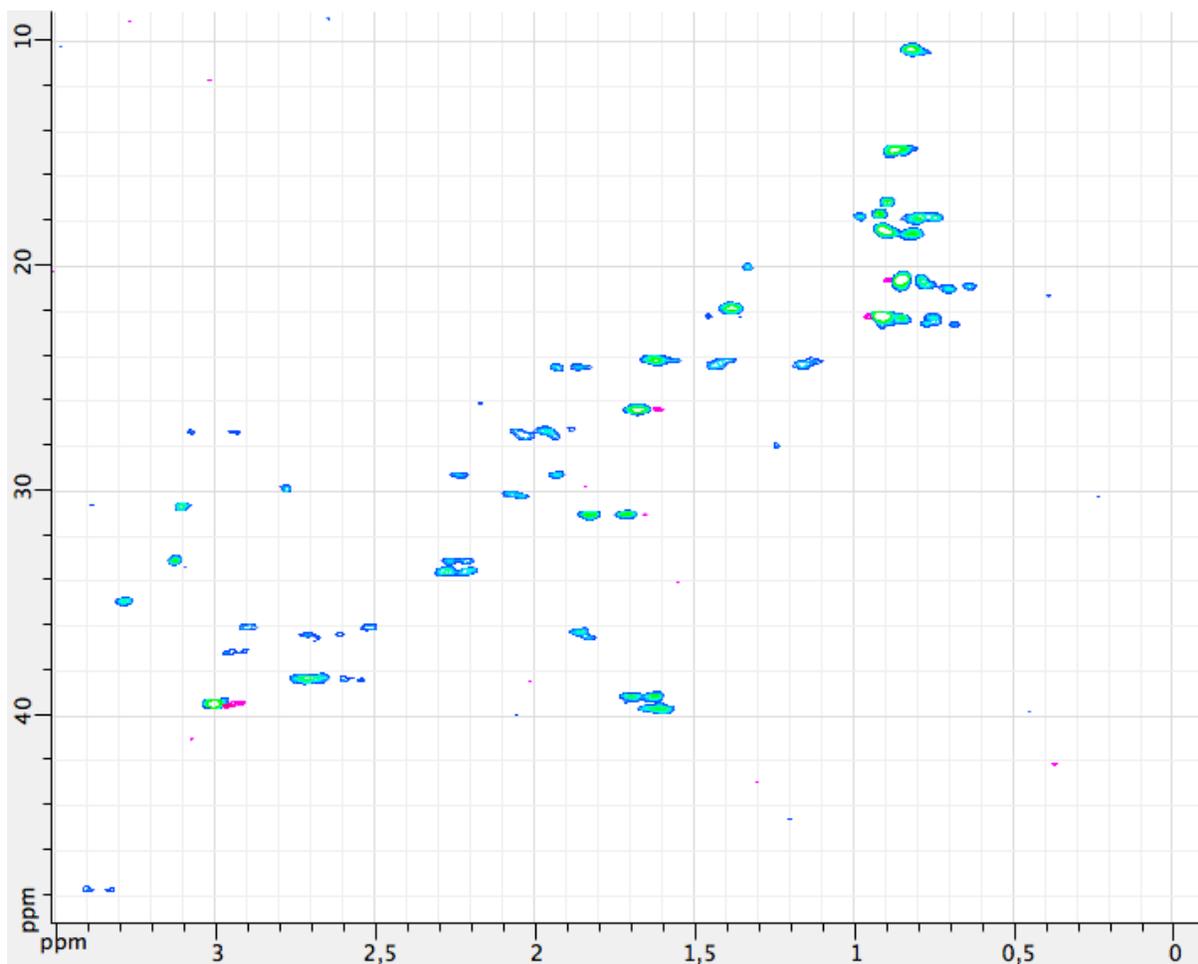


**Figure 56:** A)  $^1\text{H}$  NMR of ToFT-3, and B) temperature dependence of proton spectrum. NMR spectra of ToFT-3 were recorded at 298 K (black), 303 K (blue) and 308 K (red).



**Figure 57:** A)  $^1\text{H}$ - $^{15}\text{N}$  HSQC NMR spectrum of TofT-3, and B) 2D projection of the  $^{15}\text{N}$ -edited NOESY spectrum.

A  $^1\text{H}$ - $^{15}\text{N}$  HSQC spectrum was recorded for 4 h (NS 200 TD 128 x 2048 pts, see **Figure 57-A**). The spectrum displays clear four peaks. Two of those peaks are of similar and high intensity at 122.5 ppm and 118.5 ppm, a weaker one is at 119.5 ppm. The cross-peak corresponding to the glycine at  $^{15}\text{N}$  frequency 109.7 ppm could be identified, but its intensity is weak. In order to assign the three observed  $^1\text{H}$ - $^{15}\text{N}$  correlations, a 2D projection of the  $^{15}\text{N}$ -edited NOESY spectrum with 150 ms mixing time was recorded (**Figure 57-B**). Tentative assignment is depicted in **Figure 57-A**, where Leu39 is the broadest peak. A  $^1\text{H}$ - $^{13}\text{C}$  HSQC spectrum was also recorded and the spectrum displays a large heterogeneity of peak intensity suggesting the presence of well-structured residues and residues that undergo conformational exchange (**Figure 58**).



**Figure 58:**  $^1\text{H}$ - $^{13}\text{C}$  HSQC NMR spectrum of ToT-3.

## 4.12 Discussion

Designing small oligomers composed of amyloidogenic peptide segments mimicking transient species during amyloid formation has a strong potential to provide better understanding of the structural and functional properties of such metastable constructs. Having more insights about the different processes involved in amyloid formation is a key element to find new treatments for many neurodegenerative disorders such as Alzheimer's disease or Parkinson's disease.

In this study, we report on the design and total chemical synthesis of covalently-tethered oligomers composed of nine copies of the amyloidogenic [20-41]-fragment of human  $\beta$ 2-microglobuline ([20-41] $\beta$ 2m). The synthetic strategy employed to produce the different oligomers is highly innovative. Large peptide constructs with molecular weight of nearly 24 kDa were synthesized. They possess non-linear topologies and are composed of non-symmetric covalently-tethered architectures.

The use of Fmoc-SPPS and different orthogonal chemistries allows for efficient synthesis of different variants of the [20-41] $\beta$ 2m fragment possessing various chemical functionalities and the conjugation with the abiotic linker. The synthesis of non-symmetric "covalent trimers" was achieved by using the selective reaction of thiols with alkyl bromides and sequential native chemical ligations allowed for the tethering of the "covalent trimers" in an efficient and unambiguous way. The strategy is robust and flexible, two different oligomers with different *N*-methylation patterns and one isotopically-labeled oligomer were synthesized using this methodology.

The study of the oligomeric state of TofT-1 and TofT-2 revealed that the number of *N*-methylation on the external "covalent trimers" is crucial to prevent the structure from self-assembly into larger aggregates. TofT-1 (with one methyl group per peptide) forms aggregates as shown by SEC and NMR measurements. TEM also revealed the formation of small amyloid-like fibrils. In the case of TofT-2, a different *N*-methylation strategy with one more methyl group per peptide stopped the aggregation process. In solution, TofT-2 is monomeric and the  $^1\text{H}$  NMR spectrum shows a well dispersed region of amide resonances.  $^1\text{H}$ - $^{15}\text{N}$  HSQC measurement of TofT-3 ( $^{15}\text{N}$ -isotopically labeled TofT-2) indicates that some regions of the structure are well

defined and homogeneous, whereas others are more flexible and prone to conformational exchange. This is in remarkable agreement with data obtained using solid-state NMR for [20-41] $\beta$ 2m covalent trimer amyloids (see Chapter 3). In addition, we demonstrated that ToFT-1 and ToFT-2 are both giving fluorescence response when incubated with amyloid probes such as ThT, Nile-Red, Bis-ANS or p-HTAA. The ThT and Nile Red fluorescence intensity is higher when exposed to ToFT-1 as it is forming amyloid aggregates. ToFT-2 is responding well to Bis-ANS, probably because ToFT-2 possesses an accessible hydrophobic pocket.

The next step in the study of ToFT-2 will involve deeper NMR characterization of the construct, notably by adding different labels within the structure. Furthermore, X-ray crystallographic studies will be attempted. Moreover, the cytotoxicity of the different constructs will be evaluated *in vivo*. Based on the robustness of the methodology, larger constructs could be prepared such as “covalent tetramer of trimers” or “covalent pentamer of trimers”. A future direction in this project is to validate this concept with other amyloidogenic peptides, in particular, Alzheimer’s-related A $\beta$  peptide.

### **4.13 Conclusions**

The most innovate aspect of this work is the implementation of a new methodology to prepare large and soluble peptidic constructs that could, for instance, mimic structures of insoluble amyloid fibrils. This new approach will allow for the preparation of compositionally well-defined, homogeneous and monomeric species amenable for advanced characterization with analytical methods such as solution NMR or X-ray crystallography, and for biological assays. This strategy can be applied for the synthesis of other oligomers composed of disease-related peptides such as A $\beta$  peptide. Furthermore, this methodology will allow for screening of methylation patterns (or other modifications) to identify a combination (such as the number and position of the methylations) that are most efficient to prevent aggregation. Overall, we predict that this method will be very useful for the development of new structure-based amyloid inhibitors and diagnostics.

## 5. Towards molecular recognition and selective inhibition of distinct amyloid polymorphs

### 5.1 Introduction

Amyloid aggregates and the intermediates leading to the formation of amyloid fibrils are associated with many diseases. Inhibiting amyloid formation and the related transient species involved in amyloidosis appears to be a possible therapeutic strategy to prevent the progression of a disease caused by the presence and the accumulation of protein aggregates<sup>14,38,160</sup>.

One possible strategy consists in preventing the amyloid-related peptide or protein to adopt the misfolded structure that triggers the formation of amyloidogenic species into amyloids<sup>161</sup>. The therapeutic agents would bind to the target in its native folded state and stabilize the structure and prevent misfolding and therefore, protein aggregation<sup>161</sup>. This strategy was notably employed for the stabilization of transthyretin (TTR) familial amyloid polyneuropathy (FAP) or senile systemic amyloidosis. TTR exists both in its monomeric or tetrameric form<sup>162</sup>. TTR monomer was found to be prone to misfolding and aggregation into amyloid fibers. Moreover, mutation of TTR (associated with FAP) destabilizes the tetrameric form and accelerates amyloid formation. A compound *tafamidis meglumine* was designed based on the structure of TTR tetramer and was found to be highly effective as a kinetic stabilizer of the tetramer. It was approved for the use in the European Union as a TTR dissociation inhibitor for the treatment of FAP becoming the first medicine interfering with amyloid deposition<sup>163,164</sup>. The same approach was investigated for the stabilization of superoxide dismutase-type 1 (SOD1) mutant involved in familial amyotrophic lateral sclerosis (FALS)<sup>165</sup>. Dissociation of SOD1 dimer appears to promote misfolding and aggregation *in vitro*. *In silico* screening of small molecules allowed the discovery of several hits for further promising drug development<sup>166</sup>. Nevertheless, this strategy relies on the fact that the ligand can recognize the precursor of amyloid aggregation. It is highly dependent on the target structure. Identifying such ligands can be very challenging if the protein is partially folded or intrinsically disordered, such as  $\alpha$ -synuclein related to Parkinson's disease. In this case, it is the  $\alpha$ -synuclein protofibrils that were stabilized preventing the formation of amyloid fibrils using

catecholamines as ligands<sup>167</sup>. Later on, it was shown that natively helically folded  $\alpha$ -synuclein tetramers were not prone to fibrillation and, therefore, could be of great interest for the development of kinetic stabilizers as a new therapeutic agent against Parkinson's disease<sup>168</sup>.

Antibodies fused with the [33-42] or [18-21] fragment of A $\beta$  involved in Alzheimer's disease were prepared to inhibit amyloid formation<sup>169</sup>. In these constructs, the A $\beta$  fragments served as recognition motifs and antibodies functioned to prevent docking of the additional A $\beta$  subunits. The conjugated antibody was shown to prevent polymerization into amyloid fibrils by favoring the formation of small unstructured protein complexes with no cytotoxic properties<sup>169</sup>. Using a screening approach of libraries, Aducanumab<sup>170</sup>, a human immunoglobulin G1 antibody was found to selectively react with soluble A $\beta$  oligomers or A $\beta$  amyloid fibrils and reduce the amount of senile plaques in mouse brains. Aducanumab is currently in phase 3 clinical trials.

An approach to prevent amyloid formation is to trap or sequester the monomeric or the protofibrillary form of the misfolded peptide or protein to prevent subunit recruitment and polymerization. Contrary to kinetic stabilization of natively folded monomer with small molecule ligand binding, this approach relies on the specific recognition of aggregation-prone region of the misfolded polypeptide. Small to medium sized peptide-based inhibitors of amyloid propagation have become an important field of research<sup>171</sup>. The aim is to engineer peptide sequence that would bind the amyloidogenic partner (monomer, oligomer, protofibril) and act as a  $\beta$ -sheet breaker preventing other partners to perpetuate the self-assembly. Some of the first rational designs of peptide-based inhibitors are based on the target sequence and more precisely on the internal amyloidogenic segments of the target. For instance the [16-20]A $\beta$  segment (KLVFF) showed inhibitory effect on the formation of A $\beta$ (1-40) but was itself prone to aggregation into fibrils that were distinct from amyloids<sup>172</sup>. So, there is a need to combine the inhibitor peptide sequence (recognition motif) with additional modifications to prevent self-aggregation<sup>173</sup>. Many strategies were investigated:

- addition of charged residues at the extremity of the recognition motif such as polylysine (disrupting element) that would permit recognition of the target, enhanced solubility of the inhibitor and self-repulsion<sup>174</sup>;
- incorporation of  $\beta$ -sheet breaker amino acids such as proline<sup>175</sup> or  $\alpha,\alpha$ -disubstituted residues (e.g.  $\alpha$ -aminoisobutyric acid (Aib)<sup>42</sup>);

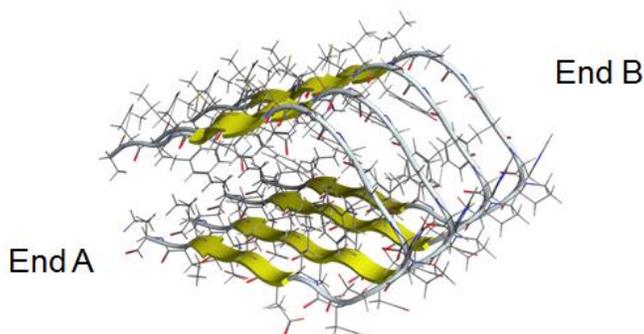
- utilization of *D*-peptides<sup>176,177</sup> able to recognize the amyloidogenic interface;
- modification of the peptide backbone by replacing amides by ester bonds leading to a weaker hydrogen bond acceptor (carbonyl) and the disappearance of a hydrogen bond donor while keeping structural similarities<sup>40</sup>;
- *N*-alkylation of the peptide backbone eliminating hydrogen-bonding interaction and creating steric hindrance that would prevent the approach of amyloidogenic subunits<sup>178,179</sup>. This method was employed for the design of peptide inhibitor of human islet amyloid polypeptide (IAPP) aggregation associated with type II diabetes<sup>178</sup> and  $\alpha$ -synuclein aggregation<sup>179</sup>.

In this study, we designed *N*-methylated peptide constructs based on the structure of the segment [20-41] of human  $\beta$ 2-microglobulin ( $\beta$ 2m). We synthesized monomers, covalent dimers and covalent trimers of [20-41] $\beta$ 2m with different *N*-methylation patterns; that is, with different numbers of *N*-methylations per peptide and at different positions within the polypeptide sequence. We decided to study *N*-methylated analogues of [20-41] $\beta$ 2m because this peptide backbone modification is highly effective to disrupt  $\beta$ -sheets and *N*-methylated peptides are known to be efficient inhibitors of amyloid growth<sup>178,179</sup> and possesses many advantageous chemical and bio-physical properties<sup>180</sup>. *N*-methylations have considerable enhancing effects on the solubility of the peptides in aqueous media and organic solvents. They are less prone to enzymatic degradation and thus present enhanced metabolic stability. Moreover, they are more membrane permeable notably regarding the transfer from the intestine to the circulatory system which makes them orally bioavailable drugs candidates<sup>180</sup>.

The chemical design of such constructs was made to better understand the impact of *N*-methylation on the structure, their physical and chemical properties and their inhibitory power against amyloid fibrillation. Designing *N*-methylated peptidic constructs with one-fold, two-fold or three-fold symmetries can permit the development of molecular probes that would selectively recognize and bind to amyloid fibrils with one-fold, two-fold or three-fold symmetries and allow for effective identification of different amyloid polymorphs in biological samples.

## 5.2 Design of the *N*-methylated peptidic constructs

Amyloid fibers possess a polar structure. In this case, polarity is related to the directionality of the fiber structure. Therefore, protofilaments do not present the same interface at both ends of the fibril (**Figure 59**). Potentially, both ends of the protofibrils or fibrils can recruit sub-units (bi-directional polymerization) and they may expand at different rates. This assumption implies that in order to fully stop the elongation of amyloid fibers, the recruitment of subunits has to be perturbed, and therefore stopped, at both ends of the growing aggregate. It means that different peptide constructs with different methylation patterns pointing at opposite directions need to be envisaged.



**Figure 59:** Molecular model of the [20-41]β2m protofibril based on the structure obtained by solid-state NMR (PDB ID 2E8D). The β-sheets are represented by the yellow arrows. Both terminal ends (end A and end B) of the structure present different interfaces with non-identical exposed backbone amide groups.

### 5.2.1 Description of the different *N*-methylated variants

Based on the structure of the [20-41]β2m segment in a protofibril, we decided to replace certain amino acids by *N*-methylated analogues. It is important to reiterate that the substitutions sites need to be in the β-strands of the peptides. According to the structure and what we learnt from the solid-state NMR and FT-IR experiments performed on the covalently tethered trimeric



The *N*-methylation-free interface of each peptide is able to recognize the target and can dock to the growing ends of protofibrils or fibrils and the *N*-methylated interface can prevent binding of other amyloidogenic peptides via disruption of hydrogen-bonding necessary for  $\beta$ -sheet formation. Combining two complementary *N*-methylated variants that display methyl groups on different interfaces could thus result in a better inhibition of amyloid growth by docking at both growing ends (if the growth is bi-directional).

### 5.3 Synthesis of the *N*-methylated peptides

The synthesis of all different variants was realized using Fmoc-SPPS either manually, or by machine-assisted synthesis or by using microwave automated synthesizer. The synthesis of *N*-methylated peptides on solid-support is often a challenging task. Coupling Fmoc-*N*-Me-amino acids is usually accomplished without any difficulty. But coupling of the next amino acid to the *N*-methylated residue on the growing peptide chain is rather difficult because of the steric hindrance caused by the methyl group and can be even more challenging if the next amino acid is  $\beta$ -branched (valine, isoleucine or threonine), due to steric hindrance. The kinetics of the coupling reaction is therefore slower and can lead to racemization reactions or to peptides containing deletions. This is why very efficient coupling reagents need to be used for the synthesis of *N*-methylated peptides such as TBTU (2-(1*H*-Benzotriazole-1-yl)-1,1,3,3-tetramethylammonium tetrafluoroborate), HOAt (1-Hydroxy-7-azabenzotriazole) with PyBOP (benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate) or HATU (1-[bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate) for instance.

#### 5.3.1 General procedure used for the synthesis of *N*-methyl peptides by Fmoc-SPPS

*Swelling*: The resin was rinsed twice with DMF. The resin was covered by DMF and was let to swell for 30 min.

Deprotection: The resin was rinsed twice with 20% piperidine in DMF. The resin was treated by the same solution for 20 min under gentle manual stirring. The resin was rinsed thoroughly with DMF.

Coupling of Fmoc-L-amino acids to non- $N^\alpha$ -methylated amino acids: Fmoc-protected amino acids (4.0 equiv.) were dissolved in 0.38 M HATU in DMF (3.8 equiv). Pre-activation was performed for 2 min by addition of DIPEA (5.7 equiv). The reaction mixture was poured onto the resin and allowed to react for 30 min under gentle manual stirring. The resin was rinsed thoroughly with DMF.

Coupling of Fmoc-L-amino acids to  $N^\alpha$ -methylated amino acids: Double couplings were performed using the same protocol as for coupling to non- $N$ -methylated amino acids. Acetylation of unreacted  $N$ -terminal amine functions was done using of acetic anhydride (10 equiv) and DIPEA (5.7 equiv) in DMF for 6 min. The resin was rinsed thoroughly with DMF. Most of the time, the Chloranil test after the coupling reaction revealed the presence of secondary amines on the resin, the capping step was therefore necessary.

Coupling of Fmoc-L- $N^\alpha$ -methylated amino acids: Fmoc-protected  $N^\alpha$ -methylated amino acids (2.0 equiv.) were dissolved in 0.38 M HATU in DMF (1.9 equiv). Pre-activation was performed for 2 min by addition of DIPEA (2.9 equiv). The reaction mixture was poured onto the resin and allowed to react for 60 min under gentle manual stirring. The resin was rinsed thoroughly with DMF. The number of equivalents of  $N^\alpha$ -methylated amino acids was divided by two and the coupling time was then doubled (in order to minimize the consumption of expensive building blocks).

Kaiser test: For the detection of unreacted  $N$ -terminal solid-phase bound primary amines, Kaiser tests were performed using standard protocol. A few beads were transferred into a glass test tube and washed with ethanol. 100  $\mu$ L of a solution of 50 mg/mL ninhydrin in ethanol and 100  $\mu$ L of a solution of 1 mM KCN<sub>(aq)</sub> (0.2 mL) in pyridine (10 mL) were added. The tube was placed in a 115°C heating block for 5 min. 1 mL of ethanol and 1 mL of water were then added. Positive test: bleu/purple color, negative test: yellow color.

Chloranil test: For the detection of unreacted  $N$ -terminal solid-phase bound secondary amines, chloranil tests were performed using standard protocol. A few beads were transferred into a glass

test tube. 100  $\mu$ L of a solution of 2% acetaldehyde (v/v) in DMF and 100  $\mu$ L of a solution of 20 mg/mL *p*-chloranil in DMF were added. It was let to react for 5 min. Positive test: bleu/green color, negative test: yellow color.

Cleavage: Prior to cleavage, the resin was washed several times with DMF, then with dichloromethane and finally dried. Peptides were cleaved and fully deprotected by treatment with the mixture TFA/TIPS/H<sub>2</sub>O (95/2.5/2.5, v/v) for 2 h (1 mL of cleavage cocktail for 100 mg of peptidyl-resin). The reaction mixtures were diluted 20-fold with -20°C diethyl ether leading to the precipitation of the crude peptides. Centrifugation was performed twice at 4500 rpm for 3 min. The supernatants were discarded and the pellets were dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN/TFA (1/1/0.001 v/v/v) and lyophilized.

Purification: The crude peptides were solubilized in 6 M guanidinium chloride solution containing 0.1% TFA and were purified by preparative RP-HPLC. The pure fractions of peptide were combined and lyophilized.

**Table 6:** Final yields of the *N*-methylated variants of [20-41] $\beta$ 2m prepared using the general Fmoc-SPPS procedure.

Name of the peptide (monomers)	Scale of the synthesis (mmol)	Isolated mass of purified peptide (mg)	Yield (%)
Inh-1	0.2	203	40
Inh-2	0.2	157	31
Inh-3	0.1	49	19
Inh-4	0.1	72	28
Inh-5	0.1	73	29
Inh-6	0.1	34	13
Inh-7	0.1	n/a	n/a
Inh-8	0.1	64	25
Inh-9	0.1	48	19

In **Table 6** are summarized all the yields obtained for the synthesis of the *N*-methylated peptides. The synthesis of Inh-7 was not successful, the coupling of the *N*-methylated glutamic acid 36 to valine 37 was very difficult and the elongation of the peptide stopped at serine 28.

### 5.3.2 General procedure for the synthesis of covalent dimers

The synthesis of the different *N*-methylated covalent dimers was realized using the same protocol as previously described for the synthesis of [20-41] $\beta$ 2m covalent dimer in chapter 3. The synthesis of disulfide-linked *N*-methylated-[20-41] $\beta$ 2m homodimers was carried out by air oxidation. The peptides were linked through the side chain of the cysteine residue at position 25. The monomer of each individual *N*-methylated variants were dissolved in a pH 9 phosphate buffer containing 6 M of guanidinium chloride as a chaotrope and the peptide concentration was around 1 mg/mL. The reactions were monitored by analytical HPLC and LC-MS and the formation of the dimeric structures was accomplished after 4 to 5 days. The yields of the synthesis are summarized in **Table 7**.

**Table 7:** Final yields of the *N*-methylated variants of [20-41] $\beta$ 2m homodimers.

Name of the peptide (monomers)	Amount of starting material ( $\mu$ mol)	Isolated mass of purified homodimer (mg)	Yield (%)
Inh-1-dimer	6.1	10.5	67
Inh-2-dimer	6.2	7.5	49
Inh-4-dimer	2.0	2.8	55
Inh-5-dimer	2.0	2.7	53
Inh-6-dimer	2.0	1.5	30
Inh-8-dimer	2.0	1.2	25
Inh-9-dimer	2.0	1.3	28

### 5.3.3 General procedure for the synthesis of covalent trimers

The synthesis of the different *N*-methylated covalent trimers was realized using the same protocol as previously described for the synthesis of [20-41] $\beta$ 2m covalent trimer in chapter 3. The tethering was performed by reacting the thiol group of the cysteine 25 of the different variants with the trivalent *N,N',N''*-(benzene-1,3,5-triyl)tris(acetamide) linker (**Figure 26** on page 77). The reactions took place in a mixture of water and acetonitrile at pH 8 in the presence of ammonium hydrocarbonate at 30°C. The reactions were monitored by analytical HPLC and LC-MS and the formation of the trimeric structures was accomplished after 2 h. The yields of the synthesis are summarized in **Table 8**.

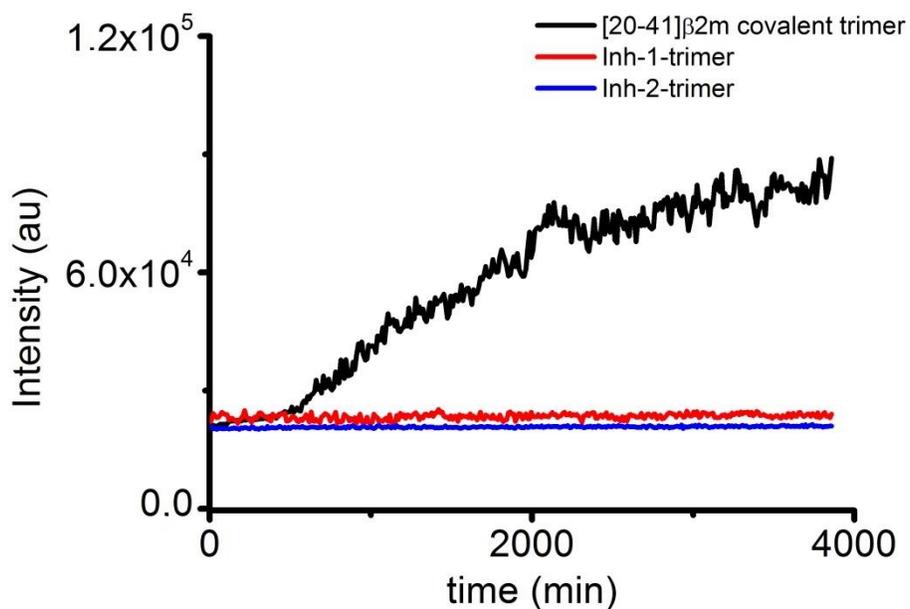
**Table 8:** Final yields of the *N*-methylated variants of [20-41] $\beta$ 2m homotrimers.

Name of the peptide (monomers)	Amount of starting material (linker, $\mu$ mol)	Isolated mass of purified homotrimer (mg)	Yield, based on the central linker (%)
Inh-1-trimer	1.5	10.9	91
Inh-2-trimer	1.5	10.5	87
Inh-3-trimer	1.1	1.5	18
Inh-4-trimer	1.5	9.0	77
Inh-5-trimer	1.5	9.1	78
Inh-6-trimer	1.65	6.2	48
Inh-8-trimer	1.25	4.8	49
Inh-9-trimer	1	3.7	47

## 5.4 Characterization of the *N*-methylated constructs

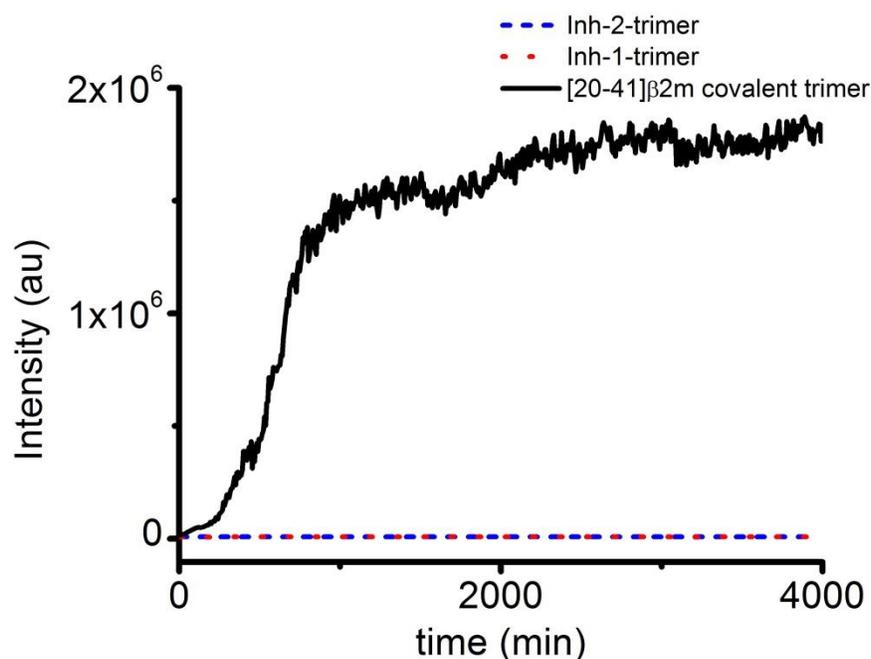
### 5.4.1 Study of the amyloidogenic properties of the constructs

In order to demonstrate that the *N*-methylated constructs do not aggregate into amyloid fibrils or prefibrillar assemblies we monitored the kinetics of amyloid growth of Inh-1-trimer and Inh-2-trimer over three days and compared them with the growth kinetic of the [20-41] $\beta$ 2m covalent trimer. 8-Anilino-naphthalene-1-sulfonate (ANS) was used as the extrinsic fluorophore for the detection of aggregates. The different peptidic constructs were solubilized at a concentration of 3.3 mM in 100% dimethyl sulfoxide (DMSO). The stock solutions in DMSO of the different trimers were diluted into 50 mM phosphate buffer (pH 7.5) with 100 mM NaCl and NaN<sub>3</sub> (0.05% w/v) at a final concentration of 33.3  $\mu$ M and incubated at 30°C. The concentration of ANS was 200  $\mu$ M and the dye concentration ratio was 2:1 (ANS:peptide as a monomer) and the final concentration of DMSO was 1% (v/v). The fluorescence was measured using a plate reader and the changes of the fluorescence intensity are shown in **Figure 60**.



**Figure 60:** Inh-1-trimer and Inh-2-trimer do not display any extrinsic fluorescence. Study of kinetics of growth determined by ANS fluorescence using a plate reader.

Both Inh-1-trimer and Inh-2-trimer did not exhibit any ANS fluorescence while [20-41] $\beta$ 2m covalent trimer clearly forms amyloid aggregates. The four *N*-methylations per peptide are clearly blocking the self-assembly into fibrils or into oligomeric species able to bind ANS. The same experiment was realized using thioflavin T (ThT) instead of ANS as the fluorescent probe. Comparing different probes can be very useful as they do not bind to their target in the same way and may have not the same affinity regarding fibrils or prefibrillar oligomeric assemblies. In this case, the concentration of ThT was 100  $\mu$ M and the dye concentration ratio was 1:1 (ThT:peptide as a monomer). The results are shown in **Figure 61** and confirm that Inh-1-trimer and Inh-2-trimer are not forming aggregates or oligomers able to bind ThT. Similar results were obtained for all the other *N*-methylated constructs (Inh-3-trimer, Inh-4-trimer, Inh-5-trimer, Inh-6-trimer, Inh-8-trimer and Inh-9-trimer).

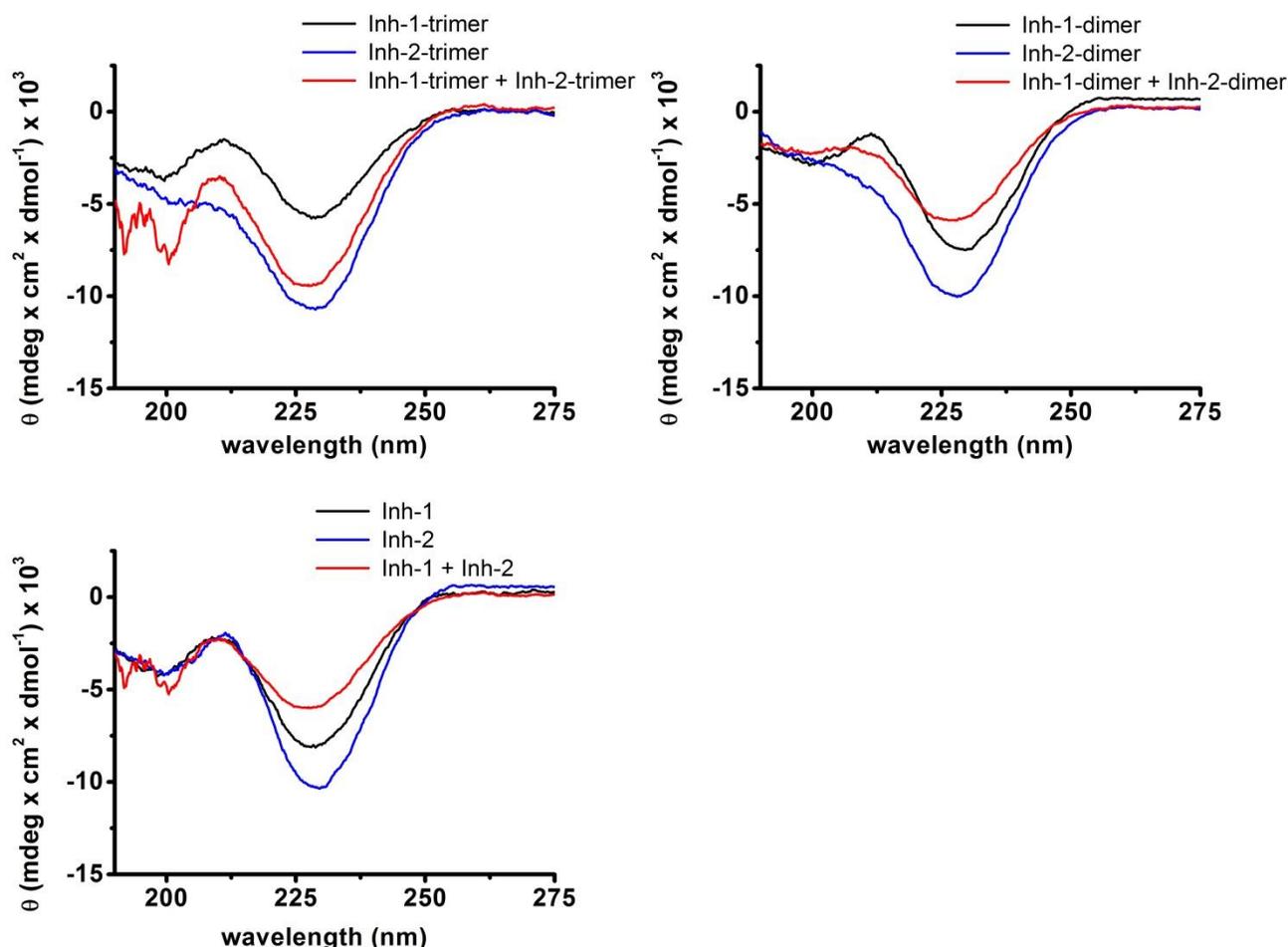


**Figure 61:** Inh-1-trimer and Inh-2-trimer do not display any extrinsic fluorescence. Study of the kinetics of growth determined by ThT fluorescence on a plate reader.

#### 5.4.2 Study of the secondary structure by circular dichroism

The secondary structure of Inh-1 and Inh-2 (monomer, homodimers and homotrimers) was studied by circular dichroism along with mixtures of each cognate structure (**Figure 62**). As

expected, the CD spectra resemble that of a  $\beta$ -sheet and suggest that the different *N*-methylated variants and structures adopt a  $\beta$ -strand conformation. It is not surprising because it is known that *N*-methylation of the peptide backbone influences the *cis-trans* equilibrium of the amide bond and affects the conformation of the adjacent residues resulting in local  $\beta$ -sheet-like structures<sup>156</sup>.



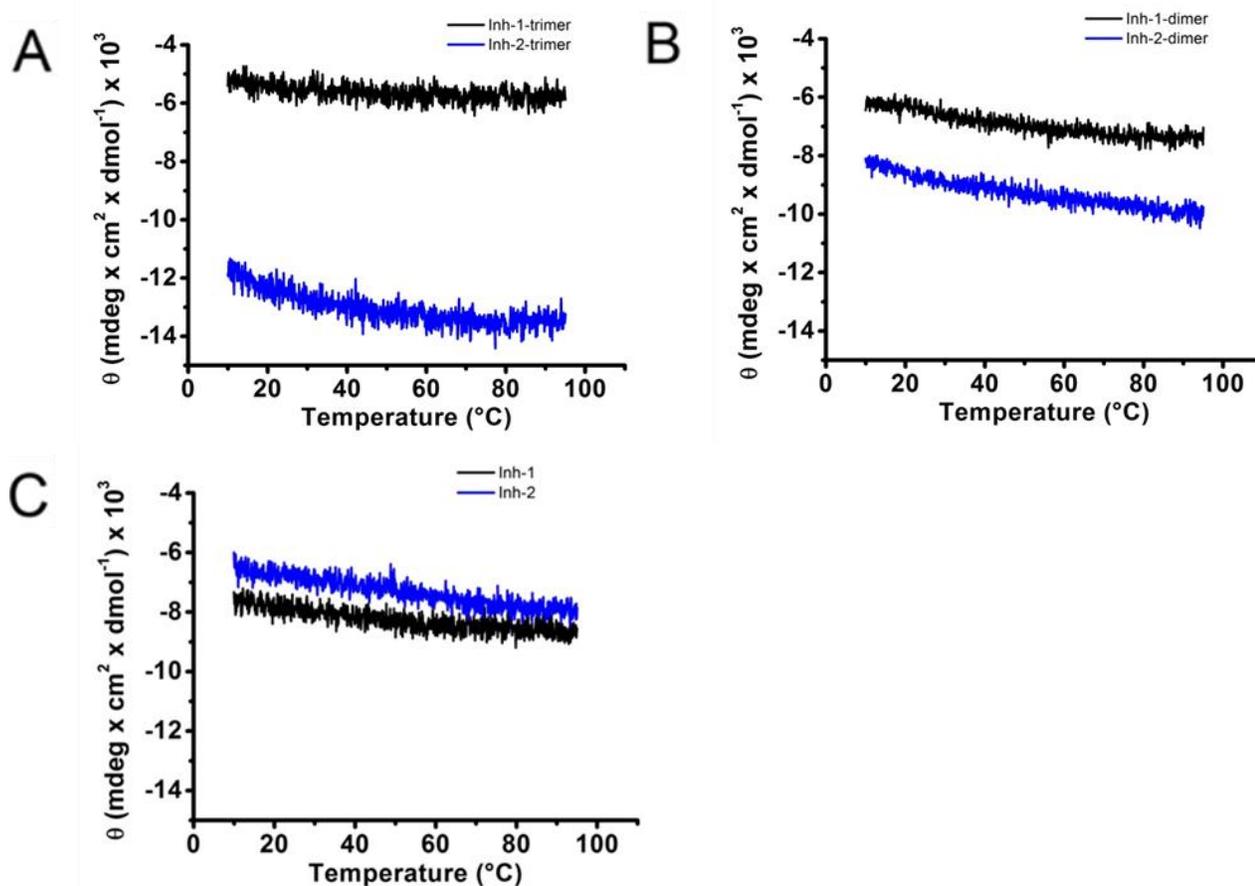
**Figure 62:** CD spectra of the different *N*-methylated constructs. For every measurement peptides were dissolved in a buffered solution (50 mM phosphate, pH 7.5) at a concentration of 25  $\mu$ M for the covalent trimers, 50  $\mu$ M for the covalent dimers and 75  $\mu$ M for the monomers. The obtained ellipticities were transformed to mean residue molar ellipticities.

The red-shifted ellipticity minimum at  $\sim$ 228 nm can be attributed to a conformational twist or distortion due to the *N*-methylation<sup>181</sup>. It appears that whatever the architecture of Inh-2 (monomer, dimer or trimer), the  $\beta$ -sheet character seems to be more pronounced compared to Inh-1. Mixtures of Inh-1 and Inh-2 (monomers, homodimers and homotrimers) were studied because we expected that both constructs would interact and form heterodimers of monomers,

dimers and trimers through  $\beta$ -sheet formation in a parallel fashion. We also anticipated that homodimers of monomers, dimers and trimers could also form through interactions in an anti-parallel way.

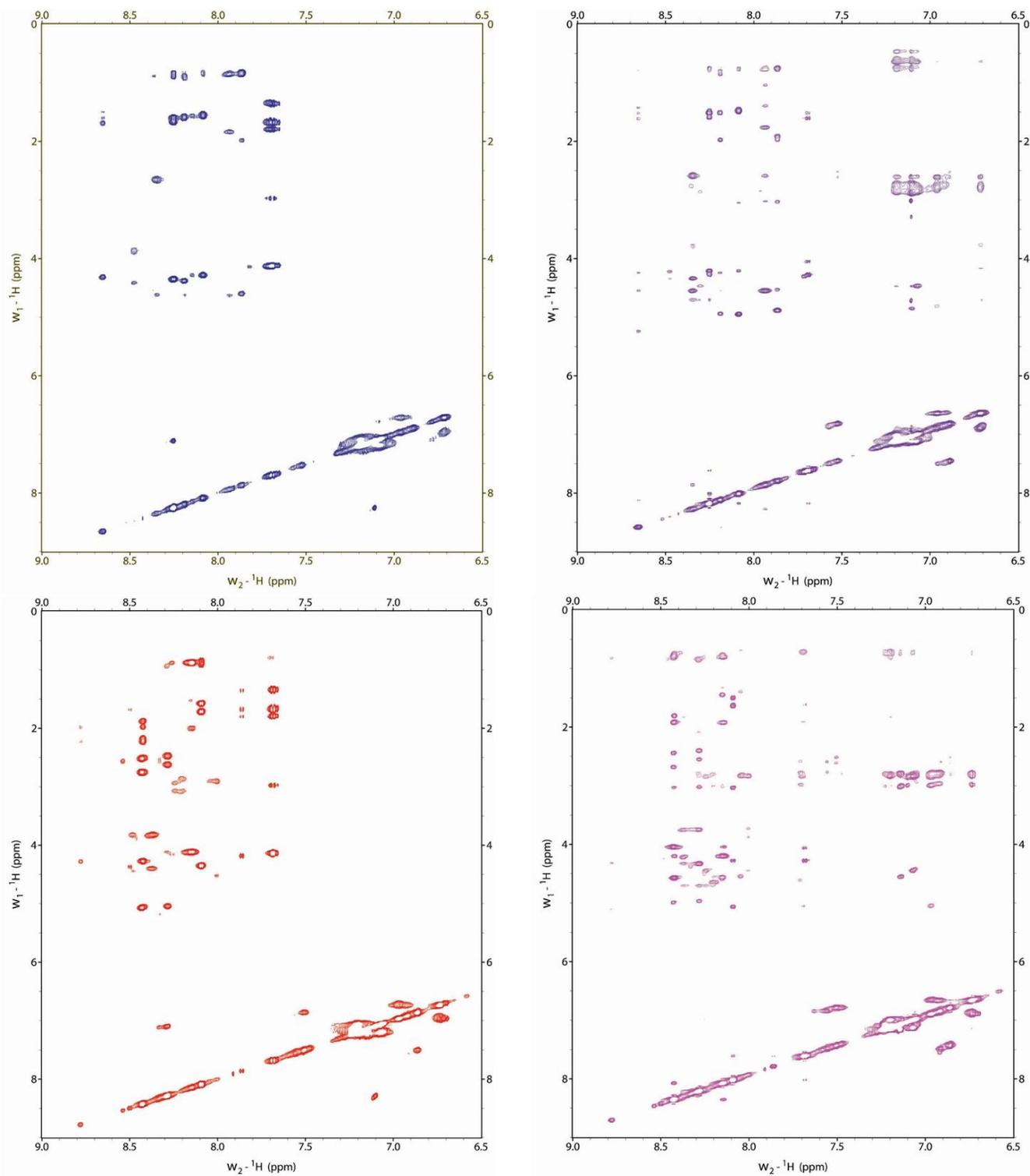
#### 5.4.3 Study of the stability of the apparent $\beta$ -stranded structures

The stability of the *N*-methylated monomers, covalent dimers and covalent trimers was assessed by melting experiments (**Figure 63**). The CD values of the different peptidic constructs were measured at the minimum of mean residue ellipticity attributed to  $\beta$ -strand conformation at different temperatures from 10 to 95°C. For every construct, the ellipticity does not vary much and even tends to decrease, which may indicate that the structure of the different *N*-methylated constructs is highly stable implying rigid  $\beta$ -strands.



**Figure 63:** Melting curves of different *N*-methylated constructs displayed as ellipticity change as a function of temperature. The temperature was increased at the rate of 1  $^\circ\text{C}/\text{min}$  and the ellipticities were measured at the minimum of intensity for every scaffolds, A) at 226 nm for the covalent trimers, B) at 229 nm for the covalent dimers and C) at 228 nm for the monomers.

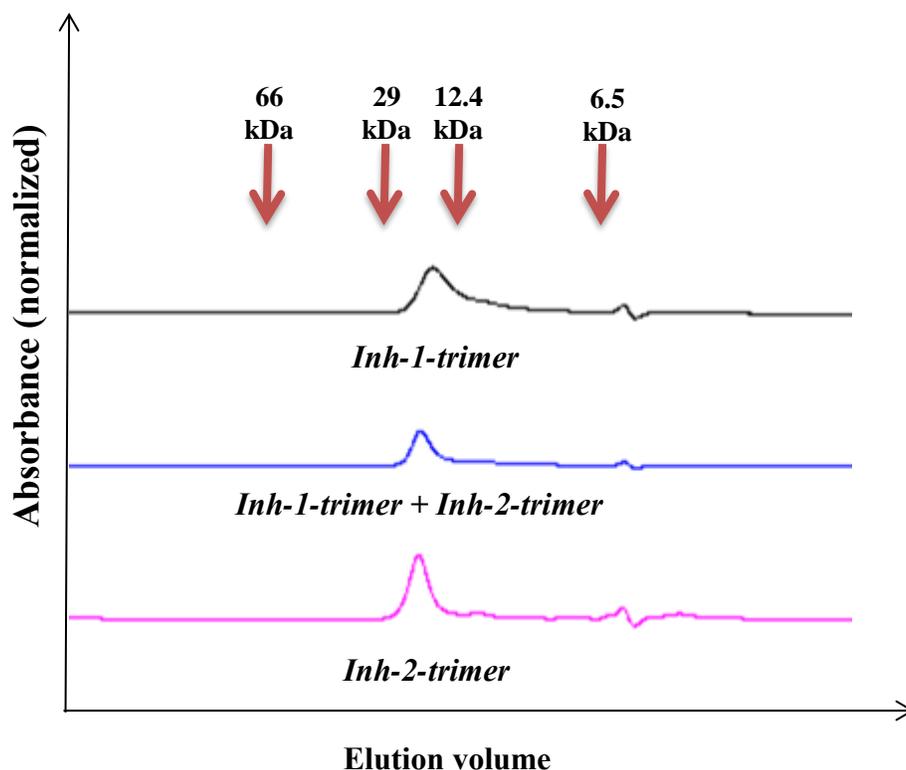
#### 5.4.4 Solution NMR measurements



**Figure 64:**  $^1\text{H}$ -TOCSY (top, left),  $^1\text{H}$ -NOESY (top, right) spectra of Inh-1-trimer and  $^1\text{H}$ -TOCSY (bottom, left),  $^1\text{H}$ -NOESY (bottom, right) spectra of Inh-2-trimer.

Tertiary structures of *N*-methylated analogs seems to be well-developed (based on well-visible NOEs especially from aromatic to alkyl groups and NH-NH cross-peaks) (**Figure 64**). Most of the amides are visible in the NMR spectra.

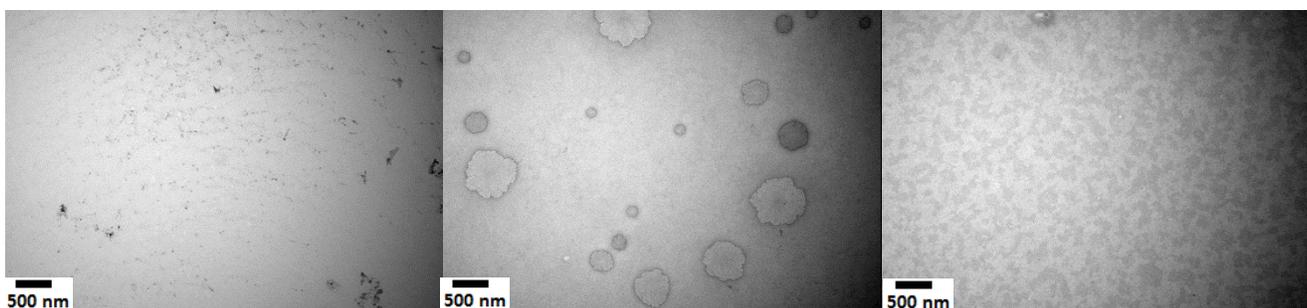
#### 5.4.5 Study of the oligomerization propensity by size-exclusion chromatography and TEM.



**Figure 65:** Size-exclusion chromatograms of Inh-1-trimer (black curve) and Inh-2-trimer (pink curve) alone and the 1:1 mixture of both trimers (blue curve). The red arrows associated with molecular weights correspond to the retention times of the related protein markers. The molecular weight of one *N*-methylated covalent trimer is around 7.9 kDa and a dimer of two such constructs would give a molecular weight around 15.8 kDa.

In order to investigate the propensity of the *N*-methylated constructs to dimerize or form even higher-order oligomers, solutions of Inh-1-trimer, Inh-2-trimer and 1:1 mixture of both trimers were analyzed by size-exclusion chromatography. The peptide concentration was 33.3  $\mu$ M and a 50 mM phosphate buffer (pH 7.5) was used as the mobile phase. Four gel filtration markers for

protein molecular weight were eluted and their retention times were used to trace a calibration curve (aprotinin: 6.5 kDa, 8.34 min; cytochrome c: 12.4 kDa, 7.21 min; carbonic anhydrase: 29 kDa, 6.63 min; bovine serum albumin: 66 kDa, 5.73 min). The corresponding elution volumes are marked with red arrows on **Figure 65**. For the three samples with four methyl groups per peptide, we observed one major peak eluting between 6.93 and 7.02 minutes. Those values indicate that the molecular weights of the different samples are around 20 kDa. At the concentration of 33.3  $\mu\text{M}$ , the constructs are not monomeric. They have a propensity to interact and form dimeric species. It is also important to mention that the final structure of the dimers of *N*-methylated covalent trimer would not resemble those of globules, which would impact on the retention times. The trimers may interact through their non-methylated interfaces. Similar results were obtained for covalent trimers containing the two methyl groups per peptide. Moreover, TEM images of the different solutions did not reveal the presence of small to big aggregates or fibrils (**Figure 66**).

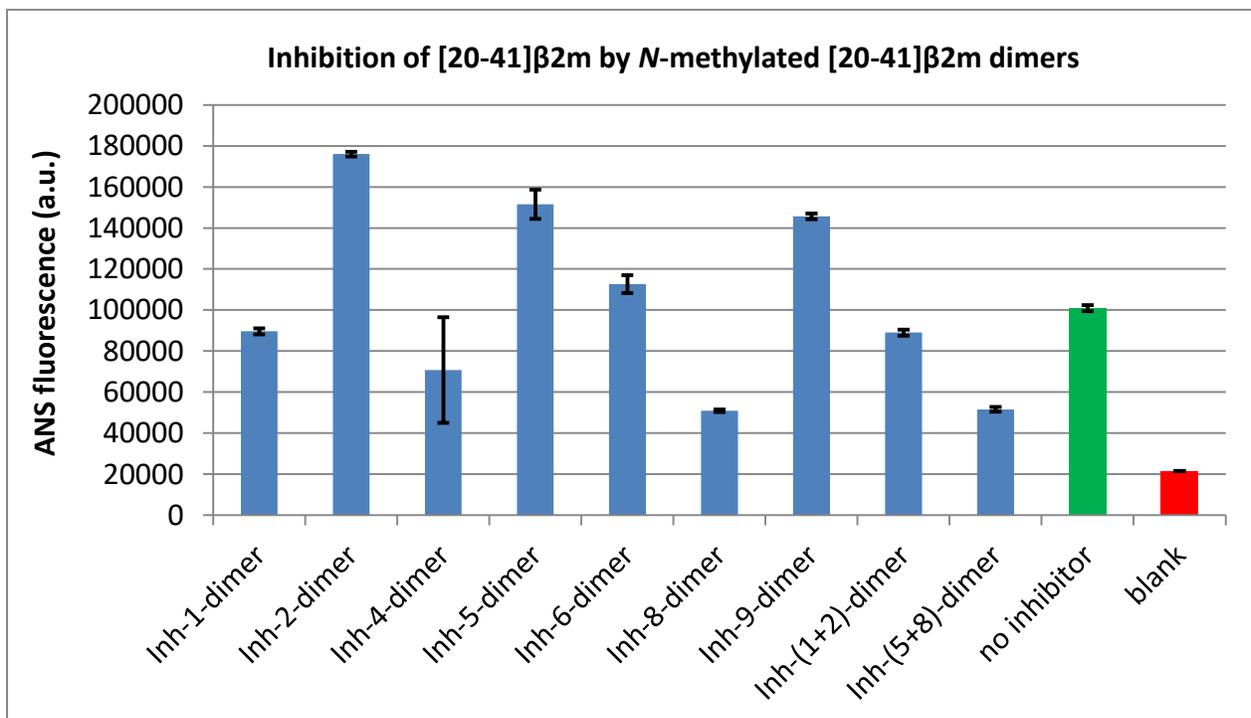
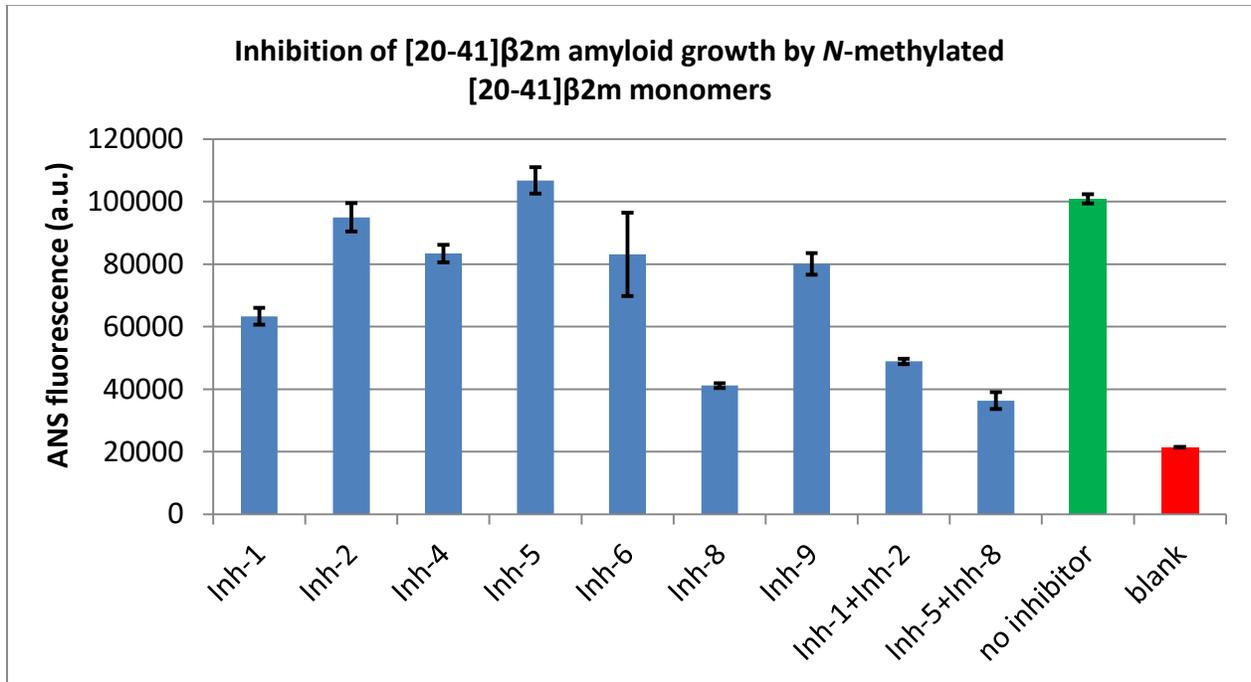


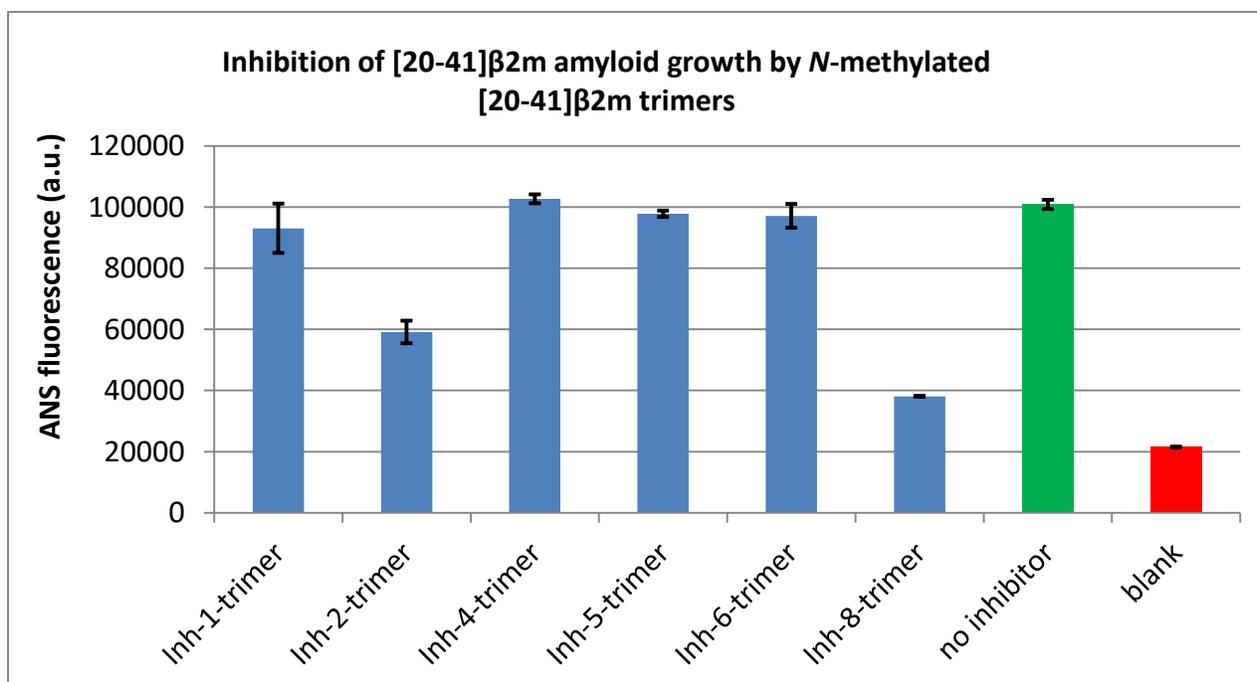
**Figure 66:** TEM images of Inh-1-trimer (left), Inh-2-trimer (center) and 1:1 mixture of Inh-1-trimer and Inh-2-trimer.

## 5.5 Study of the inhibition of amyloid growth

In order to assess the inhibition properties of the different *N*-methylated variants, the [20-41] $\beta$ 2m fragment was incubated with the *N*-methylated construct of each variants (monomeric, dimeric and trimeric). In all experiments, the concentration of [20-41] $\beta$ 2m fragment was 100  $\mu\text{M}$  and the concentration of *N*-methylated monomers, dimers and trimers was 33  $\mu\text{M}$ , 17  $\mu\text{M}$  and 11  $\mu\text{M}$ . The ANS fluorescence response was measured for 3 days and the end-point measurements are reported in the following bar charts. For the inhibition of the [20-41] $\beta$ 2m amyloid growth by

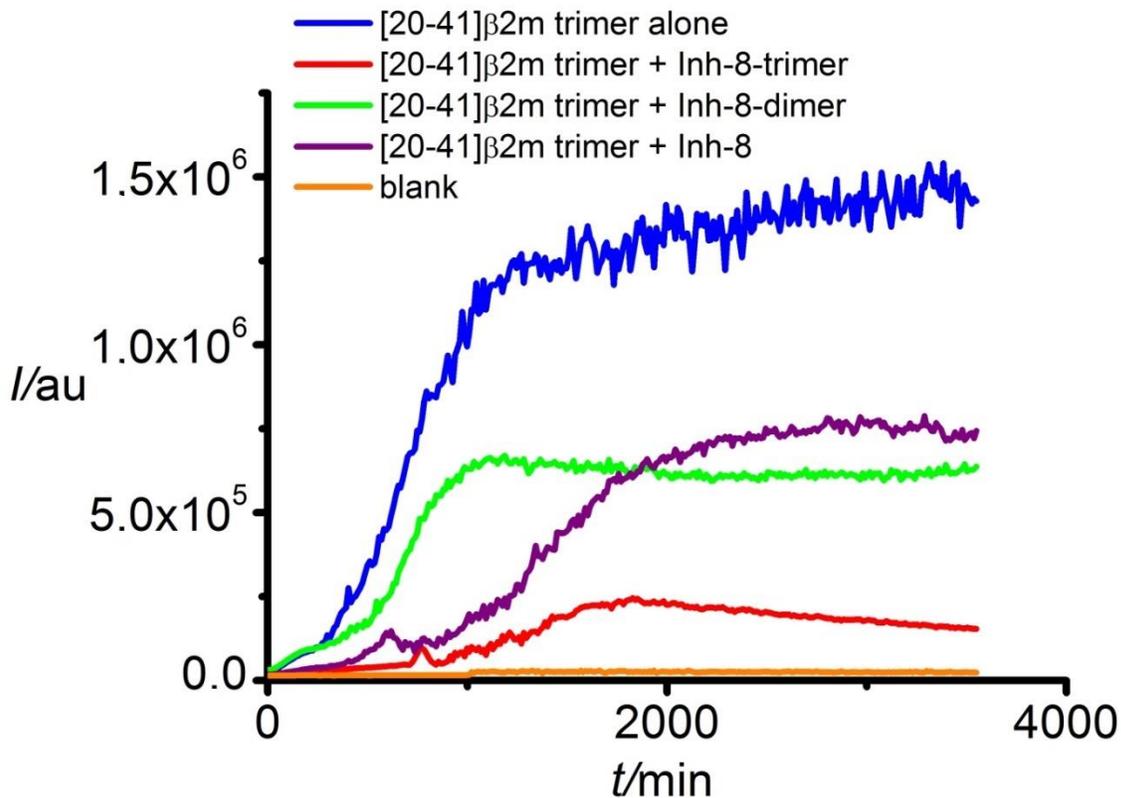
monomeric *N*-methylated peptides, we observed that one variant, Inh-8, gives the lowest ANS response after 3 days of incubation compared to the other tested. The other variants do not seem to inhibit much the amyloid growth of [20-41] $\beta$ 2m fragment.





Similar experiments were conducted with *N*-methylated [20-41]β2m dimers and *N*-methylated [20-41]β2m trimers and in both cases the best inhibitor is composed of the Inh-8 construct. Interestingly, for the inhibition of [20-41]β2m amyloid growth by some *N*-methylated constructs (Inh-2-dimer, Inh-5-dimer or Inh-9-dimer for example), the final ANS response was even higher than the ANS intensity when the [20-41]β2m fragment was incubated alone. It might be due to the presence of small to medium size oligomers or aggregates composed of methylated and non-methylated peptides that are still soluble (entirely or partially) and may give a higher fluorescence response than the [20-41]β2m amyloid fibrils.

Then, we decided to study the inhibition of [20-41]β2m “covalent trimer” by the different Inh-8 constructs to see which of the monomeric, dimeric or trimeric form of Inh-8 is a better inhibitor towards the formation of amyloid assemblies with three-fold symmetry. We analyzed kinetics of amyloid growth in each case (**Figure 67**) by time-resolved fluorometry using ThT dye as the fluorescent probe.



**Figure 67:** Kinetics of [20-41] $\beta$ 2m amyloid growth determined by ThT fluorescence using a plate reader in the presence of Inh-8-trimer, Inh-8-dimer or Inh-8 as a monomer.

The kinetics of [20-41] $\beta$ 2m amyloid growth is most affected by the presence of cognate methylated construct (Inh-8-trimer) which leads to a substantial decrease of the ThT response (red curve) compared to non-cognate constructs (Inh-8-dimer, purple curve, and Inh-8, green curve). Inh-8-trimer seems to have a higher affinity for [20-41] $\beta$ 2m “covalent trimer”. Those final results indicate a promising direction for the discovery of structure-based peptide constructs for the selective recognition of amyloid polymorphs and the subsequent inhibition of amyloid propagation.

## 5.6 Discussion

Designing structure-based peptide constructs has a strong potential to find new treatment against amyloid deposition. They would bind selectively to their target and act as amyloid growth inhibitors

and also as molecular probes. Conjugation of dyes to the designed construct could for instance permit the *in vitro* and *in vivo* visualization by staining the specific amyloids. We decided to design *N*-methylated peptide constructs because they provide many interesting features such as proteolytic stability, enhanced membrane permeability and enhanced solubility, in conjunction with  $\beta$ -sheet disrupting properties.

We designed *N*-methylated peptide constructs based on the structure of the amyloidogenic [20-41] $\beta$ 2m fragment to investigate their inhibitory potency towards amyloid formation and their selectivity towards distinct polymorphs. We also investigated their behavior in solution and showed that they are not monomeric but have a propensity to form dimeric species. The *N*-methylations are localized within the recognition motif of the [20-41] $\beta$ 2m fragment (within the  $\beta$ -stands) and they possibly induce local  $\beta$ -sheet conformations, as revealed by CD measurements. This local conformational change might facilitate and accelerate the recognition and binding of two similar *N*-methylated constructs into the formation of homodimers.

The preliminary results obtained by fluorescence measurements after three days of incubation show that many variants have few or no inhibitory properties. The fluorescence levels are almost similar to those observed without addition of *N*-methylated peptides. Only the constructs composed of the Inh-8 variant (Inh-8, Inh-8-dimer and Inh-8-trimer) show significant decrease of the fluorescence. Therefore, the position of the *N*-methyl groups within the sequence has a significant impact on the inhibition process. Moreover, adding more methyl groups does not automatically favor the inhibition. It might even lead to enhanced dimerization propensity of the *N*-methylated constructs. In fact, in the beginning of the incubation of the non-methylated with the methylated peptides, the *N*-methyl-peptides would dimerize instead of interacting with non-methylated peptides that do have yet a stable conformation. Therefore, it would be interesting to add the *N*-methylated constructs after several minutes or hours of incubation (after the formation of nuclei and small oligomers) and compare the inhibition efficiency. It would be also useful to add a disrupting element motif (e.g. poly-cationic fragment) positioned on the termini of the *N*-methylated peptide sequence to disfavor the dimerization process via self-repulsion.

It is also necessary to find other ways to evaluate the inhibition efficacy. The measurement of the fluorescence by incorporation of extrinsic dyes might not be entirely suitable for this study. The *N*-methylated constructs might in fact block the polymerization by sequestering small to medium size

soluble aggregates that would respond to the fluorescent probes with a higher intensity than do the insoluble mature fibrils. There is a need to correlate those results with other analytical techniques to better understand the mechanism of inhibition by the *N*-methylated constructs. We will investigate the content of soluble peptide after incubation by chromatographic tools and estimate the size of the soluble aggregates by SEC experiments. Regarding the selectivity of recognition of amyloid polymorphs by cognate *N*-methylated constructs, promising results were obtained for the inhibition of amyloid propagation of the “covalent trimer” of [20-41] $\beta$ 2m. If we compare the results obtained by time-resolved fluorescence experiments, we clearly see a significant impact in the kinetics profile caused by the presence of the “covalent trimer” of Inh-8 (Inh-8-trimer) leading to notable decrease of the fluorescence response. The kinetics of amyloid growth was also impacted by the presence of monomeric Inh-8 and the “covalent dimer” of Inh-8 but to a much lower extent leading to the conclusion that Inh-8-trimer has a better recognition motif towards the surfaces of aggregates. Additional studies need to be carried out for the inhibition of the “covalent dimer” of [20-41] $\beta$ 2m.

## **5.7 Conclusions**

In summary, this study demonstrated an approach for the development of amyloid growth inhibitors that are specific for distinct amyloid structures. We reported the design and synthesis of different *N*-methyl peptide constructs for the development of molecular probes that recognize different amyloid polymorphs. *N*-methylated peptides may become valuable tools as diagnostic reagents for the detection of amyloids and even report on the structural polymorphism of the fibril (two-fold versus three-fold symmetry). The characterization of the physicochemical and structural properties of *N*-methylated peptides is of high importance in order to apprehend their behavior and their interactions with the surfaces of amyloid aggregates and allow for a general approach for a rational design in development of amyloid inhibitors. Moreover, based on the results obtained for the inhibition of [20-41] $\beta$ 2m “covalent dimer” by and full length  $\beta$ 2m by their mirror-image counterpart exposed in chapter 2, we envisage the development of *N*-methylated mirror-image peptide inhibitors in order to combine heterochiral protein-protein interactions and the  $\beta$ -sheet disrupting properties of the *N*-methyl groups.

## 6. Significance and future perspectives

The structural and biological studies of amyloid fibrils and the transient species involved in amyloid formation are highly challenging. The structural polymorphism of amyloids and the metastable nature of the early-stage structures render the elucidation of the involved mechanisms in amyloid formation rather complex and complicates the finding of pathways for cytotoxicity.

Chemical protein synthesis is a powerful tool that allowed us to produce chemically modified polypeptides such as isotopically-labeled, mirror-image of *N*-methylated peptides or proteins. We produced very distinct and innovative peptide assemblies to dissect structural properties and functions of amyloids. We elaborated different synthetic methodologies for the formation of distinct amyloid fibril polymorphs for the preparation of homogeneous samples for structural and biological studies. We reported chemical strategies that enable control of the formation of distinct amyloid polymorphs with two-fold and three-fold core stoichiometries, which may serve as scaffolds for immunization experiments and screening of small-molecules to selectively inhibit fiber growth of a particular amyloid polymorph. In the future, we would like to reiterate and validate this strategy with A $\beta$  peptides to prepare different covalent dimers and covalent trimers of A $\beta$  and to produce homogeneous samples of amyloid fibrils for structural and biological studies, which will help to find new treatments of Alzheimer's disease.

We also synthesized covalently-tethered oligomers composed of nine copies of an amyloidogenic peptide segment and we were able to control the self-assembly of the structure by insertion of *N*-methylated amino-acids to obtain monomeric oligomers mimicking a portion of the structure of an amyloid fibril. The innovation in this work is the implementation of a new methodology to prepare large peptidic construct that could, for instance, mimic a cross-section of an amyloid fibril. In the future, this synthetic strategy can be implemented for the preparation of larger constructs, e.g. "covalent pentamer of covalent trimers" and for incorporation of mirror-image segments within the structure. This strategy can be applied to the synthesis of other oligomers composed of disease-related peptides such as A $\beta$  peptide. Such oligomers could serve as molecular architectures for the development of new amyloid growth inhibitors and diagnostics.

We also reported on the chiral recognition of *L*-peptides and *L*-proteins towards *D*-enantiomers during amyloid formation. Moreover, we studied the impact of *N*-methylation and mirror-image polypeptides on the inhibition of amyloid propagation. The future goal is the development of *N*-methylated mirror-image peptide inhibitors, in which  $\beta$ -sheet disrupting features, such as heterochiral protein-protein interactions and *N*-methylation, will be combined.

## 7. References

1. Knowles, T. P. J., Vendruscolo, M. & Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol.* **15**, 384–396 (2014).
2. Greenwald, J. & Riek, R. Biology of amyloid: structure, function, and regulation. *Structure* **18**, 1244–1260 (2010).
3. Nelson, R. *et al.* Structure of the cross- $\beta$  spine of amyloid-like fibrils. *Nature* **435**, 773–778 (2005).
4. Riek, R. The three-dimensional structures of amyloids. *Cold Spring Harb. Perspect. Biol.* **9**, a023572 (2017).
5. Turnell, W. G. & Finch, J. T. Binding of the dye congo red to the amyloid protein pig insulin reveals a novel homology amongst amyloid-forming peptide sequences. *J. Mol. Biol.* **227**, 1205–1223 (1992).
6. Younan, N. D. & Viles, J. H. A comparison of three fluorophores for the detection of amyloid fibers and prefibrillar oligomeric assemblies. ThT (thioflavin T); ANS (1-anilinonaphthalene-8-sulfonic Acid); and bisANS (4,4'-dianilino-1,1'-binaphthyl-5,5'-disulfonic Acid). *Biochemistry* **54**, 4297–4306 (2015).
7. Hawe, A., Sutter, M. & Jiskoot, W. Extrinsic fluorescent dyes as tools for protein characterization. *Pharm. Res.* **25**, 1487–1499 (2008).
8. Sunde, M. *et al.* Common core structure of amyloid fibrils by synchrotron X-ray diffraction11 Edited by F. E. Cohen. *J. Mol. Biol.* **273**, 729–739 (1997).
9. Fitzpatrick, A. W. P. *et al.* Atomic structure and hierarchical assembly of a cross- $\beta$  amyloid fibril. *Proc. Natl. Acad. Sci.* **110**, 5468–5473 (2013).
10. Schmidt, M. *et al.* Peptide dimer structure in an A $\beta$ (1–42) fibril visualized with cryo-EM. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 11858–11863 (2015).

11. Cobb, N. J., Sönnichsen, F. D., Mchaourab, H. & Surewicz, W. K. Molecular architecture of human prion protein amyloid: A parallel, in-register  $\beta$ -structure. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18946–18951 (2007).
12. Eisenberg, D. & Jucker, M. The amyloid state of proteins in human diseases. *Cell* **148**, 1188–1203 (2012).
13. Chiti, F. & Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **75**, 333–366 (2006).
14. Lee, S. J. C., Nam, E., Lee, H. J., Savelieff, M. G. & Lim, M. H. Towards an understanding of amyloid- $\beta$  oligomers: characterization, toxicity mechanisms, and inhibitors. *Chem. Soc. Rev.* **46**, 310–323 (2017).
15. Biancalana, M. & Koide, S. Molecular mechanism of thioflavin-T binding to amyloid fibrils. *Biochim. Biophys. Acta* **1804**, 1405–1412 (2010).
16. Zhang, H., Ma, Q., Zhang, Y. & Xu, H. Proteolytic processing of Alzheimer's  $\beta$ -amyloid precursor protein. *J. Neurochem.* **120**, 9–21 (2012).
17. Walsh, D. M. *et al.* Naturally secreted oligomers of amyloid  $\beta$  protein potently inhibit hippocampal long-term potentiation in vivo. *Nature* **416**, 535–539 (2002).
18. Jucker, M. & Walker, L. C. Self-propagation of pathogenic protein aggregates in neurodegenerative diseases. *Nature* **501**, 45–51 (2013).
19. Aguzzi, A. & Lakkaraju, A. K. K. Cell biology of prions and prionoids: A status report. *Trends Cell Biol.* **26**, 40–51 (2016).
20. Fowler, D. M., Koulov, A. V., Balch, W. E. & Kelly, J. W. Functional amyloid – from bacteria to humans. *Trends Biochem. Sci.* **32**, 217–224 (2007).
21. Chapman, M. R. *et al.* Role of Escherichia coli curli operons in directing amyloid fiber formation. *Science* **295**, 851–855 (2002).
22. Maji, S. K. *et al.* Functional amyloids as natural storage of peptide hormones in pituitary secretory granules. *Science* **325**, 328–332 (2009).

23. Fowler, D. M. *et al.* Functional amyloid formation within mammalian tissue. *PLOS Biol.* **4**, e6 (2005).
24. Kajava, A. V., Baxa, U. & Steven, A. C.  $\beta$ -arcades: recurring motifs in naturally occurring and disease-related amyloid fibrils. *FASEB J.* **24**, 1311–1319 (2010).
25. Jiménez, J. L. *et al.* The protofilament structure of insulin amyloid fibrils. *Proc. Natl. Acad. Sci.* **99**, 9196–9201 (2002).
26. Lührs, T. *et al.* 3D structure of Alzheimer's amyloid- $\beta$ (1–42) fibrils. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 17342–17347 (2005).
27. Iwata, K. *et al.* 3D structure of amyloid protofilaments of  $\beta$ 2-microglobulin fragment probed by solid-state NMR. *Proc. Natl. Acad. Sci.* **103**, 18119–18124 (2006).
28. Kajava, A. V., Baxa, U., Wickner, R. B. & Steven, A. C. A model for Ure2p prion filaments and other amyloids: The parallel superpleated  $\beta$ -structure. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 7885–7890 (2004).
29. Der-Sarkissian, A., Jao, C. C., Chen, J. & Langen, R. Structural organization of  $\alpha$ -synuclein fibrils studied by site-directed spin labeling. *J. Biol. Chem.* **278**, 37530–37535 (2003).
30. Wasmer, C. *et al.* Amyloid fibrils of the HET-s(218–289) prion form a  $\beta$ -solenoid with a triangular hydrophobic core. *Science* **319**, 1523–1526 (2008).
31. Dzwolak, W. Chirality and chiroptical properties of amyloid fibrils. *Chirality* **26**, 580–587 (2014).
32. Qiang, W., Yau, W.-M., Luo, Y., Mattson, M. P. & Tycko, R. Antiparallel  $\beta$ -sheet architecture in Iowa-mutant  $\beta$ -amyloid fibrils. *Proc. Natl. Acad. Sci.* **109**, 4443–4448 (2012).
33. Madine, J. *et al.* Structural insights into the polymorphism of amyloid-like fibrils formed by region 20–29 of amylin revealed by solid-state NMR and X-ray fiber diffraction. *J. Am. Chem. Soc.* **130**, 14990–15001 (2008).
34. Rigoldi, F., Metrangolo, P., Redaelli, A. & Gautieri, A. Nanostructure and stability of calcitonin amyloids. *J. Biol. Chem.* **292**, 7348–7357 (2017).

35. Colletier, J.-P. *et al.* Molecular basis for amyloid- $\beta$  polymorphism. *Proc. Natl. Acad. Sci.* **108**, 16938–16943 (2011).
36. Berhanu, W. M. & Hansmann, U. H. E. Structure and dynamics of amyloid- $\beta$  segmental polymorphisms. *PLoS ONE* **7**, (2012).
37. Paravastu, A. K., Leapman, R. D., Yau, W.-M. & Tycko, R. Molecular structural basis for polymorphism in Alzheimer's  $\beta$ -amyloid fibrils. *Proc. Natl. Acad. Sci.* **105**, 18349–18354 (2008).
38. Härd, T. & Lendel, C. Inhibition of amyloid formation. *J. Mol. Biol.* **421**, 441–465 (2012).
39. Reardon, S. Antibody drugs for Alzheimer's show glimmers of promise. *Nat. News* **523**, 509 (2015).
40. Gordon, D. J. & Meredith, S. C. Probing the role of backbone hydrogen bonding in  $\beta$ -amyloid fibrils with inhibitor peptides containing ester bonds at alternate positions. *Biochemistry* **42**, 475–485 (2003).
41. Yan, L.-M., Tarek-Nossol, M., Velkova, A., Kazantzis, A. & Kapurniotu, A. Design of a mimic of nonamyloidogenic and bioactive human islet amyloid polypeptide (IAPP) as nanomolar affinity inhibitor of IAPP cytotoxic fibrillogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2046–2051 (2006).
42. Gilead, S. & Gazit, E. Inhibition of amyloid fibril formation by peptide analogues modified with  $\alpha$ -aminoisobutyric acid. *Angew. Chem. Int. Ed.* **43**, 4041–4044 (2004).
43. Dzwolak, W., Ravindra, R., Nicolini, C., Jansen, R. & Winter, R. The diastereomeric assembly of polylysine is the low-volume pathway for preferential formation of  $\beta$ -sheet aggregates. *J. Am. Chem. Soc.* **126**, 3762–3768 (2004).
44. Dutta, S. *et al.* Suppression of oligomer formation and formation of non-toxic fibrils upon addition of mirror-image A $\beta$ 42 to the natural *L*-enantiomer. *Angew. Chem. Int. Ed.* **56**, 11506–11510 (2017).
45. Fischer, E. & Fourneau, E. Ueber einige Derivate des Glykocolls. *Berichte Dtsch. Chem. Ges.* **34**, 2868–2877 (1901).
46. Fischer, E. Synthese von Polypeptiden. XVII. *Berichte Dtsch. Chem. Ges.* **40**, 1754–1767 (1907).

47. du Vigneaud, V., Ressler, C., Swan, J. M., Roberts, C. W. & Katsoyannis, P. G. The synthesis of oxytocin. *J. Am. Chem. Soc.* **76**, 3115–3121 (1954).
48. Sieber, P. *et al.* Totalsynthese von Humaninsulin unter gezielter Bildung der Disulfidbindungen. Vorläufige Mitteilung. *Helv. Chim. Acta* **57**, 2617–2621 (1974).
49. Yajima, H. & Fujii, N. Totally synthetic crystalline ribonuclease A. *Biopolymers* **20**, 1859–1867 (1981).
50. Nishiuchi, Y. *et al.* Chemical synthesis of the precursor molecule of the Aequorea green fluorescent protein, subsequent folding, and development of fluorescence. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 13549–13554 (1998).
51. Cohen, S. N., Chang, A. C. Y., Boyer, H. W. & Helling, R. B. Construction of biologically functional bacterial plasmids in vitro. *Proc. Natl. Acad. Sci. U. S. A.* **70**, 3240–3244 (1973).
52. Merrifield, R. B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *J. Am. Chem. Soc.* **85**, 2149–2154 (1963).
53. Dawson, P. E., Muir, T. W., Clark-Lewis, I. & Kent, S. B. Synthesis of proteins by native chemical ligation. *Science* **266**, 776–779 (1994).
54. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. (2001).
55. Rosano, G. L. & Ceccarelli, E. A. Recombinant protein expression in *Escherichia coli*: advances and challenges. *Front. Microbiol.* **5**, (2014).
56. Weiss, G. A., Watanabe, C. K., Zhong, A., Goddard, A. & Sidhu, S. S. Rapid mapping of protein functional epitopes by combinatorial alanine scanning. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 8950–8954 (2000).
57. Xie, J. & Schultz, P. G. Adding amino acids to the genetic repertoire. *Curr. Opin. Chem. Biol.* **9**, 548–554 (2005).
58. Zhang, Y. *et al.* A semi-synthetic organism that stores and retrieves increased genetic information. *Nature* **551**, 644–647 (2017).

59. Anderson, J. C. *et al.* An expanded genetic code with a functional quadruplet codon. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 7566–7571 (2004).
60. Atherton, E., Clive, D. L. J. & Sheppard, R. C. Polyamide supports for polypeptide synthesis. *J. Am. Chem. Soc.* **97**, 6584–6585 (1975).
61. Coin, I., Beyermann, M. & Bienert, M. Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nat. Protoc.* **2**, 3247–3256 (2007).
62. Gates, Z. P., Dhayalan, B. & Kent, S. B. H. Obviation of hydrogen fluoride in Boc chemistry solid phase peptide synthesis of peptide- $\alpha$ -thioesters. *Chem. Commun.* **52**, 13979–13982 (2016).
63. Schnolzer, M. & Kent, S. B. Constructing proteins by dovetailing unprotected synthetic peptides: backbone-engineered HIV protease. *Science* **256**, 221–225 (1992).
64. Dawson, P. E., Churchill, M. J., Ghadiri, M. R. & Kent, S. B. H. Modulation of reactivity in native chemical ligation through the use of thiol additives. *J. Am. Chem. Soc.* **119**, 4325–4329 (1997).
65. Bang, D., Pentelute, B. L. & Kent, S. B. H. Kinetically controlled ligation for the convergent chemical synthesis of proteins. *Angew. Chem. Int. Ed.* **45**, 3985–3988 (2006).
66. Yan, L. Z. & Dawson, P. E. Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **123**, 526–533 (2001).
67. Malins, L. R., Giltrap, A. M., Dowman, L. J. & Payne, R. J. Synthesis of  $\beta$ -thiol phenylalanine for applications in one-pot ligation–desulfurization chemistry. *Org. Lett.* **17**, 2070–2073 (2015).
68. Sayers, J., Thompson, R. E., Perry, K. J., Malins, L. R. & Payne, R. J. Thiazolidine-protected  $\beta$ -thiol asparagine: Applications in one-pot ligation–desulfurization chemistry. *Org. Lett.* **17**, 4902–4905 (2015).
69. Haase, C., Rohde, H. & Seitz, O. Native chemical ligation at valine. *Angew. Chem. Int. Ed.* **47**, 6807–6810 (2008).
70. Zhang, Y., Xu, C., Lam, H. Y., Lee, C. L. & Li, X. Protein chemical synthesis by serine and threonine ligation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 6657–6662 (2013).

71. Loibl, S. F., Harpaz, Z. & Seitz, O. A type of auxiliary for native chemical peptide ligation beyond cysteine and glycine junctions. *Angew. Chem. Int. Ed.* **54**, 15055–15059 (2015).
72. Nadler, C., Nadler, A., Hansen, C. & Diederichsen, U. A photocleavable auxiliary for extended native chemical ligation. *Eur. J. Org. Chem.* **2015**, 3095–3102 (2015).
73. Kent, S. B. H. Total chemical synthesis of proteins. *Chem. Soc. Rev.* **38**, 338–351 (2009).
74. Camarero, J. A., Cotton, G. J., Adeva, A. & Muir, T. W. Chemical ligation of unprotected peptides directly from a solid support. *J. Pept. Res.* **51**, 303–316 (1998).
75. Blanco-Canosa, J. B. & Dawson, P. E. An efficient Fmoc-SPPS approach for the generation of thioester peptide precursors for use in native chemical ligation. *Angew. Chem. Int. Ed.* **47**, 6851–6855 (2008).
76. Blanco-Canosa, J. B., Nardone, B., Albericio, F. & Dawson, P. E. Chemical protein synthesis using a second-generation *N*-acylurea linker for the preparation of peptide-thioester precursors. *J. Am. Chem. Soc.* **137**, 7197–7209 (2015).
77. Fang, G.-M., Wang, J.-X. & Liu, L. Convergent chemical synthesis of proteins by ligation of peptide hydrazides. *Angew. Chem. Int. Ed.* **51**, 10347–10350 (2012).
78. Fang, G.-M. *et al.* Protein chemical synthesis by ligation of peptide hydrazides. *Angew. Chem. Int. Ed.* **50**, 7645–7649 (2011).
79. Ollivier, N., Dheur, J., Mhidia, R., Blanpain, A. & Melnyk, O. Bis(2-sulfanylethyl)amino native peptide ligation. *Org. Lett.* **12**, 5238–5241 (2010).
80. Dheur, J., Ollivier, N. & Melnyk, O. Synthesis of thiazolidine thioester peptides and acceleration of native chemical ligation. *Org. Lett.* **13**, 1560–1563 (2011).
81. Terrier, V. P., Adihou, H., Arnould, M., Delmas, A. F. & Aucagne, V. A straightforward method for automated Fmoc-based synthesis of bio-inspired peptide crypto-thioesters. *Chem. Sci.* **7**, 339–345 (2015).
82. Nilsson, B. L., Kiessling, L. L. & Raines, R. T. Staudinger ligation: A peptide from a thioester and azide. *Org. Lett.* **2**, 1939–1941 (2000).

83. Nilsson, B. L., Kiessling, L. L. & Raines, R. T. High-yielding Staudinger ligation of a phosphinothioester and azide to form a peptide. *Org. Lett.* **3**, 9–12 (2001).
84. Tam, A., Soellner, M. B. & Raines, R. T. Water-soluble phosphinothiols for traceless Staudinger ligation and integration with expressed protein ligation. *J. Am. Chem. Soc.* **129**, 11421–11430 (2007).
85. Pattabiraman, V. R., Ogunkoya, A. O. & Bode, J. W. Chemical protein synthesis by chemoselective  $\alpha$ -ketoacid–hydroxylamine (KAHA) ligations with 5-oxaproline. *Angew. Chem. Int. Ed.* **51**, 5114–5118 (2012).
86. Pusterla, I. & Bode, J. W. The mechanism of the  $\alpha$ -ketoacid–hydroxylamine amide-forming ligation. *Angew. Chem. Int. Ed.* **51**, 513–516 (2012).
87. He, C., Kulkarni, S. S., Thuaud, F. & Bode, J. W. Chemical synthesis of the 20 kDa heme protein nitrophorin 4 by  $\alpha$ -ketoacid–hydroxylamine (KAHA) ligation. *Angew. Chem. Int. Ed.* **54**, 12996–13001 (2015).
88. Pusterla, I. & Bode, J. W. An oxazetidine amino acid for chemical protein synthesis by rapid, serine-forming ligations. *Nat. Chem.* **7**, 668–672 (2015).
89. Chalifour, R. J. *et al.* Stereoselective interactions of peptide inhibitors with the  $\beta$ -amyloid peptide. *J. Biol. Chem.* **278**, 34874–34881 (2003).
90. Poduslo, J. F., Curran, G. L., Kumar, A., Frangione, B. & Soto, C.  $\beta$ -sheet breaker peptide inhibitor of Alzheimer's amyloidogenesis with increased blood–brain barrier permeability and resistance to proteolytic degradation in plasma. *J. Neurobiol.* **39**, 371–382 (1999).
91. Van Regenmortel, M. H. & Muller, S. D-peptides as immunogens and diagnostic reagents. *Curr. Opin. Biotechnol.* **9**, 377–382 (1998).
92. Sievers, S. A. *et al.* Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* **475**, 96–100 (2011).
93. Wiesehan, K. *et al.* Selection of D-amino-acid peptides that bind to Alzheimer's disease amyloid peptide A $\beta$ 1–42 by mirror image phage display. *ChemBioChem* **4**, 748–753 (2003).

94. Funke, S. A. *et al.* Development of a small *D*-enantiomeric Alzheimer's amyloid- $\beta$  binding peptide ligand for future in vivo imaging applications. *PLOS ONE* **7**, e41457 (2012).
95. Fuhrhop, J.-H., Krull, M. & Büldt, G. Precipitates with  $\beta$ -pleated sheet structure by mixing aqueous solutions of helical poly(*D*-lysine) and poly(*L*-lysine). *Angew. Chem. Int. Ed. Engl.* **26**, 699–700 (1987).
96. Dzwolak, W., Ravindra, R., Nicolini, C., Jansen, R. & Winter, R. The diastereomeric assembly of polylysine is the low-volume pathway for preferential formation of  $\beta$ -sheet aggregates. *J. Am. Chem. Soc.* **126**, 3762–3768 (2004).
97. Yamaoki, Y. *et al.* An FT-IR study on packing defects in mixed  $\beta$ -aggregates of poly(*L*-glutamic acid) and poly(*D*-glutamic acid): A high-pressure rescue from a kinetic trap. *J. Phys. Chem. B* **116**, 5172–5178 (2012).
98. Weissbuch, I., Illos, R. A., Bolbach, G. & Lahav, M. Racemic  $\beta$ -sheets as Templates of relevance to the origin of homochirality of peptides: Lessons from crystal chemistry. *Acc. Chem. Res.* **42**, 1128–1140 (2009).
99. Nagy, K. J., Giano, M. C., Jin, A., Pochan, D. J. & Schneider, J. P. Enhanced mechanical rigidity of hydrogels formed from enantiomeric peptide assemblies. *J. Am. Chem. Soc.* **133**, 14975–14977 (2011).
100. Swanekamp, R. J., DiMaio, J. T. M., Bowerman, C. J. & Nilsson, B. L. Coassembly of enantiomeric amphipathic peptides into amyloid-inspired rippled  $\beta$ -sheet fibrils. *J. Am. Chem. Soc.* **134**, 5556–5559 (2012).
101. Koga, T., Matsuoka, M. & Higashi, N. Structural control of self-assembled nanofibers by artificial  $\beta$ -sheet peptides composed of *D*- or *L*-isomer. *J. Am. Chem. Soc.* **127**, 17596–17597 (2005).
102. Krebs, M. R. H., Morozova-Roche, L. A., Daniel, K., Robinson, C. V. & Dobson, C. M. Observation of sequence specificity in the seeding of protein amyloid fibrils. *Protein Sci. Publ. Protein Soc.* **13**, 1933–1938 (2004).

103. Milton, R. C., Milton, S. C. & Kent, S. B. Total chemical synthesis of a *D*-enzyme: the enantiomers of HIV-1 protease show reciprocal chiral substrate specificity [corrected]. *Science* **256**, 1445–1448 (1992).
104. Kar, K., Arduini, I., Drombosky, K. W., van der Wel, P. C. A. & Wetzel, R. *D*-polyglutamine amyloid recruits *L*-polyglutamine monomers and kills cells. *J. Mol. Biol.* **426**, 816–829 (2014).
105. Esler, W. P. *et al.* Stereochemical specificity of Alzheimer's disease  $\beta$ -peptide assembly. *Biopolymers* **49**, 505–514 (1999).
106. Ban, T. *et al.* Direct observation of A $\beta$  amyloid fibril growth and inhibition. *J. Mol. Biol.* **344**, 757–767 (2004).
107. Chung, D. M. & Nowick, J. S. Enantioselective molecular recognition between  $\beta$ -sheets. *J. Am. Chem. Soc.* **126**, 3062–3063 (2004).
108. Kozhukh, G. V. *et al.* Investigation of a peptide responsible for amyloid fibril formation of  $\beta$ 2-microglobulin by *Achromobacter* protease I. *J. Biol. Chem.* **277**, 1310–1315 (2002).
109. Ohhashi, Y., Hasegawa, K., Naiki, H. & Goto, Y. Optimum amyloid fibril formation of a peptide fragment suggests the amyloidogenic preference of  $\beta$ 2-microglobulin under physiological conditions. *J. Biol. Chem.* **279**, 10814–10821 (2004).
110. Wadai, H. *et al.* Stereospecific amyloid-like fibril formation by a peptide fragment of  $\beta$ 2-microglobulin. *Biochemistry* **44**, 157–164 (2005).
111. Decatur, S. M. Elucidation of residue-level structure and dynamics of polypeptides via isotope-edited infrared spectroscopy. *Acc. Chem. Res.* **39**, 169–175 (2006).
112. Welch, W. R. W., Kubelka, J. & Keiderling, T. A. Infrared, vibrational circular dichroism, and raman spectral simulations for  $\beta$ -sheet structures with various isotopic labels, interstrand, and stacking arrangements using density functional theory. *J. Phys. Chem. B* **117**, 10343–10358 (2013).
113. Gosal, W. S. *et al.* Competing pathways determine fibril morphology in the self-assembly of  $\beta$ 2-microglobulin into amyloid. *J. Mol. Biol.* **351**, 850–864 (2005).

114. Debelouchina, G. T., Platt, G. W., Bayro, M. J., Radford, S. E. & Griffin, R. G. Intermolecular alignment in  $\beta$ 2-microglobulin amyloid fibrils. *J. Am. Chem. Soc.* **132**, 17077–17079 (2010).
115. Torbeev, V. Y. & Hilvert, D. Both the cis-trans equilibrium and isomerization dynamics of a single proline amide modulate  $\beta$ 2-microglobulin amyloid assembly. *Proc. Natl. Acad. Sci.* **110**, 20051–20056 (2013).
116. Debelouchina, G. T., Platt, G. W., Bayro, M. J., Radford, S. E. & Griffin, R. G. Magic angle spinning NMR analysis of  $\beta$ 2-microglobulin amyloid fibrils in two distinct morphologies. *J. Am. Chem. Soc.* **132**, 10414–10423 (2010).
117. Pauling, L. & Corey, R. B. Configurations of polypeptide chains with favored orientations around single bonds two new pleated sheets. *Proc. Natl. Acad. Sci.* **37**, 729–740 (1951).
118. Pauling, L. & Corey, R. B. Two rippled-sheet configurations of polypeptide chains, and a note about the pleated sheets. *Proc. Natl. Acad. Sci.* **39**, 253–256 (1953).
119. Jacques, J., Collet, A. & Wilen, S. H. *Enantiomers, racemates, and resolutions*. Krieger Publ: Malabar, Florida, 1994.
120. Tycko, R. Amyloid polymorphism: Structural basis and neurobiological relevance. *Neuron* **86**, 632–645 (2015).
121. Morris, V. K. *et al.* Solid-state NMR spectroscopy of functional amyloid from a fungal hydrophobin: A well-ordered  $\beta$ -sheet core amidst structural heterogeneity. *Angew. Chem. Int. Ed.* **51**, 12621–12625 (2012).
122. Tycko, R. Solid-state NMR studies of amyloid fibril structure. *Annu. Rev. Phys. Chem.* **62**, 279–299 (2011).
123. Lopez del Amo, J. M. *et al.* An asymmetric dimer as the basic subunit in alzheimer's disease amyloid  $\beta$  fibrils. *Angew. Chem. Int. Ed.* **51**, 6136–6139 (2012).
124. Schütz, A. K. *et al.* Atomic-resolution three-dimensional structure of amyloid  $\beta$  fibrils bearing the Osaka mutation. *Angew. Chem. Int. Ed.* **54**, 331–335 (2015).

125. Ahmed, M. *et al.* Structural conversion of neurotoxic amyloid- $\beta$ 1-42 oligomers to fibrils. *Nat. Struct. Mol. Biol.* **17**, 561–567 (2010).
126. Colvin, M. T. *et al.* Atomic resolution structure of monomorphic A $\beta$ 42 amyloid fibrils. *J. Am. Chem. Soc.* **138**, 9663–9674 (2016).
127. Wälti, M. A. *et al.* Atomic-resolution structure of a disease-relevant A $\beta$ (1–42) amyloid fibril. *Proc. Natl. Acad. Sci.* **113**, E4976–E4984 (2016).
128. Zhang, R. *et al.* Interprotofilament interactions between Alzheimer's A $\beta$ 1–42 peptides in amyloid fibrils revealed by cryoEM. *Proc. Natl. Acad. Sci.* **106**, 4653–4658 (2009).
129. Tycko, R. & Wickner, R. B. Molecular structures of amyloid and prion fibrils: Consensus versus controversy. *Acc. Chem. Res.* **46**, 1487–1496 (2013).
130. Yamaguchi, S. *et al.* Surface dynamics of bacteriorhodopsin as revealed by  $^{13}\text{C}$  NMR studies on [ $^{13}\text{C}$ ] Ala-labeled proteins: Detection of millisecond or microsecond motions in interhelical loops and C-terminal  $\alpha$ -helix. *J. Biochem.* **129**, 373–382 (2001).
131. Decatur, S. M. Elucidation of residue-level structure and dynamics of polypeptides via isotope-edited infrared spectroscopy. *Acc. Chem. Res.* **39**, 169–175 (2006).
132. Paz, M. L. de la & Serrano, L. Sequence determinants of amyloid fibril formation. *Proc. Natl. Acad. Sci.* **101**, 87–92 (2004).
133. West, M. W. *et al.* De novo amyloid proteins from designed combinatorial libraries. *Proc. Natl. Acad. Sci.* **96**, 11211–11216 (1999).
134. Huang, R. K. *et al.* Conformational switching in polyGln amyloid fibrils resulting from a single amino acid insertion. *Biophys. J.* **106**, 2134–2142 (2014).
135. Sciarretta, K. L., Gordon, D. J., Petkova, A. T., Tycko, R. & Meredith, S. C. A $\beta$ 40-lactam(D23/K28) models a conformation highly favorable for nucleation of amyloid. *Biochemistry* **44**, 6003–6014 (2005).
136. Yu, L. *et al.* Structural characterization of a soluble amyloid  $\beta$ -peptide oligomer. *Biochemistry* **48**, 1870–1877 (2009).

137. Lendel, C. *et al.* A Hexameric peptide barrel as building block of amyloid- $\beta$  protofibrils. *Angew. Chem. Int. Ed.* **53**, 12756–12760 (2014).
138. Sandberg, A. *et al.* Stabilization of neurotoxic alzheimer amyloid- $\beta$  oligomers by protein engineering. *Proc. Natl. Acad. Sci.* **107**, 15595–15600 (2010).
139. Hård, T. Protein engineering to stabilize soluble amyloid  $\beta$ -protein aggregates for structural and functional studies. *FEBS J.* **278**, 3884–3892 (2011).
140. Tu, L.-H. & Raleigh, D. P. Role of aromatic interactions in amyloid formation by islet amyloid polypeptide. *Biochemistry* **52**, 333–342 (2013).
141. Armstrong, A. H., Chen, J., McKoy, A. F. & Hecht, M. H. Mutations that replace aromatic side chains promote aggregation of the Alzheimer's A $\beta$  peptide. *Biochemistry* **50**, 4058–4067 (2011).
142. Gazit, E. Mechanisms of amyloid fibril self-assembly and inhibition. *FEBS J.* **272**, 5971–5978 (2005).
143. Van Melckebeke, H. *et al.* Atomic-resolution three-dimensional structure of HET-s(218–289) amyloid fibrils by solid-state NMR spectroscopy. *J. Am. Chem. Soc.* **132**, 13765–13775 (2010).
144. Tuttle, M. D. *et al.* Solid-state NMR structure of a pathogenic fibril of full-length human  $\alpha$ -synuclein. *Nat. Struct. Mol. Biol.* **23**, 409 (2016).
145. Fitzpatrick, A. W. P. *et al.* Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **547**, 185 (2017).
146. Gremer, L. *et al.* Fibril structure of amyloid- $\beta$ (1–42) by cryo-electron microscopy. *Science* **358**, 116–119 (2017).
147. Kaye, R. & Lasagna-Reeves, C. A. Molecular mechanisms of amyloid oligomers toxicity. *J. Alzheimers Dis.* **33**, S67–S78 (2013).
148. Guerrero-Muñoz, M. J., Castillo-Carranza, D. L. & Kaye, R. Therapeutic approaches against common structural features of toxic oligomers shared by multiple amyloidogenic proteins. *Biochem. Pharmacol.* **88**, 468–478 (2014).

149. Hayden, E. Y., Conovaloff, J. L., Mason, A., Bitan, G. & Teplow, D. B. Preparation of pure populations of covalently stabilized amyloid  $\beta$ -protein oligomers of specific sizes. *Anal. Biochem.* **518**, 78–85 (2017).
150. Ono, K., Condrón, M. M. & Teplow, D. B. Structure–neurotoxicity relationships of amyloid  $\beta$ -protein oligomers. *Proc. Natl. Acad. Sci.* **106**, 14745–14750 (2009).
151. Zhang, S., Fox, D. M. & Urbanc, B. Insights into formation and structure of A $\beta$  oligomers cross-linked via tyrosines. *J. Phys. Chem. B* **121**, 5523–5535 (2017).
152. Shinoda, K., Sohma, Y. & Kanai, M. Synthesis of chemically-tethered amyloid- $\beta$  segment trimer possessing amyloidogenic properties. *Bioorg. Med. Chem. Lett.* **25**, 2976–2979 (2015).
153. Dolphin, G. T., Dumy, P. & Garcia, J. Control of amyloid  $\beta$ -peptide protofibril formation by a designed template assembly. *Angew. Chem. Int. Ed.* **45**, 2699–2702 (2006).
154. Wuttke, A., Fischer, S. N., Nebel, A., Marsch, M. & Geyer, A. Improved  $\delta$ -valerolactam templates for the assembly of A $\beta$ -mini-amyloids by boronic ester formation. *Org. Biomol. Chem.* **14**, 5032–5048 (2016).
155. Kokkoni, N., Stott, K., Amijee, H., Mason, J. M. & Doig, A. J. *N*-methylated peptide inhibitors of  $\beta$ -amyloid aggregation and toxicity. Optimization of the inhibitor structure. *Biochemistry* **45**, 9906–9918 (2006).
156. Chatterjee, J., Rechenmacher, F. & Kessler, H. *N*-Methylation of peptides and proteins: An important element for modulating biological functions. *Angew. Chem. Int. Ed.* **52**, 254–269 (2013).
157. Zhang, J. & Germann, M. W. Characterization of secondary amide peptide bonds isomerization: Thermodynamics and kinetics from 2D NMR spectroscopy. *Biopolymers* **95**, 755–762 (2011).
158. Piriou, F. *et al.* Amino acid side chain conformation in angiotensin II and analogs: Correlated results of circular dichroism and <sup>1</sup>H nuclear magnetic resonance. *Proc. Natl. Acad. Sci.* **77**, 82–86 (1980).
159. Åslund, A. *et al.* Novel pentameric thiophene derivatives for in vitro and in vivo optical imaging of a plethora of protein aggregates in cerebral amyloidoses. *ACS Chem. Biol.* **4**, 673–684 (2009).

160. Cohen, F. E. & Kelly, J. W. Therapeutic approaches to protein-misfolding diseases. *Nature* **426**, 905-909 (2003).
161. Woods, L. A. *et al.* Ligand binding to distinct states diverts aggregation of an amyloid-forming protein. *Nat. Chem. Biol.* **7**, 730–739 (2011).
162. Hou, X., Aguilar, M.-I. & Small, D. H. Transthyretin and familial amyloidotic polyneuropathy. *FEBS J.* **274**, 1637–1650 (2007).
163. Razavi, H. *et al.* Benzoxazoles as transthyretin amyloid fibril inhibitors: Synthesis, evaluation, and mechanism of action. *Angew. Chem. Int. Ed.* **42**, 2758–2761 (2003).
164. Connelly, S., Choi, S., Johnson, S. M., Kelly, J. W. & Wilson, I. A. Structure-based design of kinetic stabilizers that ameliorate the transthyretin amyloidoses. *Curr. Opin. Struct. Biol.* **20**, 54–62 (2010).
165. Ray, S. S., Nowak, R. J., Brown, R. H. & Lansbury, P. T. Small-molecule-mediated stabilization of familial amyotrophic lateral sclerosis-linked superoxide dismutase mutants against unfolding and aggregation. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 3639–3644 (2005).
166. Nowak, R. J., Cuny, G. D., Choi, S., Lansbury, P. T. & Ray, S. S. Improving binding specificity of pharmacological chaperones that target mutant superoxide dismutase-1 linked to familial amyotrophic lateral sclerosis using computational methods. *J. Med. Chem.* **53**, 2709–2718 (2010).
167. Conway, K. A., Rochet, J.-C., Bieganski, R. M. & Lansbury, P. T. Kinetic stabilization of the  $\alpha$ -synuclein protofibril by a dopamine- $\alpha$ -synuclein adduct. *Science* **294**, 1346–1349 (2001).
168. Bartels, T., Choi, J. G. & Selkoe, D. J.  $\alpha$ -Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature* **477**, 107-110 (2011).
169. Ladiwala, A. R. A. *et al.* Rational design of potent domain antibody inhibitors of amyloid fibril assembly. *Proc. Natl. Acad. Sci.* **109**, 19965–19970 (2012).
170. Sevigny, J. *et al.* The antibody aducanumab reduces A $\beta$  plaques in Alzheimer's disease. *Nature* **537**, 50–56 (2016).

171. Sciarretta, K. L., Gordon, D. J. & Meredith, S. C. Peptide-based inhibitors of amyloid assembly. in *Methods in Enzymology* **413**, 273–312 (2006).
172. Tjernberg, L. O. *et al.* Arrest of amyloid fibril formation by a pentapeptide ligand. *J. Biol. Chem.* **271**, 8545–8548 (1996).
173. Soto, C. *et al.*  $\beta$ -sheet breaker peptides inhibit fibrillogenesis in a rat brain model of amyloidosis: Implications for Alzheimer's therapy. *Nat. Med.* **4**, 822–826 (1998).
174. Ghanta, J., Shen, C.-L., Kiessling, L. L. & Murphy, R. M. A strategy for designing inhibitors of  $\beta$ -amyloid toxicity. *J. Biol. Chem.* **271**, 29525–29528 (1996).
175. Moriarty, D. F. & Raleigh, D. P. Effects of sequential proline substitutions on amyloid formation by human amylin<sub>20-29</sub>. *Biochemistry* **38**, 1811–1818 (1999).
176. Sievers, S. A. *et al.* Structure-based design of non-natural amino-acid inhibitors of amyloid fibril formation. *Nature* **475**, 96-100 (2011).
177. Kumar, J. & Sim, V. D-amino acid-based peptide inhibitors as early or preventative therapy in Alzheimer disease. *Prion* **8**, 119–124 (2014).
178. Yan, L.-M., Tarek-Nossol, M., Velkova, A., Kazantzis, A. & Kapurniotu, A. Design of a mimic of nonamyloidogenic and bioactive human islet amyloid polypeptide (IAPP) as nanomolar affinity inhibitor of IAPP cytotoxic fibrillogenesis. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 2046–2051 (2006).
179. Madine, J., Doig, A. J. & Middleton, D. A. Design of an N-methylated peptide inhibitor of  $\alpha$ -synuclein aggregation guided by solid-state NMR. *J. Am. Chem. Soc.* **130**, 7873–7881 (2008).
180. Chatterjee, J., Gilon, C., Hoffman, A. & Kessler, H. N-methylation of peptides: A new perspective in medicinal chemistry. *Acc. Chem. Res.* **41**, 1331–1342 (2008).
181. Gordon, D. J., Tappe, R. & Meredith, S. C. Design and characterization of a membrane permeable N-methyl amino acid-containing peptide that inhibits A $\beta$ <sub>1–40</sub> fibrillogenesis. *J. Pept. Res.* **60**, 37–55 (2002).

# Experimental part

## **7.1 Materials and methods**

**Materials:** All solvents, chemicals, and reagents were purchased from commercial sources and used without further purification. Coupling reagents, Fmoc- $\alpha$ -L-amino acids, Fmoc-N-methyl-amino acids and resins for SPPS were purchased from Aapptec (Louisville, KY), Bachem (Bubendorf, Switzerland), or Iris Biotech (Marktredwitz, Germany); DTT, acetonitrile, DCM, DMF, NMP, DIPEA, TIPS, diethyl ether, piperidine, pyridine, bromoacetic acid, allyl chloroformate, methanol, Fmoc-chloride, 1,4-dioxane, MPAA, MesNa, TCEP, bromoacetamide, acetic anhydride and Boc anhydride were from Sigma-Aldrich; TFA (BioGrade) was from Halocarbon (Peachtree Corners, GA). [1- $^{13}$ C]phenylalanine, [1- $^{13}$ C]valine, [1- $^{13}$ C]glycine, [1- $^{13}$ C]leucine,  $^{15}$ N-phenylalanine,  $^{15}$ N-valine,  $^{15}$ N-glycine,  $^{15}$ N-leucine and  $^{15}$ N-isoleucine were purchased from Cambridge Isotope Laboratories (Tewksbury, MA) and Fmoc-protected in-house. All other chemicals were purchased from Roth, VWR, Sigma-Aldrich or Alfa Aesar.

**Peptide synthesis:** Peptides were prepared manually on a 0.2 mmol scale or by machine-assisted synthesis on either 0.1 or 0.2 mmol scale by Fmoc/ $^t$ Bu solid-phase peptide synthesis (Fmoc-SPPS; side-chain protection: Asn(Trt), Arg(Pbf), Asp(O $^t$ Bu), Cys(Trt), Gln(Trt), Glu(O $^t$ Bu), His(Trt), Lys(Boc), Ser( $^t$ Bu), Trp(Boc), Tyr( $^t$ Bu)). Prior to cleavage, peptidyl resin was washed with DCM and dried under vacuum. Peptides were cleaved and fully deprotected by treatment with TFA/TIPS/H $_2$ O (95:2.5:2.5, v/v/v) or TFA/DTT/TIPS/H $_2$ O (92.5:2.5:2.5:2.5, v/w/v/v) for 2 h (1 mL of cleavage reagent for 100 mg of resin). The reaction mixtures were diluted 20-fold with diethyl ether (-20°C) for precipitation of the crude peptides. After centrifugation, the precipitates were dissolved in H $_2$ O/CH $_3$ CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized.

**Analytical HPLC:** Analytical reversed-phase HPLC was performed on a Nexera XR UHPLC instrument (Shimadzu, Kyoto, Japan) equipped with LC-20AD liquid chromatograph modules and an SPD-M20A Prominence diode array detector. A Phenomenex Kinetex XB-C18 column

(50×2.1 mm, 100 Å, 2.6 µm) was used at 0.5 or 1 mL/min with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v).

**Preparative HPLC:** Purifications of peptides were performed on a preparative HPLC instrument (Shimadzu) equipped with two LC-20AP pumps and an SPD-20A Prominence UV/vis detector connected to an FRC-10A fraction collector. Phenomenex Kinetex XB-C18 columns (100×21.2 mm or 250×21.2 mm, 100 Å, 5 µm) were used at 10 mL/min with a gradient of water with TFA (0.1%) and acetonitrile with TFA (0.08%).

**LC-MS:** Peptides masses were determined on an LC/MS instrument equipped with an Accela UHPLC (Thermo Fisher Scientific; Hypers II GOLD column, 50×2.1 mm, 1.9 µm) integrated with an LCQ Fleet ion-trap (Thermo Fisher Scientific). Deconvolution of data was performed in MagTran 1.03 (Amgen, Thousand Oaks, CA).

More precise MS measurements for larger protein product have been performed using Thermo Scientific Q Exactive hybrid quadrupole-Orbitrap mass-spectrometer. Direct infusion into electrospray ionization source was performed from solutions of proteins in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing formic acid (0.1%, v/v).

**Size-exclusion chromatography:** SEC was performed on a Nexera XR UHPLC instrument (Shimadzu, Kyoto, Japan) equipped with LC-20AD liquid chromatograph modules and an SPD-M20A Prominence diode array detector. A Phenomenex Yarra 3µm SEC 2000 column was used (300×4.6 mm, 2.8 µm) with an isocratic solvent system composed of NaH<sub>2</sub>PO<sub>4</sub> (10 mM), Na<sub>2</sub>HPO<sub>4</sub> (40 mM), NaCl (100 mM) and NaN<sub>3</sub> (0.05%, w/w), pH 7.5, at a flow of 0.5 mL/min, at 27°C, detected at 280 nm. A gel filtration markers kit for protein molecular weights (6,500-66,000 Da) was used for calibration. The kit was composed of Blue Dextran (for the determination of the void volume), aprotinin from bovine lung (6.5 kDa), cytochrome *c* from equine heart (12.4 kDa), carbonic anhydrase from bovine erythrocytes (29 kDa) and albumin from bovine serum (66 kDa) and was purchased from Sigma-Aldrich.

**NMR spectroscopy:** Liquid <sup>1</sup>H NMR analysis for modified amino acids, linkers, and their synthetic intermediates were performed at room temperature on a 400 MHz Avance spectrometer (Bruker). The data were processed with TopSpin 2.0 (Bruker) and analyzed with MestReNova 10.0 (Mestrelab, Santiago de Compostela, Spain). NMR of Inh-1-trimer and Inh-2-trimer were

performed on a 600 MHz Avance spectrometer (Bruker) and NMR of TofT-3 construct on a 700 MHz Avance spectrometer (Bruker).

**Amyloid growth experiments:** Peptides stock solutions were prepared in DMSO at 10 mM (monomers), 5 mM (covalent dimer), or 3.33 mM (covalent trimer). Typically, stock solution (5  $\mu$ L) was diluted in buffer (495  $\mu$ L; buffer 1: Tris·HCl (50 mM, pH 7.5), NaCl (100 mM), NaN<sub>3</sub> (0.05%, w/v); buffer 2: sodium phosphate (50 mM, pH 7.5), NaCl (100 mM), NaN<sub>3</sub> (0.05%, w/v)) in an Eppendorf tube. Experiments showed that buffer composition did not affect peptide self-assembly into amyloids. Samples were incubated in the dark at room temperature or at 37°C, and amyloid growth was monitored by ThT fluorescence and by HPLC. Using HPLC to estimate the remaining peptide in solution, we found that conversion of [20-41] $\beta$ 2m monomer, covalent dimer, and covalent trimer into was 84, 89, and 96%, respectively, after 20 days of incubation.

**Kinetics of amyloid growth:** Continuous amyloid growth assays were performed on a Victor X5 2030-0050 multi-label HTS plate-reader (PerkinElmer) with Nunc black standard 96-well plates sealed with transparent ThermalSeal RT film (Sigma-Aldrich). Fluorescence readings were taken at 27-30°C every 5 min, with shaking for 1 s between readings; reactions were monitored for up to five days. Excitation/emission wavelengths were set with optical filters (ThT: 485 nm/535 nm; bis-ANS: 405 nm/460 nm). Each well contained ThT dye (50  $\mu$ M) with covalent trimer (33  $\mu$ M), covalent dimer (50  $\mu$ M), or monomeric [20-41] $\beta$ 2m (100  $\mu$ M) (or conjugate) in buffer (100  $\mu$ L; phosphate (50 mM, pH 7.5), NaCl (100 mM), NaN<sub>3</sub> (0.05%, w/v)). Peptides were added by 100-fold dilution of stock solutions in DMSO.

For seeding experiments, amyloids (100  $\mu$ L suspensions) of covalent trimer or dimer were submitted to three freeze-thaw cycles (45 min at -20°C, 30 min at room temperature). Then, the fragmented amyloids were added to monomeric [20-41] $\beta$ 2m peptide (100  $\mu$ M) in buffer containing ThT (50  $\mu$ M), so that the concentration of seeds represented either 20 or 10% of total peptide. Experiments were done in triplicate with a plate reader.

Complementary continuous assays for amyloid growth kinetics followed by turbidity measurements were performed on a CLARIOstar plate reader (BMG Labtech, Ortenberg, Germany) with standard 96-well plates (Greiner) sealed with transparent film (Thermo Fisher Scientific). Absorbance (450 or 635 nm) was measured for three or five replicates every 5 min

with shaking for 10 s (600 rpm) before each cycle, for up to three days at 25°C. Each well contained covalent trimer (33  $\mu$ M), covalent dimer (50  $\mu$ M), or monomeric [20-41] $\beta$ 2m peptide (100  $\mu$ M) (or conjugate) in buffer (100  $\mu$ L, phosphate (50 mM, pH 7.5), NaCl (100 mM), NaN<sub>3</sub> (0.05%, w/v)). Peptides were added by 100-fold dilution of stock solutions in DMSO.

Time point measurements by UV/Vis absorption of the earlier stages of amyloid growth were performed at 27°C on a V-670 spectrophotometer (JASCO) with a PS semi-micro cuvette (1.0 cm $\times$ 0.4 cm). Absorption (635 nm) was recorded every 15 min without shaking or pipetting. Each cuvette contained covalent trimer (33  $\mu$ M), covalent dimer (50  $\mu$ M), or monomeric [20-41] $\beta$ 2m peptide (100  $\mu$ M) in buffer (1 mL, phosphate (50 mM, pH 7.5), NaCl (100 mM), NaN<sub>3</sub> (0.05%, w/v)). Peptides were added by 100-fold dilution of stock solutions in DMSO.

Time-point measurements of amyloid growth by ThT fluorescence were performed at 27°C on a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies) with a PS semi-micro cuvette (1.0 cm $\times$ 0.4 cm). Fluorescence (442/487 nm) was recorded every 15 min without shaking or pipetting. Each cuvette contained ThT (50  $\mu$ M) and covalent trimer (33  $\mu$ M), covalent dimer (50  $\mu$ M), or monomeric [20-41] $\beta$ 2m peptide (100  $\mu$ M) in buffer (1 mL, phosphate (50 mM, pH 7.5), NaCl (100 mM), NaN<sub>3</sub> (0.05%, w/v)). Peptides were added by 100-fold dilution of stock solutions in DMSO. The photographs at different time points were taken with an EOS 100D camera (Canon) with an EF-S IS STM lens (18-135 mm, f3.5-5.6).

**Fluorescence spectroscopy:** Fluorescence spectra were recorded on a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies) with a quartz cuvette (1.0 cm $\times$ 0.2 cm). The excitation wavelength was 442 nm (ThT), 400 nm (bis-ANS and ANS), or 554 nm (Nile Red). The emission spectra were collected at 460-650 nm (ThT and bis-ANS), 420-650 nm (ANS), and 575-800 nm (Nile Red). Data are averages of three scans. The dyes (5-200  $\mu$ M) were added to suspensions of amyloid fibers grown for seven days. Peptide concentrations were kept constant at 33.3  $\mu$ M (covalent trimer), 50  $\mu$ M (covalent dimer), or 100  $\mu$ M (monomer) in sodium phosphate (50 mM, pH 7.5) with NaCl (100 mM) and NaN<sub>3</sub> (0.05%, w/v).

**UV/Vis spectroscopy:** Absorption spectra of [20-41] $\beta$ 2m-(Ellman's thiol) conjugate were recorded on a V-670 spectrophotometer (JASCO) with a SUPRASIL quartz cuvette (10 mm $\times$ 2 mm; Hellma Analytics, Müllheim, Germany). Absorption spectra (200-800 nm) were averaged

from three scans. [20-41] $\beta$ 2m-(Ellman's thiol) (100  $\mu$ M) was in sodium phosphate (50 mM, pH 7.5) with NaCl (100 mM) and NaN<sub>3</sub> (0.05%, w/v).

**Circular dichroism:** CD spectra were recorded on a J-1500 (Jasco) spectrophotometer. All samples were prepared in quartz cuvettes with a thickness of 1 mm and a volume of 300  $\mu$ L. For every measurement proteins were dissolved in a buffered solution (25 mM phosphate, pH 7.4) at a concentration range of 25  $\mu$ M to 75  $\mu$ M. For all scans the following parameters were set: scan range 280-180 nm, band width 1.00 nm, scanning speed 20 nm/min, data pitch 0.1 nm. Every CD curve was obtained by averaging 5 scans and subtracting the background signal. The obtained ellipticities were transformed to mean residue molar ellipticities using the following equation:

$$[\theta][deg \cdot cm^2 \cdot dmol^{-1}] = \frac{\theta [mdeg] \cdot 10^6}{d [mm] \cdot c [\mu M] \cdot (n - 1)}$$

$[\theta]$ : residue molar ellipticity,  $\theta$ : ellipticity,  $d = 1$  mm (path length or thickness of cuvette),  $c$ : protein concentration,  $n$ : amount of amino acids in the protein sequence.

**Transmission electron microscopy:** Carbon Type B (15-25 nm) formvar-coated Cu electron microscopy grids (400 mesh; Ted Pella, Redding, CA) were placed coated-side-down for 60 s on the amyloid sample. The grids were washed with deionized water, stained with uranyl acetate (2%, w/v) for 20 s, and then air-dried. Images were taken on a CM12 electron microscope (80 keV; Philips).

### **Cryogenic transmission electron microscopy**

*Sample preparation:* The amyloid-containing sample was diluted to 3-5  $\mu$ M in Tris (10 mM) with EDTA (1 mM) and NaCl (10 mM). An aliquot (10  $\mu$ L) was deposited onto a clean Parafilm foil. Electron microscopy grids covered with a Quantifoil holey carbon film (R2/2, 300 mesh) were placed for 60 s carbon-coated side down on the amyloid sample. The grid was then flash frozen in liquid ethane by using a Vitrobot II automated plunger (FEI, Hillsboro, OR; blotting time 4 s, blotting force 5, humidity 95%, temperature 20°C).

*Data acquisition:* Fibers of [20-41] $\beta$ 2m covalent trimer were imaged in a Titan Krios transmission electron cryo-microscope (FEI; 300 kV, and a Cs corrector). Images were recorded

under low-dose condition (total dose  $22 \text{ e}^- \text{Å}^{-2}$ ) using a EPU automated data collection software (FEI) on a  $4096 \times 4096$  Falcon II direct detector camera (FEI; magnification 96000; pixel size 0.0675 nm). A total of 17 frames with an electron dose of  $3.2 \text{ e}^- \text{Å}^{-2}$  per frame were collected. Frames 2-8 were aligned by using the optical flow protocol and averaged for further analysis. A high-dose image ( $55 \text{ e}^- \text{Å}^{-2}$ ) was generated by summing all frames, and used as reference for frame alignment.

Fibers of [20-41] $\beta$ 2m-(Ellman's thiol) were imaged on a Tecnai polara F30 transmission electron cryo-microscope (FEI; 100 kV). Images were recorded under low-dose condition (total  $22 \text{ e}^- \text{Å}^{-2}$ ) using a EPU automated data collection software (FEI) and a Falcon I  $4096 \times 4096$  direct detector camera (FEI; magnification 59000; pixel size 0.178 nm).

*Image processing:* Patches of individual amyloid fibers were selected interactively in boxes ( $96 \times 96$  pixels) avoiding overlap between patches. Images of fiber patches were normalized to a null average image density and a constant image variance. Individual images were aligned in rotation by cross-correlation methods in order to orient the fiber axes vertically. The aligned images were averaged, and a mean intensity profile was determined over the whole length of the patch by using Digital Micrograph software (Gatan, Pleasanton, CA). The width at mid-height of the density profile was used as fiber diameter. Two-dimensional image classification was performed after multi statistical analysis and hierarchical ascendant image clustering methods, according to IMAGIC-V software (Imaging Sciences, Berlin, Germany). Average images from individual classes were analyzed to identify local variations in fiber structure.

### **Solid-state NMR spectroscopy**

*Sample preparation:* Labeled peptide (0.65-1.54  $\mu\text{mol}$ ) was dissolved in buffer (sodium phosphate (50 mM, pH 7.5), NaCl (100 mM),  $\text{NaN}_3$  (0.05%, w/v)) to 33.3  $\mu\text{M}$  (trimer construct) in a 50 mL Falcon tube. The tube was incubated in the dark at room temperature for 15 days, and then centrifuged at 12857 g for 30 min. The supernatant was removed, and amyloid hydrogel was lyophilized. The fibers were stored at  $-20^\circ\text{C}$  before loading the sample into the ssNMR spinner.

*Measurements:* Solid-state MAS NMR experiments were performed on an Avance 750 MHz wide-bore spectrometer (Bruker; 188.5 MHz for  $^{13}\text{C}$ ). DP- and CP-MAS  $^{13}\text{C}$  spectra were acquired with a double-resonance MAS probe designed for 2.5 mm o.d. zirconia rotors (closed

with Kel-F caps); samples were spun at 30 kHz. For  $^{13}\text{C}$ , $^1\text{H}$  HETCOR experiment an ultra-fast triple-resonance MAS probe built for 1.3 mm o.d. zirconia rotors (with Kel-F caps) was used; samples were spun at 55 kHz. The RF fields for 1D CP/MAS modified Hartmann-Hahn (HH) conditions were 51 kHz ( $^{13}\text{C}$ ), 81 kHz ( $^1\text{H}$ ); for HETCOR experiments, the HH conditions were 62 kHz ( $^1\text{H}$ ) and 117 kHz ( $^{13}\text{C}$ ). CP contact times were chosen to be shorter for HETCOR (182  $\mu\text{s}$ ) and longer for 1D (2 ms for  $^{13}\text{C}$ ). Proton decoupling during acquisition was obtained by using SPINAL-64 at a 180 kHz RF field for 1D (214 kHz for 2D); the recycle time was 3 s for all CP spectra. Because of the spectral wide lines and in order to get undistorted line shapes we used improved versions of CP and HETCOR experiments that perform Hahn's echo prior to FID acquisition. The echo had to be synchronized with the rotation (echo time= $N$  rotation periods) and was identical for all  $^{13}\text{C}$  spectra (two rotation periods, total echo time 66.67  $\mu\text{s}$  for 1D and 36.37  $\mu\text{s}$  for 2D). 1D  $^{13}\text{C}\{^1\text{H}\}$  DP spectra were acquired directly with the original Hahn's echo sequence, with a recycling delay of 15 s. All 1D  $^{13}\text{C}$  experiments were acquired with 16.384 ms (61 Hz) time resolution with a 62.5 kHz spectral width (330 ppm). Lorentzian line broadening (100 Hz) was applied prior to Fourier transformation. A total of 256 complex data points were acquired in the  $^1\text{H}$  indirect dimension for each  $t_1$  increment, with 64 scans accumulated, thus leading to a spectral resolution of 429 Hz for  $^1\text{H}$  and 61 Hz for  $^{13}\text{C}$ . Lorentzian line broadening (150 Hz) was applied in the direct dimension prior to Fourier transformation; in the proton dimension apodization was applied with a  $90^\circ$ -shifted squared sine-bell function.  $^{13}\text{C}$  chemical shifts were calibrated by the substitution method by referencing the high field peak of adamantane standard (38.2 ppm). For protons in 2D we also used the adamantane peak setting it signal to 1.8 ppm (the mean value for the two protons).

**Electron paramagnetic resonance:** For EPR measurements, 30% (v/v) glycerol was added to the solutions containing amyloids of the nitroxide-labeled [20-41] $\beta$ 2m dimeric constructs. Once transferred to glass capillaries (3 mm Quartz tube), the samples were frozen in liquid nitrogen and the EPR spectra were measured at 120 K using a Bruker ElexSys E500 EPR spectrometer operating at 9.6 GHz. The microwave frequency was 9.55 GHz, with an attenuation of 36 Db, corresponding to a power of 50 mW. Field modulation at 100 kHz was used with amplitude of 0.2 mT. Finally, a time constant of 1.28 ms and a conversion time of 5.12 ms were used. Between 400 and 800 scans were recorded for each sample, depending on the signal-to-noise ratio.

**ATR-FT-IR:** ATR-FT-IR analyses were performed on a Vertex 70 FT-IR spectrometer (Bruker) equipped with a Meridian Diamond ATR (Harrick, Pleasantville, NY). Residual TFA was removed from the samples by HCl treatment (dissolving 1 mg of peptide in 1 mL of solvent) and lyophilization. Amyloid growth was performed in D<sub>2</sub>O equilibrated samples. Each sample was dried under argon and measured three times to obtain 256 scans after background measurement (4000-700 cm<sup>-1</sup>, resolution 4 cm<sup>-1</sup>). Only the third measurement was considered, and it was subjected to water-compensation in OPUS 6.5 software (Bruker Optik). Spectrum deconvolution was performed in OriginPro 8.5 with the peak analyzer tools.

**Molecular modeling:** Molecular graphics and analyses were performed using MOE software. Model 1 from PDB ID: 2E8D (solid-state NMR structure of [20-41]β<sub>2</sub>m amyloid protofilament) was selected. A stack of four β-arch-containing units was built along the fiber axis. Energies of the models were minimized by using the minimization force field AMBER10:EHT, with R-Field as a solvation mode. Minimizations were done stepwise: first, backbones were fixed and models were minimized, thus allowing side-chain reorganization; then backbones were tethered, and a second minimization afforded reorganization of peptide backbone and side chains simultaneously; finally, models were completely freed and minimized.

## 7.2 Chapter 2

### Synthesis of [20-41]β<sub>2</sub>m

Sequence: SNFLNCYVSGFHPSDIEVDLLK

[20-41]β<sub>2</sub>m was synthesized on a 0.2 mmol scale on Fmoc-Lys(Boc)-Wang resin (294 mg, loading : 0.68 mmol/g, mesh size: 180-333). Automated synthesis was carried out by using HATU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 930 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was

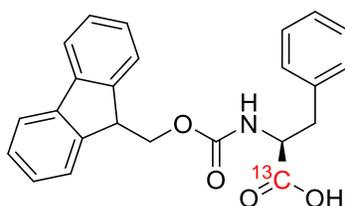
480 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 208 mg (42%). LC-MS (ESI):  $[M + 2H^+] = 1249.6 m/z$ ,  $[M + 3H^+] = 833.6 m/z$ ;  $C_{113}H_{169}N_{27}O_{35}S$ ; Average isotope calculated 2497.8 Da  $[M]$ ; found: 2497.4 Da.

### Synthesis of *D*-[20-41]β2m-amide

Sequence: snflncyvsgfhpsdievdlk- $\alpha$ CONH<sub>2</sub>

*D*-[20-41]β2m-amide was synthesized on a 0.1 mmol scale on Fmoc-Rink Amid resin (135 mg, loading: 0.74 mmol/g, mesh size: 100-200 µm). Automated synthesis was carried out by using HATU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 336 mg. The peptide was cleaved from the resin by treatment with 4 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 61 mg (24.5%). LC-MS (ESI):  $[M + 2H^+] = 1249.1 m/z$ ,  $[M + 3H^+] = 833.1 m/z$ ;  $C_{113}H_{170}N_{28}O_{34}S$ ; Average isotope calculated 2496.8 Da  $[M]$ ; found: 2496.2 Da.

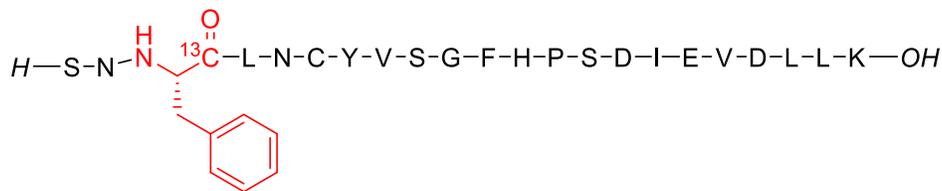
### Synthesis of Fmoc-[1-<sup>13</sup>C]Phe-OH



*L*-[1-<sup>13</sup>C]phenylalanine (1.0 g, 6.1 mmol, 1 equiv) and sodium carbonate (1.8 g, 17.4 mmol, 2.8 equiv) were dissolved in 80 mL of water. A solution of Fmoc *N*-hydroxysuccinimide ester (5 g, 15 mmol, 2.5 equiv) in 70 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react at r.t. for 17 h. The reaction mixture was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub>. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 2.1 g (89%). LC-MS (ESI): Average isotope calculated 388.4 Da [*M*]; found: 388.6 Da.

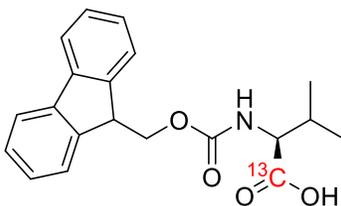
### Synthesis of [1-<sup>13</sup>C]Phe23-[21-40]β2m

Sequence: SN**F**LNCYVSGFHPSDIEVDLLK



[1-<sup>13</sup>C]Phe23-[21-40]β2m was synthesized on a 0.2 mmol scale on Fmoc-Lys(Boc)-Wang resin (303 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). [1-<sup>13</sup>C]Phe at position 23 was coupled manually (2 equiv, 1h). The mass of the dried peptidyl-resin at the end of the synthesis was 1.02 g. The peptide was cleaved from the resin by treatment with 10 mL of TFA/TIPS/H<sub>2</sub>O (95/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 444 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 111 mg (22%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1250.0 *m/z*, [*M* + 3H<sup>+</sup>] = 833.8 *m/z*; C<sub>113</sub>H<sub>169</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2498.8 Da [*M*]; found: 2498.3 Da.

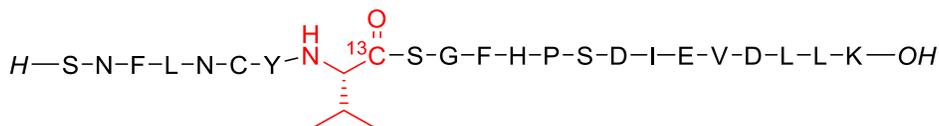
## Synthesis of Fmoc-[1-<sup>13</sup>C]Val-OH



*L*-[1-<sup>13</sup>C]valine (1.0 g, 8.5 mmol, 1 equiv) and sodium carbonate (1.8 g, 17.4 mmol, 2 equiv) were dissolved in 80 mL of water. A solution of Fmoc *N*-hydroxysuccinimide ester (7.2 g, 21.4 mmol, 2.5 equiv) in 70 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react at r.t. for 18 h. The reaction mixture was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub>. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 2.4 g (81%). <sup>1</sup>H NMR (400 MHz, [D<sub>6</sub>]DMSO): δ, ppm = 12.48 (s, 1H), 7.90 (d, 2H), 7.76 (d, 2H), 7.60 (d, 1H), 7.42 (t, 2H), 7.33 (t, 2H), 4.25 (m, 3H), 3.88 (dt, 1H), 2.08 (m, 1H), 0.91 (d, 6H); C<sub>20</sub>H<sub>21</sub>NO<sub>4</sub>; LC-MS (ESI): Average isotope calculated 340.4 Da [*M*]; found: 340.6 Da.

## Synthesis of [1-<sup>13</sup>C]Val28-[20-41]β2m

Sequence: SNFLNCYVSGFHPSDIEVDLLK



[1-<sup>13</sup>C]Val28-[20-41]β2m was synthesized on a 0.2 mmol scale on Fmoc-Lys(Boc)-Wang resin (303 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). [1-<sup>13</sup>C]Val at position 28 was coupled manually (2 equiv, 1h). The mass of the dried peptidyl-resin at the end of the synthesis was 1.03 g. The peptide was cleaved from the resin by treatment with 10 mL of TFA/TIPS/H<sub>2</sub>O (95/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v)

containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 457 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 144 mg (29%). LC-MS (ESI):  $[M + 2H^+] = 1250.0$   $m/z$ ,  $[M + 3H^+] = 833.7$   $m/z$ ;  $C_{113}H_{169}N_{27}O_{35}S$ ; Average isotope calculated 2498.8 Da  $[M]$ ; found: 2498.0 Da.

### Synthesis of disulfide linked [20-41]β2m homodimer

[20-41]β2m (16.3 mg, 6.5 µmol) were dissolved in 25 mL of buffer ( $[Na_2HPO_4] = 0.1$  M,  $[Gn \cdot HCl] = 7$  M,  $[Na_4 \cdot EDTA] = 5$  mM, pH = 7.0). The pH was adjusted to 9.0 after peptide dissolution with a 6 M  $NaOH_{(aq)}$ . It was let to react at 30°C for 4 days. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 9.1 mg (36%). LC-MS (ESI):  $[M + 3H^+] = 1665.5$   $m/z$ ,  $[M + 4H^+] = 1249.2$   $m/z$ ;  $[M + 5H^+] = 999.6$   $m/z$ ,  $[M + 6H^+] = 833.2$   $m/z$ ;  $C_{226}H_{336}N_{54}O_{70}S_2$ ; Average isotope calculated 4991.6 Da  $[M]$ ; found: 4992.5 Da.

The corresponding  $^{13}C$ -labeled covalent trimers were synthesized according to the same protocol on a scale of several tens of milligrams in order to have sufficient quantities to perform FT-IR and ssNMR studies.

**Disulfide linked [1- $^{13}C$ ]Phe23-[20-41]β2m homodimer:** LC-MS (ESI):  $[M + 3H^+] = 1665.7$   $m/z$ ,  $[M + 4H^+] = 1249.5$   $m/z$ ;  $[M + 5H^+] = 999.8$   $m/z$ ,  $[M + 6H^+] = 833.4$   $m/z$ ;  $C_{226}H_{336}N_{54}O_{70}S_2$ ; Average isotope calculated 4993.6 Da  $[M]$ ; found: 4994.1 Da.

**Disulfide linked [1- $^{13}C$ ]Val28-[20-41]β2m homodimer:** LC-MS (ESI):  $[M + 3H^+] = 1665.8$   $m/z$ ,  $[M + 4H^+] = 1249.5$   $m/z$ ;  $[M + 5H^+] = 999.8$   $m/z$ ,  $[M + 6H^+] = 833.4$   $m/z$ ;  $C_{226}H_{336}N_{54}O_{70}S_2$ ; Average isotope calculated 4993.6 Da  $[M]$ ; found: 4994.3 Da.

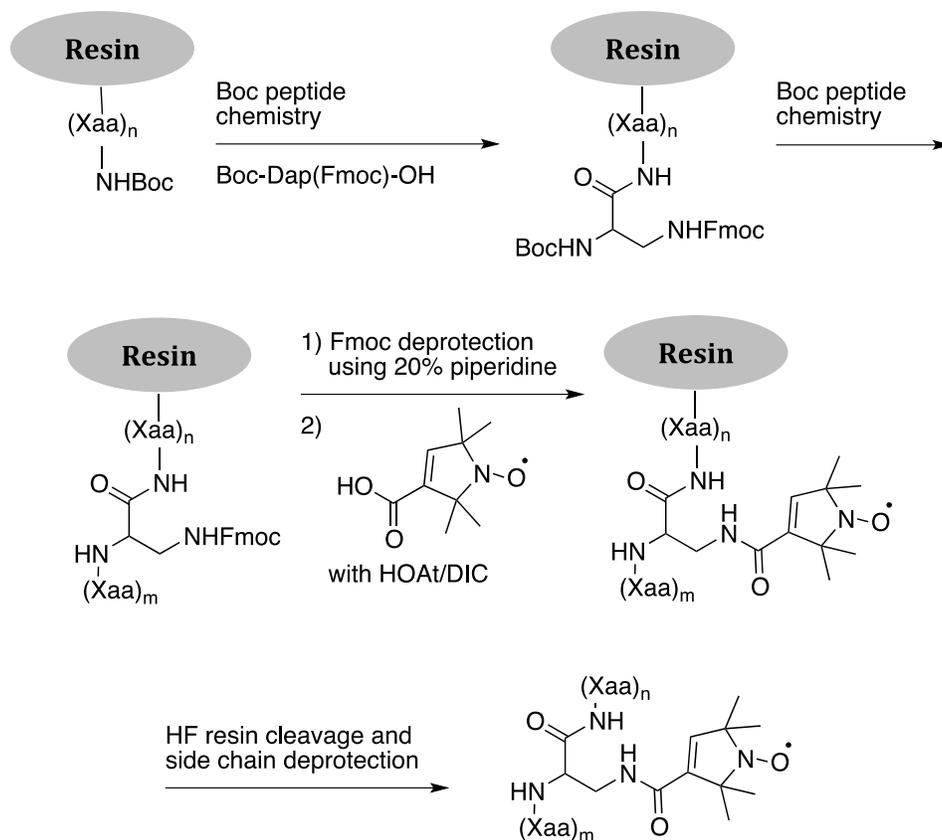
Synthesis done prior to my arrival at the laboratory (\*)

Synthesis of \*disulfide linked L/D-[20-41]β2m heterodimer

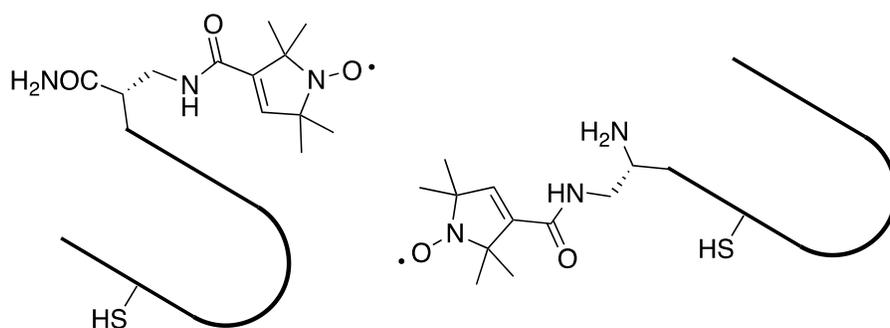
The synthesis of the L/D-[20-41]β2m heterodimer was accomplished by a two-step procedure. In the first step, an activated mixed disulfide L-[20-41]β2m-Cys(S-S-Ar) was synthesized by reaction of reduced peptide with 5,5'-dithiobis-(2-nitrobenzoic acid) (Ellman's reagent) (see **Figure 2-3** in Chapter 2 on page 53). The L-[20-41]β2m peptide (conc. 0.6 mg/mL) was dissolved in a buffer of 7 M Gn·HCl, 0.1 M phosphate at pH 7.1, and a 2-fold excess of Ellman's reagent was added. The progress of the reaction was monitored by LC-MS and analytical HPLC. Formation of the L-[20-41]β2m-Cys(S-S-Ar) mixed disulfide was complete within 2.5 h, and the product purified by preparative reverse-phase HPLC. In the second step, L-[20-41]β2m-Cys(S-S-Ar) and an equivalent amount of the reduced D-K3 peptide were dissolved at a concentration of ~ 0.8 mg/mL in 7 M Gn·HCl, 0.1 M phosphate at pH 7.1. Displacement of 2-nitro-5-thiobenzoate by the cysteine thiol group of the D-[20-41]β2m peptide afforded L/D-[20-41]β2m heterodimer after overnight incubation (~14 h). The reaction mixture was separated by preparative HPLC to yield pure L/D-[20-41]β2m heterodimer.

Synthesis of \*nitroxide-labeled-[20-41]β2m and the corresponding disulfide linker heterodimer

Synthesis of nitroxide-labeled [20-41]β2m peptides was carried out by the strategy outlined in **Figure S1**. Boc-3-(Fmoc-amino)-L-alanine was coupled as an extra residue at either the N- or C-terminus of the [20-41]β2m peptide sequence using Boc-SPPS (**Figure S2**). After chain assembly was complete, the Fmoc group was removed with 20% piperidine and 2,2,5,5-tetramethyl-3-pyrrolin-1-oxyl-3-carboxylic acid (nitroxide label) was coupled with the DIC/HOAt protocol. Finally, the peptide was cleaved from the resin and simultaneously deprotected by treatment with HF/p-cresol and the product purified by preparative HPLC. Air oxidation (as described above) was used to produce the corresponding N-terminal or C-terminal nitroxide-labeled L/L-[20-41]β2m homodimers.



**Figure S1:** Synthetic strategy showing incorporation of Dap-spin-label (Dap-SL) using the orthogonally protected Boc-3-(Fmoc-amino)-*L*-alanine building block. (Xaa)<sub>n</sub> and (Xaa)<sub>m</sub> refer to the n and m amino acids preceding and following Dap-SL, respectively.



**Figure S2:** Models of Dap-SL-containing [20-41]β<sub>2</sub>m peptides. A C-terminally labeled monomer on the left and an N-terminally labeled monomer on the right. The parent 22-residue β-arch of [20-41]β<sub>2</sub>m is highlighted in bold. The thiol group (SH) shows the position of the cysteine 25 residue.

### Synthesis of \*L-β2m and \*D-β2m

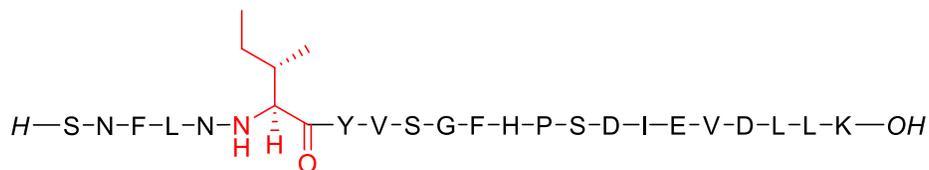
The *D*-enantiomer of β2m was synthesized by ‘one pot’ native chemical ligation of three *D*-peptide fragments according to a previously published protocol for *L*-β2m<sup>1</sup>. The purity of the final product was verified by reverse-phase HPLC and LC-MS.

1. Torbeev, V.Y., Hilvert D. Both the cis-trans equilibrium and isomerization dynamics of a single proline amide modulate β2-microglobulin amyloid assembly. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 18119-18124 (2006).

## 7.3 Chapter 3

### Synthesis of Cys25Ile-[20-41]β2m

Sequence: SNFLN**I**YVSGFHPSDIEVDLLK

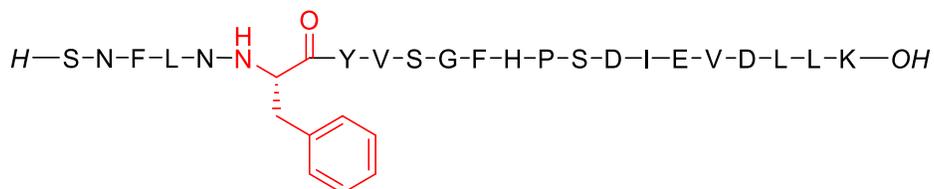


Cys25Ile-[20-41]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 444 mg. The peptide was cleaved from the resin by treatment with 4.5 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 240 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 40 mg (16%). LC-MS (ESI):  $[M + 2H^+] =$

1254.7  $m/z$ ,  $[M + 3H^+] = 837.0 m/z$ ;  $C_{116}H_{175}N_{27}O_{35}$ ; Average isotope calculated 2507.8 Da  $[M]$ ; found: 2507.6 Da.

### Synthesis of Cys25Phe-[20-41] $\beta$ 2m

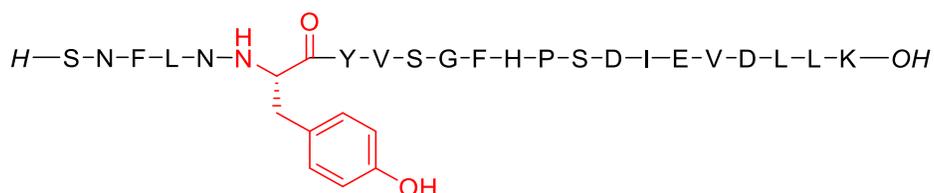
Sequence: SNFLN**F**YVSGFHPSDIEVDLLK



Cys25Phe-[20-41] $\beta$ 2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 340 mg. The peptide was cleaved from the resin by treatment with 3.5 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1,  $v/v$ ) containing TFA (0.1%,  $v/v$ ) and lyophilized. Mass of crude peptide was 190 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 55 mg (22%). LC-MS (ESI):  $[M + 2H^+] = 1271.5 m/z$ ,  $[M + 3H^+] = 848.2 m/z$ ;  $C_{119}H_{173}N_{27}O_{35}$ ; Average isotope calculated 2541.9 Da  $[M]$ ; found: 2541.2 Da.

### Synthesis of Cys25Tyr-[20-41] $\beta$ 2m

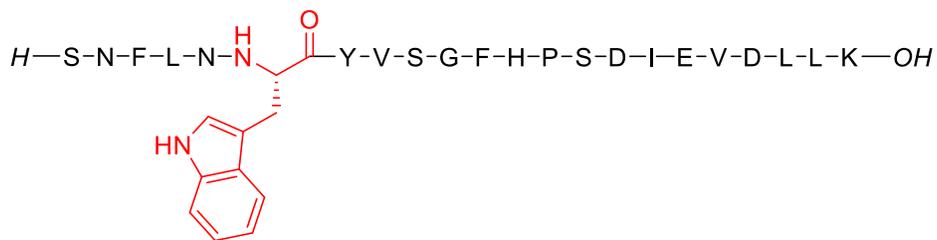
Sequence: SNFLN**Y**YVSGFHPSDIEVDLLK



Cys25Tyr-[20-41] $\beta$ 2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 413 mg. The peptide was cleaved from the resin by treatment with 4.5 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 225 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 79 mg (31%). LC-MS (ESI):  $[M + 2H^+] = 1279.6 m/z$ ,  $[M + 3H^+] = 853.4 m/z$ ; C<sub>119</sub>H<sub>173</sub>N<sub>27</sub>O<sub>36</sub>; Average isotope calculated 2557.9 Da  $[M]$ ; found: 2557.2 Da.

### Synthesis of Cys25Trp-[20-41] $\beta$ 2m

Sequence: SNFLN**W**YVSGFHPSDIEVDLLK

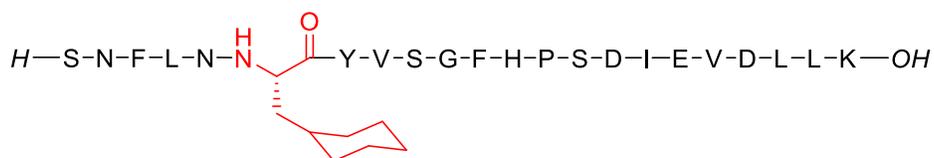


Cys25Trp-[20-41] $\beta$ 2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 488 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 238 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-

C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 29 mg (11%). LC-MS (ESI):  $[M + 2H^+] = 1290.9$   $m/z$ ,  $[M + 3H^+] = 861.1$   $m/z$ ;  $C_{121}H_{174}N_{28}O_{36}$ ; Average isotope calculated 2580.9 Da  $[M]$ ; found: 2580.0 Da.

### Synthesis of Cys25Cha-[21-40]β2m

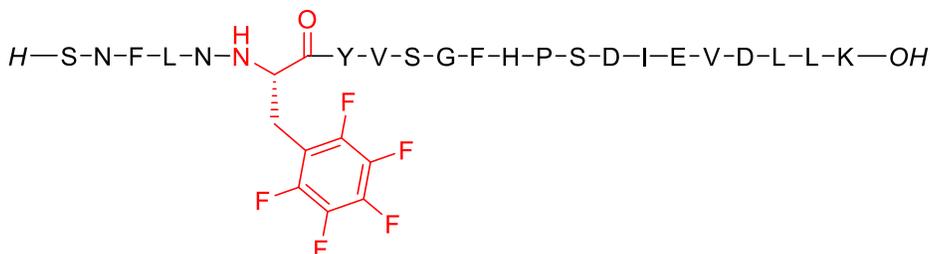
Sequence: SNFLN**Cha**YVSGFHPSDIEVDLLK, **Cha** = *L*-β-cyclohexylalanine



Cys25Cha-[21-40]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). *L*-β-cyclohexylalanine at position 26 was coupled manually (4 equiv, 45 min). The mass of the dried peptidyl-resin at the end of the synthesis was 440 mg. The peptide was cleaved from the resin by treatment with 4.5 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 300 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 37 mg (15%). LC-MS (ESI):  $[M + 2H^+] = 1274.8$   $m/z$ ,  $[M + 3H^+] = 850.3$   $m/z$ ;  $C_{119}H_{178}N_{27}O_{35}$ ; Average isotope calculated 2546.9 Da  $[M]$ ; found: 2547.7 Da.

### Synthesis of Cys25pfPhe-[20-41]β2m

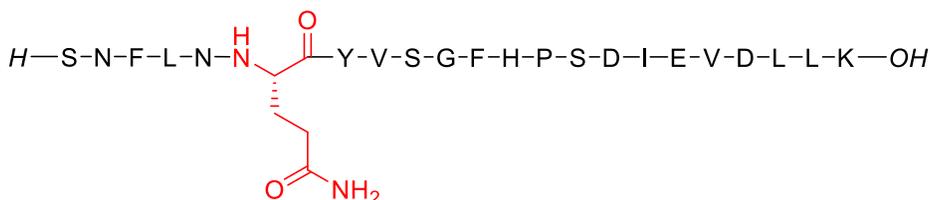
Sequence: SNFLN**pfPhe**YVSGFHPSDIEVDLLK, **pfPhe** = L-2,3,4,5,6-pentafluorophenylalanine



Cys25pfPhe-[20-41]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). L-2,3,4,5,6-pentafluorophenylalanine at position 26 was coupled manually (4 equiv, 45 min). The mass of the dried peptidyl-resin at the end of the synthesis was 449 mg. The peptide was cleaved from the resin by treatment with 4.5 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide: 236 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 36 mg (14%). LC-MS (ESI): [M + 2H<sup>+</sup>] = 1316.8 m/z, [M + 3H<sup>+</sup>] = 878.4 m/z; C<sub>119</sub>H<sub>168</sub>F<sub>5</sub>N<sub>27</sub>O<sub>35</sub>; Average isotope calculated 2631.8 Da [M]; found: 2631.7 Da.

### Synthesis of Cys25Gln-[20-41]β2m

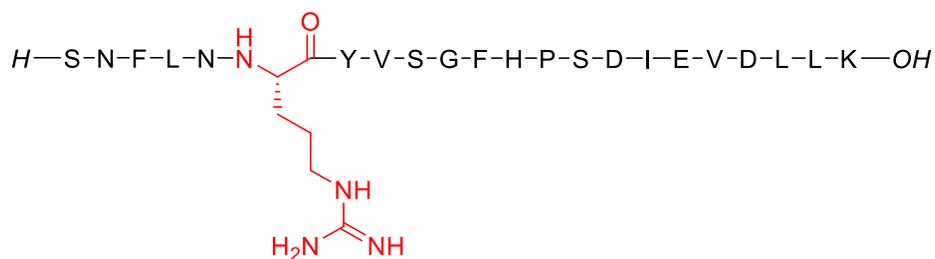
Sequence: SNFLN**Q**YVSGFHPSDIEVDLLK



Cys25Gln-[20-41] $\beta$ 2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 540 mg. The peptide was cleaved from the resin by treatment with 5.5 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 290 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 44 mg (17%). LC-MS (ESI):  $[M + 2H^+] = 1261.9$  m/z,  $[M + 3H^+] = 841.7$  m/z; C<sub>115</sub>H<sub>172</sub>N<sub>28</sub>O<sub>36</sub>; Average isotope calculated 2522.8 Da [M]; found: 2521.9 Da.

### Synthesis of Cys25Arg-[20-41] $\beta$ 2m

Sequence: SNFLN**R**YVSGFHPSDIEVDLLK

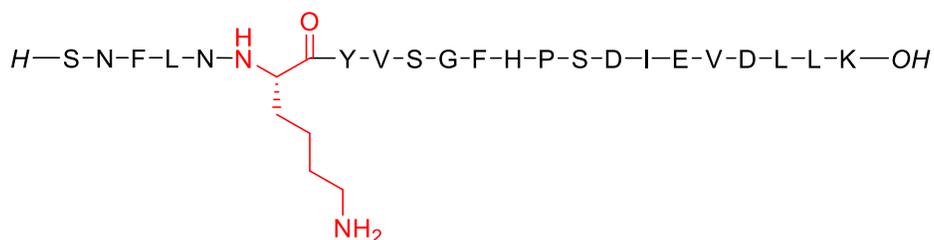


Cys25Gln-[20-41] $\beta$ 2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 510 mg. The peptide was cleaved from the resin by treatment with 5.0 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 274 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-

C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 41 mg (16%). LC-MS (ESI):  $[M + 2H^+] = 1276.1$   $m/z$ ,  $[M + 3H^+] = 851.2$   $m/z$ ,  $[M + 4H^+] = 638.7$ ; C<sub>116</sub>H<sub>176</sub>N<sub>30</sub>O<sub>35</sub>; Average isotope calculated 2550.9 Da  $[M]$ ; found: 2550.4 Da.

### Synthesis of Cys25Lys-[20-41]β2m

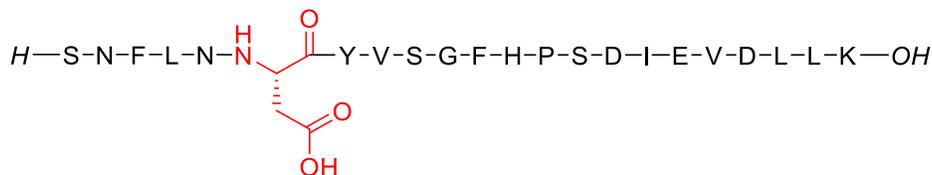
Sequence: SNFLN**K**YVSGFHPSDIEVDLLK



Cys25Gln-β2m[20-41] was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 500 mg. The peptide was cleaved from the resin by treatment with 5.0 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 285 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 55 mg (22%). LC-MS (ESI):  $[M + 2H^+] = 1262.0$   $m/z$ ,  $[M + 3H^+] = 841.9$   $m/z$ ,  $[M + 4H^+] = 631.9$   $m/z$ ; C<sub>116</sub>H<sub>176</sub>N<sub>28</sub>O<sub>35</sub>; Average isotope calculated 2522.9 Da  $[M]$ ; found: 2522.7 Da.

### Synthesis of Cys25Asp-[20-41]β2m

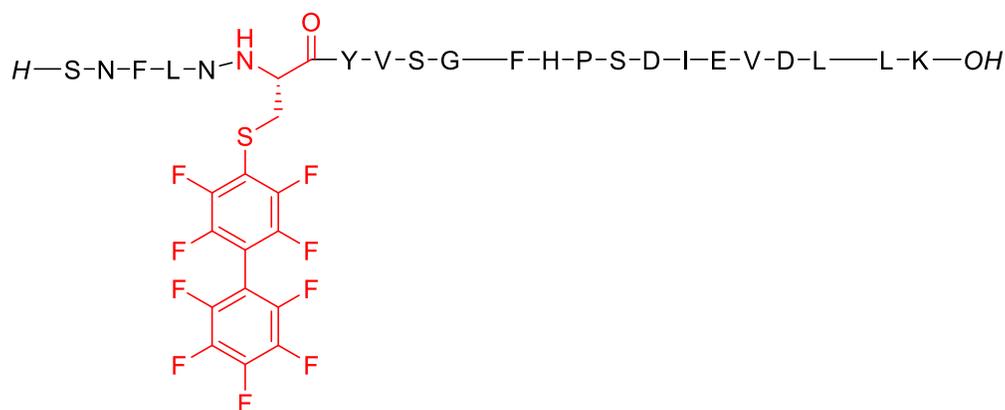
Sequence: SNFLNDYVSGFHPSDIEVDLLK



Cys25Gln-[20-41]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (152 mg, loading: 0.66 mmol/g, mesh size: 100-200). Automated synthesis was carried out by using HBTU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 499 mg. The peptide was cleaved from the resin by treatment with 5.0 mL of TFA/TIPS/H<sub>2</sub>O (90/5/5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 265 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 51 mg (20%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1255.6 *m/z*, [*M* + 3H<sup>+</sup>] = 837.3 *m/z*; C<sub>114</sub>H<sub>169</sub>N<sub>27</sub>O<sub>37</sub>; Average isotope calculated 2509.8 Da [*M*]; found: 2509.1 Da.

### Synthesis of [20-41]β2m-(perfluoro-[1,1'-biphenyl]-4-yl) conjugate

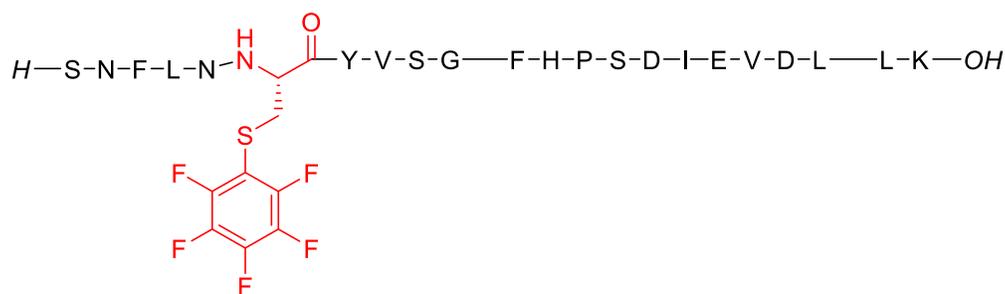
Sequence: SNFLNC(pfbis-Ph)YVSGFHPSDIEVDLLK



[20-41] $\beta$ 2m peptide (20 mg, 8  $\mu$ mol) was dissolved in tris(hydroxymethyl)aminomethane (Tris; 0.8 mL, 50 mM in DMF). Decafluorobiphenyl (26.7 mg, 80  $\mu$ mol) was added, and the mixture was stirred at 25°C for 2.5 h. Then, the solvent was removed under reduced pressure, and the resulting solid was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. The crude solid was dissolved in Gn·HCl (6 M) and purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on peptide was 12.5 mg (56%). LC-MS (ESI): [ $M + 2H^+$ ] = 1406.8  $m/z$ , [ $M + 3H^+$ ] = 938.3  $m/z$ ; C<sub>125</sub>H<sub>168</sub>F<sub>9</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2811.9 Da [ $M$ ]; found: 2811.4 Da.

### Synthesis of $\beta$ 2m[20-41]-pentafluorophenyl conjugate

Sequence: SNFLN**C(F<sub>5</sub>Ph)**YVSGFHPSDIEVDLLK

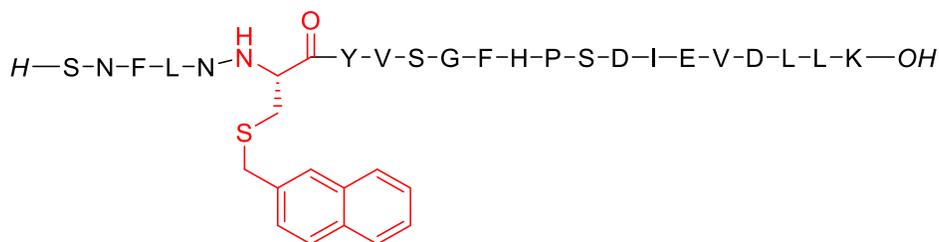


[21-40] $\beta$ 2m peptide (10 mg, 8  $\mu$ mol) was dissolved in tris(hydroxymethyl)aminomethane (Tris; 0.8 mL, 50 mM in DMF). Hexafluorobenzene (0.8 mL, 100 mM in DMF, 400  $\mu$ mol) was added,

and the mixture was diluted with DMF (3.4 mL). The reaction mixture was shaken on a Thermomixer (400 rpm) at room temperature for 24 h, and then diluted with 35 mL of H<sub>2</sub>O/CH<sub>3</sub>CN (9:1, v/v) containing TFA (0.1%, v/v). It was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on peptide was 1.6 mg (15%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1332.8 *m/z*, [*M* + 3H<sup>+</sup>] = 889.0 *m/z*; C<sub>119</sub>H<sub>168</sub>F<sub>5</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2663.9 Da [*M*]; found: 2663.7 Da.

### Synthesis of [20-41]β2m-(naphthalene-2-ylmethyl) conjugate

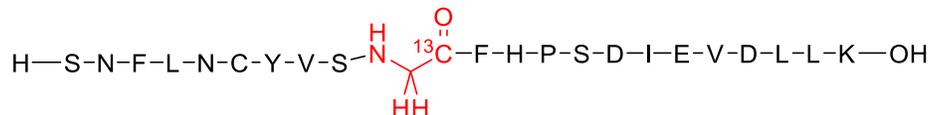
Sequence: SNFLN**C(Napht)**YVSGFHPSDIEVDLLK



[20-41]β2m peptide (15.3 mg, 6 µmol) was dissolved in aqueous NH<sub>4</sub>HCO<sub>3</sub> (6 mL, 200 mM, pH 8) with EDTA (5 mM). 2-(bromomethyl)-naphthalene (2 mg, 9 µmol) in acetonitrile (1.24 mL) was added. The reaction mixture was shaken on a thermomixer (500 rpm) at 30°C for 1 h. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on peptide was 4.5 mg (28%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1319.6 *m/z*, [*M* + 3H<sup>+</sup>] = 880.1 *m/z*; C<sub>119</sub>H<sub>168</sub>F<sub>5</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2638.0 Da [*M*]; found: 2637.1 Da.

### Synthesis of [1-<sup>13</sup>C]Gly29-[20-41]β2m

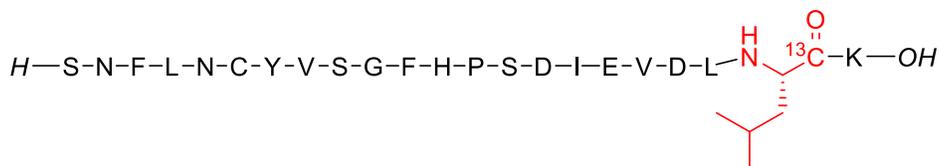
Sequence: SNFLNCYV**S**GFHPSDIEVDLLK



[1-<sup>13</sup>C]Gly29-[20-41]β2m was synthesized on a 0.2 mmol scale on Fmoc-Lys(Boc)-Wang resin (303 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). [1-<sup>13</sup>C]Gly at position 29 was coupled manually (2 equiv, 1h). The mass of the dried peptidyl-resin at the end of the synthesis was 1.11 g. The peptide was cleaved from the resin by treatment with 10 mL of TFA/TIPS/H<sub>2</sub>O (95/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 489 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 136 mg (27%). LC-MS (ESI): [M + 2H<sup>+</sup>] = 1250.3 m/z, [M + 3H<sup>+</sup>] = 833.9 m/z; C<sub>113</sub>H<sub>169</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2498.8 Da [M]; found: 2498.6 Da.

### Synthesis of [1-<sup>13</sup>C]Leu40-[20-41]β2m

Sequence: SNFLNCYVSGFHPSDIEVDL**L**K

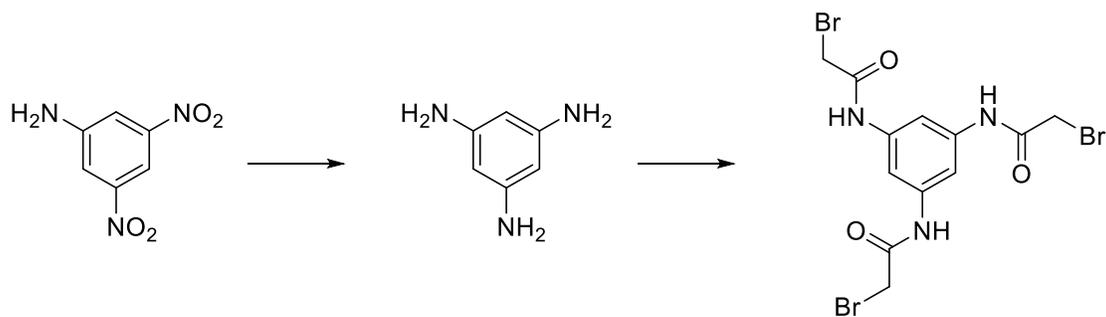


[1-<sup>13</sup>C]Leu40-[20-41]β2m was synthesized on a 0.2 mmol scale on Fmoc-Lys(Boc)-Wang resin (303 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). [1-<sup>13</sup>C]Leu at position 40 was coupled manually (2 equiv, 1h). The mass of the dried peptidyl-resin at the end of the synthesis was 1.15 g. The peptide was cleaved from the resin by treatment with 10 mL of TFA/TIPS/H<sub>2</sub>O (95/2.5/2.5) for 2 h. The cleavage

cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 513 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 201 mg (27%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1250.1 *m/z*, [*M* + 3H<sup>+</sup>] = 833.8 *m/z*; C<sub>113</sub>H<sub>169</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2498.8 Da [*M*]; found: 2498.2 Da.

### Synthesis of *N,N,N'*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) linker

Scheme of the synthesis:



*Benzene-1,3,5-triamine*: 3,5-Dinitroaniline (1 g, 5.3 mmol) was dissolved in methanol (25 mL), then Pd/C (10%, 282 mg, 0.26 mmol) was added. The dark suspension was bubbled with H<sub>2</sub> for 10 min. The mixture was stirred at room temperature for 16 h under H<sub>2</sub>, then filtered through a Celite pad, and the filter cake was washed with methanol. The filtrate was evaporated to give a brownish solid of crude benzene-1,3,5-triamine, which was used in the following step without purification.

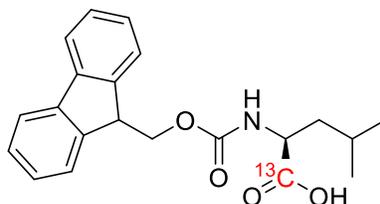
*N,N,N'*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide): Benzene-1,3,5-triamine (61.6 mg, 0.5 mmol) was dissolved in anhydrous DMF (1 mL) under argon. Bromoacetic acid (289 mg, 2.02 mmol) was added, followed by *N,N'*-diisopropylcarbodiimide (309 µL, 2.02 mmol). The suspension was shaken at room temperature for 3 h. The reaction mixture was diluted with 10 mL

of H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) and directly injected into preparative HPLC and purified using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on 3,5-dinitroaniline was 53 mg (22%). <sup>1</sup>H NMR (400 MHz, [D<sub>6</sub>]DMSO): δ, ppm = 10.49 (s, 3H), 7.71 (s, 3H), 4.02 (s, 6H); LC-MS (ESI): [*M* + H<sup>+</sup>] = 483.91 *m/z*, 485.88 *m/z*, 487.86 *m/z* and 489.86 *m/z*; [*M* + H<sub>2</sub>O]<sup>+</sup> = 501.10 *m/z*, 502.88 *m/z*, 505.02 *m/z* and 506.79 *m/z*; Average isotope calculated 482.8 Da [*M*]; found: 482.9 Da.

### Synthesis of [20-41]β2m covalent trimer

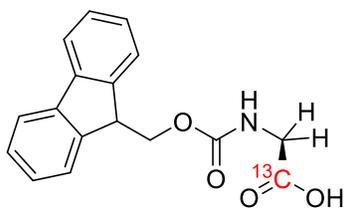
β2m[20-41] peptide (15.3 mg, 6 μmol) was suspended in NH<sub>4</sub>HCO<sub>3</sub> (6 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (4 mL, 1 mM in acetonitrile, 4 μmol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn·HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions of [20-41]β2m covalent trimer were combined and lyophilized to give a white solid. The isolated yield based on the linker was 15.2 mg (49%). LC-MS (ESI): [*M* + 5H<sup>+</sup>] = 1547.9 *m/z*; [*M* + 6H<sup>+</sup>] = 1290.2 *m/z*; [*M* + 7H<sup>+</sup>] = 1106.0 *m/z*; [*M* + 8H<sup>+</sup>] = 967.9 *m/z*; [*M* + 9H<sup>+</sup>] = 860.6 *m/z*; Average isotope calculated 7736.6 Da [*M*]; found: 7735.2 Da.

### Synthesis of Fmoc-[1-<sup>13</sup>C]Leu-OH



*L*-[1-<sup>13</sup>C]leucine (1.0 g, 7.56 mmol, 1 equiv) and sodium carbonate (1.6 g, 15.2 mmol, 2 equiv) were dissolved in 80 mL of water. A solution of Fmoc-OSu (6.4 g, 19 mmol, 2.5 equiv) in 70 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react at r.t. for 17h. The reaction mixture was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub>. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 2.2 g (83%). <sup>1</sup>H NMR (400 MHz, CDCl<sub>3</sub>): δ, ppm = 7.82 (d, 2H), 7.65 (d, 2H), 7.45 (t, 2H), 7.36 (t, 2H), 5.17 (d, 1H), 4.49 (m, 3H), 4.28 (t, 1H), 1.70 (m, 2H), 1.64 (m, 1H), 1.02 (m, 6H); C<sub>21</sub>H<sub>23</sub>NO<sub>4</sub>; LC-MS (ESI): Average isotope calculated 354.4 Da [*M*]; found: 354.6 Da.

### Synthesis of Fmoc-[1-<sup>13</sup>C]Gly-OH



[1-<sup>13</sup>C]glycine (1.0 g, 13.3 mmol, 1 equiv) and sodium carbonate (2.9 g, 27 mmol, 2 equiv) were dissolved in 80 mL of water. A solution of Fmoc-OSu (11.2 g, 33 mmol, 2.5 equiv) in 70 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react at r.t. for 17h. The reaction mixture was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub>. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 3.1 g (79%). <sup>1</sup>H NMR (400 MHz, [D<sub>6</sub>]DMSO): δ, ppm = 12.44 (s, 1H), 7.89 (d, 2H), 7.73 (d, 2H), 7.65 (t, 1H), 7.41 (t, 2H), 7.33 (t, 2H), 4.30 (d, 2H), 4.23 (t, 1H), 3.67 (d, 2H); C<sub>17</sub>H<sub>15</sub>NO<sub>4</sub>; LC-MS (ESI): Average isotope calculated 354.4 Da [*M*]; found: 354.6 Da.

The corresponding <sup>13</sup>C-labeled covalent trimers were synthesized according to the same protocol on a scale of several tens of milligrams in order to have sufficient quantities to perform FT-IR and ssNMR studies.

**[1-<sup>13</sup>C]Phe22-[20-41]β2m covalent trimer:** LC-MS (ESI):  $[M + 5H^+] = 1548.6 m/z$ ;  $[M + 6H^+] = 1290.6 m/z$ ;  $[M + 7H^+] = 1106.5 m/z$ ;  $[M + 8H^+] = 968.4 m/z$ ;  $[M + 9H^+] = 860.9 m/z$ ;  
Average isotope calculated 7739.6 Da  $[M]$ ; found: 7738.3 Da.

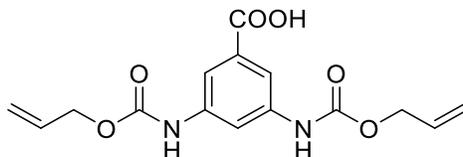
**[1-<sup>13</sup>C]Val27-[20-41]β2m covalent trimer:** LC-MS (ESI):  $[M + 5H^+] = 1548.6 m/z$ ;  $[M + 6H^+] = 1290.6 m/z$ ;  $[M + 7H^+] = 1106.5 m/z$ ;  $[M + 8H^+] = 968.4 m/z$ ;  $[M + 9H^+] = 860.7 m/z$ ;  
Average isotope calculated 7739.6 Da  $[M]$ ; found: 7738.3 Da.

**[1-<sup>13</sup>C]Gly29-[20-41]β2m covalent trimer:** LC-MS (ESI):  $[M + 5H^+] = 1548.6 m/z$ ;  $[M + 6H^+] = 1290.7 m/z$ ;  $[M + 7H^+] = 1106.5 m/z$ ;  $[M + 8H^+] = 968.4 m/z$ ;  $[M + 9H^+] = 861.1 m/z$ ;  
Average isotope calculated 7739.6 Da  $[M]$ ; found: 7738.5 Da.

**[1-<sup>13</sup>C]Leu40-[20-41]β2m covalent trimer :** LC-MS (ESI):  $[M + 5H^+] = 1548.6 m/z$ ;  $[M + 6H^+] = 1290.6 m/z$ ;  $[M + 7H^+] = 1106.5 m/z$ ;  $[M + 8H^+] = 968.3 m/z$ ;  $[M + 9H^+] = 861.0 m/z$ ;  
Average isotope calculated 7739.6 Da  $[M]$ ; found: 7738.3 Da.

## 7.4 Chapter 4

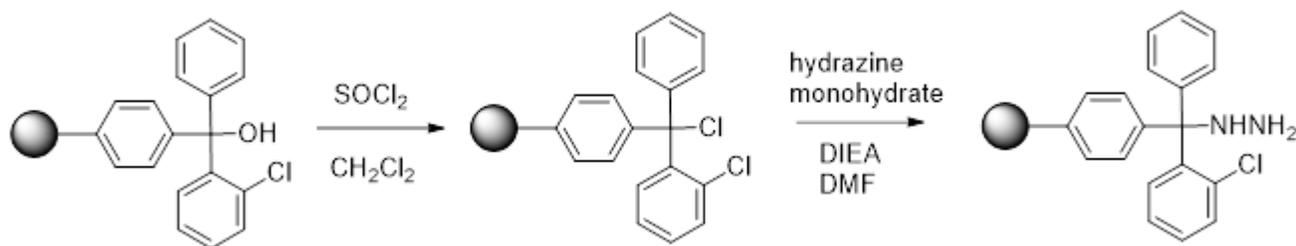
### Synthesis of the central linker: 3,5-bis(((allyloxy)carbonyl)amino)benzoic acid



Potassium hydroxide (1.09 g, 19.5 mmol, 3 equiv) was added to a mixture of 3,5-diaminobenzoic acid (1.05 g, 6.5 mmol, 1 equiv) in water (40 mL). It was stirred until a clear solution was obtained. 1,4-dioxane (30 mL) was added to the reaction mixture followed by a dropwise addition of a solution of allyl chloroformate (1.52 mL, 14.3 mmol, 2.2 equiv) in 1,4-dioxane (10 mL). It was stirred at r.t. overnight. The volatiles were removed under reduced pressure. Water (100 mL) was added to the resulting paste and the pH was adjusted to 8-9 with sat.  $\text{Na}_2\text{HCO}_3(\text{aq})$ . The aqueous solution was washed with ethyl acetate (2 x 50 mL). The aqueous layer was then treated with 6M  $\text{HCl}(\text{aq})$  to lower the pH to 3 and extracted with ethyl

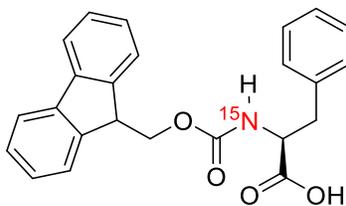
acetate (3 x 50 mL). The organic layers were gathered, washed with brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 2.0 g (95%).  $^1\text{H}$  NMR (400 MHz,  $[\text{D}_6]\text{DMSO}$ ):  $\delta$ , ppm = 12.92 (s, 1H), 9.91 (s, 2H), 7.94 (t, 1H), 7.73 (d, 2H), 5.99 (m, 2H), 5.36 (ddd, 2H), 5.24 (ddd, 2H), 4.62 (ddd, 4H);  $\text{C}_{15}\text{H}_{16}\text{N}_2\text{O}_6$ ; LC-MS (ESI): Average isotope calculated 320.3 Da [ $M$ ]; found: 320.8 Da.

### Preparation of the hydrazide resin



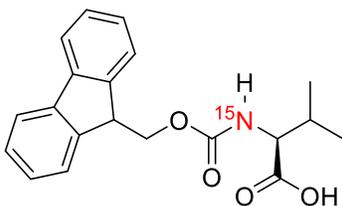
The 2-CT-OH resin (3 g, 4.7 mmol/g) was placed in a three-necked round bottom flask. It was flushed several times with argon. Dry dichloromethane (28 mL) was added. The mixture was gently stirred allowing the resin to swell. Thionyl chloride (450  $\mu\text{L}$ , 6.2 mmol, 1.3 equiv) was slowly added at  $0^\circ\text{C}$ . The suspension was stirred under argon for 2 h allowing it to slowly warm up to room temperature. The solvent was removed and the resin was washed with DCM and DMF. The 2-CT-Cl resin was then swollen for 20 min in DMF (18 mL). A mixture of hydrazine monohydrate (700  $\mu\text{L}$ , 14.4 mmol, 3 equiv) and DIEA (2 mL, 11.5 mmol, 2.4 equiv) in DMF (4 mL) was added to the resin at  $0^\circ\text{C}$ . The suspension was stirred at r.t. for 90 min. The reaction was quenched by the addition of methanol (300  $\mu\text{L}$ ). The resin was washed with DMF, water, DMF, methanol and diethyl ether and dried under reduced pressure. Final mass of the resin: 3g.

### Synthesis of Fmoc- $^{15}\text{N}$ Phe-OH



*L*-[<sup>15</sup>N]-phenylalanine (1.0 g, 6.1 mmol, 1 equiv) and sodium carbonate (1.8 g, 17.4 mmol, 2.8 equiv) were dissolved in 80 mL of water. A solution of Fmoc-OSu (5 g, 15 mmol, 2.5 equiv) in 70 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react at r.t. for 17h. The reaction mixture was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub> to allow precipitation of the Fmoc-amino acid. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 1.7 g; C<sub>17</sub>H<sub>15</sub>NO<sub>4</sub>; LC-MS (ESI): Average isotope calculated 388.4 Da [*M*]; found: 388.6 Da.

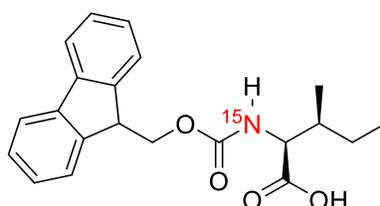
### Synthesis of Fmoc-[<sup>15</sup>N]Val-OH



*L*-[<sup>15</sup>N]-valine (0.5 g, 4.3 mmol, 1 equiv) and sodium carbonate (1.4 g, 12.9 mmol, 3 equiv) were dissolved in 40 mL of water then 30 mL of 1,4-dioxane were added. A solution of Fmoc chloride (1.2 g, 4.7 mmol, 1.1 equiv) in 10 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react overnight at r.t. The volatiles were removed under reduced pressure. The resulting paste was dissolved in 100 mL of water. The aqueous solution was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub> to allow precipitation of the Fmoc-amino acid. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 1.2 g (83%). <sup>1</sup>H NMR (400 MHz, [D<sub>6</sub>]DMSO): δ, ppm = 12.59 (s, 1H), 7.90 (d, 2H), 7.75 (d, 2H), 7.51 (d, 1H), 7.42 (t, 2H),

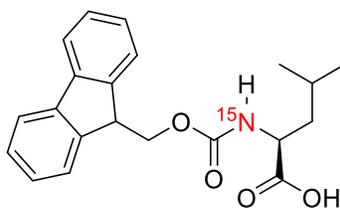
7.33 (t, 2H), 4.27 (m, 2H), 4.23 (m, 1H), 3.92 (t, 1H), 1.81 (m, 1H), 0.87 (m, 6H); C<sub>20</sub>H<sub>21</sub>NO<sub>4</sub>; LC-MS (ESI): Average isotope calculated 340.4 Da [M]; found: 340.9 Da.

### Synthesis of Fmoc-<sup>15</sup>N]Ile-OH



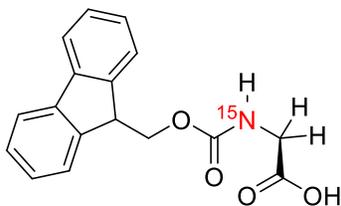
L-[<sup>15</sup>N]-isoleucine (0.5 g, 3.8 mmol, 1 equiv) and sodium carbonate (1.2 g, 11.4 mmol, 3 equiv) were dissolved in 40 mL of water then 30 mL of 1,4-dioxane were added. A solution of Fmoc chloride (1.1 g, 4.2 mmol, 1.1 equiv) in 10 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react overnight at r.t. The volatiles were removed under reduced pressure. The resulting paste was dissolved in 100 mL of water. The aqueous solution was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub> to allow precipitation of the Fmoc-amino acid. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 1.0 g (75%). <sup>1</sup>H NMR (400 MHz, [D<sub>6</sub>]DMSO): δ, ppm = 7.89 (d, 2H), 7.74 (d, 2H), 7.50 (t, 1H), 7.42 (t, 2H), 7.33 (t, 2H), 4.24 (m, 3H), 3.92 (t, 1H), 1.81 (m, 1H), 1.41 (m, 1H), 1.23 (m, 1H), 0.86 (m, 6H); C<sub>21</sub>H<sub>23</sub>NO<sub>4</sub>; LC-MS (ESI): Average isotope calculated 354.4 Da [M]; found: 354.8 Da.

### Synthesis of Fmoc-<sup>15</sup>N]Leu-OH



*L*-[<sup>15</sup>N]-eucine (0.5 g, 3.8 mmol, 1 equiv) and sodium carbonate (1.2 g, 11.4 mmol, 3 equiv) were dissolved in 40 mL of water then 30 mL of 1,4-dioxane were added. A solution of Fmoc chloride (1.1 g, 4.2 mmol, 1.1 equiv) in 10 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react overnight at r.t. The volatiles were removed under reduced pressure. The resulting paste was dissolved in 100 mL of water. The aqueous solution was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub> to allow precipitation of the Fmoc-amino acid. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 1.1 g (81%). <sup>1</sup>H NMR (400 MHz, [D<sub>6</sub>]DMSO): δ, ppm = 12.55 (s, 1H), 7.90 (d, 2H), 7.73 (d, 2H), 7.51 (t, 1H), 7.42 (t, 2H), 7.33 (t, 2H), 4.31 (d, 2H), 4.28 (t, 1H), 3.98 (m, 1H), 1.57 (m, 3H), 0.88 (dd, 6H); C<sub>21</sub>H<sub>23</sub>NO<sub>4</sub>; LC-MS (ESI): Average isotope calculated 354.4 Da [*M*]; found: 354.8 Da.

### Synthesis of Fmoc-[<sup>15</sup>N]Gly-OH



[<sup>15</sup>N]-glycine (0.5 g, 6.7 mmol, 1 equiv) and sodium carbonate (2.7 g, 20.0 mmol, 3 equiv) were dissolved in 30 mL of water then 20 mL of 1,4-dioxane were added. A solution of Fmoc chloride (1.9 g, 7.3 mmol, 1.1 equiv) in 10 mL of 1,4-dioxane was added dropwise to the reaction mixture. It was let to react overnight at r.t. The volatiles were removed under reduced pressure. The resulting paste was dissolved in 100 mL of water. The aqueous solution was washed with 3 x 50 mL of ethyl acetate. The pH of the aqueous layer was adjusted to 2 with 6M HCl<sub>(aq)</sub> to allow precipitation of the Fmoc-amino acid. It was extracted with 3 x 50 mL of ethyl acetate. The organic layers were combined, washed with 50 mL of brine, dried over anhydrous sodium sulfate and concentrated under reduced pressure. The isolated yield was 1.7 g (86%). <sup>1</sup>H NMR (400 MHz, [D<sub>6</sub>]DMSO): δ, ppm = 12.58 (s, 1H), 7.90 (d, 2H), 7.72 (d, 2H), 7.52 (t, 1H), 7.42 (t, 2H),

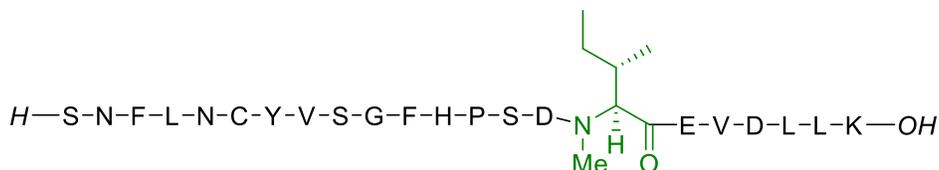
7.33 (t, 2H), 4.30 (d, 2H), 4.23 (t, 1H), 3.67 (d, 2H); C<sub>17</sub>H<sub>15</sub>NO<sub>4</sub>; LC-MS (ESI): Average isotope calculated 298.3 Da [*M*]; found: 298.7 Da.

**List of peptides, covalent trimers, covalent dimers of covalent trimers and covalent trimers of covalent trimers:**

- peptide **A**: *N*-Me-Ile35-[20-41]β2m
- peptide **B**: Cys25Orn(linker),*N*-Me-Ile35-[20-41]β2m
- peptide **C**: [20-41]β2m
- peptide **D**: Thz-GGG-Cys25Orn(linker)-[20-41]β2m
- peptide **E**: *N*-Me-Glu36-[20-41]β2m
- peptide **F**: Thz-GGG-Cys25Orn(linker),*N*-Me-Glu36-[20-41]β2m
- peptide **G**: *N*-Me-MeAsn24,*N*-Me-MeVal37-[20-41]β2m
- peptide **H**: Cys25Orn(linker),*N*-Me-Asn24,*N*-Me-Val37-[20-41]β2m
- peptide **I**: *N*-Me-Leu23,*N*-Me-Asp38-[20-41]β2m
- peptide **J**: Thz-GGG-Cys25Orn(linker),*N*-Me-Leu23,*N*-Me-Asp38-[20-41]β2m
- peptide **K**: (<sup>15</sup>N-labeled)-Thz-GGG-Cys25Orn(linker)-[20-41]β2m
- covalent **trimer-1-A**
- covalent **trimer-2-A**
- covalent **trimer-3-A**
- covalent **trimer-1-B**
- covalent **trimer-3-B**
- covalent **trimer-2-B**
- covalent dimer of trimers 1: **DofT-1**
- covalent dimer of trimers 2: **DofT-2**
- covalent dimer of trimers 3: **DofT-3**
- covalent trimer of trimers 1: **TofT-1**
- covalent trimer of trimers 1: **TofT-2**
- covalent trimer of trimers 1: **TofT-3**

### Synthesis of peptide A: *N*-Me-Ile35-[20-41]β2m

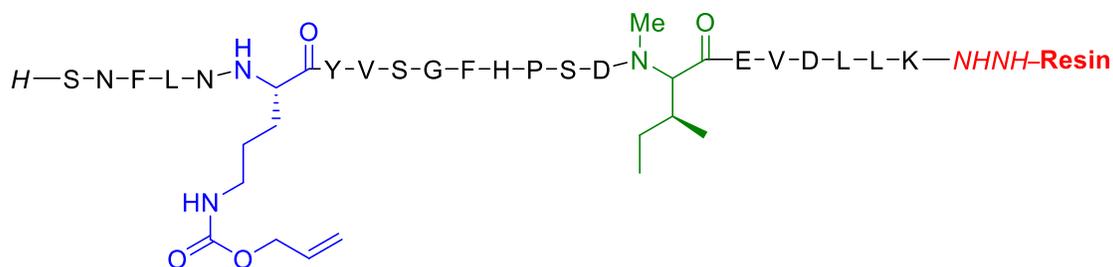
Sequence: SNFLNCYVSGFHPSD*N-Me*IEVDLLK



*N*-Me-Ile35-[20-41]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (147 mg, loading: 0.68 mmol/g, mesh size: 180-333). Automated synthesis was carried out by using HATU activation (5 equiv). *N*-Me-Ile at position 35 was coupled manually (3 equiv, 1 h). Kaiser test was performed to detect the presence of unreacted primary amines. No acetylation step was required. The next amino acid Asp34 was coupled twice (5 equiv, 40 min). Chloranil test was performed to detect the presence of unreacted secondary amines. An acetylation step was required and applied. The mass of the dried peptidyl-resin at the end of the synthesis was 447 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 215 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 94 mg (37%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1256.9 *m/z*; [*M* + 3H<sup>+</sup>] = 838.3 *m/z*; C<sub>114</sub>H<sub>171</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2511.8 Da [*M*]; found: 2511.9 Da.

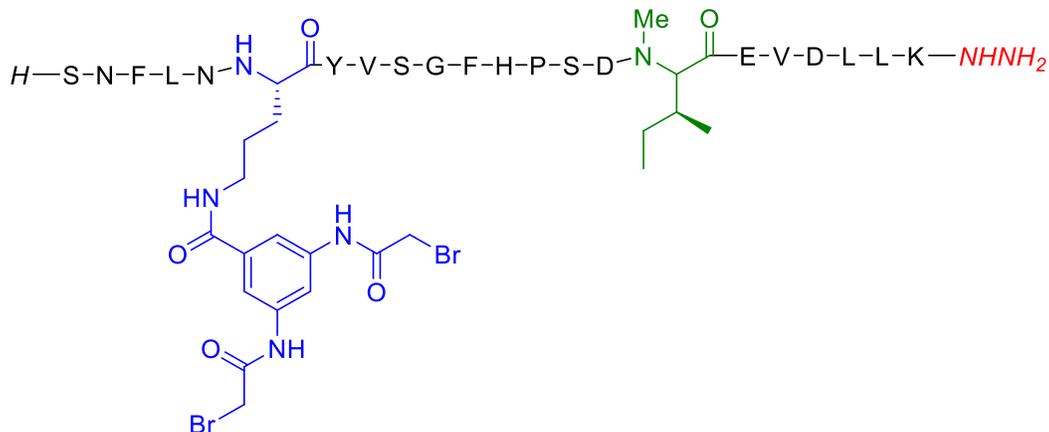
### Synthesis of peptide B: Cys25Orn(linker),*N*-Me-Ile35-[20-41]β2m

Sequence: SNFLNOrn(alloc)YVSGFHPSD*N-Me*IEVDLLK-*NHNH*-Resin



Cys25Orn(alloc),*N*-Me-Ile35-[20-41] $\beta$ 2m was synthesized on a 0.1 mmol scale on 2-ChloroTrityl-hydrazide resin (250 mg, loading: 0.40 mmol/g, mesh size: 200-400). Automated synthesis was carried out by using HATU activation (5 equiv). *N*-Me-Ile at position 35 was coupled manually (3 equiv, 1 h). Kaiser test was performed to detect the presence of unreacted primary amines. No acetylation step was required and applied. The next amino acid Asp34 was coupled twice (5 equiv, 40 min). Chloranil test was performed to detect the presence of unreacted secondary amines. An acetylation step was required. Ornithine 25 was introduced as Fmoc-Orn(alloc)-OH (5 equiv, 40 min).

Sequence: SNFLNOrn(linker)YVSGFHPSD*N*-MeIEVDLLK-*NHNH*<sub>2</sub>



*Boc protection:* The resin was swollen in DMF. N-terminal Boc protection of serine 20 was performed with the addition of Boc anhydride (87.5 mg, 0.4 mmol, 4 equiv, in 1 mL of DMF) and DIEA (68  $\mu$ L, 0.4 mmol, 4 equiv). It was let to react at r.t. for 30 min with manual stirring. Kaiser test was negative.

*Alloc deprotection of ornithine:* The resin was swollen in DCM purged with argon. A solution of phenylsilane (296  $\mu$ L, 2.4 mmol, 24 equiv) in DCM (1 mL) purged with argon was added to the

resin. Alloc deprotection was carried out upon the addition of a solution of Pd(PPh<sub>3</sub>)<sub>4</sub> (58 mg, 0.05 mmol, 0.5 equiv) in DCM (2 mL) purged with argon. It was let to react at r.t. for 30 min with manual stirring. Kaiser test was positive.

*Linker coupling to Orn25 side chain:* The resin was swollen in DMF. A solution of 3,5-bis(((allyloxy)carbonyl)amino)benzoic acid (128 mg, 0.4 mmol, 4 equiv), HATU (144 mg, 0.38 mmol, 3.8 equiv) and DIEA (100  $\mu$ L, 0.6 mmol, 6 equiv) in 1 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 30 min. Kaiser test was negative.

*Alloc deprotection of the linker:* Same procedure as before (repeated twice). Chloranil test was positive.

*Bromoacetylation of the linker:* The resin was swollen in DMF. A solution of bromoacetic acid (167 mg, 1.2 mmol, 12 equiv) and *N,N'*-diisopropylcarbodiimide (86  $\mu$ L, 1.1 mmol, 11.2 equiv) in 2 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 60 min. Chloranil test was negative.

The mass of the dried peptidyl-resin at the end of the synthesis was 340 mg. The peptide was cleaved from the resin by treatment with 4 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 210 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 84 mg (29%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1457.9 *m/z*, 1457.1 *m/z* and 1456.4 *m/z*; [*M* + 3H<sup>+</sup>] = 972.6 *m/z*, 971.9 *m/z* and 971.1 *m/z*; [*M* + 4H<sup>+</sup>] = 729.6 *m/z*, 728.9 *m/z* and 728.3 *m/z*; C<sub>137</sub>H<sub>198</sub>Br<sub>2</sub>N<sub>34</sub>O<sub>42</sub>S; Average isotope calculated 2912.9 Da [*M*]; found: 2910.3 Da, 2912.6 Da and 2915.4 Da.

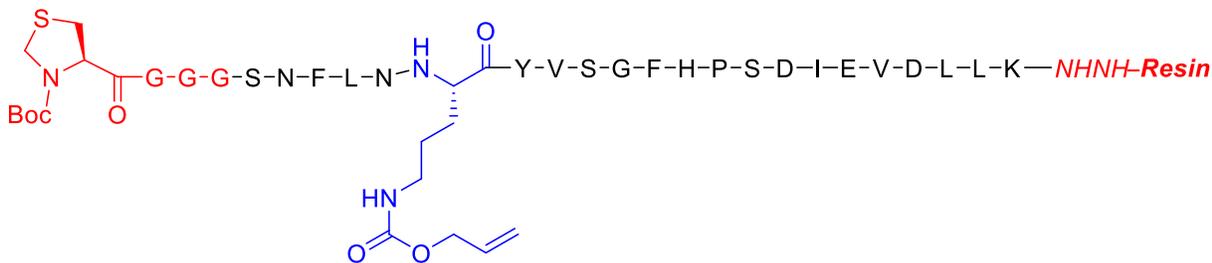
### Synthesis of peptide C: [20-41] $\beta$ 2m

Sequence: SNFLNCYVSGFHPSDIEVDLLK

[20-41] $\beta$ 2m was synthesized on a 0.2 mmol scale on Fmoc-Lys(Boc)-Wang resin (294 mg, loading: 0.68 mmol/g, mesh size: 180-333). Automated synthesis was carried out by using HATU activation (5 equiv). The mass of the dried peptidyl-resin at the end of the synthesis was 930 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 480 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 208 mg (42%). LC-MS (ESI): [ $M + 2H^+$ ] = 1249.6  $m/z$ ; [ $M + 3H^+$ ] = 833.6  $m/z$ ; C<sub>113</sub>H<sub>169</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2497.8 Da [ $M$ ]; found: 2497.4 Da.

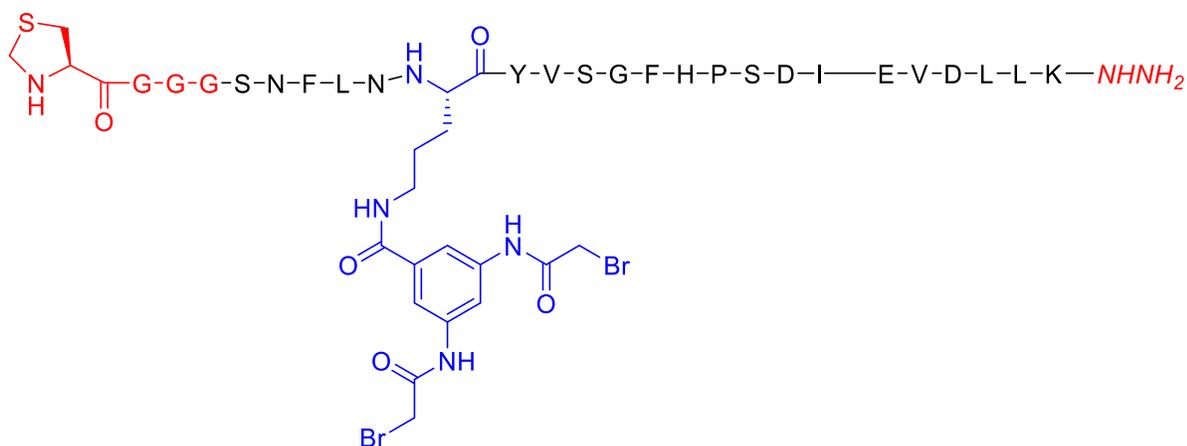
#### Synthesis of peptide D: Thz-GGG-Cys25Orn(linker)-[20-41] $\beta$ 2m

Sequence: Boc-ThzGGG-SNFLNOrn(alloc)YVSGFHPSDIEVDLLK-NHNH-Resin



Thz-GGG-Cys25Orn(alloc)- $\beta$ 2m[20-41] was synthesized on a 0.2 mmol scale on 2-ChloroTrityl-hydrazide resin (500 mg, loading: 0.40 mmol/g, mesh size: 200-400). Automated synthesis was carried out by using HATU activation (5 equiv). Ornithine 25 was introduced as Fmoc-Orn(alloc)-OH (5 equiv, 40 min). N-terminal thiazolidine was introduced as Boc-Thz-OH (5 equiv, 40 min).

Sequence: H-ThzGGG-SNFLNOrn(linker)YVSGFHPSDIEVDLLK-NHNH<sub>2</sub>



*Alloc deprotection of ornithine:* the resin was swollen in DCM purged with argon. A solution of phenylsilane (592  $\mu$ L, 2.4 mmol, 24 equiv) in DCM (2 mL) purged with argon was added to the resin. Alloc deprotection initiated with the addition of a solution of Pd(PPh<sub>3</sub>)<sub>4</sub> (116 mg, 0.1 mmol, 0.5 equiv) in DCM (4 mL) purged with argon. It was let to react at r.t. for 30 min with manual stirring. Kaiser test was positive.

*Linker coupling to Orn25 side chain:* the resin was swollen in DMF. A solution of 3,5-bis(((allyloxy)carbonyl)amino)benzoic acid (256 mg, 0.8 mmol, 4 equiv), HATU (288 mg, 0.76 mmol, 3.8 equiv) and DIEA (200  $\mu$ L, 1.2 mmol, 6 equiv) in 2 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 30 min. Kaiser test was negative.

*Alloc deprotection of the linker:* same procedure as before (repeated twice). Chloranil test was positive.

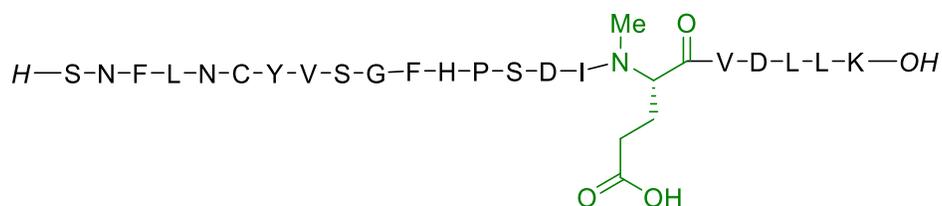
*Bromoacetylation of the linker:* the resin was swollen in DMF. A solution of bromoacetic acid (333 mg, 2.4 mmol, 12 equiv) and *N,N'*-diisopropylcarbodiimide (172  $\mu$ L, 2.2 mmol, 11.2 equiv) in 4 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 60 min. Chloranil test was negative.

The mass of the dried peptidyl-resin at the end of the synthesis was 1.54 g. The peptide was cleaved from the resin by treatment with 4 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v)

containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 827 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 280 mg (44%). LC-MS (ESI):  $[M + 2H^+] = 1594.1 m/z$ ,  $1593.2 m/z$  and  $1592.4 m/z$ ;  $[M + 3H^+] = 1063.6 m/z$ ,  $1062.7 m/z$  and  $1062.1 m/z$ ;  $[M + 4H^+] = 797.7 m/z$ ,  $797.0 m/z$  and  $796.4 m/z$ ;  $C_{137}H_{198}Br_2N_{34}O_{42}S$ ; Average isotope calculated 3185.2 Da  $[M]$ ; found: 3183.1 Da, 3185.1 Da and 3187.8 Da.

### Synthesis of peptide E: *N*-Me-Glu36-[20-41]β2m

Sequence: SNFLNCYVSGFHPSDIN-MeEVDLLK

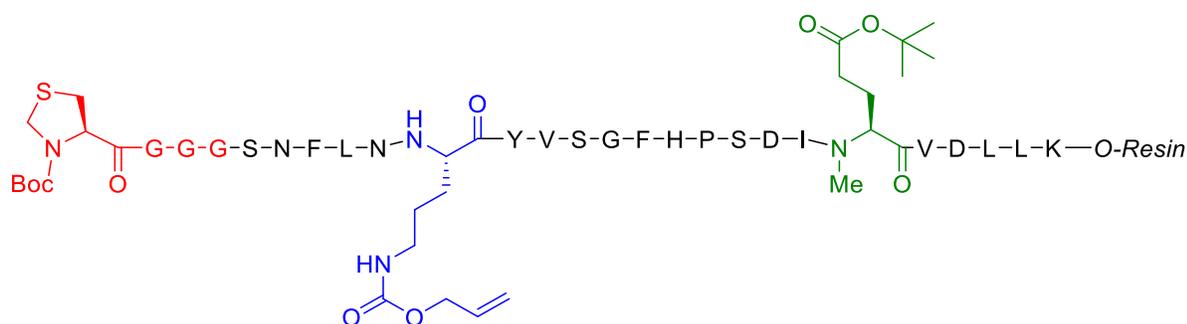


*N*-Me-Glu36-[21-40]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (147 mg, loading: 0.68 mmol/g, mesh size: 180-333). Automated synthesis was carried out by using HATU activation (5 equiv). *N*-Me-Glu at position 36 was coupled manually (3 equiv, 1 h). Kaiser test was performed to detect the presence of unreacted primary amines. No acetylation step was required. The next amino acid Ile35 was coupled twice (5 equiv, 40 min). Chloranil test was performed to detect the presence of unreacted secondary amines. An acetylation step was required and applied. The mass of the dried peptidyl-resin at the end of the synthesis was 460 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude

peptide was 220 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 96 mg (38%). LC-MS (ESI):  $[M + 2H^+] = 1256.8$  m/z;  $[M + 3H^+] = 838.2$  m/z;  $C_{114}H_{171}N_{27}O_{35}S$ ; Average isotope calculated 2511.8 Da  $[M]$ ; found: 2511.5 Da.

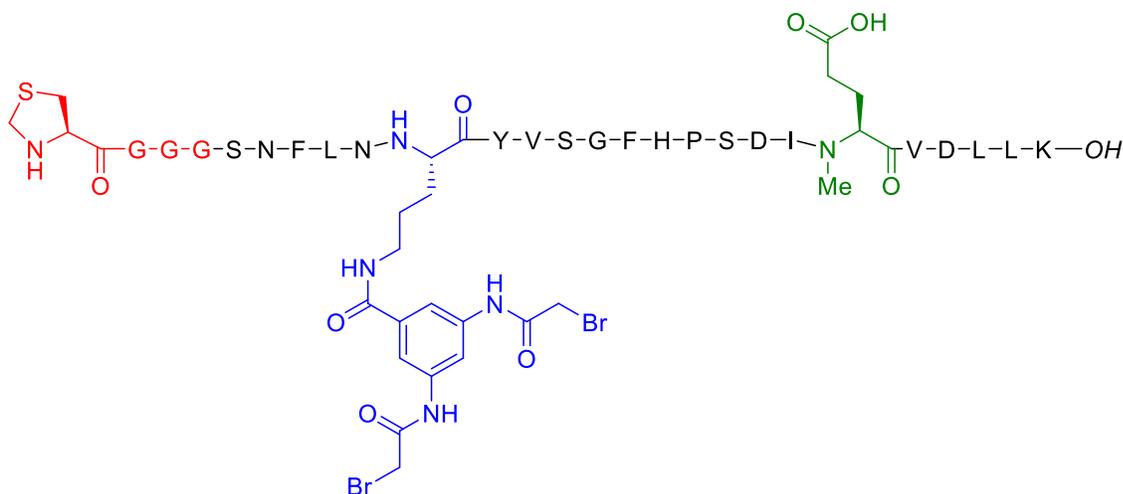
### Synthesis of peptide F: Thz-GGG-Cys25Orn(linker),N-Me-Glu36-[20-41]β2m

Sequence: *Boc-ThzGGG-SNFLNOrn(alloc)YVSGFHPSDI-N-MeEVDLLK-OH*



Thz-GGG-Cys25Orn(linker),N-Me-Glu36-[20-41]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (147 mg, loading: 0.68 mmol/g, mesh size: 180-333). Automated synthesis was carried out by using HATU activation (5 equiv). N-Me-Glu at position 36 was coupled manually (3 equiv, 1 h). Kaiser test was performed to detect the presence of unreacted primary amines. No acetylation step was required. The next amino acid Ile35 was coupled twice (5 equiv, 40 min). Chloranil test was performed to detect the presence of unreacted secondary amines. An acetylation step was required and applied. Ornithine 25 was introduced as Fmoc-Orn(alloc)-OH (5 equiv, 40 min). N-terminal thiazolidine was introduced as Boc-Thz-OH (5 equiv, 40 min).

Sequence: *H*-ThzGGG-SNFLNOrn(linker)YVSGFHPSDI*N*-MeEVDLLK-*OH*



*Alloc deprotection of ornithine:* The resin was swollen in DCM purged with argon. A solution of phenylsilane (296  $\mu$ L, 2.4 mmol, 24 equiv) in DCM (1 mL) purged with argon was added to the resin. Alloc deprotection was performed with the addition of a solution of Pd(PPh<sub>3</sub>)<sub>4</sub> (58 mg, 0.05 mmol, 0.5 equiv) in DCM (2 mL) purged with DCM. It was let to react at r.t. for 30 min with manual stirring. Kaiser test was positive.

*Linker coupling to Orn25 side chain:* The resin was swollen in DMF. A solution of 3,5-bis(((allyloxy)carbonyl)amino)benzoic acid (128 mg, 0.4 mmol, 4 equiv), HATU (144 mg, 0.38 mmol, 3.8 equiv) and DIEA (100  $\mu$ L, 0.6 mmol, 6 equiv) in 1 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 30 min. Kaiser test was negative.

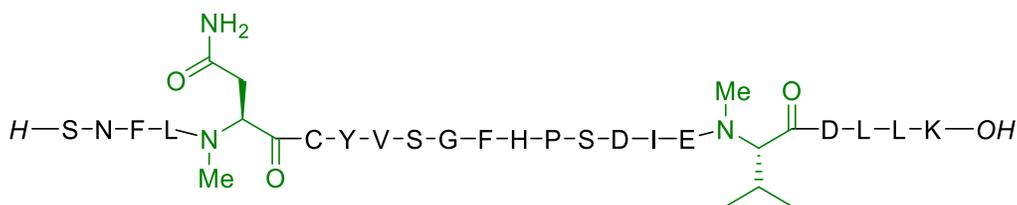
*Alloc deprotection of the linker:* Same procedure as before (repeated twice). Chloranil test was positive.

*Bromoacetylation of the linker:* The resin was swollen in DMF. A solution of bromoacetic acid (167 mg, 1.2 mmol, 12 equiv) and *N,N'*-diisopropylcarbodiimide (86  $\mu$ L, 1.1 mmol, 11.2 equiv) in 2 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 60 min. Chloranil test was negative.

The mass of the dried peptidyl-resin at the end of the synthesis is 490 mg. The peptide was cleaved from the resin by treatment with 4 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 260 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 85 mg (30%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1594.2 *m/z*, 1593.6 *m/z* and 1593.0 *m/z*; [*M* + 3H<sup>+</sup>] = 1063.8 *m/z*, 1062.8 *m/z* and 1062.0 *m/z*; [*M* + 4H<sup>+</sup>] = 798.0 *m/z*, 797.1 *m/z* and 796.3 *m/z*; C<sub>137</sub>H<sub>198</sub>Br<sub>2</sub>N<sub>34</sub>O<sub>42</sub>S; Average isotope calculated 3185.2 Da [*M*]; found: 3182.7 Da, 3185.3 Da and 3188.1 Da.

### Synthesis of peptide G: *N*-Me-MeAsn<sub>24</sub>,*N*-Me-MeVal<sub>37</sub>-[20-41]β<sub>2</sub>m

Sequence: SNFL*N*-MeNCYVSGFHPSDIEN-*Me*VDLLK

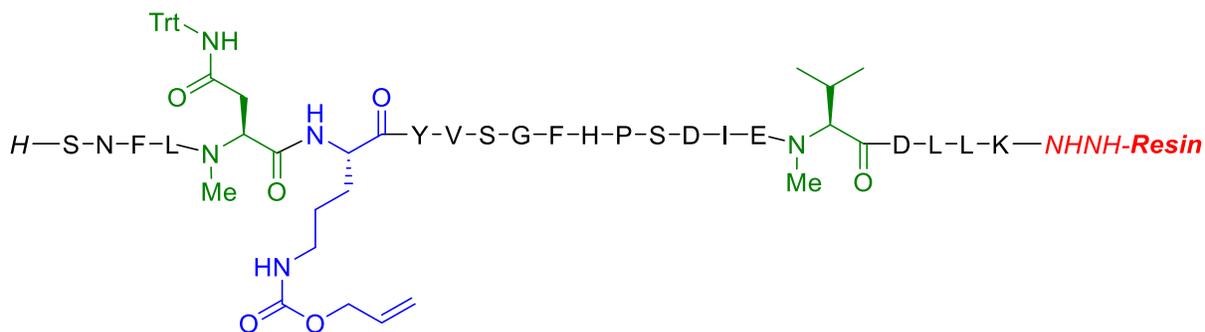


*N*-Me-Asn<sub>24</sub>,*N*-Me-Val<sub>37</sub>-[20-41]β<sub>2</sub>m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (147 mg, loading: 0.68 mmol/g, mesh size: 180-333). Automated microwave synthesis was carried out by using DIC/OxymaPure activation (5 equiv). Leu<sub>24</sub> and Glu<sub>36</sub> were coupled twice. The mass of the dried peptidyl-resin at the end of the synthesis was 455 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 250 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and

acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 76 mg (30%). LC-MS (ESI):  $[M + 2H^+] = 1263.4 m/z$ ;  $[M + 3H^+] = 842.8 m/z$ ;  $C_{115}H_{173}N_{27}O_{35}S$  Average isotope calculated 2525.9 Da  $[M]$ ; found: 2525.3 Da.

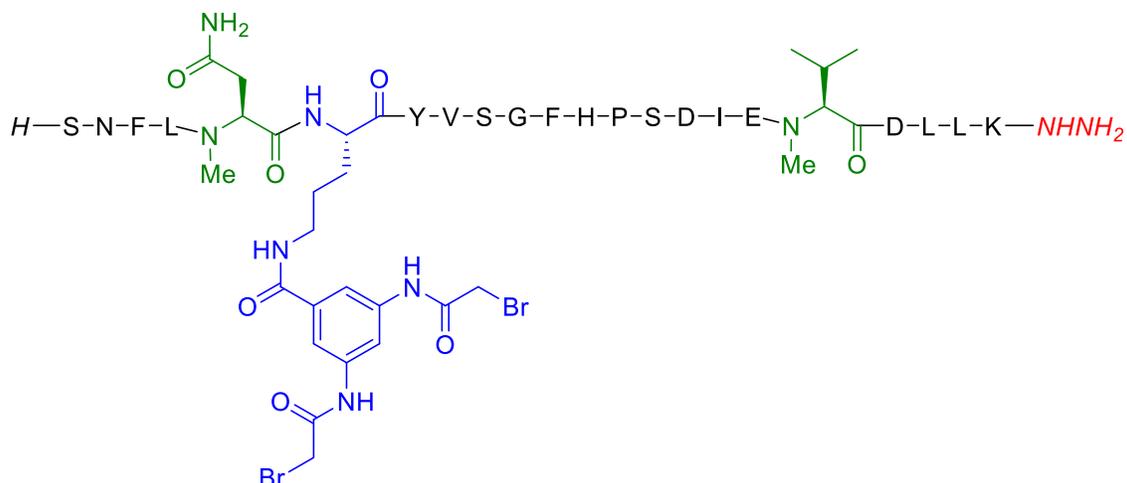
**Synthesis of peptide H: Cys25Orn(linker),N-Me-Asn24,N-Me-Val37-[20-41]β2m**

Sequence: SNFL*N-Me*NOrn(alloc)YVSGFHPSDIEN-*Me*VDLLK-*NHNH-Resin*



Cys25Orn(linker),N-Me-Asn24,N-Me-Val37-[20-41]β2m was synthesized on a 0.1 mmol scale on 2-ChloroTrityl-hydrazide resin (250 mg, loading: 0.40 mmol/g, mesh size: 200-400). Automated microwave synthesis was carried out by using DIC/OxymaPure activation (5 equiv). Leu24 and Glu36 were coupled twice. Ornithine 25 was introduced as Fmoc-Orn(alloc)-OH.

Sequence: SNFL*N-Me*NOrn(linker)YVSGFHPSDIEN-*Me*VDLLK-*NHNH<sub>2</sub>*



*Boc protection:* The resin was swollen in DMF. N-terminal Boc protection of serine 20 was performed with the addition of Boc anhydride (87.5 mg, 0.4 mmol, 4 equiv, in 1 mL of DMF) and DIEA (68  $\mu$ L, 0.4 mmol, 4 equiv). It was let to react at r.t. for 30 min with manual stirring. Kaiser test was negative.

*Alloc deprotection of ornithine:* The resin was swollen in argon purged DCM. A solution of phenylsilane (296  $\mu$ L, 2.4 mmol, 24 equiv) in argon purge DCM (1 mL) was added to the resin. Alloc deprotection was performed with the addition of a solution of Pd(PPh<sub>3</sub>)<sub>4</sub> (58 mg, 0.05 mmol, 0.5 equiv) in argon purged DCM (2 mL). It was let to react at r.t. for 30 min with manual stirring. Kaiser test was positive.

*Linker coupling to Orn26 side chain:* The resin was swollen in DMF. A solution of 3,5-bis(((allyloxy)carbonyl)amino)benzoic acid (128 mg, 0.4 mmol, 4 equiv), HATU (144 mg, 0.38 mmol, 3.8 equiv) and DIEA (100  $\mu$ L, 0.6 mmol, 6 equiv) in 1 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 30 min. Kaiser test was negative.

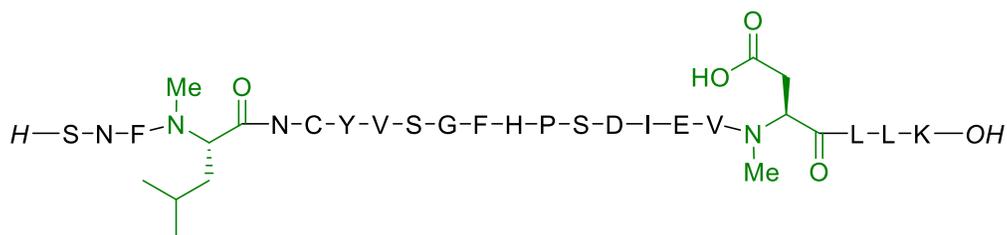
*Alloc deprotection of the linker:* Same procedure as before (repeated twice). Chloranil test was positive.

*Bromoacetylation of the linker:* The resin was swollen in DMF. A solution of bromoacetic acid (167 mg, 1.2 mmol, 12 equiv) and *N,N'*-diisopropylcarbodiimide (86  $\mu$ L, 1.1 mmol, 11.2 equiv) in 2 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 60 min. Chloranil test was negative.

The mass of the dried peptidyl-resin at the end of the synthesis was 1 g. The peptide was cleaved from the resin by treatment with 10 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing (0.1%, v/v) and lyophilized. Mass of crude peptide was 694 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 150 mg (33%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1465.5 *m/z*, 1464.5 *m/z* and 1465.5 *m/z*; [*M* + 3H<sup>+</sup>] = 975.7 *m/z*, 976.4 *m/z* and 977.2 *m/z*; [*M* + 4H<sup>+</sup>] = 732.0 *m/z*, 732.7 *m/z* and 733.3 *m/z*; C<sub>128</sub>H<sub>188</sub>Br<sub>2</sub>N<sub>32</sub>O<sub>37</sub>; Average isotope calculated 2926.9 Da [*M*]; found: 2923.9 Da, 2926.3 Da and 2929.3 Da.

### Synthesis of peptide I: *N*-Me-Leu23,*N*-Me-Asp38-[20-41]β2m

Sequence: SNFN-*Me*LNCYVSGFHPSDIEV*N*-*Me*DLLK

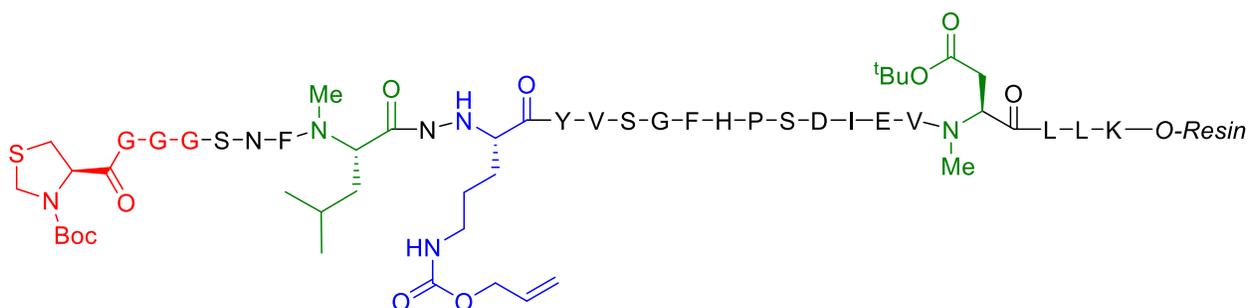


*N*-Me-Leu23,*N*-Me-Asp38-[20-41]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (147 mg, loading: 0.68 mmol/g, mesh size: 180-333). Automated microwave synthesis was carried out by using DIC/OxymaPure activation (5 equiv). Phe22 and Val37 were coupled twice. The mass of the dried peptidyl-resin at the end of the synthesis was 475 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 225 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-

C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 77 mg (30%). LC-MS (ESI):  $[M + 2H^+] = 1263.6$  m/z;  $[M + 3H^+] = 842.6$  m/z;  $C_{115}H_{173}N_{27}O_{35}S$ ; Average isotope calculated 2525.9 Da  $[M]$ ; found: 2525.0 Da.

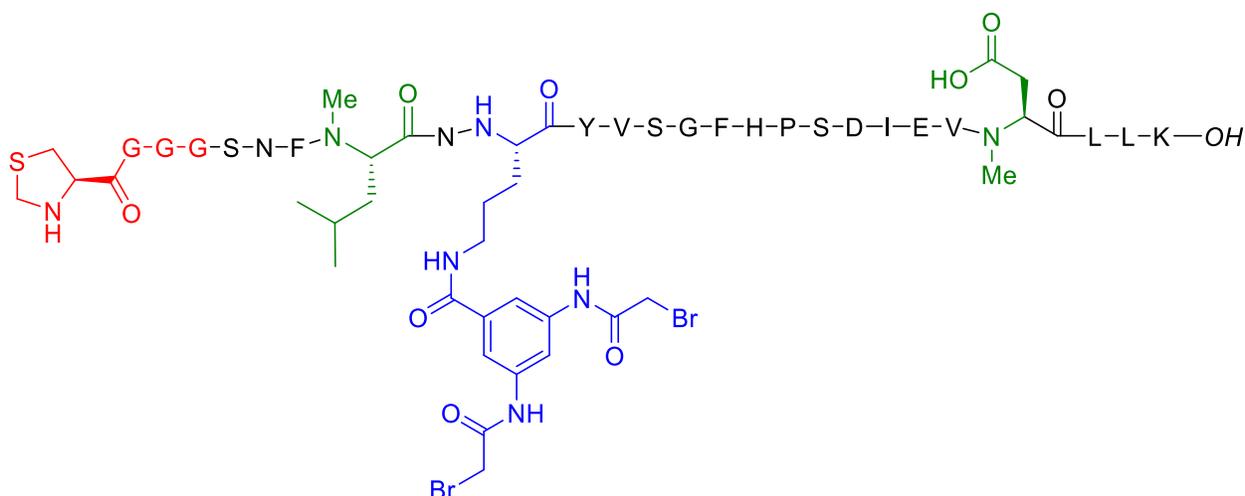
**Synthesis of peptide J: Thz-GGG-Cys25Orn(linker),N-Me-Leu23,N-Me-Asp38-[20-41]β2m**

Sequence: *Boc-ThzGGG-SNFN-MeLNOrn(alloc)YVSGFHPSDIEVN-MeDLLK-OH*



Thz-GGG-Cys25Orn(linker),N-Me-Leu23,N-Me-Asp38-[20-41]β2m was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (147 mg, loading: 0.68 mmol/g, mesh size: 180-333). Automated microwave synthesis was carried out by using DIC/OxymaPure activation (5 equiv). Phe22 and Val37 were coupled twice. Ornithine 25 was introduced as Fmoc-Orn(alloc)-OH and N-terminal thiazolidine was introduced as Boc-Thz-OH.

Sequence: *H-ThzGGG-SNFN-MeLNOrn(linker)YVSGFHPSDIEVN-MeDLLK-OH*



*Alloc deprotection of ornithine:* The resin was swollen in DCM purged with argon. A solution of phenylsilane (296  $\mu$ L, 2.4 mmol, 24 equiv) in DCM (1 mL) purged with argon was added to the resin. Alloc deprotection was performed with the addition of a solution of Pd(PPh<sub>3</sub>)<sub>4</sub> (58 mg, 0.05 mmol, 0.5 equiv) in argon purged DCM (2 mL). It was let to react at r.t. for 30 min with manual stirring. Kaiser test was positive.

*Linker coupling to Orn25 side chain:* The resin was swollen in DMF. A solution of 3,5-bis(((allyloxy)carbonyl)amino)benzoic acid (128 mg, 0.4 mmol, 4 equiv), HATU (144 mg, 0.38 mmol, 3.8 equiv) and DIEA (100  $\mu$ L, 0.6 mmol, 6 equiv) in 1 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 30 min. Kaiser test was negative.

*Alloc deprotection of the linker:* Same procedure as before (repeated twice). Chloranil test was positive.

*Bromoacetylation of the linker:* The resin was swollen in DMF. A solution of bromoacetic acid (167 mg, 1.2 mmol, 12 equiv) and *N,N'*-diisopropylcarbodiimide (86  $\mu$ L, 1.1 mmol, 11.2 equiv) in 2 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 60 min. Chloranil test was negative.

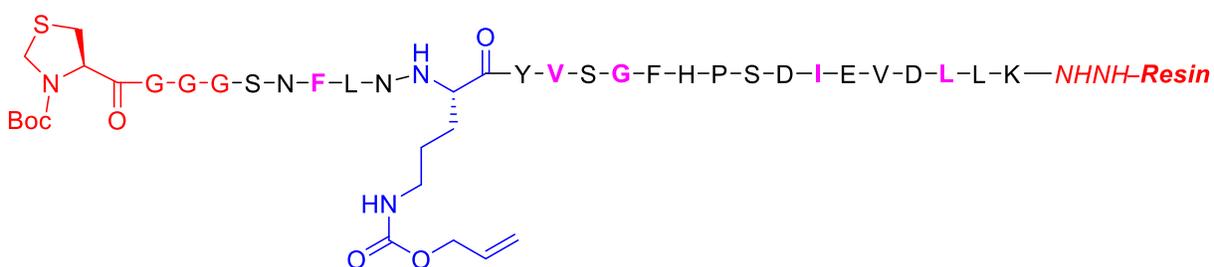
The mass of the dried peptidyl-resin at the end of the synthesis is 498 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of

the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 277 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 71 mg (22%). LC-MS (ESI): [M + 2H<sup>+</sup>] = 1600.5 m/z, 1599.9 m/z and 1599.3 m/z; [M + 3H<sup>+</sup>] = 1068.1 m/z, 1067.3 m/z and 1066.6 m/z; C<sub>138</sub>H<sub>200</sub>Br<sub>2</sub>N<sub>34</sub>O<sub>42</sub>S; Average isotope calculated 3199.2 Da [M]; found: 3196.8 Da, 3199.2 Da and 3201.7 Da.

### Synthesis of peptide K: (<sup>15</sup>N-labeled)-Thz-GGG-Cys25Orn(linker)-[20-41]β2m

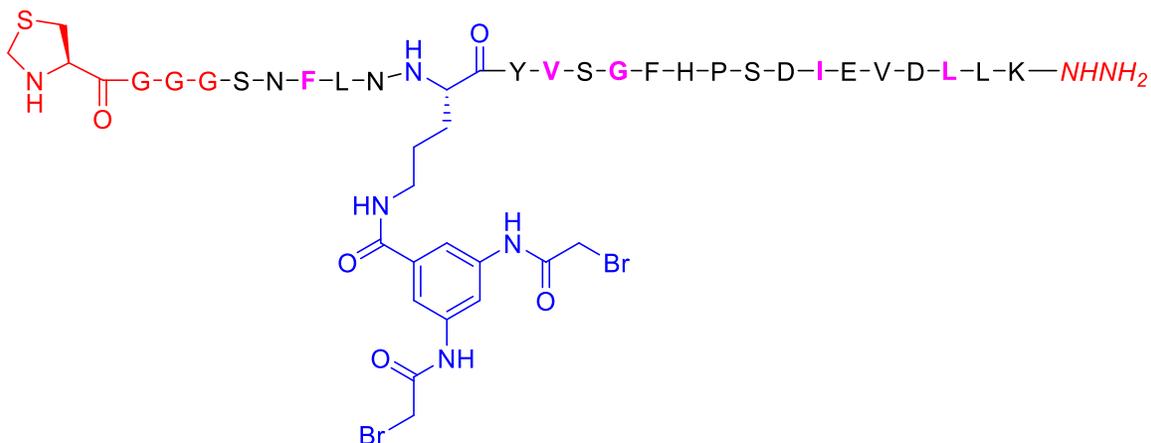
Sequence: Boc-ThzGGG-SN<sup>FLN</sup>Orn(alloc)Y<sup>V</sup>S<sup>G</sup>FHPSD<sup>IEVD</sup>LLK-NHNH-Resin

The amino acid positions highlighted in magenta are <sup>15</sup>N-labeled.



(<sup>15</sup>N-labeled)-Thz-GGG-Cys25Orn(alloc)-β2m[20-41] was synthesized on a 0.2 mmol scale on 2-ChloroTrityl-hydrazide resin (500 mg, loading: 0.40 mmol/g, mesh size: 200-400). Manual synthesis was carried out by using HATU activation (5 equiv). Ornithine 25 was introduced as Fmoc-Orn(alloc)-OH (5 equiv, 40 min). N-terminal thiazolidine was introduced as Boc-Thz-OH (5 equiv, 40 min). Phe22, Val27, Gly29, Ile35 and Leu39 were introduced as <sup>15</sup>N-labeled amino acids.

Sequence: *H*-ThzGGG-SN**F**LNOrn(linker)**Y**V**S**GFHP**S**D**I**EV**D**LL**K**-*NHNH*<sub>2</sub>



*Alloc deprotection of ornithine:* The resin was swollen in DCM purged with argon. A solution of phenylsilane (592  $\mu$ L, 4.8 mmol, 24 equiv) in DCM (2 mL) purged with argon was added to the resin. Alloc deprotection was performed with the addition of a solution of Pd(PPh<sub>3</sub>)<sub>4</sub> (116 mg, 0.1 mmol, 0.5 equiv) in argon purged DCM (4 mL). It was let to react at r.t. for 30 min with manual stirring. Kaiser test was positive.

*Linker coupling to Orn25 side chain:* The resin was swollen in DMF. A solution of 3,5-bis(((allyloxy)carbonyl)amino)benzoic acid (256 mg, 0.8 mmol, 4 equiv), HATU (288 mg, 0.76 mmol, 3.8 equiv) and DIEA (200  $\mu$ L, 1.2 mmol, 6 equiv) in 4 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 30 min. Kaiser test was negative.

*Alloc deprotection of the linker:* Same procedure as before (repeated twice). Chloranil test was positive.

*Bromoacetylation of the linker:* The resin was swollen in DMF. A solution of bromoacetic acid (333 mg, 2.4 mmol, 12 equiv) and *N,N'*-diisopropylcarbodiimide (172  $\mu$ L, 2.2 mmol, 11.2 equiv) in 4 mL of DMF was prepared (2 min of pre-activation) and added to the resin. It was let to react at r.t. for 60 min. Chloranil test was negative.

The mass of the dried peptidyl-resin at the end of the synthesis was 1.60 g. The peptide was cleaved from the resin by treatment with 16 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of

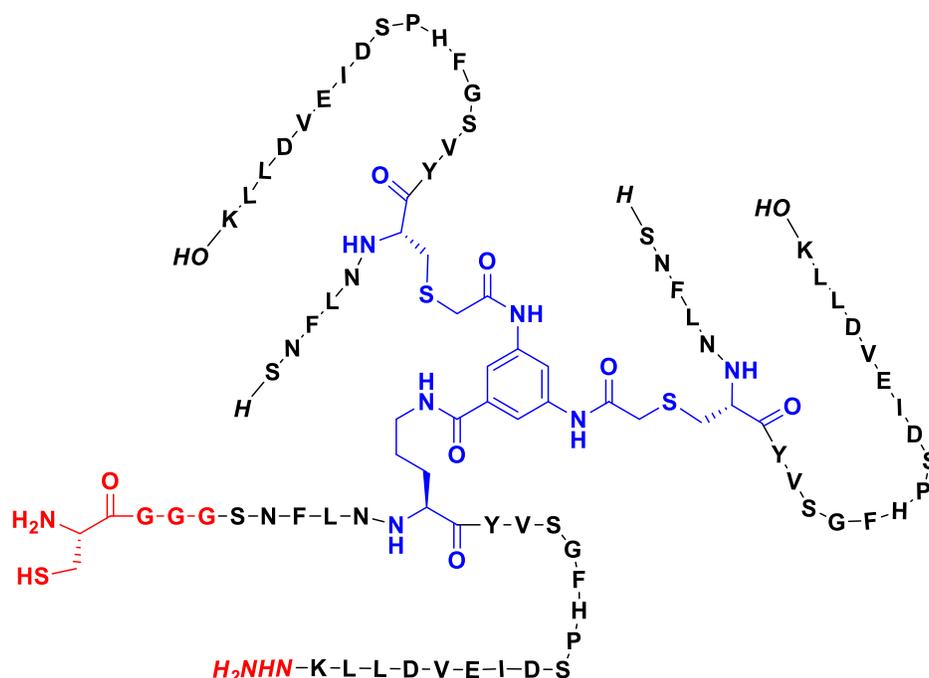
the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 568 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 107 mg (17%). LC-MS (ESI):  $[M + 2H^+] = 1596.9$   $m/z$ , 1596.1  $m/z$  and 1595.4  $m/z$ ;  $[M + 3H^+] = 1065.7$   $m/z$ , 1064.8  $m/z$  and 1064.2  $m/z$ ;  $[M + 4H^+] = 799.2$   $m/z$ , 798.6  $m/z$  and 798.1  $m/z$ ; C<sub>137</sub>H<sub>198</sub>Br<sub>2</sub>N<sub>34</sub>O<sub>42</sub>S; Average isotope calculated 3190.1 Da  $[M]$ ; found: 3189.1 Da, 3191.3 Da and 3194.1 Da.

### Synthesis of covalent trimer-1-A

Description: ligation of 2 equivalents of peptide **A** with peptide **B** through bromide displacement reaction → formation of a trimeric structure. Oxidation of peptide hydrazide to peptide azide followed by thioesterification through addition/elimination of the azide by sodium 2-mercaptoethanesulfonate.



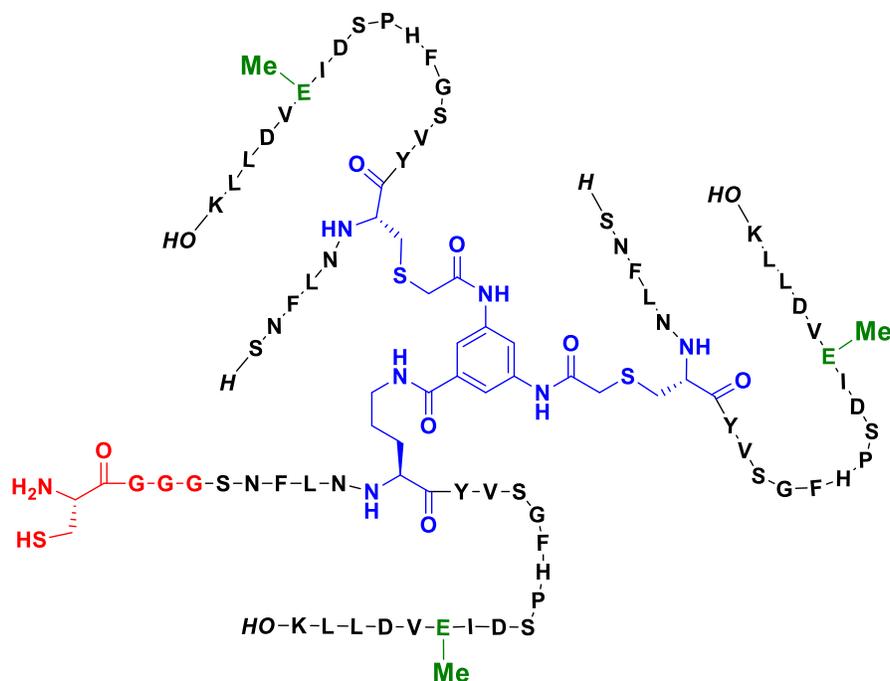
Description: Ligation of 2 equivalents of **C** with **D** through bromide displacement reaction → formation of a trimeric structure. The ligation is followed by the conversion of thiazolidine to cysteine.



Peptide **C** (34.8 mg, 13.9  $\mu\text{mol}$ ) and peptide **D** (22.0 mg, 6.9  $\mu\text{mol}$ ) were dissolved in 9.2 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{Na}_4\cdot\text{EDTA}] = 5 \text{ mM}$ ,  $\text{pH} = 8.0$ ). The pH was adjusted to 8.0 with a 1 M  $\text{Na}_2\text{CO}_{3(\text{aq})}$  solution. It was let to react at 30°C for 2 h. Methoxyammonium chloride (154 mg, 16.7 mg/mL) was added to the reaction mixture. The pH was then lowered to 4 with 6 M  $\text{HCl}(\text{aq})$ . It was let to react at 30°C for 2 h. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 22.5 mg (41%). LC-MS (ESI):  $[M + 5 \text{H}^+] = 1601.9 \text{ m/z}$ ;  $[M + 6\text{H}^+] = 1335.2 \text{ m/z}$ ;  $[M + 7\text{H}^+] = 1144.8 \text{ m/z}$ ;  $[M + 8\text{H}^+] = 1001.7 \text{ m/z}$ ;  $[M + 9\text{H}^+] = 890.6 \text{ m/z}$ ;  $\text{C}_{361}\text{H}_{534}\text{N}_{90}\text{O}_{111}\text{S}_3$ ; Average isotope calculated 8006.9 Da  $[M]$ ; found: 8006.1 Da.

### Synthesis of covalent trimer-3-A

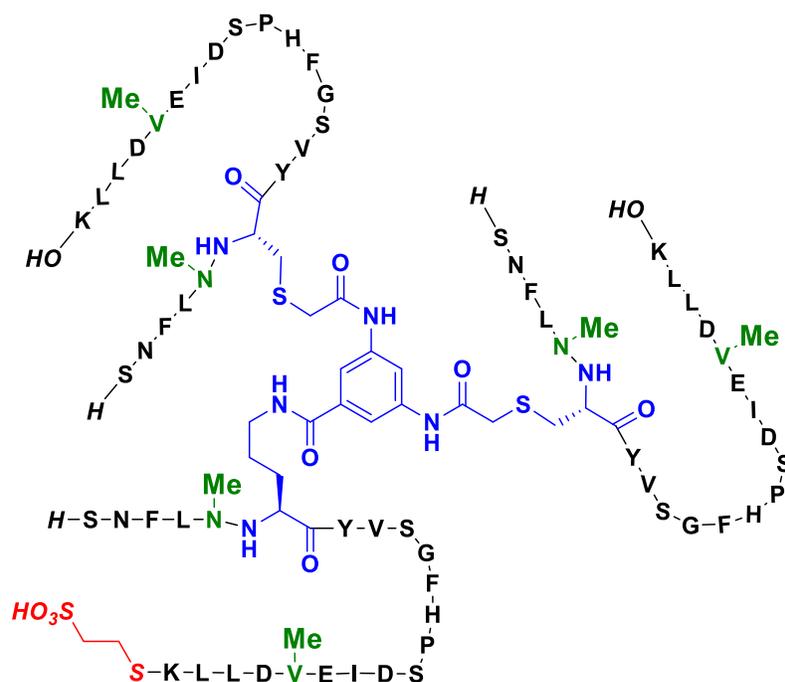
Description: Ligation of 2 equivalents of **E** with **F** through bromide displacement reaction → formation of a trimeric structure. The ligation is followed by the conversion of thiazolidine to cysteine.



Peptide **E** (23.6 mg, 9.4  $\mu\text{mol}$ ) and peptide **F** (14.0 mg, 4.7  $\mu\text{mol}$ ) were dissolved in 4 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{Na}_4\cdot\text{EDTA}] = 5 \text{ mM}$ ,  $\text{pH} = 8.0$ ). The pH was adjusted to 8.0 with a 1 M  $\text{Na}_2\text{CO}_{3(\text{aq})}$  solution. It was let to react at 30°C for 2 h. Methoxyammonium chloride (67 mg, 16.7 mg/mL) was added to the reaction mixture. The pH was then lowered to 4 with 6 M  $\text{HCl}(\text{aq})$ . It was let to react at 30°C for 2 h. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 20.3 mg (54%). LC-MS (ESI):  $[M + 5\text{H}^+] = 1608.0 \text{ m/z}$ ;  $[M + 6\text{H}^+] = 1340.3 \text{ m/z}$ ;  $[M + 7\text{H}^+] = 1149.0 \text{ m/z}$ ;  $[M + 8\text{H}^+] = 1005.5 \text{ m/z}$ ;  $[M + 9\text{H}^+] = 894.1 \text{ m/z}$ ;  $\text{C}_{364}\text{H}_{538}\text{N}_{88}\text{O}_{112}\text{S}_3$ ; Average isotope calculated 8035.8 Da  $[M]$ ; found: 8035.0 Da.

### Synthesis of covalent trimer-1-B

Description: ligation of 2 equivalents of peptide **G** with peptide **H** through bromide displacement reaction → formation of a trimeric structure. Oxidation of peptide hydrazide to peptide azide followed by thioesterification through addition/elimination of the azide by sodium 2-mercaptoethanesulfonate.

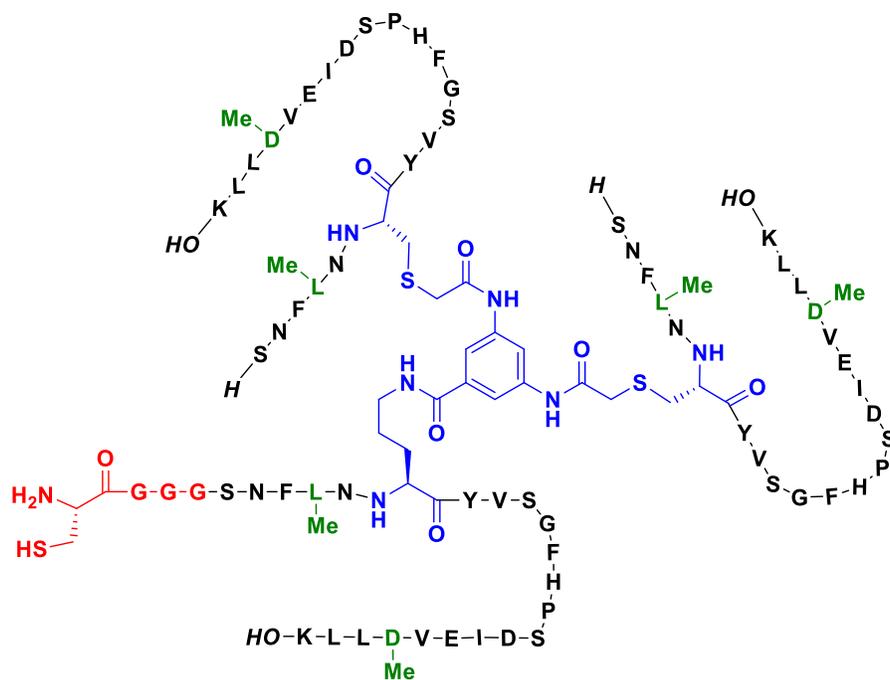


Peptide **G** (35.4 mg, 14  $\mu\text{mol}$ ) and peptide **H** (20.5 mg, 7  $\mu\text{mol}$ ) were dissolved in 5.8 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{Na}_4\cdot\text{EDTA}] = 5 \text{ mM}$ ,  $\text{pH} = 8.0$ ). The pH was adjusted to 8.0 with a 1 M  $\text{Na}_2\text{CO}_{3(\text{aq})}$  solution. It was let to react at  $30^\circ\text{C}$  for 2 h. The pH was then lowered to 3 with 6 M  $\text{HCl}(\text{aq})$ . Sodium nitrite, solubilized in argon purged pH 3 buffer (35  $\mu\text{mol}$ ,  $c = 0.2 \text{ M}$ , 175  $\mu\text{L}$ , 5 equiv), was added to the reaction mixture. It was let to react at  $-15^\circ\text{C}$  for 15 min under argon. Sodium 2-mercaptoethanesulfonate, solubilized in argon purged pH 7 buffer (350  $\mu\text{mol}$ ,  $c = 0.2 \text{ M}$ , 1.75 mL, 50 equiv). It was let to react at  $-15^\circ\text{C}$  for 2 h under argon. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100  $\text{\AA}$ , 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 27.1 mg (49%). LC-MS (ESI):  $[M + 4\text{H}^+] = 1982.6 \text{ m/z}$ ;  $[M + 5\text{H}^+] = 1586.2 \text{ m/z}$ ;  $[M + 6\text{H}^+] = 1322.0 \text{ m/z}$ ;  $[M + 7\text{H}^+] = 1133.4 \text{ m/z}$ ;

$[M + 8H^+] = 991.7 m/z$ ;  $C_{360}H_{534}N_{84}O_{110}S_4$ ; Average isotope calculated 7926.9 Da  $[M]$ ; found: 7926.2 Da.

### Synthesis of covalent trimer-3-B

Description: Ligation of 2 equivalents of **I** with **J** through bromide displacement reaction → formation of a trimeric structure. The ligation is followed by the conversion of thiazolidine to cysteine.



Peptide **I** (35.1 mg, 13.9  $\mu\text{mol}$ ) and peptide **J** (22.1 mg, 6.9  $\mu\text{mol}$ ) were dissolved in 9.2 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{Na}_4\cdot\text{EDTA}] = 5 \text{ mM}$ ,  $\text{pH} = 8.0$ ). The pH was adjusted to 8.0 with a 1 M  $\text{Na}_2\text{CO}_{3(\text{aq})}$  solution. It was let to react at 30°C for 2 h. Methoxyammonium chloride (154 mg, 16.7 mg/mL) was added to the reaction mixture. The pH was then lowered to 4 with 6 M  $\text{HCl}(\text{aq})$ . It was let to react at 30°C for 2 h. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 36.8 mg (66%). LC-MS (ESI):  $[M + 5H^+] = 1616.4 m/z$ ;  $[M + 6H^+]$



= 1336.3  $m/z$ ;  $[M + 7H^+] = 1145.5 m/z$ ;  $[M + 8H^+] = 1002.2 m/z$ ;  $[M + 9H^+] = 891.0 m/z$ ;  
 $C_{361}H_{534}N_{90}O_{111}S_3$ ; Average isotope calculated 8011.0 Da  $[M]$ ; found: 8011.4 Da.

### Synthesis of the covalent dimer of trimers 1: DofT-1

Description: Native chemical ligation between **trimer-1-A** and **trimer-2-A**  $\rightarrow$  formation of a covalent dimer of covalent trimers. The ligation is followed by the alkylation of the cysteine at the ligation site and by the conversion of the hydrazide moiety into MesNa thioester localized on peptide **D** (trimer-2-A).

**Trimer-1-A** (5.0 mg, 0.63  $\mu$ mol) and **trimer-2-A** (5.1 mg, 0.63  $\mu$ mol) were dissolved in 500  $\mu$ L of argon purged buffer 1 ( $[Na_2HPO_4] = 0.2$  M,  $[Gn \cdot HCl] = 6$  M, pH = 7.0) containing MPAA (13 mM, 6.3  $\mu$ mol, 10 equiv) and TCEP (20 mM, 10.1  $\mu$ mol, 16 equiv). The pH was adjusted to 7.0 with a 1 M  $Na_2CO_{3(aq)}$  solution after peptide dissolution. It was let to react at 40°C for 2 h under argon. Bromoacetamide (79  $\mu$ L, 0.2 M, 16  $\mu$ mol, 25 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react at 40°C for 30 min. The pH was adjusted to 3.0 with a 6 M  $HCl_{(aq)}$  solution. At -10°C, sodium nitrite (32  $\mu$ L, 0.2 M, 6.3  $\mu$ mol, 10 equiv) was solubilized in buffer 2 ( $[Na_2HPO_4] = 0.2$  M,  $[Gn \cdot HCl] = 6$  M, pH = 3.0) and was added to the reaction mixture. It was let to react for 20 min. At -10°C, sodium 2-mercaptoethanesulfonate (320  $\mu$ L, 0.2 M, 63  $\mu$ mol, 100 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react for 2 h. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 4.2 mg (54%). LC-MS (ESI):  $[M + 8H^+] = 1990.8 m/z$ ;  $[M + 9H^+] = 1769.6 m/z$ ;  $[M + 10H^+] = 1592.6 m/z$ ;  $[M + 11H^+] = 1448.1 m/z$ ;  $[M + 12H^+] = 1327.6 m/z$ ;  $[M + 13H^+] = 1225.4 m/z$ ;  $[M + 14H^+] = 1138.1 m/z$ ;  $[M + 15H^+] = 1062.4 m/z$ ;  $[M + 16H^+] = 996.3 m/z$ ; Average isotope calculated 15916.8 Da  $[M]$ ; found: 15919.2 Da. There is a great error between the calculated and experimental mass due to an incorrect calibration of the LC-MS apparatus.

### Synthesis of the covalent dimer of trimers 2: DofT-2

Description: Native chemical ligation between **trimer-1-B** and **trimer-2-A** → formation of a covalent dimer of covalent trimers. The ligation is followed by the alkylation of the cysteine at the ligation site and by the conversion of the hydrazide moiety into MesNa thioester localized on peptide **D** (trimer-2-A).

**Trimer-1-B** (10.0 mg, 1.3  $\mu\text{mol}$ ) and **trimer-2-A** (10.1 mg, 1.3  $\mu\text{mol}$ ) were dissolved in 1.0 mL of argon purged buffer 1 ( $[\text{Na}_2\text{HPO}_4] = 0.2 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $\text{pH} = 7.0$ ) containing MPAA (13 mM, 12.6  $\mu\text{mol}$ , 10 equiv) and TCEP (20 mM, 20.2  $\mu\text{mol}$ , 16 equiv). The pH was adjusted to 7.0 with a 1 M  $\text{Na}_2\text{CO}_{3(\text{aq})}$  solution after peptide dissolution. It was let to react at 40°C for 2 h under argon. Bromoacetamide (158  $\mu\text{L}$ , 0.2 M, 32.5  $\mu\text{mol}$ , 25 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react at 40°C for 30 min. The pH was adjusted to 3.0 with a 6 M  $\text{HCl}_{(\text{aq})}$  solution. At -10°C, sodium nitrite (64  $\mu\text{L}$ , 0.2 M, 12.6  $\mu\text{mol}$ , 10 equiv) was solubilized in buffer 2 ( $[\text{Na}_2\text{HPO}_4] = 0.2 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $\text{pH} = 3.0$ ) and was added to the reaction mixture. It was let to react for 20 min. At -10°C, sodium 2-mercaptoethanesulfonate (640  $\mu\text{L}$ , 0.2 M, 126  $\mu\text{mol}$ , 100 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react for 2 h. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 7.0 mg (34%). LC-MS (ESI):  $[M + 9\text{H}^+] = 1773.9 \text{ m/z}$ ;  $[M + 10\text{H}^+] = 1596.6 \text{ m/z}$ ;  $[M + 11\text{H}^+] = 1451.6 \text{ m/z}$ ;  $[M + 12\text{H}^+] = 1330.8 \text{ m/z}$ ;  $[M + 13\text{H}^+] = 1228.5 \text{ m/z}$ ;  $[M + 14\text{H}^+] = 1140.8 \text{ m/z}$ ;  $[M + 15\text{H}^+] = 1064.9 \text{ m/z}$ ;  $[M + 16\text{H}^+] = 998.3 \text{ m/z}$ ;  $\text{C}_{723}\text{H}_{1067}\text{N}_{173}\text{O}_{222}\text{S}_7$ ; Average isotope calculated 15958.90 Da  $[M]$ ; found: 15957.0 Da.

### Synthesis of the covalent dimer of trimers 3: DofT-3

Description: Native chemical ligation between **trimer-1-B** and **trimer-2-B** → formation of a covalent dimer of covalent trimers. The ligation is followed by the alkylation of the cysteine at the ligation site and by the conversion of the hydrazide moiety into MesNa thioester localized on peptide **K** (trimer-2-A).

**Trimer-1-B** (8.3 mg, 1.08  $\mu\text{mol}$ ) and **trimer-2-B** (8.3 mg, 1.08  $\mu\text{mol}$ ) were dissolved in 860  $\mu\text{L}$  of argon purged buffer 1 ( $[\text{Na}_2\text{HPO}_4] = 0.2 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $\text{pH} = 7.0$ ) containing MPAA (13 mM, 10.8  $\mu\text{mol}$ , 10 equiv) and TCEP (20 mM, 17.3  $\mu\text{mol}$ , 16 equiv). The pH was adjusted to 7.0 with a 1 M  $\text{Na}_2\text{CO}_{3(\text{aq})}$  solution after peptide dissolution. It was let to react at 40°C for 2 h under argon. Bromoacetamide (135  $\mu\text{L}$ , 0.2 M, 27  $\mu\text{mol}$ , 25 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react at 40°C for 30 min. The pH was adjusted to 3.0 with a 6 M  $\text{HCl}_{(\text{aq})}$  solution. At -10°C, sodium nitrite (55  $\mu\text{L}$ , 0.2 M, 10.8  $\mu\text{mol}$ , 10 equiv) was solubilized in buffer 2 ( $[\text{Na}_2\text{HPO}_4] = 0.2 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $\text{pH} = 3.0$ ) and was added to the reaction mixture. It was let to react for 20 min. At -10°C, sodium 2-mercaptoethanesulfonate (550  $\mu\text{L}$ , 0.2 M, 108  $\mu\text{mol}$ , 100 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react for 2 h. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 5.0 mg (29%). LC-MS (ESI):  $[M + 8\text{H}^+] = 1996.5 \text{ m/z}$ ;  $[M + 9\text{H}^+] = 1774.5 \text{ m/z}$ ;  $[M + 10\text{H}^+] = 1597.4 \text{ m/z}$ ;  $[M + 11\text{H}^+] = 1452.1 \text{ m/z}$ ;  $[M + 12\text{H}^+] = 1331.3 \text{ m/z}$ ;  $[M + 13\text{H}^+] = 1229.0 \text{ m/z}$ ;  $[M + 14\text{H}^+] = 1141.2 \text{ m/z}$ ;  $[M + 15\text{H}^+] = 1065.2 \text{ m/z}$ ;  $[M + 16\text{H}^+] = 998.7 \text{ m/z}$ ;  $\text{C}_{723}\text{H}_{1067}\text{N}_{173}\text{O}_{222}\text{S}_7$ ; Average isotope calculated 15964.0 Da  $[M]$ ; found: 15963.5 Da.

### Synthesis of the covalent trimer of trimers 1: **TofT-1**

Description: Native chemical ligation between the covalent dimer of trimers 1 (**DofT-1**) and **trimer-3-A**  $\rightarrow$  formation of a covalent trimer of covalent trimers. The ligation is followed by the alkylation of the cysteine at the ligation site.

**Trimer-3-A** (8.2 mg, 1.06  $\mu\text{mol}$ ) and **DofT-1** (8.2 mg, 0.53  $\mu\text{mol}$ ) were dissolved in 250  $\mu\text{L}$  of argon purged buffer 1 ( $[\text{Na}_2\text{HPO}_4] = 0.2 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $\text{pH} = 7.0$ ) containing MPAA (11 mM, 5.7  $\mu\text{mol}$ , 5 equiv) and TCEP (34 mM, 2.02  $\mu\text{mol}$ , 16 equiv). The pH was adjusted to 6.8 with a 1 M  $\text{Na}_2\text{CO}_{3(\text{aq})}$  solution after peptide dissolution. It was let to react at 40°C for 15 min under argon. 500  $\mu\text{L}$  of buffer 1 was added. Bromoacetamide (66  $\mu\text{L}$ , 0.2 M, 13.3  $\mu\text{mol}$ , 25 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react at r.t. for

30 min. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 3.2 mg (25%). LC-MS (ESI): [ $M + 12H^+$ ] = 1989.8  $m/z$ ; [ $M + 13H^+$ ] = 1837.7  $m/z$ ; [ $M + 14H^+$ ] = 1706.2  $m/z$ ; [ $M + 15H^+$ ] = 1592.3  $m/z$ ; [ $M + 16H^+$ ] = 1492.8  $m/z$ ; [ $M + 17H^+$ ] = 1405.1  $m/z$ ; [ $M + 18H^+$ ] = 1327.1  $m/z$ ; [ $M + 19H^+$ ] = 1257.5  $m/z$ ; [ $M + 20H^+$ ] = 1194.5  $m/z$ ; [ $M + 21H^+$ ] = 1137.8  $m/z$ ; [ $M + 22H^+$ ] = 1086.0  $m/z$ ; [ $M + 23H^+$ ] = 1038.9  $m/z$ ;  $C_{1084}H_{1596}N_{262}O_{332}S_8$ ; Average isotope calculated 23866.7 Da [ $M$ ]; found: 23870.6 Da. There is a great error between the calculated and experimental mass due to an incorrect calibration of the LC-MS apparatus. ESI-MS-Orbitrap: [ $M + 12H^+$ ] = 1989.9  $m/z$ ; [ $M + 13H^+$ ] = 1836.9  $m/z$ ; [ $M + 14H^+$ ] = 1705.8  $m/z$ ; [ $M + 15H^+$ ] = 1592.0  $m/z$ ; [ $M + 16H^+$ ] = 1492.7  $m/z$ ; [ $M + 17H^+$ ] = 1404.9  $m/z$ ; [ $M + 18H^+$ ] = 1326.9  $m/z$ ; [ $M + 19H^+$ ] = 1257.1  $m/z$ ; [ $M + 20H^+$ ] = 1194.3  $m/z$ ; [ $M + 21H^+$ ] = 1137.6  $m/z$ ;  $C_{1084}H_{1596}N_{262}O_{332}S_8$ ; Average isotope calculated 23866.7 Da [ $M$ ]; found: 23866.4 Da.

### Synthesis of the covalent trimer of trimers 2: **TofT-2**

Description: Native chemical ligation between the covalent dimer of trimers 2 (**DofT-2**) and **trimer-3-B** → formation of a covalent trimer of covalent trimers. The ligation is followed by the alkylation of the cysteine at the ligation site.

**Trimer-3-B** (6.5 mg, 0.8 µmol) and **DofT-2** (6.6 mg, 0.4 µmol) were dissolved in 200 µL of argon purged buffer 1 ( $[Na_2HPO_4] = 0.2$  M,  $[Gn \cdot HCl] = 6$  M, pH = 7.0) containing MPAA (11 mM, 2 µmol, 5 equiv) and TCEP (34 mM, 6.6 µmol, 16 equiv). The pH was adjusted to 6.8 with a 1 M  $Na_2CO_{3(aq)}$  solution after peptide dissolution. It was let to react at 40°C for 15 min under argon. 400 µL of buffer 1 was added. Bromoacetamide (51 µL, 0.2 M, 13.3 µmol, 25 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react at r.t. for 30 min. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 3.7 mg (38%). LC-MS (ESI): [ $M + 13H^+$ ] = 1843.2

$m/z$ ;  $[M + 14H^+] = 1711.5 m/z$ ;  $[M + 15H^+] = 1597.7 m/z$ ;  $[M + 16H^+] = 1497.4 m/z$ ;  $[M + 17H^+] = 1409.7 m/z$ ;  $[M + 18H^+] = 1331.6 m/z$ ;  $[M + 19H^+] = 1261.4 m/z$ ;  $[M + 20H^+] = 1198.6 m/z$ ;  $[M + 21H^+] = 1141.6 m/z$ ;  $[M + 22H^+] = 1089.7 m/z$ ;  $[M + 23H^+] = 1042.6 m/z$ ;  $[M + 24H^+] = 999.4 m/z$ ;  $C_{1090}H_{1608}N_{262}O_{332}S_8$ ; Average isotope calculated 23950.8 Da  $[M]$ ; found: 23950.3 Da. ESI-MS-Orbitrap:  $[M + 12H^+] = 1996.9 m/z$ ;  $[M + 13H^+] = 1843.3 m/z$ ;  $[M + 14H^+] = 1711.8 m/z$ ;  $[M + 15H^+] = 1597.6 m/z$ ;  $[M + 16H^+] = 1497.9 m/z$ ;  $[M + 17H^+] = 1409.9 m/z$ ;  $[M + 18H^+] = 1331.6 m/z$ ;  $[M + 19H^+] = 1261.6 m/z$ ;  $[M + 20H^+] = 1198.5 m/z$ ;  $[M + 21H^+] = 1141.5 m/z$ ;  $C_{1084}H_{1596}N_{262}O_{332}S_8$ ; Average isotope calculated 23950.8 Da  $[M]$ ; found: 23950.5 Da.

### Synthesis of the covalent trimer of trimers 3: **ToFT-3**

Description: Native chemical ligation between the covalent dimer of trimers 3 (**DoFT-3**) and **trimer-3-B** → formation of a covalent trimer of covalent trimers. The ligation is followed by the alkylation of the cysteine at the ligation site.

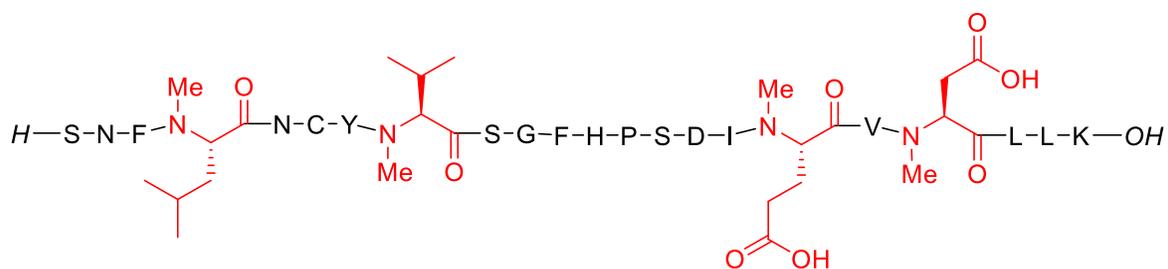
**Trimer-3-B** (3.5 mg, 0.44  $\mu$ mol) and **DoFT-3** (3.6 mg, 0.22  $\mu$ mol) were dissolved in 110  $\mu$ L of argon purged buffer 1 ( $[Na_2HPO_4] = 0.2$  M,  $[Gn-HCl] = 6$  M, pH = 7.0) containing MPAA (11 mM, 1.1  $\mu$ mol, 5 equiv) and TCEP (34 mM, 3.5  $\mu$ mol, 16 equiv). The pH was adjusted to 6.8 with a 1 M  $Na_2CO_{3(aq)}$  solution after peptide dissolution. It was let to react at 40°C for 15 min under argon. 400  $\mu$ L of buffer 1 was added. Bromoacetamide (28  $\mu$ L, 0.2 M, 5.5  $\mu$ mol, 25 equiv) was solubilized in buffer 1 and added to the reaction mixture. It was let to react at r.t. for 30 min. The reaction mixture was diluted and directly injected into preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 2.1 mg (40%). LC-MS (ESI):  $[M + 12H^+] = 1997.0 m/z$ ;  $[M + 13H^+] = 1843.8 m/z$ ;  $[M + 14H^+] = 1712.3 m/z$ ;  $[M + 15H^+] = 1597.9 m/z$ ;  $[M + 16H^+] = 1498.1 m/z$ ;  $[M + 17H^+] = 1410.1 m/z$ ;  $[M + 18H^+] = 1331.9 m/z$ ;  $[M + 19H^+] = 1261.8 m/z$ ;  $[M + 20H^+] = 1198.9 m/z$ ;  $[M + 21H^+] = 1141.9 m/z$ ;  $[M + 22H^+] = 1089.9 m/z$ ;  $[M + 23H^+] = 1042.5 m/z$ ;  $C_{1090}H_{1608}N_{262}O_{332}S_8$ ; Average isotope calculated 23955.9 Da  $[M]$ ; found: 23955.8 Da. ESI-MS-Orbitrap:  $[M + 12H^+] = 1997.1 m/z$ ;  $[M + 13H^+] = 1843.8 m/z$ ;  $[M + 14H^+] = 1712.1 m/z$ ;  $[M + 15H^+] = 1597.9 m/z$ ;  $[M + 16H^+] = 1498.2 m/z$ ;  $[M + 17H^+] = 1410.2 m/z$ ;  $[M$

+ 18H<sup>+</sup>] = 1331.9 *m/z*; [*M* + 19H<sup>+</sup>] = 1261.8 *m/z*; [*M* + 20H<sup>+</sup>] = 1198.8 *m/z*; [*M* + 21H<sup>+</sup>] = 1141.7 *m/z*; C<sub>1090</sub>H<sub>1608</sub>N<sub>262</sub>O<sub>332</sub>S<sub>8</sub>; Average isotope calculated 23955.9 Da [*M*]; found: 23955.5 Da.

## 7.5 Chapter 5

### Synthesis of Inh-1

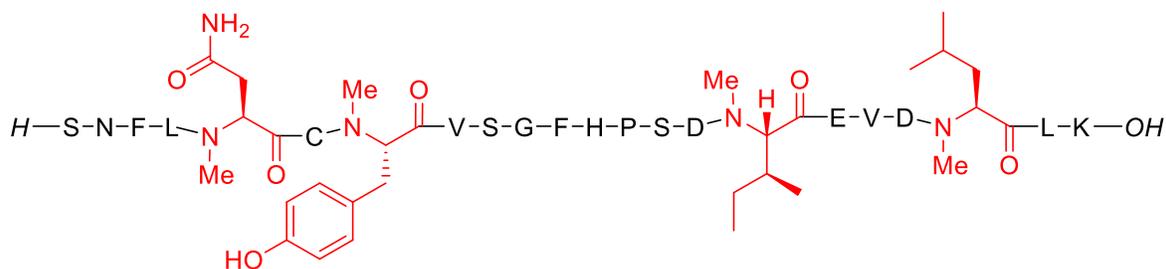
Sequence: SNF*N*-MeLNCY*N*-MeVSGFHPSDI*N*-MeEV*N*-MeDLLK



Inh-1 was synthesized on a 0.2 mmol scale on Fmoc-Lys(Boc)-Wang resin (303 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). For the coupling of Fmoc-*N*-methyl-amino acids, two equivalents were used. Phe22, Tyr26, Ile35 and Val37 had to be coupled twice. The mass of the dried peptidyl-resin at the end of the synthesis was 1020 mg. The peptide was cleaved from the resin by treatment with 10 mL of TFA/TIPS/H<sub>2</sub>O (90/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, *v/v*) containing TFA (0.1%, *v/v*) and lyophilized. Mass of crude peptide was 466 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, *v/v*) and acetonitrile with TFA (0.08%, *v/v*). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 157 mg (31%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1277.3 *m/z*; [*M* + 3H<sup>+</sup>] = 851.9 *m/z*; C<sub>117</sub>H<sub>177</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2553.9 Da [*M*]; found: 2553.0 Da.

### Synthesis of **Inh-2**

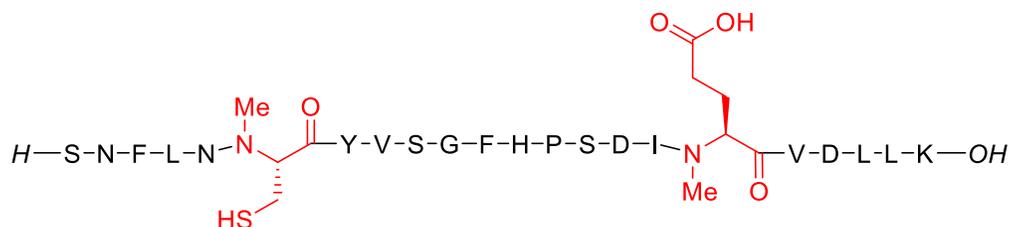
Sequence: SNFLN-MeNCN-MeYVSGFHPSDN-MeIEVDN-MeLLK



Inh-2 was synthesized on a 0.2 mmol scale on Fmoc-Lys(Boc)-Wang resin (303 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). For the coupling of Fmoc-*N*-methyl-amino acids, two equivalents were used. Leu23, Cys25 and Asp39 had to be coupled twice and Asp34 required triple coupling. The mass of the dried peptidyl-resin at the end of the synthesis was 1130 mg. The peptide was cleaved from the resin by treatment with 10 mL of TFA/TIPS/H<sub>2</sub>O (90/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 552 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 203 mg (40%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1277.4 *m/z*; [*M* + 3H<sup>+</sup>] = 852.0 *m/z*; C<sub>117</sub>H<sub>177</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2553.9 Da [*M*]; found: 2553.0 Da.

### Synthesis of **Inh-3**

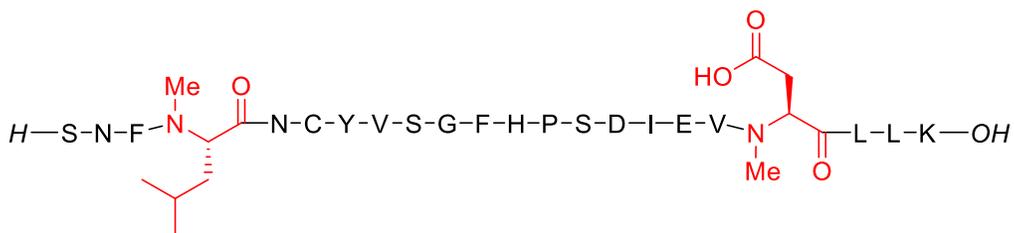
Sequence: SNFLN-MeCYVSGFHPSDIN-MeEVDLLK



Inh-3 was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (157 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). For the coupling of Fmoc-*N*-methyl-amino acids, three equivalents were used. Asp34 and *N*-MeGlu36 had to be coupled twice and Asn24 and Ile35 required triple coupling and an acetylation step was needed. The mass of the dried peptidyl-resin at the end of the synthesis was 439 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 236 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 49 mg (19%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1263.4 *m/z*; [*M* + 3H<sup>+</sup>] = 842.8 *m/z*; C<sub>115</sub>H<sub>173</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2525.8 Da [*M*]; found: 2525.2 Da.

### Synthesis of **Inh-4**

Sequence: SNF**N-Me**LNCYVSGFHPSDIEV**N-Me**DLLK

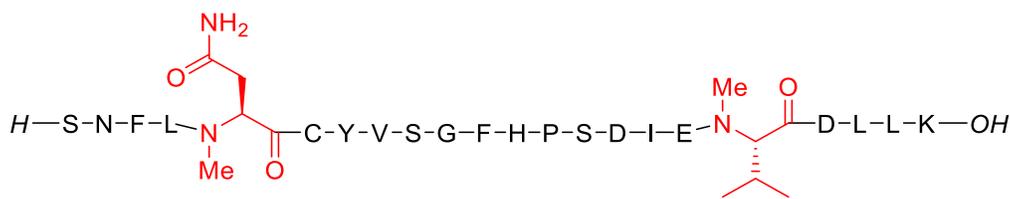


Inh-4 was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (157 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation

(4 equiv). For the coupling of Fmoc-*N*-methyl-amino acids, three equivalents were used. Phe22 and Val37 required triple coupling and an acetylation step was needed. The mass of the dried peptidyl-resin at the end of the synthesis was 436 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 238 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 71 mg (28%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1263.4 *m/z*; [*M* + 3H<sup>+</sup>] = 842.8 *m/z*; C<sub>115</sub>H<sub>173</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2525.8 Da [*M*]; found: 2525.3 Da.

### Synthesis of **Inh-5**

Sequence: SNFL*N*-MeNCYVSGFHPSDIE*N*-MeVDLLK

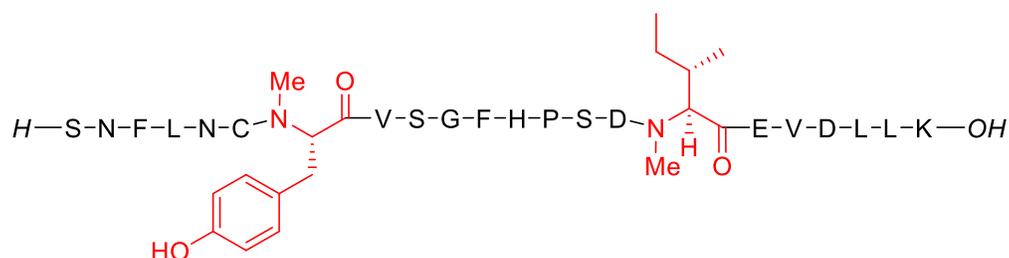


Inh-5 was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (157 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). For the coupling of Fmoc-*N*-methyl-amino acids, three equivalents were used. Glu36 had to be coupled twice and Leu23 required triple coupling and an acetylation step was needed. The mass of the dried peptidyl-resin at the end of the synthesis was 462 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20 °C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 227 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2

mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 73 mg (29%). LC-MS (ESI):  $[M + 2H^+] = 1263.4$   $m/z$ ;  $[M + 3H^+] = 842.8$   $m/z$ ;  $C_{115}H_{173}N_{27}O_{35}S$ ; Average isotope calculated 2525.8 Da  $[M]$ ; found: 2525.0 Da.

### Synthesis of Inh-6

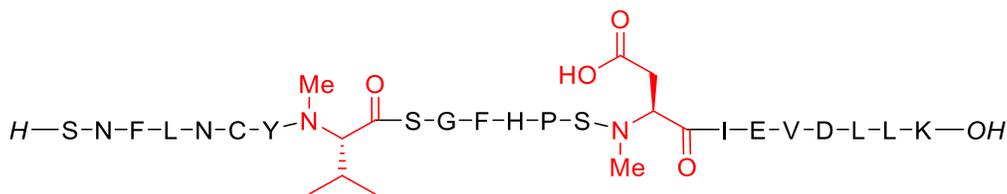
Sequence: SNFLNC*N*-MeYVSGFHPSD*N*-MeIEVDLLK



Inh-6 was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (157 mg, loading: 0.66 mmol/g, mesh size: 100-200). Manual synthesis was carried out by using HBTU activation (4 equiv). For the coupling of Fmoc-*N*-methyl-amino acids, three equivalents were used. Cys25 had to be coupled three times and Asp34 required quadruple coupling and an acetylation step was needed. The mass of the dried peptidyl-resin at the end of the synthesis was 460 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide was 192 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 34 mg (13%). LC-MS (ESI):  $[M + 2H^+] = 1263.3$   $m/z$ ;  $[M + 3H^+] = 842.8$   $m/z$ ;  $C_{115}H_{173}N_{27}O_{35}S$ ; Average isotope calculated 2525.8 Da  $[M]$ ; found: 2525.1 Da.

### Synthesis of Inh-8

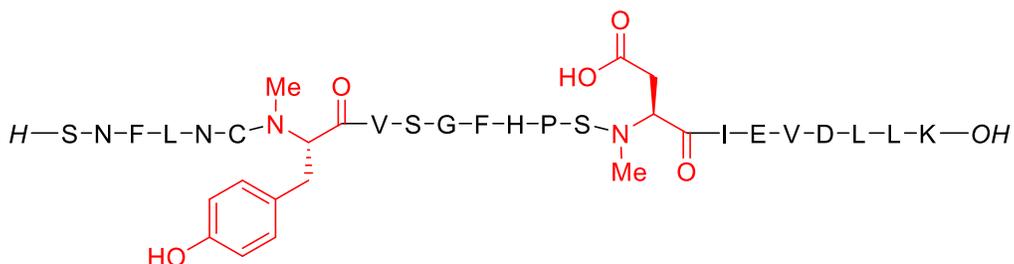
Sequence: SNFLNCY*N-Me*VSGFHPS*N-Me*DIEVDLLK



Inh-8 was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (147 mg, loading: 0.68 mmol/g, mesh size: 100-200). Automated microwave synthesis was carried out by using OxymaPure/DIC activation (4 equiv). Tyr26 and Ser33 had to be coupled twice. The mass of the dried peptidyl-resin at the end of the synthesis was 485 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 μm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 64 mg (25%). LC-MS (ESI): [*M* + 2H<sup>+</sup>] = 1263.6 *m/z*; [*M* + 3H<sup>+</sup>] = 842.9 *m/z*; C<sub>115</sub>H<sub>173</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2525.8 Da [*M*]; found: 2525.4 Da.

### Synthesis of Inh-9

Sequence: SNFLNC*N-Me*YVSGFHPS*N-Me*DIEVDLLK



Inh-9 was synthesized on a 0.1 mmol scale on Fmoc-Lys(Boc)-Wang resin (147 mg, loading: 0.68 mmol/g, mesh size: 100-200). Automated microwave synthesis was carried out by using

OxymaPure/DIC activation (4 equiv). Cys25 and Ser33 had to be coupled twice. The mass of the dried peptidyl-resin at the end of the synthesis was 499 mg. The peptide was cleaved from the resin by treatment with 5 mL of TFA/DTT/TIPS/H<sub>2</sub>O (92.5/2.5/2.5/2.5) for 2 h. The cleavage cocktail was diluted 20-fold with chilled diethyl ether (-20°C) for precipitation of the crude peptide. After centrifugation, the crude peptide was dissolved in H<sub>2</sub>O/CH<sub>3</sub>CN (1:1, v/v) containing TFA (0.1%, v/v) and lyophilized. Mass of crude peptide: 352 mg. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (250 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the resin loading was 48 mg (19%). LC-MS (ESI): [ $M + 2H^+$ ] = 1264.0  $m/z$ ; [ $M + 3H^+$ ] = 843.0  $m/z$ ; C<sub>115</sub>H<sub>173</sub>N<sub>27</sub>O<sub>35</sub>S; Average isotope calculated 2525.8 Da [ $M$ ]; found: 2526.0 Da.

### Synthesis of **Inh-1-dimer**

Inh-1 (15.5 mg, 6.1 µmol) were dissolved in 17.5 mL of buffer ([Na<sub>2</sub>HPO<sub>4</sub>] = 0.1 M, [Gn·HCl] = 6 M, [EDTA.Na<sub>4</sub>] = 5 mM, pH = 7.0). The pH was adjusted to 9.0 after peptide dissolution with a 6 M NaOH<sub>(aq)</sub>. It was let to react at 30°C for 4 days. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5 µm) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 7.5 mg (30%). LC-MS (ESI): [ $M + 3H^+$ ] = 1702.7  $m/z$ ; [ $M + 4H^+$ ] = 1277.3  $m/z$ ; [ $M + 5H^+$ ] = 1022.0  $m/z$ ; [ $M + 6H^+$ ] = 851.9  $m/z$ ; C<sub>234</sub>H<sub>368</sub>N<sub>54</sub>O<sub>70</sub>S<sub>2</sub>; Average isotope calculated 5105.6 Da [ $M$ ]; found: 5105.1 Da.

### Synthesis of **Inh-2-dimer**

Inh-2 (15.8 mg, 6.2  $\mu\text{mol}$ ) were dissolved in 4 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{EDTA}\cdot\text{Na}_4] = 5 \text{ mM}$ ,  $\text{pH} = 7.0$ ). The pH was adjusted to 9.0 after peptide dissolution with a 6 M  $\text{NaOH}_{(\text{aq})}$ . It was let to react at  $30^\circ\text{C}$  for 5 days. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100  $\text{\AA}$ , 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 10.5 mg (67%). LC-MS (ESI):  $[M + 3\text{H}^+] = 1702.3 \text{ m/z}$ ;  $[M + 4\text{H}^+] = 1277.3 \text{ m/z}$ ;  $[M + 5\text{H}^+] = 1022.1 \text{ m/z}$ ;  $[M + 6\text{H}^+] = 851.9 \text{ m/z}$ ;  $\text{C}_{234}\text{H}_{368}\text{N}_{54}\text{O}_{70}\text{S}_2$ ; Average isotope calculated 5105.6 Da  $[M]$ ; found: 5105.5 Da.

#### Synthesis of **Inh-4-dimer**

Inh-4 (5.0 mg, 2.0  $\mu\text{mol}$ ) were dissolved in 4 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{EDTA}\cdot\text{Na}_4] = 5 \text{ mM}$ ,  $\text{pH} = 7.0$ ). The pH was adjusted to 9.0 after peptide dissolution with a 6 M  $\text{NaOH}_{(\text{aq})}$ . It was let to react at  $30^\circ\text{C}$  for 5 days. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100  $\text{\AA}$ , 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 2.8 mg (55%). LC-MS (ESI):  $[M + 3\text{H}^+] = 1684.5 \text{ m/z}$ ;  $[M + 4\text{H}^+] = 1263.3 \text{ m/z}$ ;  $[M + 5\text{H}^+] = 1010.9 \text{ m/z}$ ;  $[M + 6\text{H}^+] = 842.7 \text{ m/z}$ ;  $\text{C}_{230}\text{H}_{344}\text{N}_{54}\text{O}_{70}\text{S}_2$ ; Average isotope calculated 5049.7 Da  $[M]$ ; found: 5050.3 Da.

#### Synthesis of **Inh-5-dimer**

Inh-5 (5.0 mg, 2.0  $\mu\text{mol}$ ) were dissolved in 4 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{EDTA}\cdot\text{Na}_4] = 5 \text{ mM}$ ,  $\text{pH} = 7.0$ ). The pH was adjusted to 9.0 after peptide dissolution with a 6 M  $\text{NaOH}_{(\text{aq})}$ . It was let to react at  $30^\circ\text{C}$  for 5 days. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100  $\text{\AA}$ , 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 2.7 mg (53%). LC-MS (ESI):  $[M + 3\text{H}^+] = 1684.0 \text{ m/z}$ ;  $[M + 4\text{H}^+] = 1263.5 \text{ m/z}$ ;  $[M + 5\text{H}^+] = 1011.2 \text{ m/z}$ ;  $[M + 6\text{H}^+] = 842.9 \text{ m/z}$ ;  $\text{C}_{230}\text{H}_{344}\text{N}_{54}\text{O}_{70}\text{S}_2$ ; Average isotope calculated 5049.7 Da  $[M]$ ; found: 5050.7 Da.

### Synthesis of **Inh-6-dimer**

Inh-6 (5.0 mg, 2.0  $\mu\text{mol}$ ) were dissolved in 4 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{EDTA}\cdot\text{Na}_4] = 5 \text{ mM}$ ,  $\text{pH} = 7.0$ ). The pH was adjusted to 9.0 after peptide dissolution with a 6 M  $\text{NaOH}_{(\text{aq})}$ . It was let to react at  $30^\circ\text{C}$  for 5 days. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100  $\text{\AA}$ , 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 1.5 mg (30%). LC-MS (ESI):  $[M + 3\text{H}^+] = 1684.2 \text{ m/z}$ ;  $[M + 4\text{H}^+] = 1263.6 \text{ m/z}$ ;  $[M + 5\text{H}^+] = 1011.2 \text{ m/z}$ ;  $[M + 6\text{H}^+] = 842.8 \text{ m/z}$ ;  $\text{C}_{230}\text{H}_{344}\text{N}_{54}\text{O}_{70}\text{S}_2$ ; Average isotope calculated 5049.7 Da  $[M]$ ; found: 5050.8 Da.

### Synthesis of **Inh-8-dimer**

Inh-8 (5.0 mg, 2.0  $\mu\text{mol}$ ) were dissolved in 4 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{EDTA}\cdot\text{Na}_4] = 5 \text{ mM}$ ,  $\text{pH} = 7.0$ ). The pH was adjusted to 9.0 after peptide dissolution with a 6 M  $\text{NaOH}_{(\text{aq})}$ . It was let to react at  $30^\circ\text{C}$  for 5 days. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100  $\text{\AA}$ , 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions were combined and lyophilized to give a white solid. The isolated yield was 1.2 mg (25%). LC-MS (ESI):  $[M + 3\text{H}^+] = 1684.2 \text{ m/z}$ ;  $[M + 4\text{H}^+] = 1263.5 \text{ m/z}$ ;  $[M + 5\text{H}^+] = 1010.8 \text{ m/z}$ ;  $[M + 6\text{H}^+] = 842.7 \text{ m/z}$ ;  $\text{C}_{230}\text{H}_{344}\text{N}_{54}\text{O}_{70}\text{S}_2$ ; Average isotope calculated 5049.7 Da  $[M]$ ; found: 5050.0 Da.

### Synthesis of **Inh-9-dimer**

Inh-9 (5.0 mg, 2.0  $\mu\text{mol}$ ) were dissolved in 4 mL of buffer ( $[\text{Na}_2\text{HPO}_4] = 0.1 \text{ M}$ ,  $[\text{Gn}\cdot\text{HCl}] = 6 \text{ M}$ ,  $[\text{EDTA}\cdot\text{Na}_4] = 5 \text{ mM}$ ,  $\text{pH} = 7.0$ ). The pH was adjusted to 9.0 after peptide dissolution with a 6 M  $\text{NaOH}_{(\text{aq})}$ . It was let to react at  $30^\circ\text{C}$  for 5 days. The crude peptide was purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100  $\text{\AA}$ , 5  $\mu\text{m}$ ) column with a gradient of water with TFA (0.1%,  $v/v$ ) and acetonitrile with TFA (0.08%,  $v/v$ ). Pure fractions

were combined and lyophilized to give a white solid. The isolated yield was 1.3 mg (28%). LC-MS (ESI):  $[M + 3H^+] = 1684.0 m/z$ ;  $[M + 4H^+] = 1263.4 m/z$ ;  $[M + 5H^+] = 1010.9 m/z$ ;  $[M + 6H^+] = 842.6 m/z$ ;  $C_{230}H_{344}N_{54}O_{70}S_2$ ; Average isotope calculated 5049.7 Da  $[M]$ ; found: 5049.4 Da.

### Synthesis of **Inh-1-trimer**

Inh-1 (15.3 mg, 6  $\mu$ mol) was suspended in  $NH_4HCO_3$  (6 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (1.5 mL, 1 mM in acetonitrile, 1.5  $\mu$ mol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn-HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the linker was 10.9 mg (91%). LC-MS (ESI):  $[M + 5H^+] = 1581.4 m/z$ ;  $[M + 6H^+] = 1318.1 m/z$ ;  $[M + 7H^+] = 1130.0 m/z$ ;  $[M + 8H^+] = 989.0 m/z$ ;  $[M + 9H^+] = 879.5 m/z$ ;  $C_{363}H_{540}N_{84}O_{108}S_3$ ; Average isotope calculated 7904.9 Da  $[M]$ ; found: 7903.3 Da.

### Synthesis of **Inh-2-trimer**

Inh-2 (15.3 mg, 6  $\mu$ mol) was suspended in  $NH_4HCO_3$  (6 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (1.5 mL, 1 mM in acetonitrile, 1.5  $\mu$ mol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn-HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the linker was 10.4 mg (87%). LC-

MS (ESI):  $[M + 5H^+] = 1581.7 m/z$ ;  $[M + 6H^+] = 1318.1 m/z$ ;  $[M + 7H^+] = 1130.0 m/z$ ;  $[M + 8H^+] = 988.9 m/z$ ;  $[M + 9H^+] = 879.2 m/z$ ;  $C_{363}H_{540}N_{84}O_{108}S_3$ ; Average isotope calculated 7904.9 Da  $[M]$ ; found: 7903.3 Da.

### Synthesis of **Inh-3-trimer**

Inh-3 (11.1 mg, 4.4  $\mu$ mol) was suspended in  $NH_4HCO_3$  (4 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (1.1 mL, 1 mM in acetonitrile, 1.1  $\mu$ mol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn·HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the linker was 10.4 mg (87%). LC-MS (ESI):  $[M + 4H^+] = 1955.6 m/z$ ;  $[M + 5H^+] = 1564.9 m/z$ ;  $[M + 6H^+] = 1304.2 m/z$ ;  $[M + 7H^+] = 1118.2 m/z$ ;  $[M + 8H^+] = 978.6 m/z$ ;  $C_{357}H_{528}N_{84}O_{108}S_3$ ; Average isotope calculated 7820.8 Da  $[M]$ ; found: 7819.7 Da.

### Synthesis of **Inh-4-trimer**

Inh-4 (15.2 mg, 6  $\mu$ mol) was suspended in  $NH_4HCO_3$  (6 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (1.5 mL, 1 mM in acetonitrile, 1.5  $\mu$ mol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn·HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the linker was 9 mg (77%). LC-MS

(ESI):  $[M + 4H^+] = 1956.1 m/z$ ;  $[M + 5H^+] = 1564.9 m/z$ ;  $[M + 6H^+] = 1304.2 m/z$ ;  $[M + 7H^+] = 1118.2 m/z$ ;  $[M + 8H^+] = 978.7 m/z$ ;  $[M + 9H^+] = 870.4 m/z$ ;  $C_{357}H_{528}N_{84}O_{108}S_3$ ; Average isotope calculated 7820.8 Da  $[M]$ ; found: 7820.1 Da.

### Synthesis of **Inh-5-trimer**

Inh-5 (15.3 mg, 6  $\mu$ mol) was suspended in  $NH_4HCO_3$  (6 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (1.5 mL, 1 mM in acetonitrile, 1.5  $\mu$ mol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn·HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the linker was 9 mg (77%). LC-MS (ESI):  $[M + 4H^+] = 1955.8 m/z$ ;  $[M + 5H^+] = 1564.8 m/z$ ;  $[M + 6H^+] = 1304.2 m/z$ ;  $[M + 7H^+] = 1118.2 m/z$ ;  $[M + 8H^+] = 978.6 m/z$ ;  $[M + 9H^+] = 871.0 m/z$ ;  $C_{357}H_{528}N_{84}O_{108}S_3$ ; Average isotope calculated 7820.8 Da  $[M]$ ; found: 7820.2 Da.

### Synthesis of **Inh-6-trimer**

Inh-6 (11.1 mg, 4.4  $\mu$ mol) was suspended in  $NH_4HCO_3$  (4 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (1.1 mL, 1 mM in acetonitrile, 1.1  $\mu$ mol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn·HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the linker was 3.1 mg (48%). LC-

MS (ESI):  $[M + 4H^+] = 1956.1 m/z$ ;  $[M + 5H^+] = 1564.9 m/z$ ;  $[M + 6H^+] = 1304.2 m/z$ ;  $[M + 7H^+] = 1118.2 m/z$ ;  $[M + 8H^+] = 978.6 m/z$ ;  $[M + 9H^+] = 870.0 m/z$ ;  $C_{357}H_{528}N_{84}O_{108}S_3$ ; Average isotope calculated 7820.8 Da  $[M]$ ; found: 7820.3 Da.

### Synthesis of **Inh-8-trimer**

Inh-8 (12.6 mg, 5  $\mu$ mol) was suspended in  $NH_4HCO_3$  (5 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (1.25 mL, 1 mM in acetonitrile, 1.25  $\mu$ mol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn·HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the linker was 4.8 mg (49%). LC-MS (ESI):  $[M + 4H^+] = 1956.1 m/z$ ;  $[M + 5H^+] = 1565.2 m/z$ ;  $[M + 6H^+] = 1304.4 m/z$ ;  $[M + 7H^+] = 1118.3 m/z$ ;  $[M + 8H^+] = 978.6 m/z$ ;  $[M + 9H^+] = 870.0 m/z$ ;  $C_{357}H_{528}N_{84}O_{108}S_3$ ; Average isotope calculated 7820.8 Da  $[M]$ ; found: 7820.6 Da.

### Synthesis of **Inh-9-trimer**

Inh-9 (10.1 mg, 4  $\mu$ mol) was suspended in  $NH_4HCO_3$  (4 mL, 200 mM, pH 8) containing EDTA (5 mM). Then, *N,N',N''*-(benzene-1,3,5-triyl)-tris(2-bromoacetamide) (1.0 mL, 1 mM in acetonitrile, 1.0  $\mu$ mol) was added. The suspension was vortexed and sonicated until a clear solution was obtained. The reaction mixture was shaken on a ThermoMixer (600 rpm) at 30°C for 2 h and then lyophilized. The crude solid was dissolved in guanidine hydrochloride (Gn·HCl, 6 M, 1 mL per 10 mg of crude solid) and immediately purified by preparative HPLC using a Phenomenex Kinetex XB-C18 (100 x 21.2 mm, 100 Å, 5  $\mu$ m) column with a gradient of water with TFA (0.1%, v/v) and acetonitrile with TFA (0.08%, v/v). Pure fractions were combined and lyophilized to give a white solid. The isolated yield based on the linker was 3.7 mg (47%). LC-

MS (ESI):  $[M + 4H^+] = 1956.3 \text{ m/z}$ ;  $[M + 5H^+] = 1565.2 \text{ m/z}$ ;  $[M + 6H^+] = 1304.5 \text{ m/z}$ ;  $[M + 7H^+] = 1118.4 \text{ m/z}$ ;  $[M + 8H^+] = 978.6 \text{ m/z}$ ;  $[M + 9H^+] = 870.1 \text{ m/z}$ ;  $C_{357}H_{528}N_{84}O_{108}S_3$ ; Average isotope calculated 7820.8 Da  $[M]$ ; found: 7821.2 Da.

## 8. Curriculum vitae

**Régis BOEHRINGER**

**PhD student in chemical sciences at the University of Strasbourg, France**

Date of birth: April 13<sup>th</sup>, 1991

E-mail : regis.boehringer@etu.unistra.fr

### *Skills*

---

Chemical synthesis	organic synthesis, organometallic synthesis, peptide synthesis (Fmoc-SPPS and Boc-SPPS), chemical modifications of peptides and (bio)-orthogonal chemistry.
Analytic sciences	Analytical and preparative HPLC, size-exclusion chromatography, NMR spectroscopy, ATR-FTIR spectroscopy, UV-Vis spectroscopy, fluorescence spectroscopy, mass-spectrometry (LC-MS, MALDI-TOF), electron microscopy.
Biology	Purification of biomolecules, structural biology, biology of amyloids, nano- and micro-biotechnologies
Languages	French (mother tongue), English (fluent), German (intermediate).

### *Education*

---

2017	4 <sup>th</sup> year PhD student in chemical sciences, Université de Strasbourg.
2014	Master's degree in Chemistry and Biology, Université de Strasbourg.
2012	Bachelor degree in Chemistry, Université de Strasbourg.
2011	DUT de chimie (two-year technical degree), IUT Robert Schuman, Illkirch-Graffenstaden.
2009	High-School diploma, lycée Alphonse Heinrich, Haguenau.

### *Work experience*

---

2014 - Now	<b>PhD</b> Laboratoire de Chimie des Protéines, UMR 7006, Strasbourg <i>Chemical protein synthesis to study protein misfolding and aggregation.</i>
2014	<b>Master 2 Internship</b> Laboratoire de Chimie Organique Synthétique, UMR 7177, Strasbourg <i>Organocatalysis and synthesis of cycloalkanes with electrophilic, <math>\omega</math>-unsaturated side-chains.</i>

- 2013                    **Master 1 Internship**  
Laboratoire de Chimie Organique Synthétique, UMR 7177, Strasbourg  
*Synthesis of polycyclic compound possessing biological activities.*
- 2012                    **Quality controller**  
Gunther Tools Inc., Soultz-Sous-Forêts
- 2011                    **Chemical technician and quality controller**  
Menicon Pharma Inc., Illkirch-Graffenstaden  
*Validation of expiry dates of reagents and raw materials for the production of ophthalmic solutions.*

### *Publications*

---

- 2016                    Ruiz J\*, **Boehringer R\***, Grogg M, Raya J, Schirer A, Crucifix C, Hellwig P, Schultz P, Torbeev V. Covalent tethering and residues with bulky hydrophobic side chains enable self-assembly of distinct amyloid structures. *ChemBioChem* **2016**, *17*, 2274-2285  
[\* = co-first authors]
- 2016                    Torbeev V, Grogg M, Ruiz J, **Boehringer R**, Schirer A, Hellwig P, Jeschke G, Hilvert D. Chiral recognition in amyloid fiber growth. *J. Pept. Sci.* **2016**, *22*, 290-304
- 2015                    **Boehringer R**, Geoffroy P, Miesch M. Base catalyzed synthesis of bicyclo[3.2.1]octane scaffolds. *Org. Biomol. Chem.* **2015**, *13*, 6940-6943.

### *Scientific communication*

---

- Poster presentation
  - Boehringer R, Ruiz J, Torbeev V. Towards the synthesis of a covalently-tethered oligomer mimicking amyloid structure.  
**Bordeaux Foldamers 2016 Symposium**  
26-28<sup>th</sup> September, 2016, Bordeaux, France
  - Boehringer R, Torbeev V. Towards the synthesis of a covalently-tethered oligomer mimicking amyloid structure.  
**20<sup>ème</sup> Congrès du GFPP** (Groupe Français des Peptides et de Protéines)  
26-30<sup>th</sup> March, 2017, Arcachon, France
- Oral presentation
  - Boehringer R, Torbeev V. Synthesis of a covalently-tethered oligomer mimicking amyloid structure.  
**Symposium “Diverse Facets of Chemistry”**  
11-12<sup>th</sup> May, 2017, Strasbourg, France



# Synthèse chimique de protéines pour l'étude structurale et fonctionnelle de fibres amyloïdes

## Résumé

Les fibres amyloïdes sont souvent à l'origine de nombreuses maladies dégénératives telles que la maladie d'Alzheimer ou la maladie de Parkinson. La formation de ces plaques insolubles est due à une agrégation anormale de protéines. Les études structurales et biologiques des amyloïdes sont hautement complexes du fait de leur organisation sous forme de superstructures unidirectionnelles composées d'une infinité d'unités peptidiques ou protéiques, mais aussi à cause de leur hétérogénéité conformationnelle et polymorphique. Au cours de ces différents travaux de thèse en collaboration avec différents laboratoires d'analyses structurales, nous avons développé plusieurs outils de synthèse tant pour la formation de différents polymorphes de fibres amyloïdes que pour la formation d'espèces oligomériques de tailles conséquentes qui sont un challenge du point de vue synthétique et méthodologique mais aussi pour leur caractérisation. Ces différentes avancées permettront de mieux comprendre les mécanismes de formation de fibres amyloïdes et de préparer des échantillons homogènes pour les analyses structurales et biologiques. L'étude de modifications chimiques telles que la *N*-méthylation ou les polypeptides *D* est également un enjeu important pour l'élucidation des interactions protéine-protéine vis-à-vis des structures amyloïdogéniques et ainsi permettre l'élaboration de nouveaux composés inhibant la formation de plaques amyloïdes.

Mots-clefs : fibres amyloïdes, synthèse chimique de protéines, reconnaissance chirale, polymorphisme, oligomère, inhibition.

## Résumé en anglais

Amyloid fibrils are associated with many human disorders including Alzheimer's or Parkinson's diseases. The formation of insoluble plaques is the result of protein misfolding and aggregation due to abnormal conformational isomerization of the involved protein. The structural and biological studies of amyloids are highly complex. In this thesis, we report on the development of different synthetic methodologies for the preparation of distinct amyloid fibril polymorphs as homogeneous samples for structural and biological studies. We also synthesized covalently-tethered oligomers composed of nine copies of an amyloidogenic peptide segment, where we were able to control the self-assembly of the structure by insertion of *N*-methylated amino-acids and to obtain monomeric oligomers mimicking a cross section of an amyloid fibril. We also report on the chiral recognition of *L*-peptides and *L*-proteins towards corresponding *D*-enantiomers during amyloid formation. Moreover, we studied various *N*-methylated peptide analogues to suppress amyloid growth. Overall, the results obtained in this thesis pave the way towards rational design of peptide-based inhibitors and diagnostics against amyloid propagation.

Key words: amyloid fibrils, chemical protein synthesis, chiral recognition, polymorphism, oligomers, inhibition.