Towards Bionic Proteins

Irène Jeanne Marie Arrata

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

University of Leeds

School of Chemistry

March 2017

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

The work reported in Chapter 3 constituted the basis for the research article: 'Interfacing native and non-native peptides: using Affimers to recognise a-helix mimicking foldamers', I. Arrata, A. Barnard, D. C. Tomlinson, A. J. Wilson, Chem. Commun., 2017, 53, 2834-2837. The contributions of the authors are as follows: IA (candidate) was the lead author on this piece of work and drafted the original manuscript. AB synthesised most of the library of foldamers used. AJW and DCT designed the research and edited the manuscript into its present form (see attached copy).

Acknowledgements

companion, for letting me pester you with my bad jokes whenever I make a coffee; and Captain Emma Kirk, for basically being me, but younger; Sonja, for being a new gig-companion; Marcella, for always offering help, and reminding me that the world isn't that black; and our newest recruits, Martin and James, it's been a pleasure to have you around in the last couple of months, I can only wish you joined us earlier.

A massive thank you to the BioScreening Technology Group for welcoming me and helping me a lot, in particular Anna, Christian, Janice, Lia, Sophie, Rob, Norman, it was really nice having you around and inviting me to your Christmas market outings. Tom, little Tom, my brother from another mother, I think you know how glad I am you happened to be my lab-sitter, for all the time spent in the lab, for every "heure du café", for the wine shopping sprees, out at the crag, or you annihilating me at Mario Kart – it would not have been as good without you.

To the Climbing Geek Squad: Vivien, the aweusome couch; Laura, for the hilarity of saying such wrong things with such a posh accent; Izar, porqué mi español… en la leche, perro tú Ingles esta muy bien; Steven, the weird guy from stores who also happens to be a legend (who knew?); Matt, Mr Olympic Champion all categories; Erwann, the insanely hyperactive fake French; Marvin, the real French who doesn't get the Yorkshire accent; Paul, for taking us out and teaching me so much. In particular, Vivien and Meryem, thanks for your amazing support, for sharing with me your experience, and always welcoming me with so much nice food all the time.

To my bandmates Mark, Dan, and Rick, for giving me a life outside of Uni, making me believe in a career plan B, and allowing me to let some steam out by banging on the drums… We were the absolute best, it's a shame no one else seemed to agree on that one.

To the Frenchies in my life: à mes ECPMiens: Catalina et Quentin, mes deux amours, pour nos rendez-vous impromptus aux quatre coins du globe et tous vos encouragements; Isa, Aziza, Marion, Charlotte, Vivi, pour chaque rencontre pleine de souvenirs et de potins; aux Féneloniens: Eloi, Margaux, Seb et Vincent, pour nos meetings au Croco; et aux Montaignards: Charly, Romain, et Alex, pour votre présence dépaysante, vous, mes seuls amis non-scientifiques; et enfin Didier et Emmanuelle, pour toujours me recevoir à Nanterres et me preparer un festin, sans exception.

A René Chavignon, Jean-Luc Nicolas, Sabine Choppin et Daniel Uguen, pour avoir cru en moi aux moments cruciaux et m'avoir communiqué une réelle passion pour la chimie.

A ma Maman, qui m'a appris à ne jamais abandoner.

And to Roo, my incredible lazy genius, thank you for picking the pieces and building me back so many times, loving me unconditionally, and always telling me I'm ace. For being my best friend, best team mate, and best belay.

Abstract

De novo design of foldamers is a current challenge in chemical biology. Overcoming it is essential in order to expand the protein toolbox and access "bionic proteins", *i.e.* proteins comprising non-natural segments, with enhanced biological functions. Since the α-helix represents the most abundant motif in protein structure, α-helix mimicry is a key approach to building foldamers and a stepping-stone towards generating bionic proteins. The current state-of-the-art on bottom-up foldamer synthesis for the mimicry of α-helices is described in Chapter 1.

The Wilson Group focuses on aromatic oligoamide proteomimetics for the modulation of protein-protein interactions (PPIs) involved in known diseases. The work presented in this thesis aimed at building a basis towards the elaboration of bionic proteins, by proving that such proteomimetics can be used to build novel 3-dimensional constructs.

3-*O*-alkylated and *N*-alkylated oligobenzamides with complementary charged side-chains were designed to produce self-assembling foldamers. The synthesis and conformational study of dimers and trimers is reported in Chapter 2. This forms the groundwork towards generating proteomimetic-based coiled coils.

An Affimer is a non-antibody-based scaffold, used in tandem with phage display. A small library of biotinylated *N*-alkylated proteomimetic inhibitors of p53/*h*DM2 was screened against an Affimer library. The investigation of the Affimer/foldamer interactions are reported in Chapter 3.

Taken together, these results demonstrate that aromatic oligoamides are suitable building blocks for producing non-natural peptide sequences in order to generate bionic proteins.

vi

Table of contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

ACHC	Aminocyclohexane Carboxylic Acid
ACPC	Aminocyclopentane Carboxylic Acid
Ar	Aromatic
BAK	Bcl-2 homologous antagonist/killer
BB	Blocking Buffer
Bcl	B-cell lymphoma
Bim	Bcl-2 Interacting Mediator of cell death
Bn	Benzyl
Boc	*tert*-Butyloxycarbonyl
iBu	Isobutyl
CaM	Calmodulin
CAP	Cationic Antimicrobial Peptide
carb	Carbenicillin
CD	Circular Dichroism
COSY	Correlation Spectroscopy
DCC	*N*,*N'*-Dicyclohexylcarbodiimide
DCM	Dichloromethane
DIPEA	*N*,*N*-Diisopropylethylamine
DMAP	4-Dimethylaminopyridine
DMF	*N*,*N*-Dimethylformamide
DMS	Dimethylsulfide
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic Acid
EC50	Half maximal effective concentration
ELISA	Enzyme-linked immunosorbent assay
FA	Fluorescence Anisotropy
FITC	Fluorescein-isothiocyanate
Flu	Fluorescein
Fmoc	Fluorenylmethyloxycarbonyl
GB1	B1 domain of Streptococcal protein G
gp	Glycoprotein
H-bond	Hydrogen Bonding
HBS	Hydrogen Bond Surrogate
*h*DM2	Human Double Minute 2
HPLC	High Performance Liquid Chromatography
*h*IL	Human Interleukin

HIF-1α	Hypoxia-Inducible Factor 1 alpha
HIV	Human Immunodeficiency Virus
HMBC	Heteronuclear Multiple-Bond Correlation
HMQC	Heteronuclear Multiple-Quantum Correlation
HRMS	High Resolution Mass Spectrometry
IC50	Half Maximal Inhibitory Concentration
K_d	Dissociation constant
LC-MS	Liquid phase Chromatography-Mass spectrometry
MCMM	Monte Carlo Multiple Minima
MMFF	Merck Molecular Force Field
mRNA	Messenger RiboNucleic Acid
NHS	*N*-Hydroxysuccinimide
NMP	*N*-Methyl-2-pyrrolidone
NMR	Nuclear Magnetic Resonance
NOESY	Nuclear Overhauser Effect Spectroscopy
p300	E1A binding protein 300
p53	Tumour protein 53
PBS	Phosphate Buffered Saline
PBST	Phosphate Buffered Saline + 0.1% Tween 20
PDB	Protein Data Bank
PEG	Polyethyleneglycol
PPI	Protein-Protein Interaction
ppm	Part Per Million
RCM	Ring-Closing Metathesis
ROESY	Rotating Frame Nuclear Overhauser Effect Spectroscopy
rpm	Revolution Per Minute
r.t.	Room Temperature
smMLCK	Smooth Muscle Myosin Light-Chain Kinase
SPPS	Solid-Phase Peptide Synthesis
THF	Tetrahydrofuran
TFA	Trifluoroacetic acid
TMB	3,3',5,5'-Tetramethylbenzidine
tRNA	Transfer Ribonucleic Acid
UAA	Unnatural Amino Acid
UV	Ultra Violet

Amino Acids Abbreviations

Amino acid name	Three letter code	One letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1.   Introduction

1.1. Expanding Nature's toolbox

Nature uses a stunning array of nanoscale molecular level architectures to carry out complex tasks essential to life, e.g. maintenance of the cytoskeleton, catalysis and cell signalling.[1] Biopolymers such as proteins are able to perform those intricate tasks because they self-organise to present distinct functional motifs through a precise 3-dimensional orientation that is defined by their primary structure alone. The whole human proteome, so far estimated to over 18,000 different proteins,[2] is the product of the combination of only 20 different amino acids.[1] A long-term challenge in chemical and synthetic biology is to question whether this astonishing complexity is confined to primary sequences of α-amino acids, or whether the protein toolbox can be expanded to other building blocks.[3] Building "bionic" proteins, *i.e.* proteins with non-natural constituents, is a stepping stone towards generating elaborate bio-macromolecules with enhanced or orthogonal properties.[4]

1.1.1. The α-helix

The 3-dimensional structure of proteins arises from compact conformations of helices, sheets, loops and turns, and disordered domains.[1] As the α-helix is a major component in protein secondary structure, α-helix mimicry is an important angle to approach protein mimicry.[5] A first step towards generating bionic proteins is to establish to what extent the structural and functional role of the α-helix can be reproduced with non-natural molecules in a predictable manner. It is therefore imperative to fully understand what parameters play a role in shaping the structural folding of an α-helix.

1.1.1.1. Structural analysis

The three mains types of helices are 3_{10}, α, and π, which are stabilized by *i, i+3*; *i, i+4*; and *i, i+5* hydrogen bonding interactions respectively (Figure 1.1). The α-helix (Figure 1.1a) is the most abundant motif in natural proteins, accounting for about 30% of all amino acid secondary conformations (3_{10} helix (Figure 1.1b) account for about 4% and π-helix (Figure 1.1c) is even more rare as it is less stable).[6]
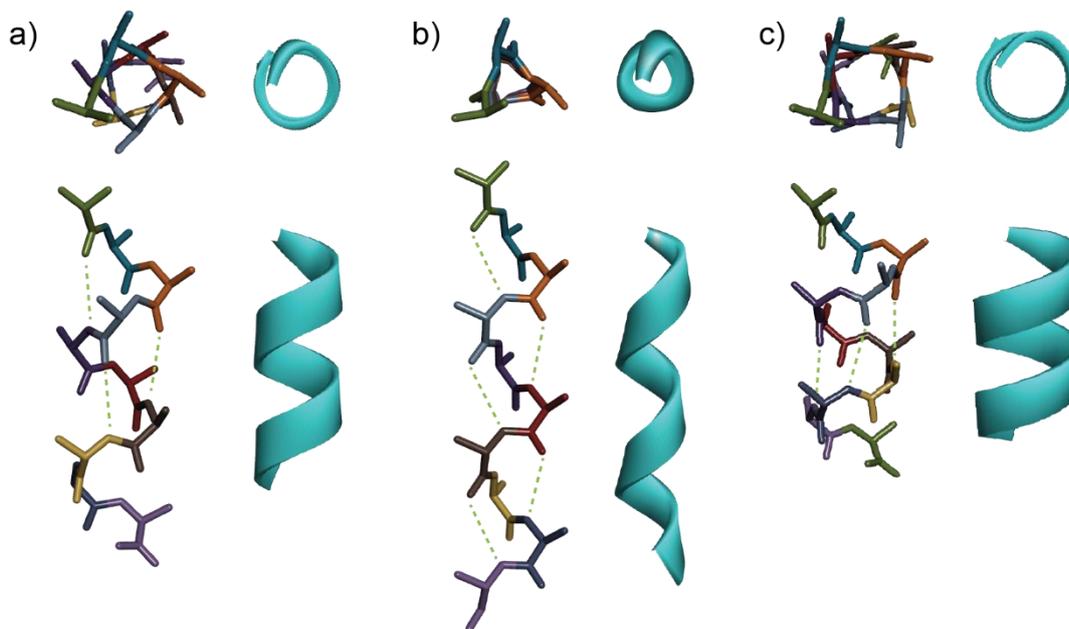
Figure 1.1: Comparison of the three types of helices, based on an alanine peptide. Top and side view of a) an α-helix, b) a 3_{10}-helix and c) a π-helix. Each colour corresponds to one alanine amino acid. Stabilising H-bonds are shown in green (for clarity purposes, only selected H-bonds are shown).

Typically, an α-helix has one turn every 3.6 amino acids. The ideal backbone dihedral angles are $\Psi = -41°$ and $\Phi = -62°$ (Figure 1.2a), and the helix rises 1.5 Å for each residue and 5.4 Å for each turn (Figure 1.2c).[7] Interactions between side-chain residues thus occur every 3-4 residues (Figure 1.2b), resulting in three distinct faces of the helix.
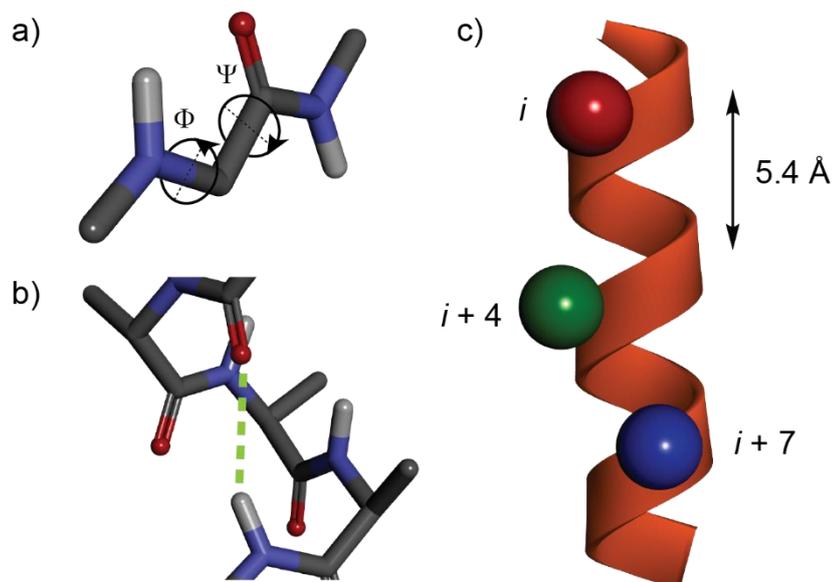
Figure 1.2: Detailed structure of an α-helix showing a) the dihedral angles Ψ and Φ, b) the stabilising H-bond between two turns, and c) the side-chains projecting at positions i, $i+4$ and $i+7$ on one face of the helix.

1.1.1.2. Mimicking α-helices

Different approaches for mimicking α-helices have been explored, e.g. using small molecules or peptides.[5, 8] Three main families of mimetics have been identified to classify these ligand types:[9]

- Type I (peptidomimetics) are short oligomers that mimic the local topology of the helix. They usually match the peptide backbone atom-for-atom, and additional covalent constraints can be used to promote a helical conformation (see paragraph 1.2.1).
- Type II (functional mimetics) are small non-peptide molecules that bind to a receptor, but don't necessarily mimic the local structure of the protein. Nutlin is a well known example of a type II mimetic.[10]
- Type III (proteomimetics)[11] are topographical mimetics that match the spatial orientation and composition of key functional residues of the original α-helix (see paragraphs 1.2.2 and 1.2.3).

1.1.1.3. α-Helix mimetics as Protein-Protein Interaction modulators

α-Helices play a fundamental recognition role in Protein-Protein Interactions (PPIs):[7, 8] despite being large and relatively flat (around 1500-3000 Å),[12, 13] PPIs do not require a full coverage of the protein-protein contact surface but only interaction with a few amino acids where most of the PPI binding energy is localised.[14] This portion of protein is called a "hot spot", and it represents a template upon which to design a synthetic ligand to compete with the original protein partner of the PPI, resulting in a modulation of the interaction (Figure 1.3).[15]
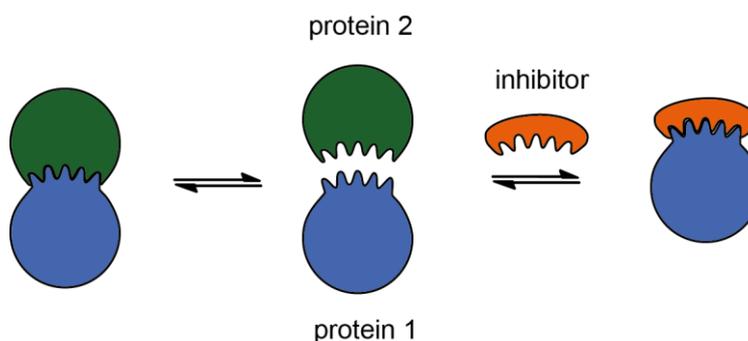
Figure 1.3: Recognition and inhibition of PPIs.

Since 62% of hot spots feature an α-helix,[16] α-helix mimicry has been an important approach to exploring PPI inhibition for therapeutic purposes.[5, 17] Therefore, in the context of generating and using non-natural amino acids for mimicking α-helices, it makes sense to look at developments of PPIs inhibitors, and this is the outlook chosen for the writing of this manuscript.

1.2. Bottom-up design of α-helix mimetics: foldamers

One main approach to α-helix mimicry is the bottom-up design of molecules that adopt a 3-dimensional structure replicating the natural α-helix: those molecules belong to the foldamer family. More generally, Gellman defined foldamers as non-natural polymers that tend to adopt a well-defined compact secondary structure.[18] They can be classified into two main groups: biotic foldamers, whose design and development are analogous to biopolymers; and abiotic foldamers, featuring backbones and folding modes different from those of biopolymers.[19]

1.2.1. Stabilisation of short peptides

1.2.1.1. Overview of the most common stabilisation techniques

In order to replicate an α-helical segment of a protein, an obvious approach is to use peptides. Nevertheless, short peptides have little ordered structure in solution:[20] for a peptide to adopt a stable helical conformation in aqueous solution, it is often necessary to introduce a stabilising group into its backbone. Various stabilisation techniques have been reported,[21] including salt bridges, metal chelates and covalent linkages such as lactam[22] or disulfide bridges[23] (Figure 1.4).

| α-helix | disulfide bridge | Glu-Lys lactam bridge | hydrocarbon stapling | hydrogen bond surrogate |
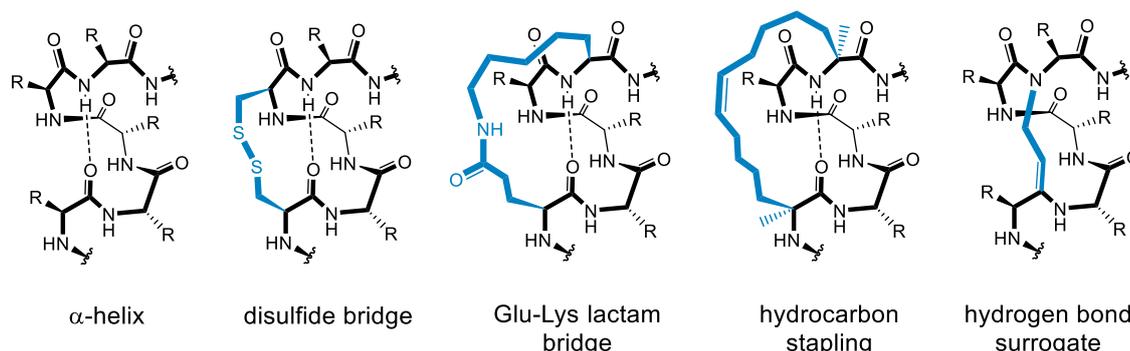
Figure 1.4: Common approaches for stabilisation of short peptides. Adapted from Azzarito *et al.*[5]

Early studies demonstrated that disulfide bridges could be used to stabilise helical conformations in short peptides:[24, 25] it has been shown that disulfide-linked peptides can inhibit nuclear receptor co-activator interactions.[26, 27] Lactam bridges have also been proven to successfully stabilise α-helical structures[28, 29] and were used to obtain a potent inhibitor of the assembly of the hexameric gp41 core that leads to HIV virus host-cell fusion.[30]

Grubbs and co-workers introduced hydrocarbon stapling by ring-closing metathesis (RCM) with *O*-allyl serine residues,[31] and this approach was later improved upon by Verdine and co-workers who extended it to unnatural α-methylated amino acids.[32] Recent work within the Wilson group ascertained that α-alkenyl substituted amino acids are effective for peptide stapling of Bcl-2 family peptides, and that the staple does not introduce a steric constraint in the binding cleft.[33, 34]

The Hydrogen Bond Surrogate (HBS) approach involves replacing a native *i-i*+4 hydrogen bond by a covalent linkage. One advantage of HBS compared to other stabilisation techniques is that it does not compromise the recognition surface of a constrained structure.[35] Solid phase peptide synthesis (SPPS) of HBS-constrained helical peptides has been described and peptides of biologically relevant conformation have been identified.[28-36]

1.2.1.2. Application to phage display screening

Generating small peptides by SPPS is limiting in terms of time of synthesis and purification; reverse screening methods can be used to avoid such limitations. The general principle of those techniques relies on rapidly generating and screening diversified libraries of compounds, and later identifying hits to biologically relevant systems.

Commonly used as a screening technique, phage display[37, 38] has been shown to be an efficient tool to study PPIs, and has therefore been extensively used for this purpose in the past.[39-42] Phage display relies on the use of bacteriophage that infect the standard recombinant bacteria host DNA and the encoded foreign peptide is expressed as part of a fusion to one of the coat proteins on the surface of the phage.[43] A phage library can be constituted by a mixture of phage displaying different peptides from different DNA inserts. As illustrated in Figure 1.5a, the randomised library of phage is screened against an immobilised target of choice. The best binders are isolated using affinity purification techniques that require the immobilisation of the receptor.[44] Unbound phage are washed away, and the binding ones are then eluted and further amplified, and after at least 3 rounds of selection, the finally eluted phage are individually characterised.[43]

Although phage display classically generates a randomised peptide whose conformation cannot be predicted, it is possible to tune the genetic code of the phage in order to produce a specific coat protein, including exclusively α-helical peptides.[45] Diderich *et al.* have recently combined phage display to peptide stapling methods to screen libraries of stabilised α-helical peptides.[46] The peptides generated by the phage consistently contained two cysteines, respectively positioned at *i* and *i*+4, to enable stabilisation with thiol-reactive linkers (Figure 1.5b).
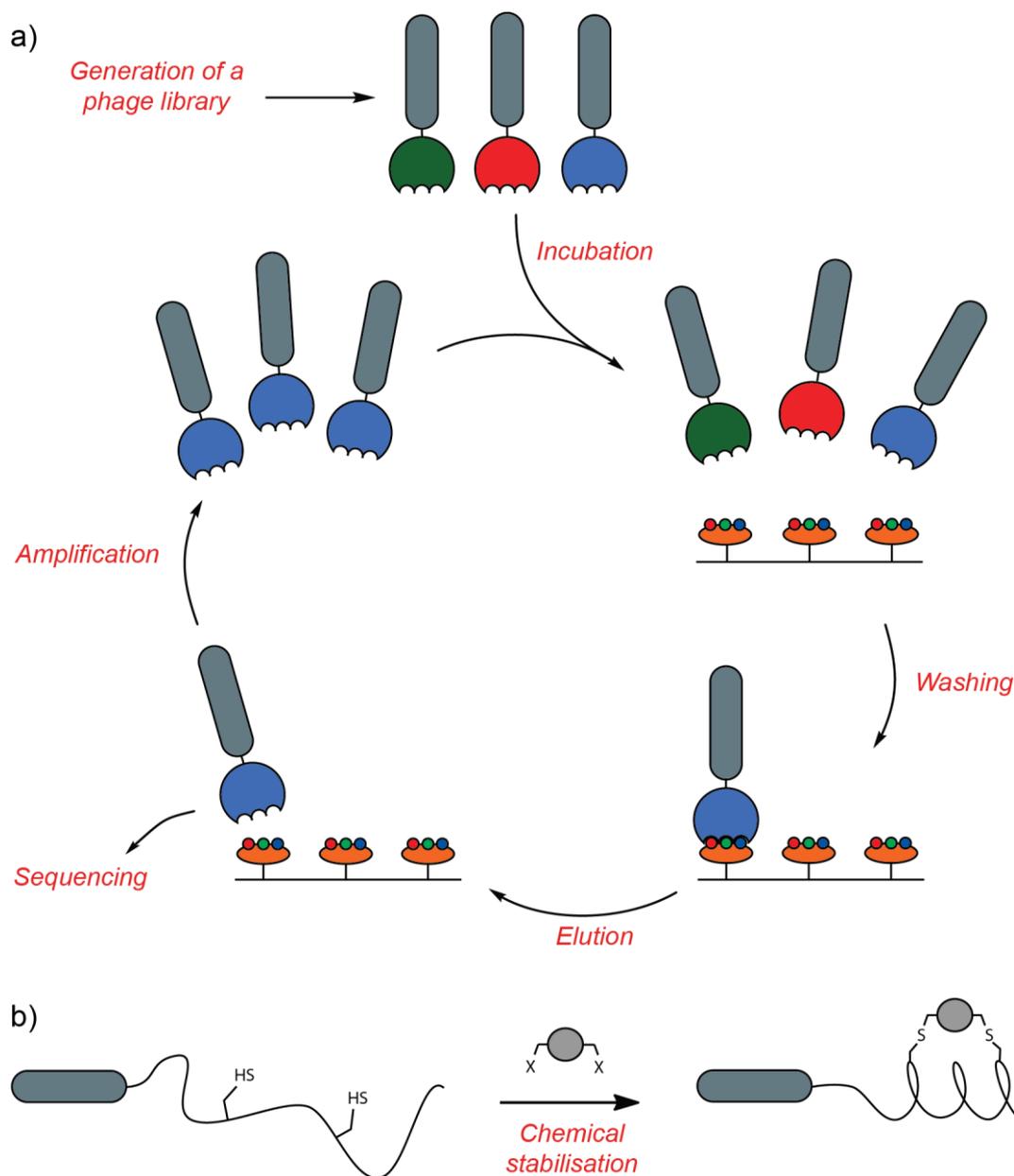
Figure 1.5: a) Cartoon representation of the principle of phage display. The generated library is tested for binding against an immobilised target. Unbound phage are washed away and bound phage are eluted, amplified and bound again. After at least three cycles, the amplified phage are characterised. b) Chemical stabilisation of peptides for building a library of α-helical peptides by phage display. The sequence of the peptide is randomised except for two cysteines at respective positions *i* and *i*+4 to allow the stabilisation.[46]

1.2.2. Backbone modifications

In order to improve the folding properties of small peptides without requiring external stabilisation, backbone modifications have been investigated. Indeed, it is possible to replace natural α-amino acids (Figure 1.6a) by their β- or γ- equivalents (Figure 1.6b and c): Gellman and co-workers studied the competition between possible intramolecular H-bonding within β- and γ-

peptide monomers.[47] In the case of β-peptides, nearest neighbour interactions are unfavourable (Figure 1.6b, red arrow), which means the only favoured intramolecular H-bonding occurs between distant residues leading to compact folding (Figure 1.6b, green arrows). On the other hand, their studies highlighted that nearest neighbour interactions within ϒ-peptides *via* formation of a 9-membered hydrogen-bond ring are favoured, which leads to a non-folded conformation of the molecule (Figure 1.6c): this results in a competition between the potential conformations adopted by the molecule, and suggests that polymers composed exclusively of ϒ-amino acids are less likely to adopt compact and specific folding patterns.
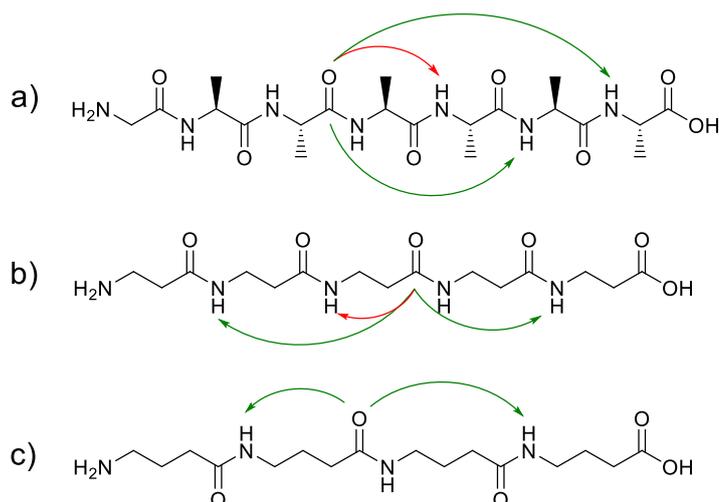
Figure 1.6: H-bond interactions within a) α-, b) β- and c) ϒ-peptides. Green arrows show favourable interactions whereas red arrows show unfavourable interactions. Adapted from Gellman *et al.*[47]

One focus of the group's work was identifying backbones that favoured helical secondary structure,[47] and to understand the consequence of altering the nature of the residues in order to modulate the β-peptide structure. A first study showed that β-peptides derived from *trans*-2-aminocyclohexanecarboxylic acid (*trans*-ACHC, Figure 1.7a) adopt a 14-helix, *i.e.* forming a 14-membered hydrogen-bonding ring (Figure 1.7b and c).[48] It was then established that *trans*-2-aminocyclopentane-carboxylic acid (*trans*-ACPC, Figure 1.7d) forms a 12-helix (Figure 1.7e and f).[49] Recent work has demonstrated that the substitution of an α-amino acid for a β-amino acid such as ACPC will not only preserve the helicity of the peptide, but also enhance its resistance to proteolysis.[50]
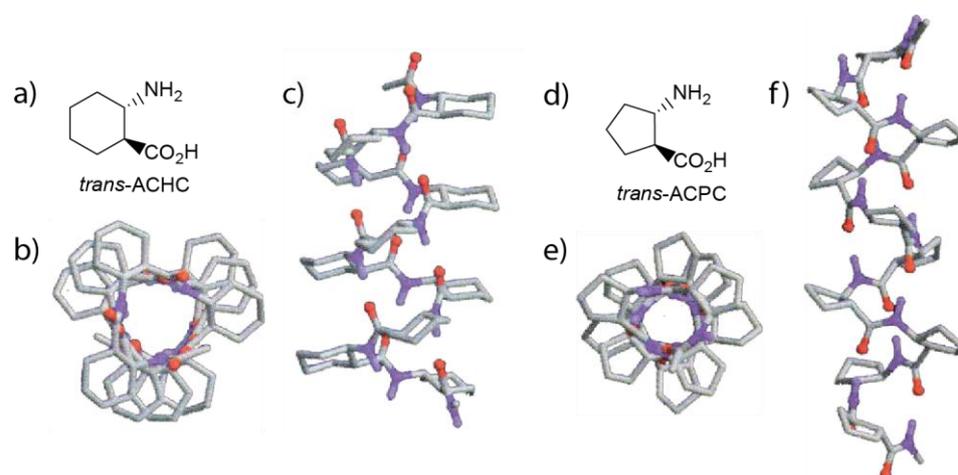
Figure 1.7: Folding properties of *trans*-ACHC and *trans*-ACPC. a) Chemical structure of the *trans*-ACHC monomer. b) Top view and c) side view of the 14-helix folding of *trans*-ACHC.[48] d) Chemical structure of the *trans*-ACPC monomer. d) Top view and e) side view of the 12-helix folding of *trans*-ACPC. Adapted with permission from Appella *et al*, *Nature*, **1997**, 387, 381-384. Copyright 1997 Nature Publishing Group.

It has been established that the introduction of constraints allows β-peptides to adopt a helical conformation and retain activity compared to the natural α-peptide.[51-53] Yet, no general conclusion can be drawn and the design of foldamers needs to be considered on a case by case basis,[54] as the β-peptide helix does not perfectly reproduce the spatial arrangements of residues found in α-helices.[55] Beck-Sickinger *et al.* showed the structural difference between α- and β-peptides by replacing secondary structure elements of the C-terminal α-helix of human interleukin 8 (hIL-8) by artificial segments.[56] Modifications were performed on the residues 61-77 of hIL-8 and the residues 105-121 of the cationic antimicrobial peptide (CAP18). Using circular dichroism (CD) spectroscopy, the β-peptides configurations were determined to be a 3_{14}-helix (Figure 1.8c, right), which is left-handed as opposed to the α-helix (Figure 1.8c, left). Yet, modelling of the peptides showed good agreement between the faces of the native α-peptides (Figure 1.8a and b, left) and the β-peptides (Figure 1.8a and b, right), and the activities of the modified proteins were comparable to the ones of the native α-peptides.
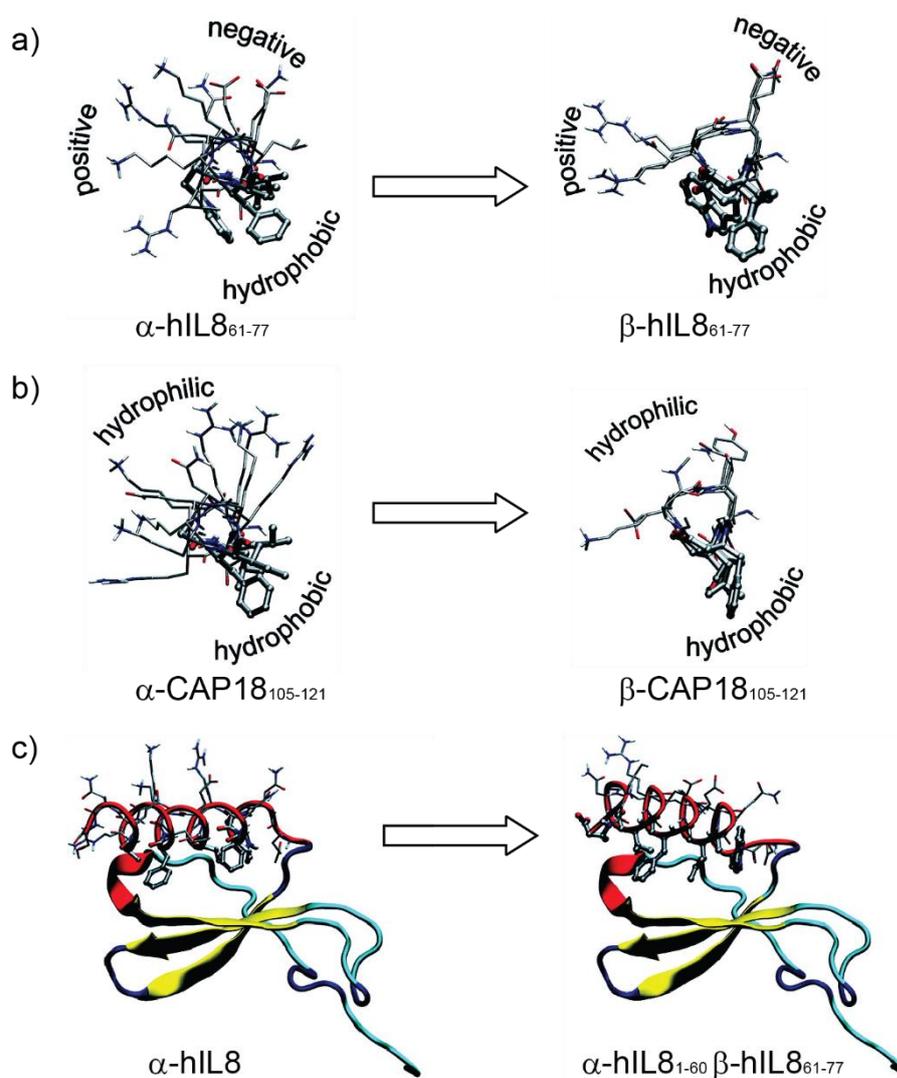
Figure 1.8: Molecular modelling comparing the folding conformations of native α-helices and their equivalent β-peptides. a) Top view of the 61-77 residues of hIL8 (left) and its corresponding unnatural 3_{14}-helical β-peptide (right). b) Top view of the 105-121 residues of CAP18 (left) and its corresponding unnatural 3_{14}-helical β-peptide (right). c) Comparison between natural α-hIL8 highlighting the native right-handed α-helix (left) and the α-hIL81-60-β-hIL861-77, highlighting the left-handed 314-helix. Adapted with permission from David *et. al.*, *J. Am. Chem. Soc.* **2008**, *130*, 15311. Copyright 2007 American Chemical Society.

Although a full change from α- to β-peptide has consequences on the helicity and can also modify the handedness, it has been demonstrated that small alterations on an oligoamide backbone can have limited influence on the folding behaviour:[57] a series of analogues have been prepared from the wild-type sequence of the B1 domain of *Streptococcal* protein G (GB1), applying different backbone modifications, such as incorporating $β^3$-residues (carbon insertion), D-α-residues (inversion of configuration), and Cα-methyl or *N*-methyl α-residues (alkylation) as shown in Figure 1.9. Results have shown that helix, loop, sheet, and turn secondary structure of GB1 could individually be modified with very little alteration of the folding properties, and these modifications could be combined with near-additive effects on the overall structure.

9
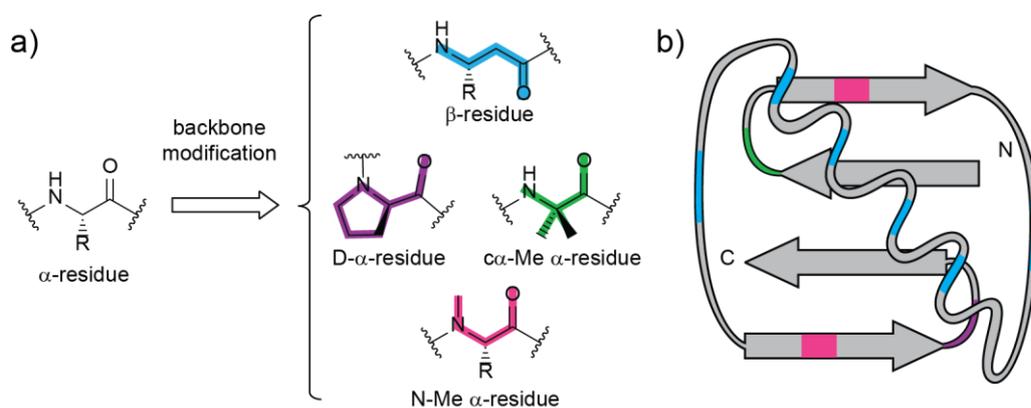
Figure 1.9: Study of the effect of backbone modification on foldamer folding properties. a) Backbone modifications applied included β³-residues, D-α-residues, and Cα/N-methyl α-residues. b) Tertiary folding topology of GB1 showing where the modifications were introduced (corresponding colors).[57]

One main challenge concerning β-peptides is to obtain a stable folded structure in aqueous solution.[58] Schepartz and coworkers have demonstrated that the 14-helix macrodipole plays a significant role on β³-peptide stability: minimising this macrodipole by introducing charged side chains at relevant positions on the helix enhanced the structural stability in water.[59] Using these results, stable 14-helical β³-oligopeptides were designed (Figure 1.10a), that self-assemble into a well-defined, thermostable, octameric quaternary structure, named Zwit-1F (Figure 1.10b).[60, 61] The spontaneous assembly into the octamer resulted from non-covalent inter-residue interactions, and the final structure possessed a hydrophobic core (Figure 1.10c).
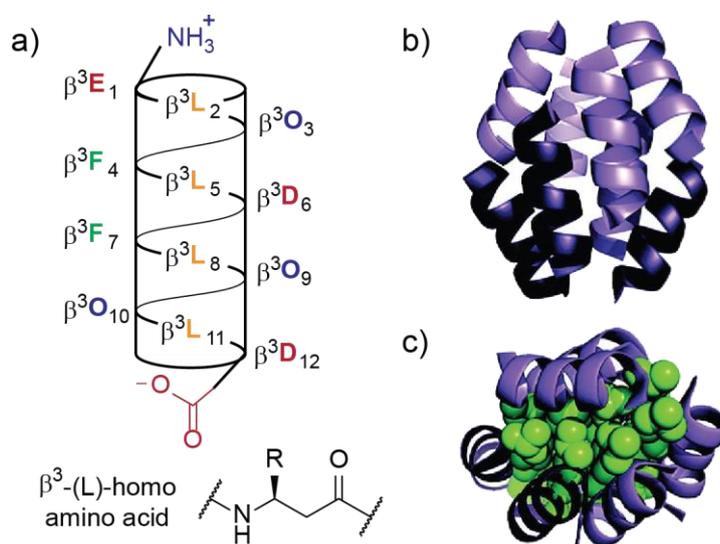
Figure 1.10: β³-peptide quaternary structure. a) 14-helical conformation was enhanced by minimising the macrodipole by alternating positively (blue) and negatively (red) charged side-chains. b) X-ray structure of the octameric quaternary structure Zwit-1F. c) Highlight of the hydrophobic core (green). Adapted with permission from Daniels *et al., J. Am. Chem. Soc.,* **2007**, *129*, 1532. Copyright 2007 American Chemical Society.

1.2.3. *De novo* design of foldamers: aromatic oligoamides

Backbone modification through α-amino acid replacement is not the only approach to foldamer synthesis: a different way to address the topic is the complete *de novo* design of structures with folding properties. For this purpose, aromatic oligoamides are very attractive, as their folded conformations are predictable and tunable, and they demonstrate high conformational stability. Indeed, it is possible to control the number of units per turn of the helix, based on the three main parameters that have a strong impact on the helix diameter:[62] the orientation of the amide linkages,[63] the size of the units,[64] and the position of any hydrogen bond. Moreover, electrostatic repulsions, steric effects and intramolecular hydrogen bonding[65] cause a restriction to the aryl-amide bond rotation, which is strong enough to determine the structure of the folded conformation. Extra stability is added to this structure from π-π interactions between aromatic rings.[62]

Examples of stability and tunability of aromatic oligoamides have been established by Huc and co-workers, by studying the racemisation between M and P helices of the 2-quinolinecarboxylic acid foldamer.[66] Racemisation rate was measured in different solvents, and the foldamer was shown to be particularly stable in polar solvents, with a racemisation rate of around $10^{-4}\,min^{-1}$ in methanol-water solutions. Derivatives of the quinolone scaffold were studied, and the design of stable conformations for which side chains formed a linear array enabled the synthesis of amphiphilic compounds (Figure 1.11).[67, 68] This work was later improved to obtain hybrid compounds with the capacity to fold in water.[69]
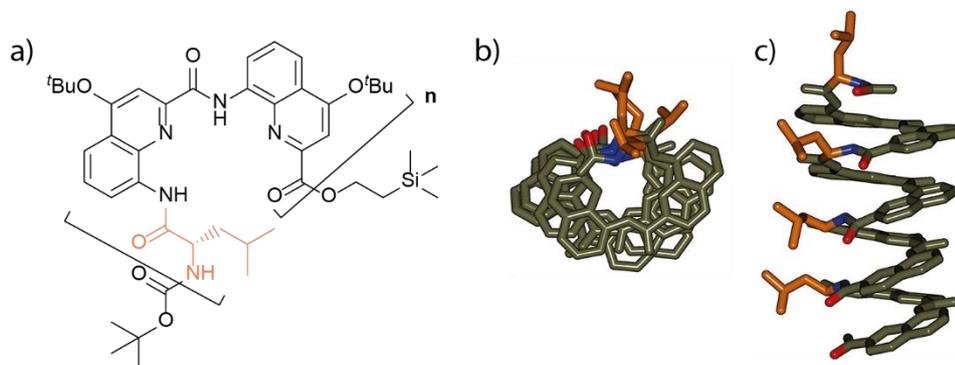
Figure 1.11: Conformation of oligoaromatic foldamers. a) Chemical structure of the $(LQ_2)_n$ main building block, where two 8-amino-2-quinolinonecarboxylic acid (Q, grey) are linked to a Leucine (L, beige). b) Top view and c) side view of the X-ray structure show that extra stability is induced by π-π interactions between the aromatic rings. These interactions allowed orientation of the functionalised side chains to form a linear array. Adapted with permission from Kudo *et al.*, *J. Am. Chem. Soc.*, **2013**, 135, 9628-9631. Copyright 2013 American Chemical Society.

Quinoline-based foldamers were further studied on the basis that positively charged quinoline-based oligomers could selectively bind to quadruplex DNA.[70] Biotinylated quinoline-based octamers were screened against a wide range of DNA targets, and their high selectivity for G-quadruplexes was established.[71, 72]

The applications linked to aromatic oligoamide foldamers are not only limited to binding to biologically relevant targets, and recent work on encapsulation has shown the versatility of that scaffold. Initial research on encapsulating water molecules[73] has led to diversified targets, such as fructose[74] and tartaric acid (Figure 1.12a),[75] as well as encapsulation driven self-assembly with citric acid (Figure 1.12b).[76]
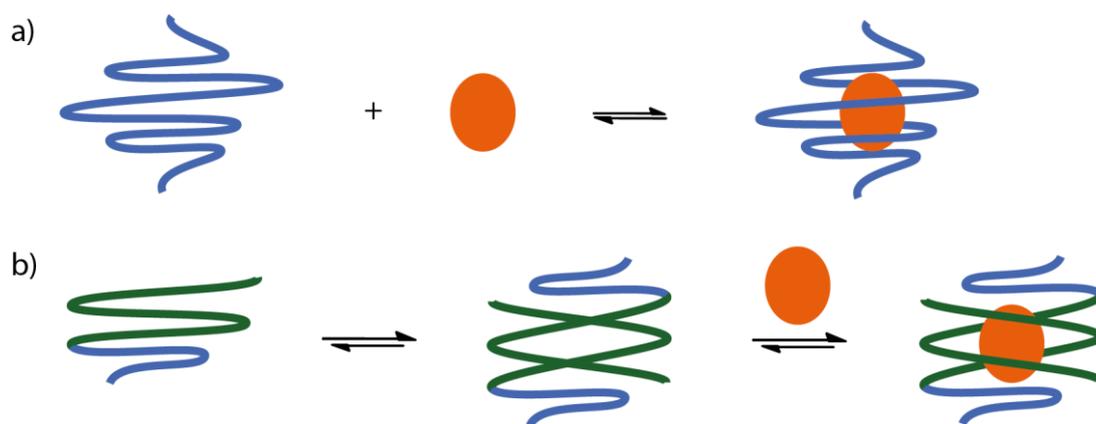
Figure 1.12: Encapsulation of targets by a) single helix[73-75] and b) self-assembling double helix[76] foldamers.

Although foldamers replicating full helices have proven their potency in terms of binding or encapsulating, smaller molecules are of interest for their ease of synthesis, purification and characterisation. The proteomimetic approach for α-helix mimicry focuses not on replicating the whole protein structure but only the spatial and angular projections of the side chain features that play a key role in PPIs. Proteomimetics are therefore small scaffolds, not necessarily peptidic. Amongst the variety of motifs used for building proteomimetics, linear aromatic oligoamides are often preferred, as the distance between two monomers matches the distance between two consecutive side chains along one face of an α-helix (Figure 1.13a).[77] Stabilising H-bonding can occur on either or both sides of the molecular strand. In the first case, a facial polarity is generated, which leads to a slight curvature of the strand which can be used for molecular recognition[77] or to give rise to facial amphiphilicity.[78] The aforementioned properties make proteomimetics an important approach for use as inhibitors of PPIs (Figure 1.13b).
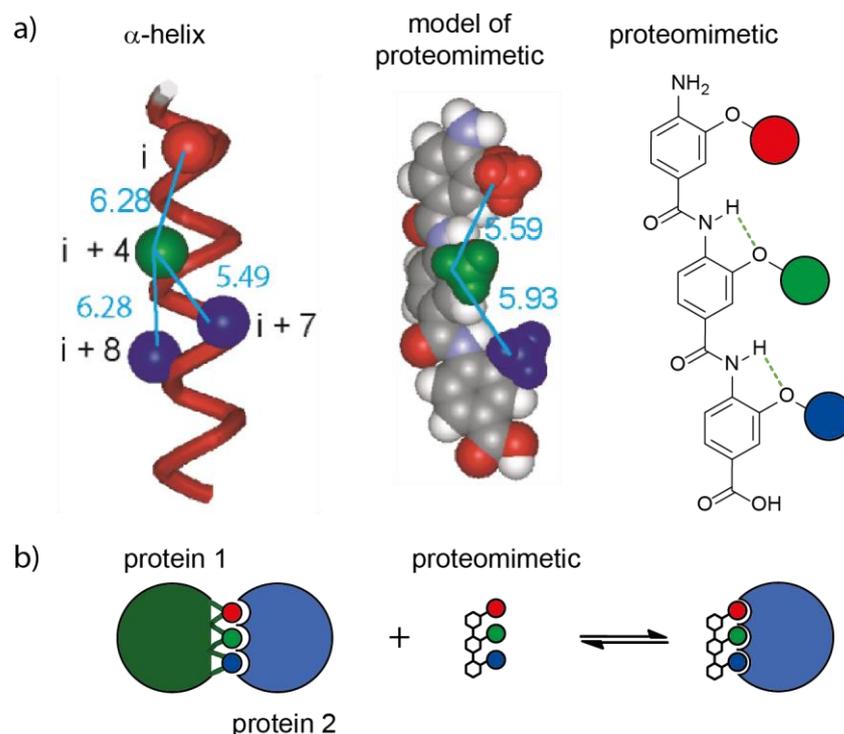
Figure 1.13: a) Schematic representation of a proteomimetic designed to mimic an α-helix, and reproducing the spatial disposition of its key functions. b) Inhibition of PPIs with proteomimetics.

Hamilton *et al*. described the terphenyl scaffold as the first proteomimetic (Figure 1.14),[11] which mimicked two turns of the smooth muscle myosin light-chain kinase (smMLCK) α-helix, in order to inhibit its interaction with calmodulin (CaM). The results indicated an accurate functional mimicry of smMLCK in binding with high affinity to CaM. Derivatives of this scaffold were successfully elaborated to target the gp41 complex[79] and the Bcl-x$_L$/BAK interaction.[80]
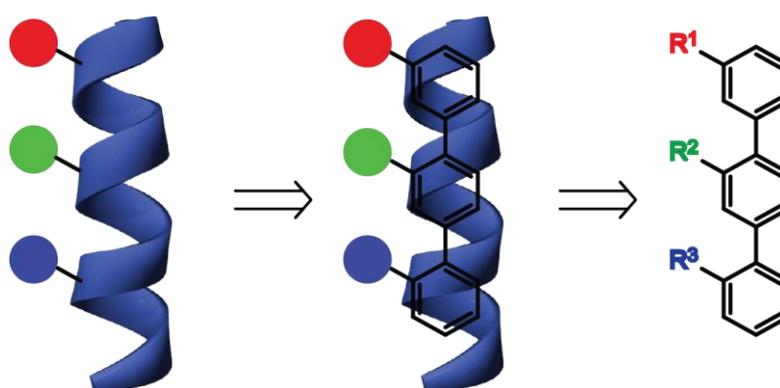
Figure 1.14: The terphenyl scaffold was the first described proteomimetic.[11]

On the basis of Hamilton's work, various proteomimetic scaffolds have been described, exhibiting different properties. The enaminone scaffold was developed as an improvement to the terphenyl scaffold, especially in terms of solubility in polar solvents (Figure 1.15a).[81, 82] Aromatic oligoamides have also been extensively used to target PPIs such as p53/*h*DM2,[83] or Bcl-x$_L$/BAK.[84]

They are particularly attractive due to their ease of synthesis and also offer the possibility to access double-sided scaffolds (Figure 1.15c),[85, 86] as many known interactions require binding to more than one face of the helix.[87] Benzoylurea-based compounds (Figure 1.15d) were also used to mimic double sided helices[88] and showed good inhibition properties of the Bcl-x$_L$/BAK interaction.[89] 1,2-Diphenylacetylene (Figure 1.15e) was used to mimic the projections on the two faces of the helical region of Bim-BH3.[90] Recently, a new scaffold based on an imidazole-phenyl-thiazole combination was elaborated (Figure 1.15f)[91] as an inhibitor of the interaction between Cdc42 and Dbs, which has been linked to tumour growth, cardiovascular and neurodegenerative diseases, and diabetes.[92]
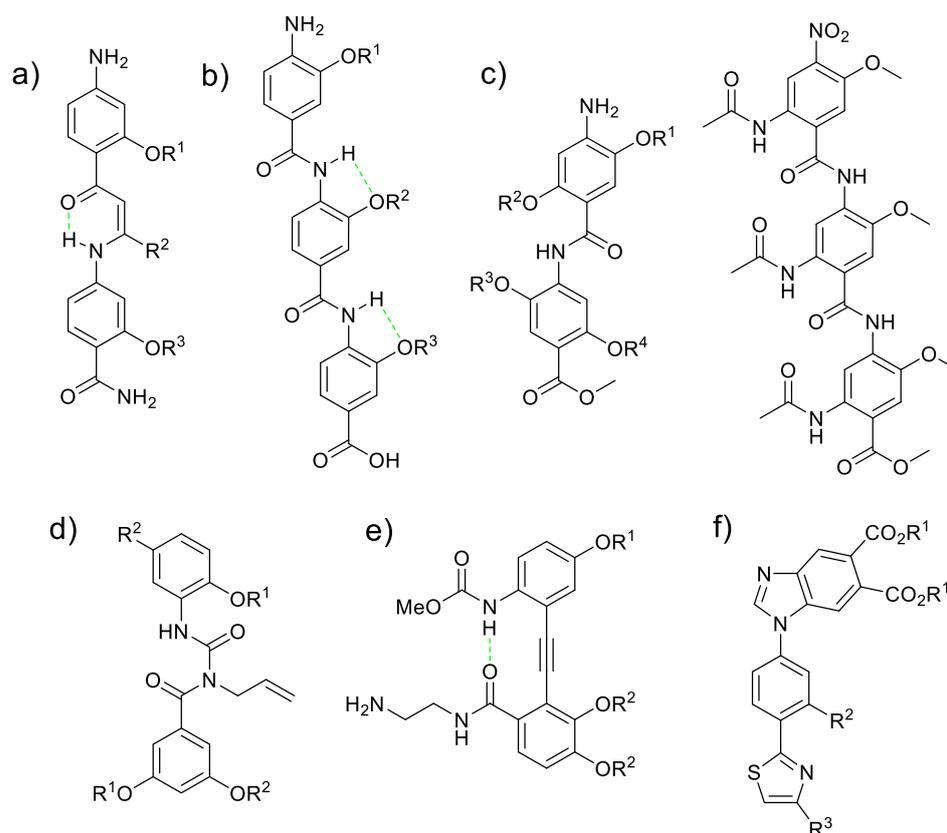
Figure 1.15: Examples of reported proteomimetic scaffolds such as a) enaminones;[82] b) aromatic oligoamides[83, 84] and their application to c) double sided helical structures[85, 86]; d) benzoylurea[88]; e) 1,2-diphenylacetylene[90] and the latest f) imidazole-phenyl-thiazole combination.[91]

Nature exhibits many examples of more complex multi-helix proteins. Mimicking two or more helices instead of one is therefore of interest for accurately replicating complex biological constructs. This was accomplished by Hamilton *et al.*, who developed double α-helix mimetics (Figure 1.16) by attaching two 3-*O*-alkylated oligoamides together with various rigid linkers.[93]
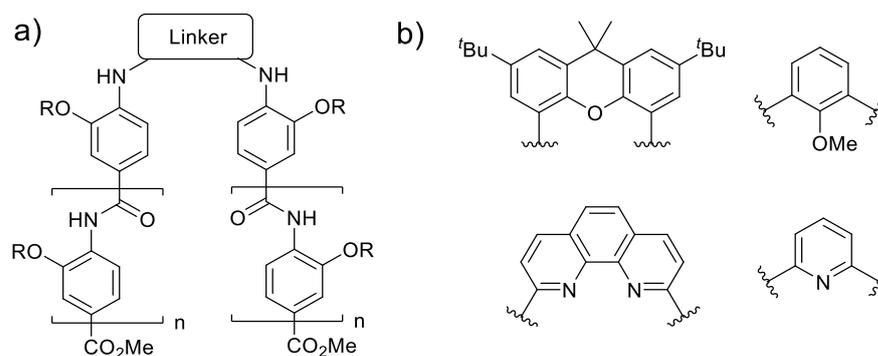
Figure 1.16: a) General structure of the double helix mimetics. b) Example of applied linkers.

The approaches for designing α-helix mimetics are varied and enable access to many different types of mimetics. Nevertheless, the design and synthesis of such molecules can be limiting in terms of time consumed by design, synthesis, troubleshooting and optimisation, compound characterisation, and testing.

1.2.4. Incorporation of unnatural building blocks into proteins

In order to use non-natural amino acids to build a bionic protein, it is necessary to incorporate a foldamer sequence into a protein. Different approaches can be undertaken depending on the size of the protein and the required type of modifications.

SPPS allows a complete freedom of the choice of amino acids, and therefore opens the door to a great variety of new proteins. However, it is limited in size to peptides and small proteins, given that the ease of synthesis, yield and purity decrease as the size of the protein increases.[94, 95] Currently, the easiest way of artificially preparing a full protein is by chemical ligation of a synthesised peptide to an expressed fragment of protein.[96] For the purpose of modifying existing proteins in an easy and rapid manner, another approach is desirable.

Although challenging, the expansion of the genetic code to new unnatural amino acids (UAAs) has already yielded several successes.[97-99] A common way of doing so is site-directed mutagenesis, which permits a single mutation of an amino acid residue located at any site within a protein to be performed.[100] Using adequate primers, this technique involves the generation of an oligodeoxyribonucleotide complementary to the desired methylated wild-type sequence, except for the altered residue. Rounds of denaturation followed by annealing synthetic nucleotides and DNA synthesis yield the mutant plasmid (Figure 1.17), and the parent wild type templates are digested with DpnI so that only the desired mutant plasmid is further transformed.[101]
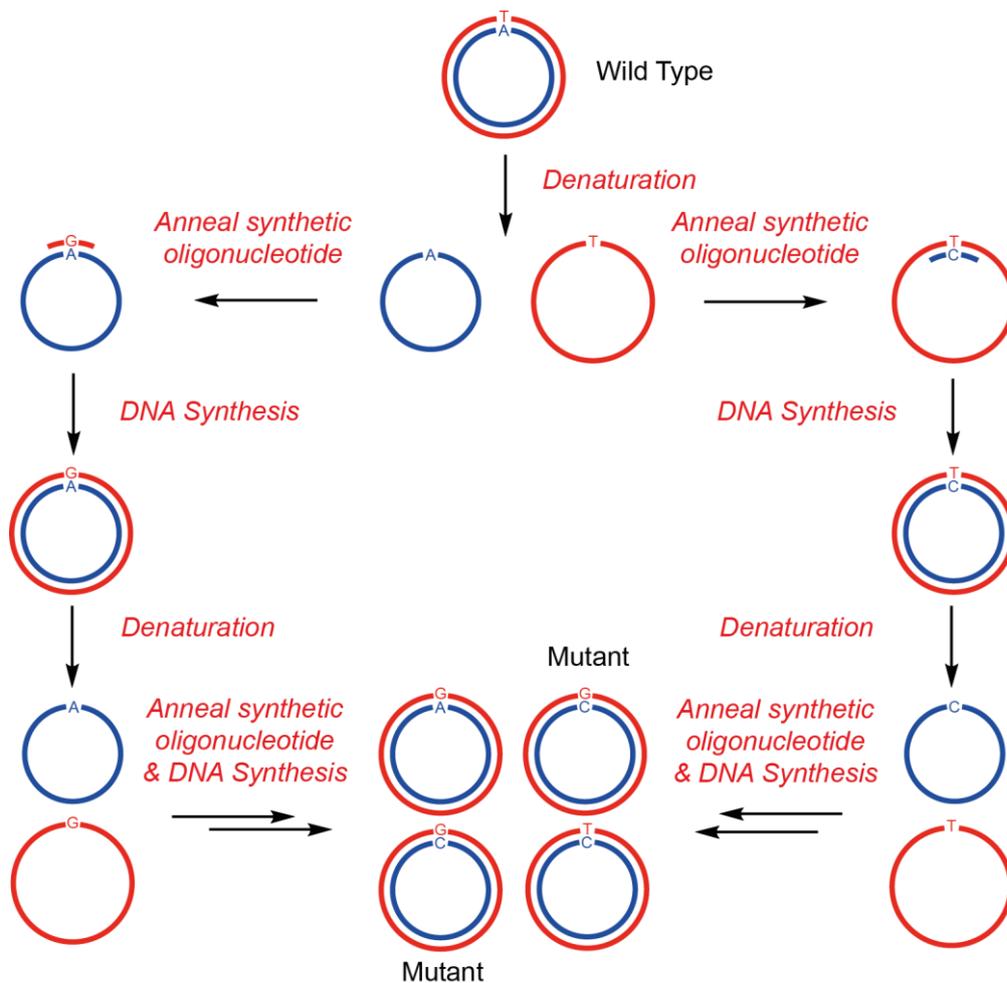
Figure 1.17: Principle of site-directed mutagenesis. The wild-type plasmid is denatured and a primer with the corresponding sequence but for one residue is used to synthesise complementary DNA. Several rounds are performed until the desired mutant is obtained.

A common application of this technique is the mutation of a chosen residue to an alanine group, which has minimal contribution to the free energy of a PPI, in order to determine the importance of that residue.[102] Site directed mutagenesis is also essential in order to use the Amber Suppression method (Figure 1.18).[103, 104] In this method, Amber suppressor tRNAs are aminoacylated with the desired non-natural amino acids through chemical aminoacylation, and the obtained aminoacyl-tRNAs (aa-tRNAs) are added to an *in vitro* translation system with a mRNA or DNA containing an Amber stop codon (UAG) at a desired position. The incorporation of the non-natural amino acid at the desired position results from the suppression of the Amber codon, which occurs on the mRNA.[105]
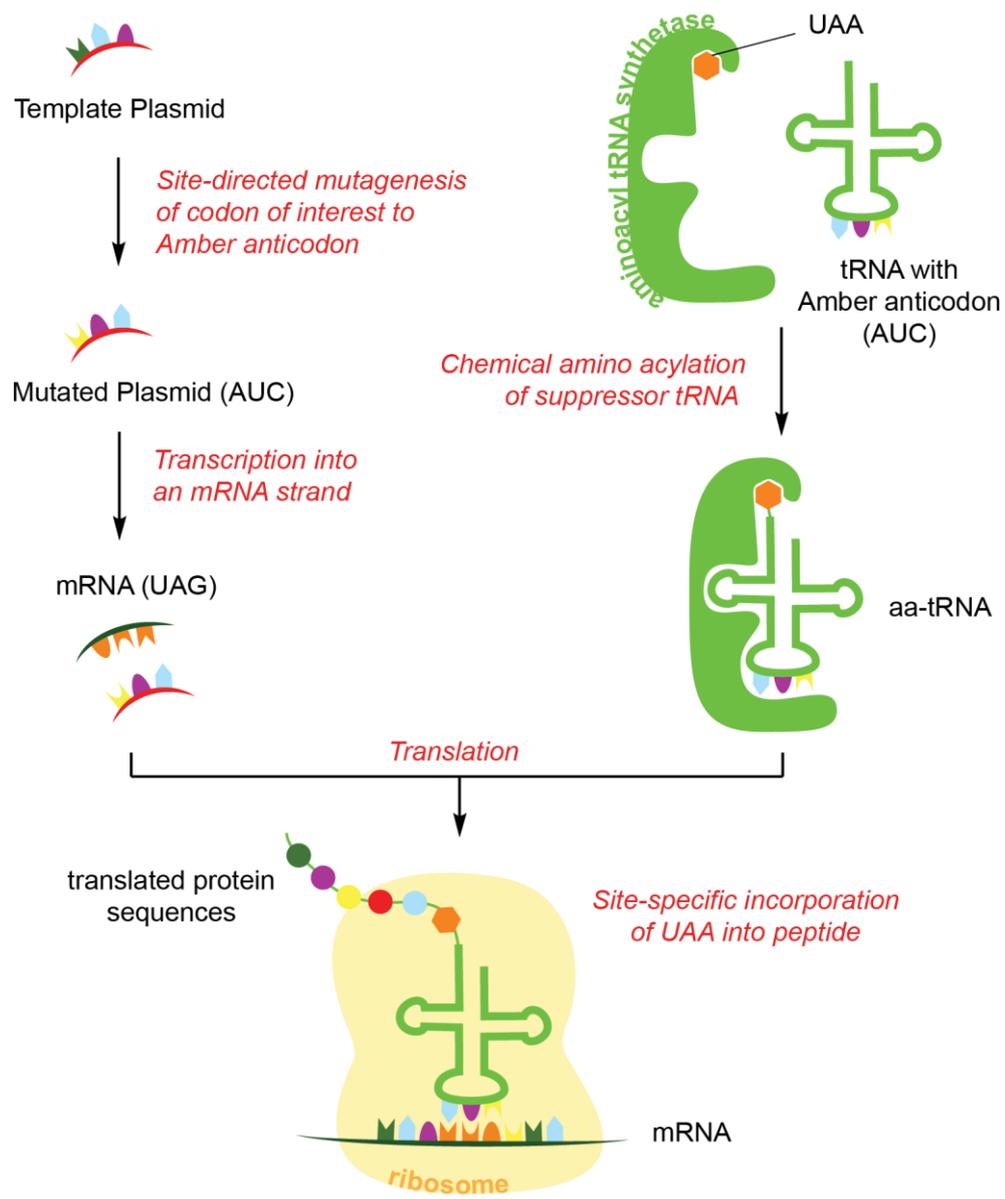
Figure 1.18: Amber suppression method for incorporation of unnatural amino acids into a protein.[103, 104]

The Amber Suppression method has been applied to *Escherichia Coli*[106] and *Saccharomyces cerevisiae*.[107] Applications include incorporation of fluorescent labels into proteins in order to study their folding and ligand binding[108] and protein therapeutics.[109] Recent studies aimed at multiple unnatural amino acid insertions into mammalian cells in order to potentially combine those two applications.[110] Indeed, by varying the codon combinations with 4-base codons, it is possible to insert more than one different UAA in a protein.[111, 112]

Though, theoretically, any insertion is possible, the size of the UAA is a limiting parameter as the peptide sequence must be small enough to get out of the ribosome. It is therefore most likely to be possible to produce bionic proteins by replacing sequences of amino acids with proteomimetic equivalents.

1.3. Previous work from the Wilson group

The Wilson group has focused on designing aromatic oligoamide scaffolds as α-helix mimetics for PPI modulation. Inhibitors of the p53/hDM2 interaction with micromolar activity were identified from libraries of the 2-O-, 3-O-[113] and N-alkylated[114] series (Figure 1.19). Further work from the group aimed at improving water solubility of the 3-O-alkylated inhibitors by introducing a "wet-edge" on the non-binding face of the compound.[115] Fluorescence anisotropy titration showed that this "wet-edge" led to very little decrease of activity, with values of IC_{50} remaining in the micromolar range, while improvement of solubility was observed (Figure 1.19).

3-O-alkylated	3-O-alkylated with "wet edge"	2-O-alkylated	N-alkylated
IC_{50} = 6.3 ± 0.3 μM	7.5 ± 0.4 μM	4.8 ± 0.43 μM	2.8 ± 0.8 μM
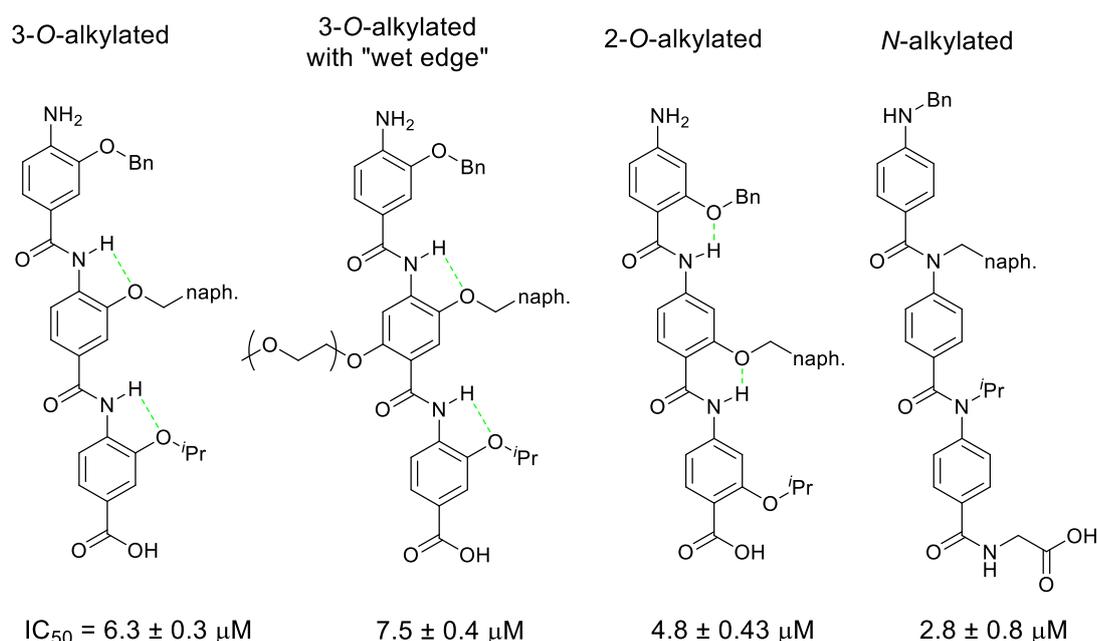
Figure 1.19: Potent inhibitors of the p53/hDM2 interactions and their IC_{50} values, from left to right: 3-O-alkylated series, 3-O-alkylated compound improved with a "wet-edge", 2-O-alkylated series, and N-alkylated series.

The projection of the side chains along one side of the helix relies on the rigidity of the scaffold, which is mainly ensured, for the O-alkylated scaffolds, by the hydrogen bonds between the 2-O and the amide below, or 3-O and the amide above (Figure 1.19, green bonds). In the case of the N-alkylated scaffold, no such interaction is possible due to the functionalization of the amine and the cis-conformation is known to be preferred.[116] While the rigidity enables the projection of the side chains on one face, it can also account for poor affinity in some cases, as a very rigid structure could be restricted in fitting into a binding pocket. In order to break the rigidity of the O-alkylated compounds, a new hybrid scaffold substituting the middle aryl-unit with an α-amino acid has recently been designed (Figure 1.20a).[117] A library of 35 hybrids was prepared using SPPS, and tested against p53/hDM2. For equivalent side chains, three combinations of compounds were compared (Figure 1.20b): 2-O-alkylated, α-amino acid, 3-O-alkylated (left); 3-O-alkylated, α-

amino acid, 3-*O*-alkylated (middle); *N*-alkylated, α-amino acid, 3-*O*-alkylated (right). Out of those three combinations, the first scaffold exhibited the best inhibition potency, as well as high selectivity over the Mcl-1/NOXA B interaction.[117]
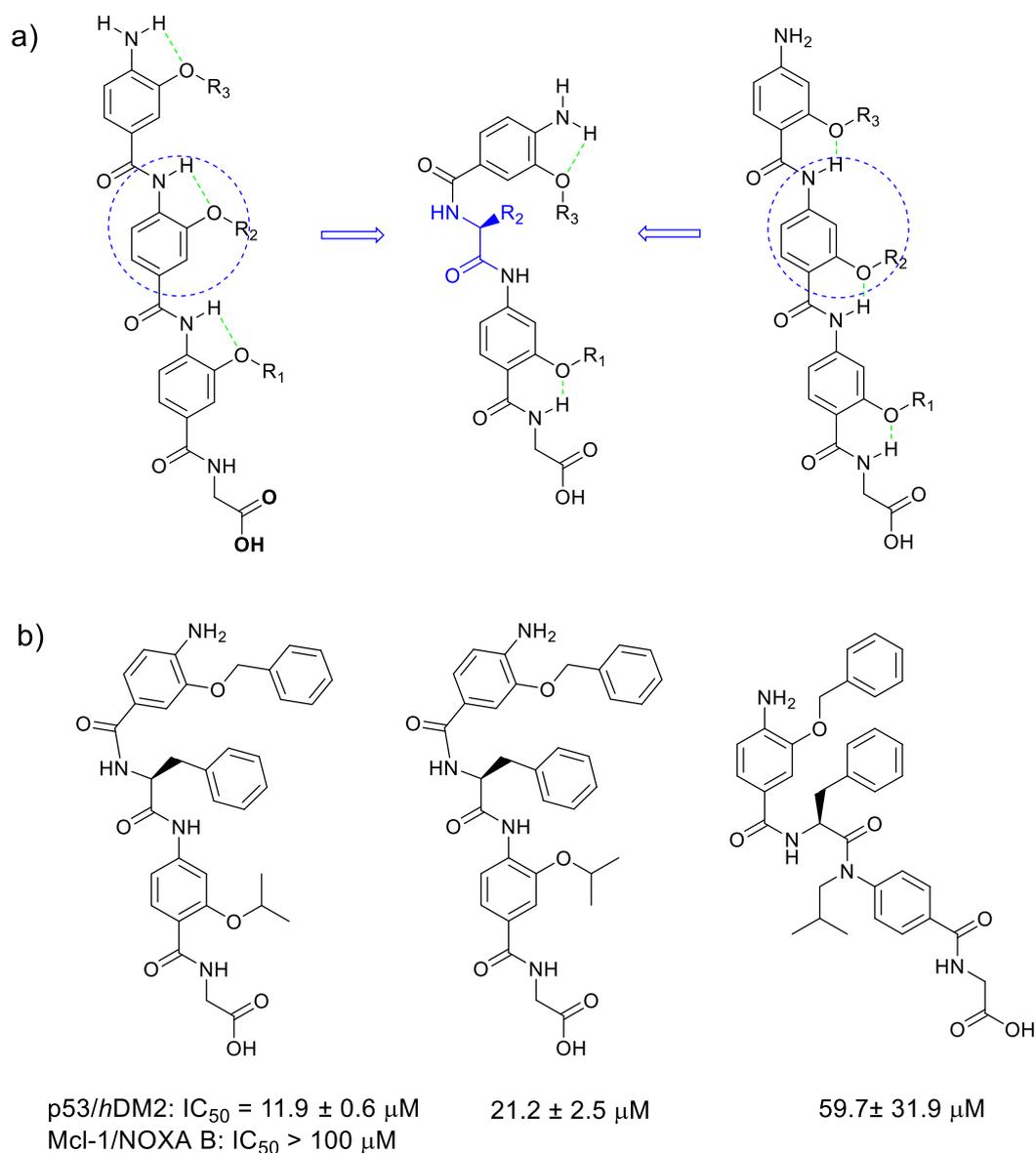
Figure 1.20: a) General design of hybrid mimetics, replacing the central aryl group by a more flexible amino acid (blue). Stabilising hydrogen bonds are highlighted in green.[117] b) Example of potent hybrid inhibitors of the p53/*h*DM2 interaction. Adapted from Azzarito.[117]

The versatility of proteomimetics was established in cell assays. A library of *N*-alkylated oligoamide mimetics of p53 was prepared and tested in a High-Content imaging Screen in U2OS osteosarcoma cells, where four endpoints were assessed after 48 hours of incubation: cell count, identification of apoptotic cells, autophagy, and arrangement of actin filaments.[118] The results highlighted that the interplay between biophysical and cellular potency is complex and that not all *in vitro* hits are potent in cells: out of 77 compounds screened, 2 exhibited potency in all

biophysical and cellular assays. This is nevertheless proof that proteomimetics can effectively inhibit a PPI in cells.

The Wilson group also reported the first peptide-proteomimetic hybrid. Based on preliminary work that demonstrated the ability for a HIF-1α proteomimetic to bind to p300,[119] a truncated HIF-1α protein was synthesised where HIF-1α$_{816-826}$ was replaced by the corresponding proteomimetic (Figure 1.21).[120] The obtained bionic peptide bound to p300 with micromolar activity, which is comparable to the original protein, but 10 fold less than the proteomimetic alone. It was further shown that the proteomimetic on its own had limited selectivity, as it exhibited micromolar activity towards hDM2, whereas the hybrid peptide was selective towards p300.[120]
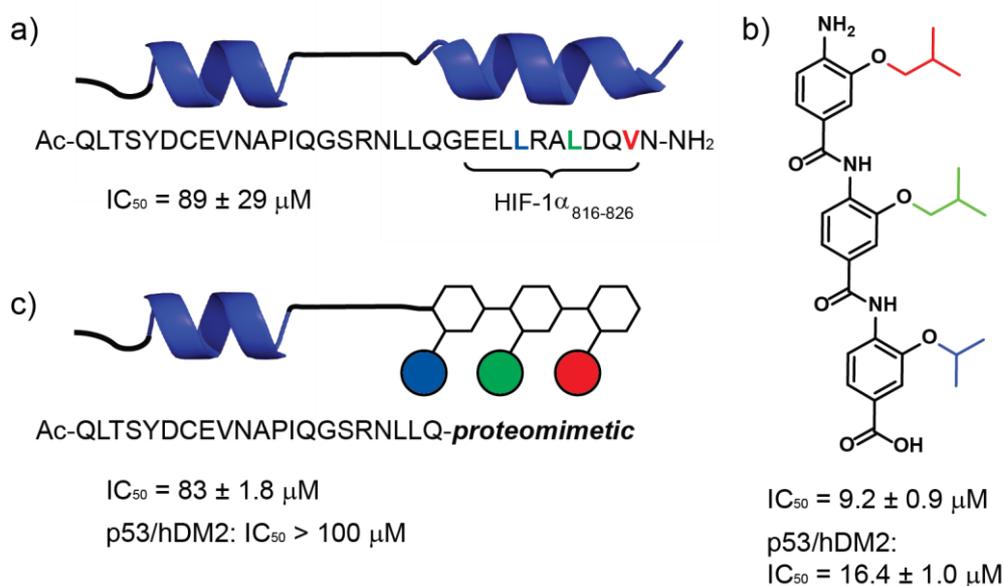
Figure 1.21: First peptide-helix mimetic hybrid.[120] a) Truncated wild type sequence of HIF-1α. b) Structure of the proteomimetic of HIF-1α$_{816-826.}$ c) Bionic peptide of the truncated HIF-1α where residues 816-826 were replaced by the proteomimetic illustrated in b).

1.4. Project Aims

Proteomimetics can accurately mimic the spatial projection of the side chains of an α-helix. Their viability in cellular environment (Figure 1.22a) and insertion as part of a longer peptide was established (Figure 1.22b). However, proteomimetic design and synthesis is, as yet, insufficiently advanced to conceive of re-engineering nature to the extent of replacing whole segments of protein backbone with non-natural prostheses as proposed here (Figure 1.22c).
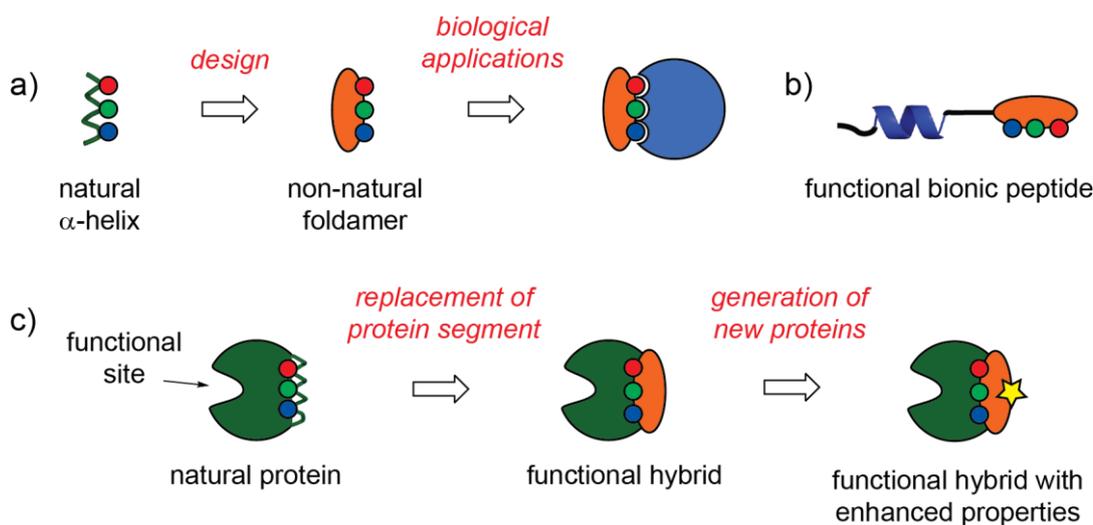
Figure 1.22: Recent developments and future aims on proteomimetics: a) proteomimetics can be designed based on natural α-helices to generate binders for specific proteins and b) can be incorporated to a peptide to form a functional bionic peptide. c) The replacement of protein segments with proteomimetics is yet to be achieved in order to expand Nature's toolbox and generate functional hybrids with enhanced properties.

The aim of this project was to investigate novel applications using proteomimetics, leading to new 3-dimensional assemblies, in the framework of further developing bionic proteins. Two approaches were explored:

1) Design and synthesis of coiled coil mimetics (Chapter 2)

Coiled coils are an essential motif in protein assemblies involved in many biological pathways. Attempting to design coiled coils with proteomimetics is therefore interesting in order to establish the ability of proteomimetics to mimic α-helical supramolecular assemblies. Indeed, although double-sided single α-helices,[85-90, 121] or double single-sided α-helices[93] were studied, no coiled coil mimetic was ever reported. Succeeding in this exercise would prove the versatility of proteomimetics as building block for the construction of bigger, protein-sized complexes.

The synthesis of proteomimetics equipped with polar hydrophilic side chains, designed to enable side-chain/side-chain interactions leading to self-assembly into a coiled coil mimetic, will be described.

2) Identifying high affinity binders to proteomimetics using Affimer display (Chapter 3)

Affimers are a novel artificial binding protein scaffold developed within the Astbury Centre.[122] Phage display screening using Affimers has been successfully used on protein targets,[123] although so far, attempts to screen against non-stabilised natural short peptides failed, due to their lack of structure. By using this technique on proteomimetics, it will be possible to prove their ability to replicate native α-helices in a broader application than previous studies, *i.e.* specific design against

one chosen interaction. This will further validate the value of proteomimetics as potential protein prosthesis building block.

Chapter 2.   Design and synthesis of coiled coil mimetics

2.1. Introduction

2.1.1. Coiled coils

The coiled coil is a naturally abundant protein motif, characterised by the distinctive packing of amino acid sidechains in a heptad repeat, meaning that two turns comprise exactly seven residues forming an a-b-c-d-e-f-g pattern, where a and d are hydrophobic and b, c, e, f, g are polar residues (Figure 2.1a).[124, 125] This structure enables the assembly of the helices into bundles, in which one side chain from one helix is surrounded by side-chains from another helix, to achieve a "knobs-into-holes" packing (Figure 2.1b).[126, 127]
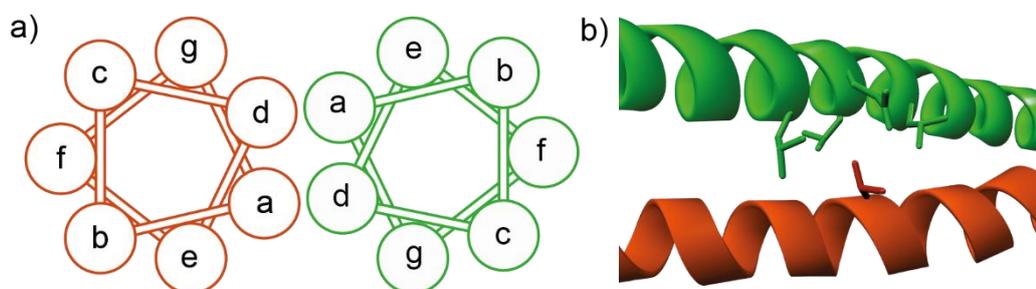
Figure 2.1: Structural characteristics of coiled-coils. a) Cartoon representation of the heptad repeat. b) Knobs-into-holes packing as observed in the heterodimeric bZIP transcription factor c-fos-c-jun leucine zipper (PDB ID: 1FOS).

Coiled coil motifs vary widely, depending on their orientation, length and spatial projection of side-chains. As 3D structure is a central feature of protein function, the coil coiled assembly is key to overall protein folding and mode of action. Examples of protein functions associated with coiled coil structures include DNA-binding,[128-133] ion-channel assembly,[134, 135] and proteasomes.[136] Complex assembly of coiled coils can also lead to architectures such as tentacles (Figure 2.2a, b)[137-139] tubes (Figure 2.2c, d),[140, 141] or stalks (Figure 2.2e).[142, 143]
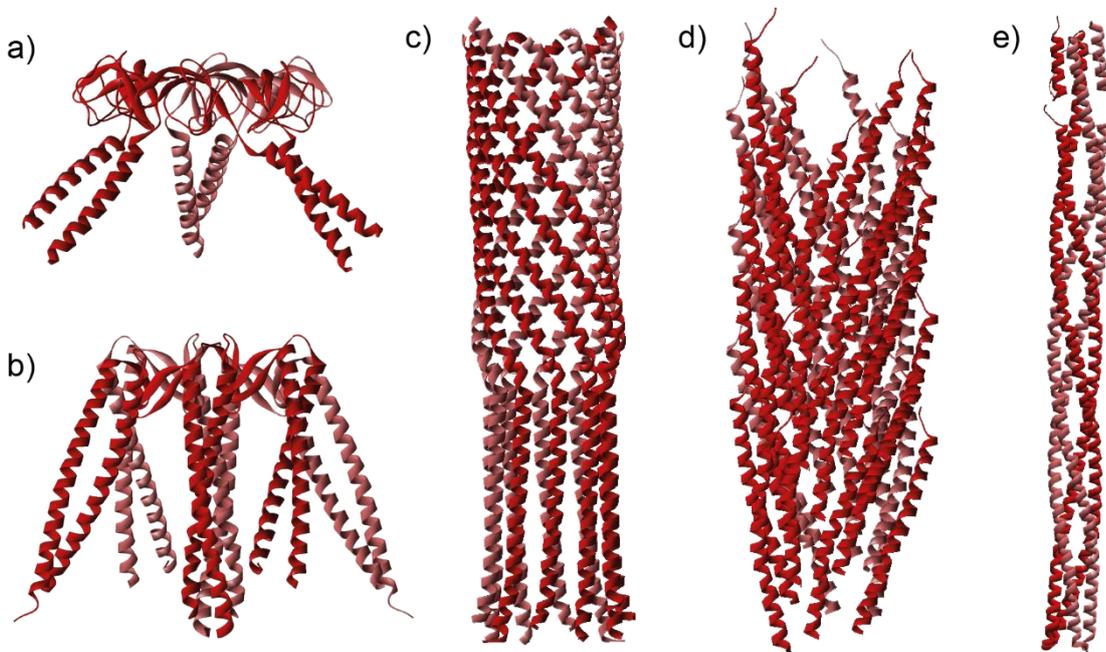
Figure 2.2: Examples of complex assemblies of coiled coils. a) Proteasome-activating nucleotidase N- domain fused to GCNA (PDB ID: 2wg5), b) Archaeal Prefoldin (PDB ID: 1FXK), c) Bacteriophage φX174 tail protein (PDB ID: 4JPP), d) Bacteriophage Pf1 coat protein B (PDB ID: 1QL2), and e) Hamp(AF1503)-Tsr fusion (PDB ID: 3zx6).

Examples of coiled coil mediated membrane-fusion complexes,[144-146] e.g. Influenza Hemagglutinin (HA),[147-150] Ebola Virus (EBOV) GP2,[151-153] and Human Immunodeficiency Virus (HIV-1) GP41[154-157] have underscore the therapeutic relevance of such structures (Figure 2.3).[158]
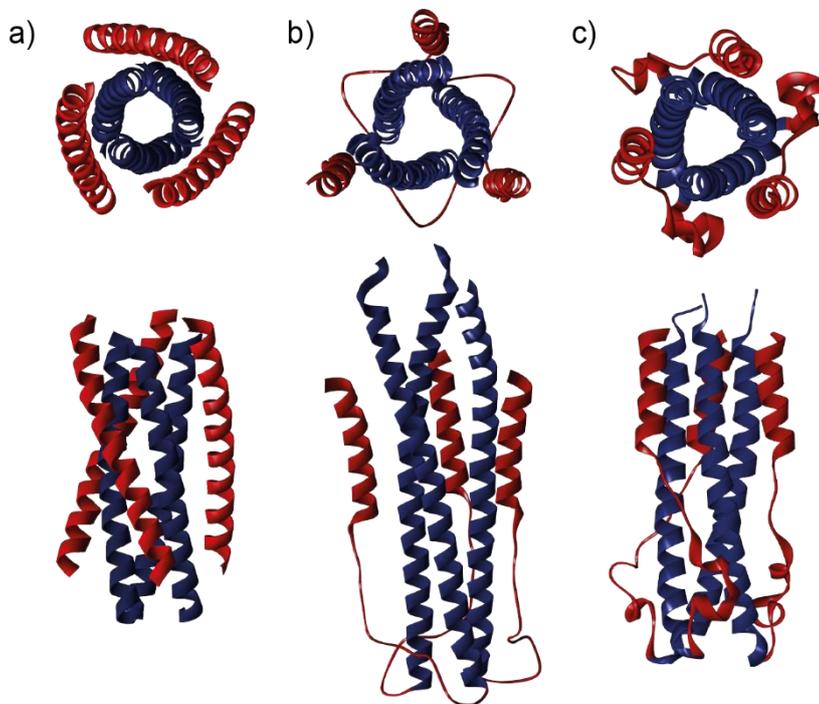
Figure 2.3: Coiled coils involved in viral membrane fusion, top and side view of: a) HIV-GP41 (PDB ID: 1AIK), b) HA (PDB ID: 1HGF) and c) EBOV-GP2 (PDB ID: 2EBO).

24

2.1.2. The proteomimetic approach

In the last two decades, efforts have focused on understanding the fundamental principles of *de novo* coiled coil design.[159-162] Successful design has so far led to assembly of fractal structures,[163] nanotubes[164, 165] and nanofibers,[166, 167] and biochemically active barrels.[168, 169] From a foldamer point of view, it has been shown that it is possible to produce coiled coils with peptides containing β and γ amino acids.[170-172] This exercise was successfully extended to the replacement of two successive α-amino acids by one non-natural 1,2,3-triazole ε²-amino acid.[173] Arora reported the first short coiled coil mimetic by covalent linking of short dimers consisting of heptad repeats.[174]

Though lots of work has been done on the modification and *de novo* design of coiled coils and foldamers, little is known about the properties of aromatic oligomers as coiled coil mimetics. Furthermore, very little synthetic work has been reported on proteomimetics bearing hydrophilic side chains.[175] If it is known that proteomimetics can inhibit a coiled coil mediated interaction,[79] it is worth considering if forming proteomimetic-based coiled coils is possible. In this work, proteomimetics with complementary charged side-chains in order to generate self-assembling foldamers were designed and synthesised. The targets comprised both dimers and trimers as a starting point, to be later extrapolated to longer oligoamides (Figure 2.4).
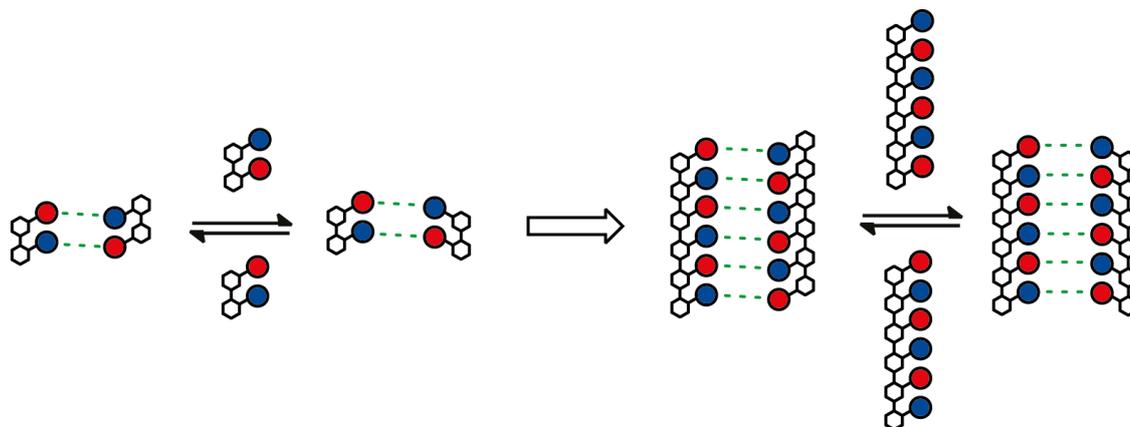
Figure 2.4: Cartoon representation of self-assembling dimers as precursors for designing full coiled coil mimetics. Complementary polar side chains are shown in red and blue. Homo-dimers are expected to self-assemble, and an equilibrium with hetero-dimer formation is expected upon addition of the reverse foldamer sequence.

2.2. Synthesis of self-assembling foldamers

Amongst the proteomimetic scaffolds developed in the Wilson group, the *O*-alkylated[83, 176, 177] and *N*-alkylated series[114, 178] were chosen. The foldamers were synthesised on solid phase on a CEM Liberty Blue® automated microwave assisted peptide synthesizer (Scheme 2.1). The Fmoc-strategy was employed, and Wang resin was chosen following previously optimised synthetic methods.[178]
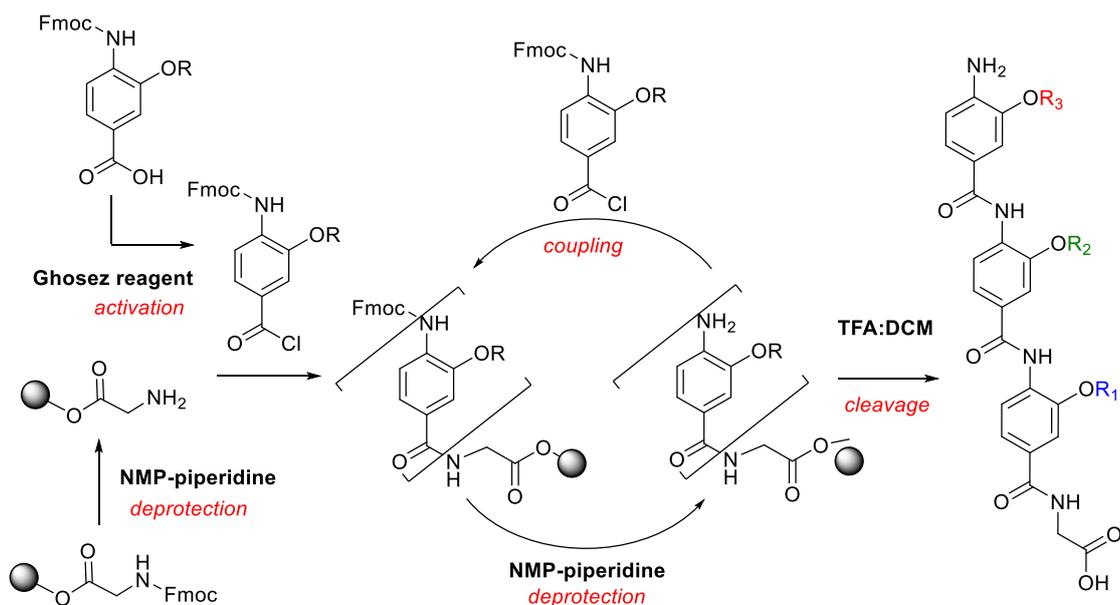
Scheme 2.1: General scheme for the solid phase synthesis of *O*- alkylated oligomers. A similar strategy is applicable for *N*-alkylated oligomer synthesis.

The foldamers were designed using a combination of three types of side chains: a carboxylic acid, an amine, and a neutral group. The synthesis of the building blocks was planned accordingly.

2.2.1. Building block synthesis

The general building block synthesis is described in Scheme 2.2.
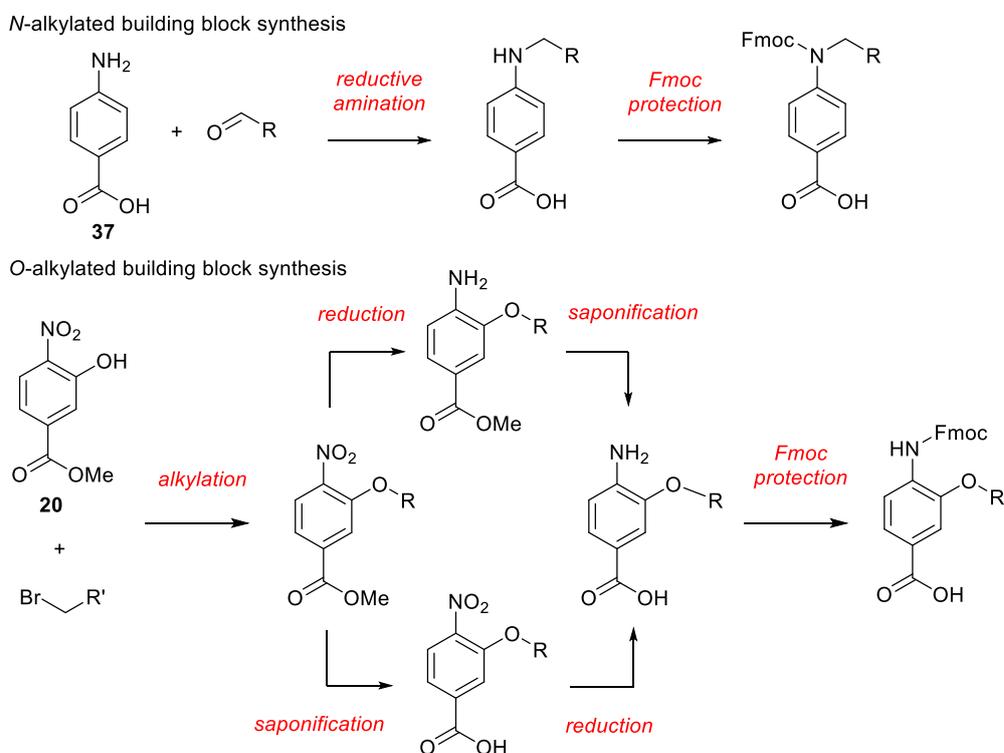
Scheme 2.2: General synthesis for *N*- and *O*-Alkylated monomers.

26

Typically, *N*-alkylated monomers were built by reductive amination of an aldehyde on 4-aminobenzoic acid **37** followed by Fmoc protection, and *O*-alkylated monomers required alkylation on methyl 3-hydroxy-4-nitrobenzoate **20** followed by nitro reduction or saponification of the ester (or the other way round, depending on the nature of the side chain), and, finally Fmoc protection. Aldehydes (*N*-alkylated) or bromides (*O*-alkylated) were synthesised when not commercially available.

2.2.1.1. Preparation of the side chains

The synthesis of the *tert*-butyl ester monomer **3** was originally described using *tert*-butyl 2-bromoacetate **1**, for direct coupling onto *O*-alkylated monomers and as a starting material to make the aldehyde for the *N*-alkylated scaffold (Scheme 2.3).[178] Unfortunately, the formation of the aldehyde from the nitro-compound **2** was unpredictably irreproducible, as the final yield obtained only reached up to 26% and the reaction would often yield no product. Troubleshooting of the reaction (conditions, quality of the reagents and solvents) did not provide any information as to the origin of this result.
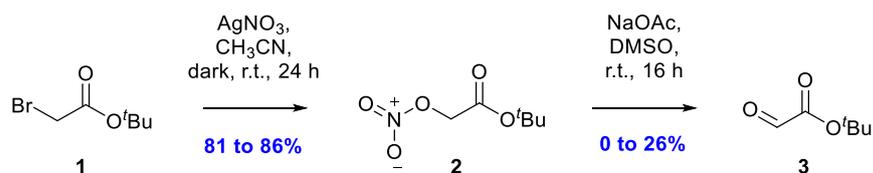
Scheme 2.3: Synthesis of the *tert*-butyl ester aldehyde **3** as previously described in the Wilson group.[178]

Other methods to obtain the same aldehyde were investigated (Scheme 2.4). Direct esterification[179] of 2-oxoacetic acid **9** or L-tartaric acid **4** were ineffective; the formation of the ester was observed with the use of CuCl,[180] but the metal salt could not be removed. Then, the aim was shifted towards obtaining di-*tert*-butyl fumarate **8** in order to form the aldehyde **3** by ozonolysis. Metathesis on *tert*-butyl acrylate was successful,[181] but the obtained yields (between 2 and 8%) were limiting. Finally, fumaryl chloride **7** in the presence of *n*-butyl lithium and *tert*-butyl alcohol formed di-*tert*-butyl fumarate **8** with 38% yield,[182] which was further ozonolysed to obtain the desired aldehyde.[183]

Scheme 2.4: Alternative synthetic ways to obtain the *t*Bu ester aldehyde.

Considering the low overall yield, the synthesis of a *tert*-butyl ester monomer with a slightly longer side chain was investigated. The synthesis described in Scheme 2.5 illustrates how both bromide **14** and aldehyde **13** required for coupling on respectively *O*- and *N*-alkylated scaffolds were obtained in a cheaper and safer manner, in only 4 steps from succinic anhydride. Forming the bromide was preferred to using the alcohol directly for Mitsunobu coupling on methyl 3-hydroxy-4-nitrobenzoate **20**, as the purification of the Mitsunobu mixture was difficult and time-consuming because of the presence of an excess of triphenylphosphine.

Scheme 2.5: Synthesis of bromide **14** and aldehyde **13** side-chains for incorporation on respectively *O*- and *N*- alkylated monomers from succinic anhydride **10**.

28

The amine side-chain monomer **16** was prepared following the previously described procedure (Scheme 2.6): 3-bromopropylamine was boc-protected in one step; 3-aminopropan-1-ol was boc-protected and oxidized into the aldehyde **19** using Swern conditions.
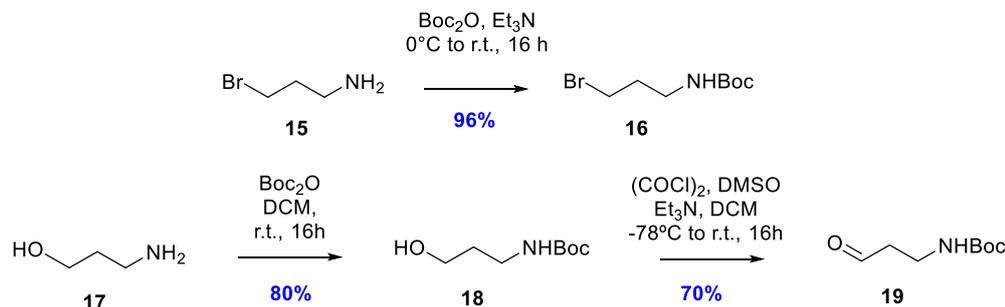
Scheme 2.6: Synthesis of the bromide **16** (top) and aldehyde **19** (bottom) boc-amine side-chains.

2.2.1.2. Synthesis of the monomers

The *O*-alkylated monomer synthesis (Scheme 2.7) started with the alkylation of methyl-4-nitro-3-hydroxybenzoate **20** with the relevant bromide. For isobutyl and amine functionalised groups, the hydrogenation was performed first, followed by the hydrolysis of the methyl ester under basic conditions and Fmoc protection (Scheme 2.7).
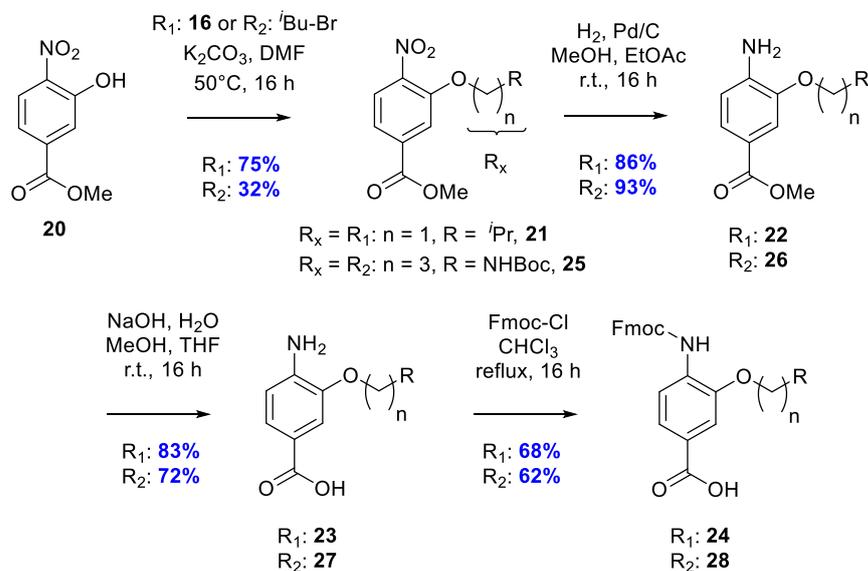
Scheme 2.7: Formation of amine (**24**) and isobutyl (**28**) *O*-alkylated monomers.

For the *tert*-butyl ester side-chain, the saponification was performed first: in order to avoid hydrolysing the *tert*-butyl ester, it was necessary to tune the conditions and use a weaker base than with the other side-chains. Such bases are not strong enough to hydrolyse the methyl ester if the nitro group is reduced to an amine, because of its electron donating properties: indeed, combined with the electron withdrawing properties of the methyl ester, the overall electron density is shifted towards the ester, making it less likely for the negative HO⁻ group to attack there

(Scheme 2.8). Instead, the hydrolysis happened in the presence of the nitro group, which is electron withdrawing.
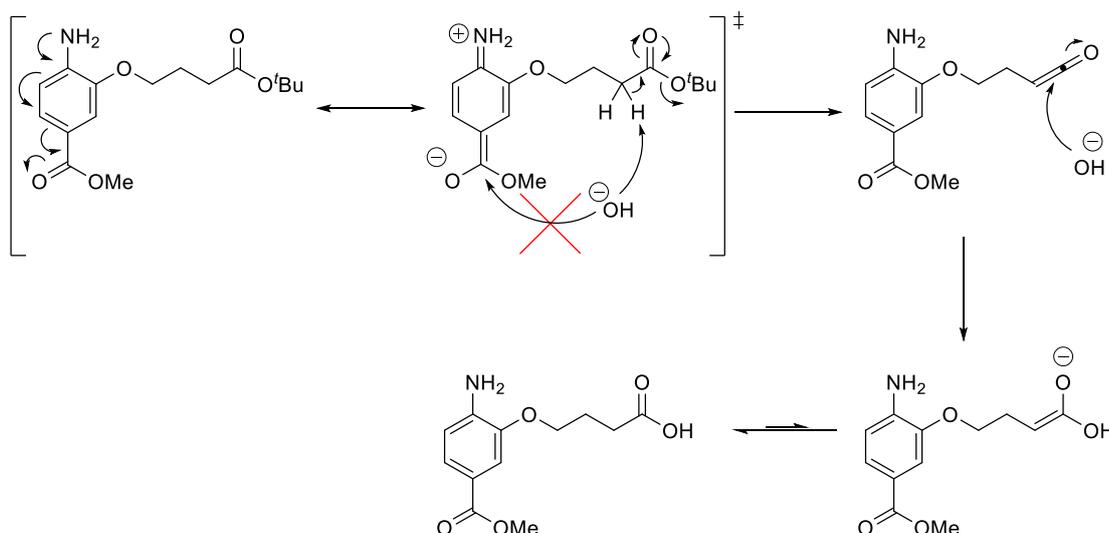
Scheme 2.8: The electron density shifted towards the ester makes it harder for the base to attack the methyl ester, favouring the deprotection of the *tert*-butyl ester.

Following the method previously described in the group, lithium hydroxide was used for n = 1; for n = 3, a 0.25 M solution of NaOH in ice was used (Scheme 2.9).
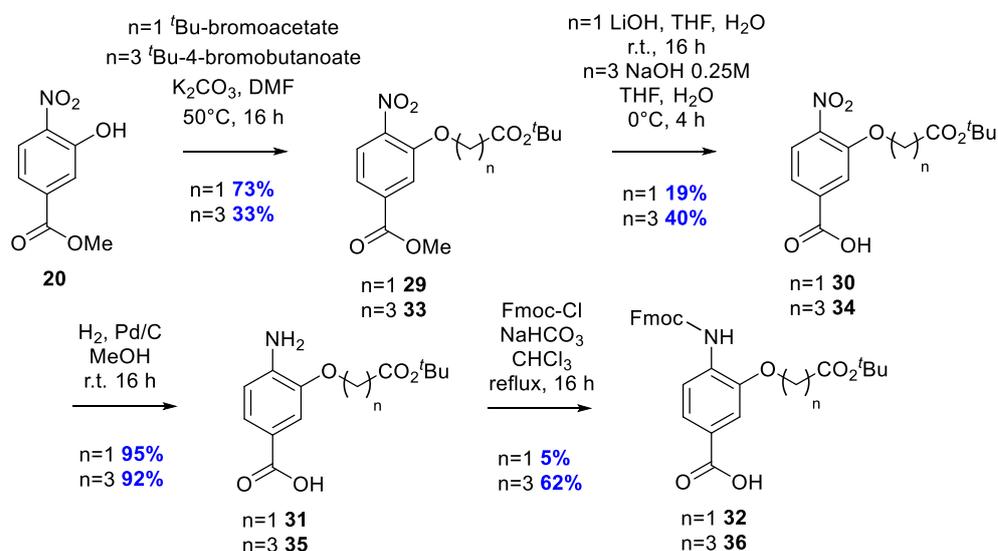
Scheme 2.9: Formation carboxylic acid *O*-alkylated monomers.

Full conversion of the starting material was never reached and any attempt to improve it inexorably lead to the deprotection of the *tert*-butyl group. Ultimately, using a diluted NaOH solution in ice gave a better yield. The ester and the carboxylic acid could in either case be separated by column chromatography, and the ester could be hydrolysed again. The hydrogenation was in both cases high-yielding, and finally, the Fmoc protection gave 5% yield for n = 1, compared to 62% for n = 3. Further investigation by Dr Ludwig Pilsl on the preparation

30

of **32** established that the formation of the Fmoc-product was effective after 2 hours, but longer reflux led to a degradation of the product and formation of an unidentified by-product.

In the case of the *N*-alkylated monomers, the length of the side chain had no impact on the yield and the protected monomers were obtained in two steps: reductive amination of 4-aminobenzoic acid **37**, followed by Fmoc protection (Scheme 2.10).

R = ^{i}Pr, n = 1, **38**, **91%**
R = CO_2^tBu, n = 1, **40**, **51%,**
n = 3, **42**, **51%**
R = NHBoc, n = 3, **44**, **55%**

R = ^{i}Pr, n = 1, **39**, **59%**
R = CO_2^tBu, n = 1, **41**, **51%**
n = 3, **43**, **52%**
R = NHBoc, n = 3, **45**, **35%**
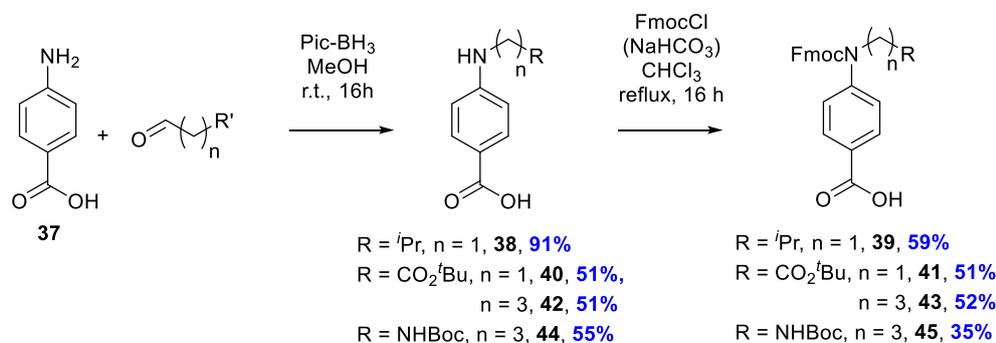
Scheme 2.10: Synthesis *N*-alkylated monomers in two steps.

2.2.2. Dimer and trimer synthesis

Dimers and trimers were prepared by SPPS following Scheme 2.1. The solvent of choice for the synthesis of the trimers was *N*-methyl-2-pyrrolidone. Indeed, in the case of the *O*-alkylated scaffold, it was previously observed that the monomers were poorly soluble in chloroform, and *N*,*N*-dimethylformamide reversibly reacts on the acyl chloride to generate a Vilsmeier intermediate which can irreversibly cap the aniline (Scheme 2.11).[177]
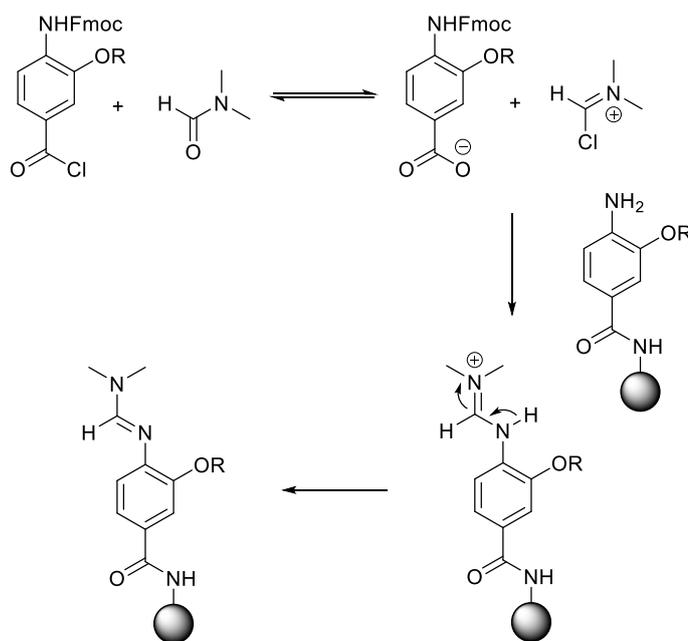
Scheme 2.11: Mechanism for capping of anilines during solid phase synthesis *via* a Vilsmeier intermediate.

On acquisition of a CEM Liberty Blue Peptide Synthesizer, it was also decided to switch from chloroform to *N*-methyl-2-pyrrolidone for the *N*-alkylated scaffolds: despite optimisation efforts, the chloroform would partially evaporate upon heating and nitrogen bubbling, which resulted in a non-homogeneous distribution of the resin in the reaction vessel and therefore non-homogeneous coupling. Inconsistent deprotections were also observed, which were attributed to the partial recondensation of the chloroform in the tubes, leading to an involuntary addition of a small amount of piperidine in the reaction vessel during the coupling steps. Using *N*-methyl-2-pyrrolidone, no evaporation occurred, allowing a homogeneous distribution of the resin, and the deprotections were consistent, suggesting the problem was indeed linked to the use of chloroform.

After synthesis using the optimised parameters, the trimers were cleaved off the resin with a 50% trifluoroacetic acid solution in dichloromethane and purified on semi-preparative HPLC using acetonitrile and water.

Two series of molecules, *O*-alkylated dimers and *N*-alkylated dimers and trimers (Figure 2.6), were synthesised, for two purposes:

1) Conformational analysis:
   - Previous studies on 3-*O*-alkylated dimers have shown that the side chain orientation in the solid phase depends mainly on the packing of the molecules in the crystal lattice,[184, 185] and that rotation around the ArNH bond is possible depending on the nature of the side chains. In solution phase, both *anti* and *syn* conformations were shown to coexist (Figure 2.5a).[184]
   - *N*-alkylated trimers were shown to prefer adopting a *cis* conformation (Figure 2.5b)[116] and were hypothesised to change to a *trans* conformation upon protein binding.[186]
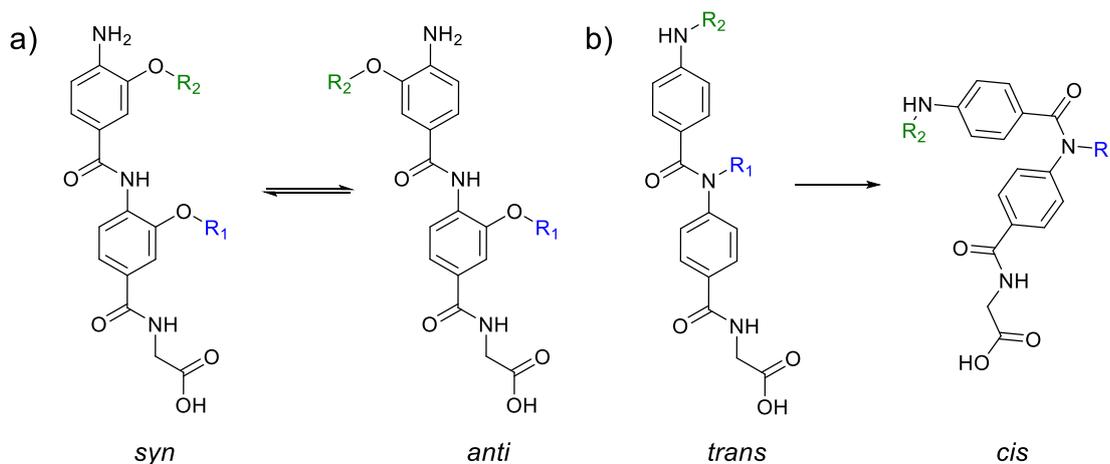
Figure 2.5: Conformational analyses on aromatic oligoamides showed a) the coexistence of *syn* and *anti* conformations for *O*-alkylated series and b) a preference for the *cis* conformation *N*-alkylated series.

The first objective was to determine the effect of the nature of the side chain on the conformation of the *O*-alkylated dimers and *N*-alkylated dimers and trimers, and establish whether it is possible to control the conformation of the foldamer by designing hydrophilic side chains that can form a stabilising intramolecular interaction.

2) Observe self-assembly of polar *O*-alkylated dimers and *N*-alkylated dimers and trimers as a precursor for coiled coil proteomimetics.

The nomenclature for the molecules is given from top to bottom, where each monomer is numbered and the nature of the side chain is indicated in subscript, considering A = carboxylic acid, and B = base (amine), and any non-polar side chain is denominated by its name (in this case *i*Bu). The molecule names start by the letter *O* or *N* depending whether it is *O*- or *N*- alkylated, and end with the letter G to denominate the amino acid obtained from the resin (in that instance, Glycine). In the case of same side-chain with different lengths, the long chain is considered as the default size and the shorter chain is indicated with an "s". The list of target molecules is presented in Figure 2.6.
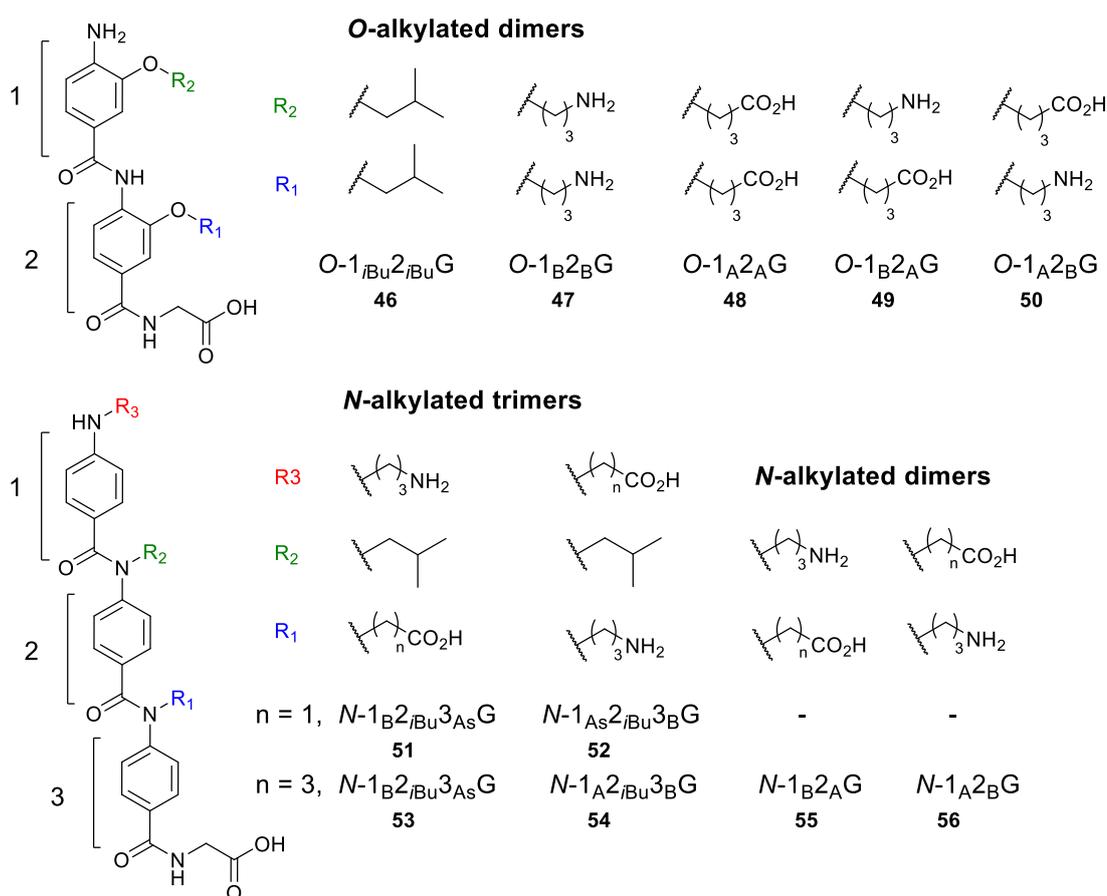
Figure 2.6: Synthetic targets for conformational analyses and self-assembly.

Because of the hydrophilic nature of the side-chains, some of the foldamers prepared degraded. In a few cases, the difference of mass corresponded to a cyclisation, supposedly of the carboxylic

acid monomer on the top alanine (Scheme 2.12), *O*-alkylated cyclisation observed by Dr Ludwig Pilsl). In other instances, the reason for the degradation was not clear.
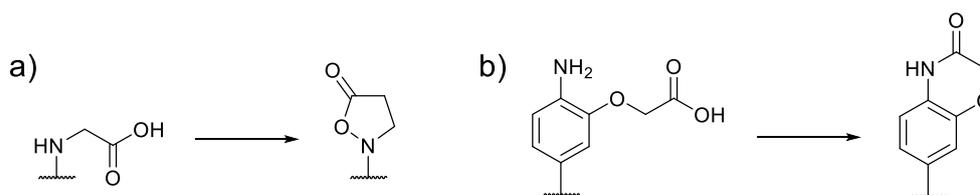
Scheme 2.12: Cyclisation of the top acid monomer a) for *N*- and b) for *O*-alkylated series.

Attempts at synthesising *N*-1$_{As}$2$_{iBu}$3$_B$G (**52**) led to a dimer/trimer mixture that co-eluted and required several rounds of purification. Cyclisation of the top monomer happened during purification, leading to a mixture of dimer, trimer and cyclised dimer that could not be further purified (Figure 2.7). The cyclisation was considered to be caused by the presence of 0.1% of formic acid in the purification gradient. Once this was established, the purification procedure was changed to a gradient of water and acetonitrile only, without using any acid. The synthesis of *N*-1$_{As}$2$_{iBu}$3$_B$G (**52**) was not repeated due to insufficient amounts of monomer available. The synthesis efforts were thus focused on the n = 3 monomer (A) as it was easier to prepare in greater amounts.
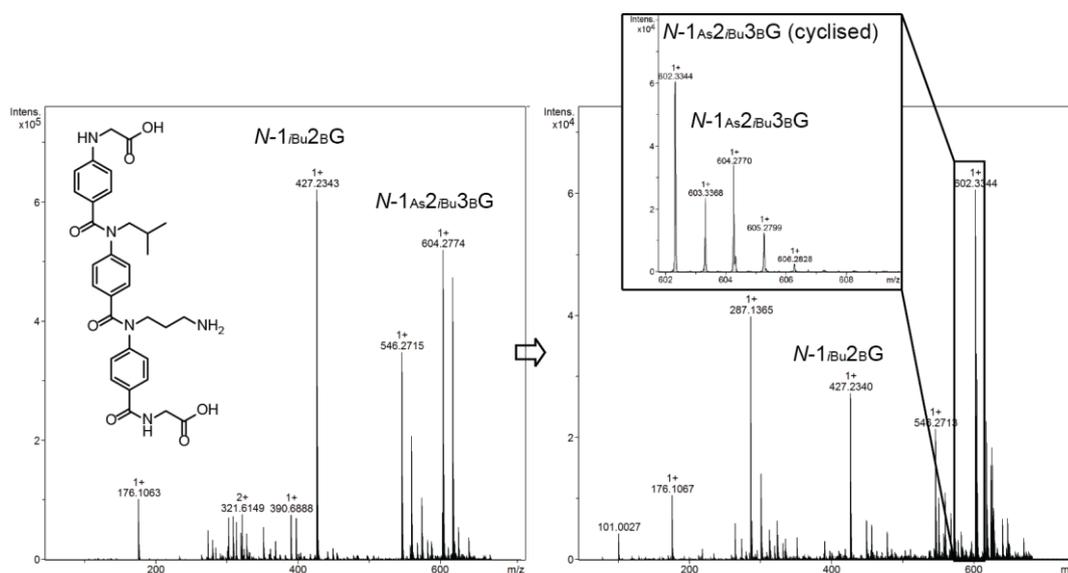
Figure 2.7: HRMS traces of *N*-1$_{As}$2$_{iBu}$3$_B$G (**52**) before purification, as a mixture with dimer *N*-1$_{iBu}$2$_B$G (left) and observed cyclisation after purification (right).

O-1$_B$2$_B$G (**47**) and *O*-1$_B$2$_A$G (**49**) degraded respectively after or during purification (Figure 2.8). By-products could not be identified.

Attempts to synthesise *N*-alkylated dimers afforded *N*-1$_B$2$_A$G (**55**) but *N*-1$_A$2$_B$G (**56**) could not be synthesised (not observed by test-cleavage of the crude product). Finally, the control dimer *O*-1$_{iBu}$2$_{iBu}$ (**46**) was prepared, and although it did not degrade, several impurities appeared after

34

purification, which led to an insufficient amount of product to obtain a clean NMR spectra. While it is understandable that charged dimers or trimers could degrade, for example due to the observed cyclisation, it is unclear why any trouble would arise from purification of a dimer containing only isobutyl side-chains. It is therefore believed that the semi-preparative column used on the HPLC might have been contaminated with residual traces of acid.

Dimers O-1_A2_BG (**50**), O-1_A2_AG (**48**), N-1_A2_BG (**55**) and trimers N-$1_B2_{iBu}3_AG$ (**53**) and N-$1_A2_{iBu}3_BG$ (**54**) were obtained. For O-alkylated dimers, the observations linked to compound degradation showed a correlation with the presence of an amine group at the top position. There is no obvious explanation why these molecules would be more prone to degradation than the others, but an answer could lie in the fact that the amine monomer did not always couple fully during the synthesis, even with efforts to increase the yield (longer reaction time or triple couplings were attempted with little improvement). Because of this, the amount of product in the crude mixture was limited, making the purification more difficult and leading to no or too little final product to be isolated and used.
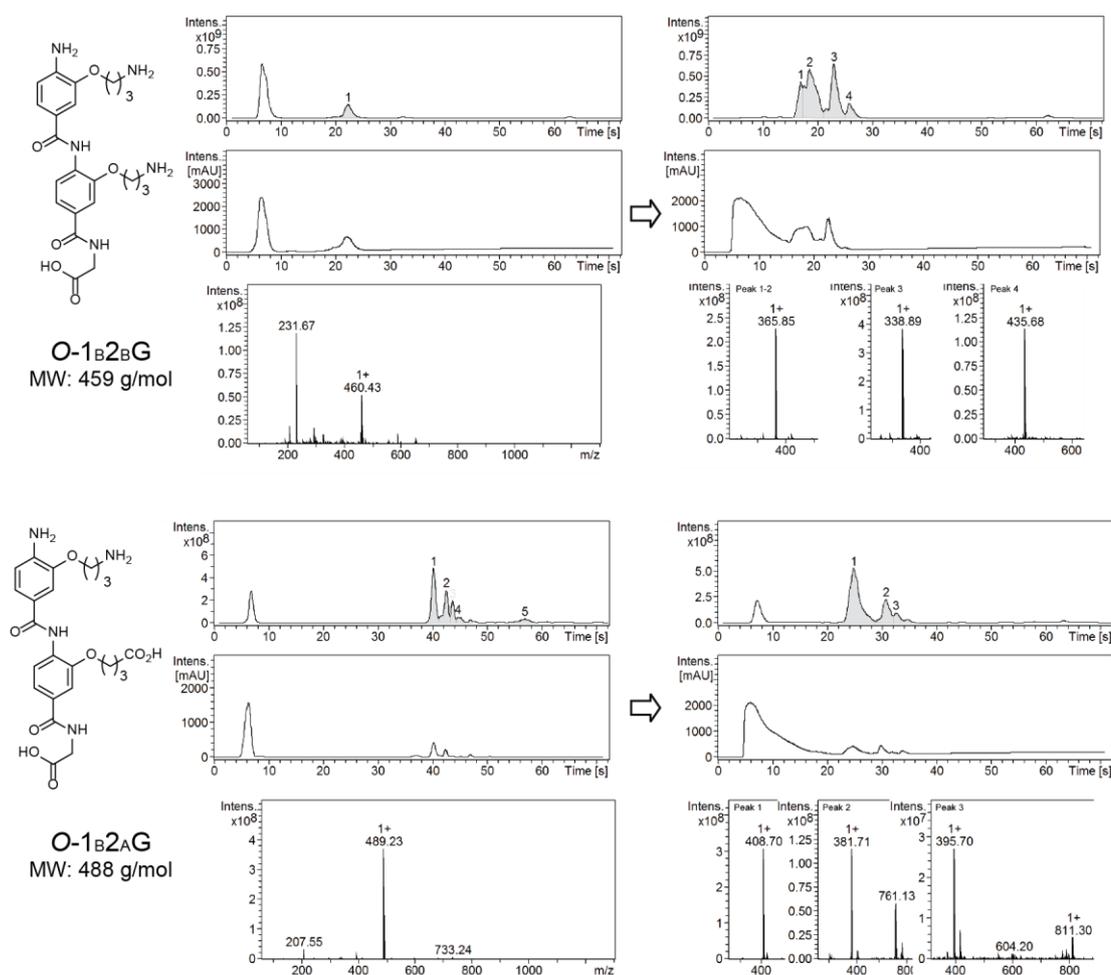
Figure 2.8: LC-MS traces showing the degradation of O-1_B2_BG (**47**) and O-1_B2_AG (**49**) respectively after or during purification.

35

Due to time restrictions, further efforts to understand the problems better and repeat those syntheses were not accomplished, and the NMR analyses were focused on the obtained molecules: O-1_A2_AG (**48**), O-1_A2_BG (**50**), N-1_A2_BG (**55**), N-$1_A2_{iBu}3_BG$ (**54**), and N-$1_B2_{iBu}3_AG$ (**53**). Instead of O-$1_{iBu}2_{iBu}$ (**46**), a readily available dimer, O-$1_{sBu}2_{sBu}G$ (**57**), prepared previously by Dr Sílvia Rodriguez-Marin, was used.

2.3. Molecular modelling predictions

Molecular modelling was carried out in order to generate ideas as to the behaviour of the dimers and the trimers. A Monte Carlo Multiple Minima (MCMM) molecular mechanics conformational search was carried out using MacroModel (Schrödinger) and the Merck Molecular Force Field (MMFF) forcefield, without restraints and with a cutoff of 1Å in octanol. Considering the pKa of the relevant groups (Table 2.1), it was established that the carboxylic acid would be deprotonated, and the amine protonated, and this was taken into account for the modelling.

Table 2.1: pKa and ionisation states of the hydrophilic side-chains.

Group	pKa	Ionisation state
Primary amine (ammonium)	11-12	NH_3^+
Carboxylic acid (including Gly)	4-5	CO_2^-
Aromatic amine	5	NH_2

The conformational landscapes were largely dominated by the structures shown in Figure 2.9. For clarity, only the compounds that were obtained are discussed in this section: the modelling of the molecules that degraded or could not be synthesised are presented in Appendix I. For O-alkylated dimers (Figure 2.9a), the conformations obtained depended on the nature of the side-chains. For sBu groups (O-$1_{sBu}2_{sBu}G$, **57**), an equal distribution of *syn* and *anti* conformations was observed; for O-1_A2_AG (**48**), the *anti* conformation was predominant; for O-1_A2_BG (**50**), a *syn* conformation was preferred, stabilised by side-chain/side-chain H-bonding. In the case of N-alkylated dimer N-1_B2_AG (**55**) (Figure 2.9b), a mixture of *trans* and *cis* conformations was obtained, each stabilised by a ammonium/carboxylate H-bond respectively between side-chains or with the bottom glycine. Finally, N-alkylated trimers N-$1_B2_{iBu}3_AG$ (**53**) and N-$1_A2_{iBu}3_BG$ (**54**) (Figure 2.9c) exhibited a preference for the *cis* conformation at each amide bond, with a bridge between the amine side-chain and both carboxylic acids.
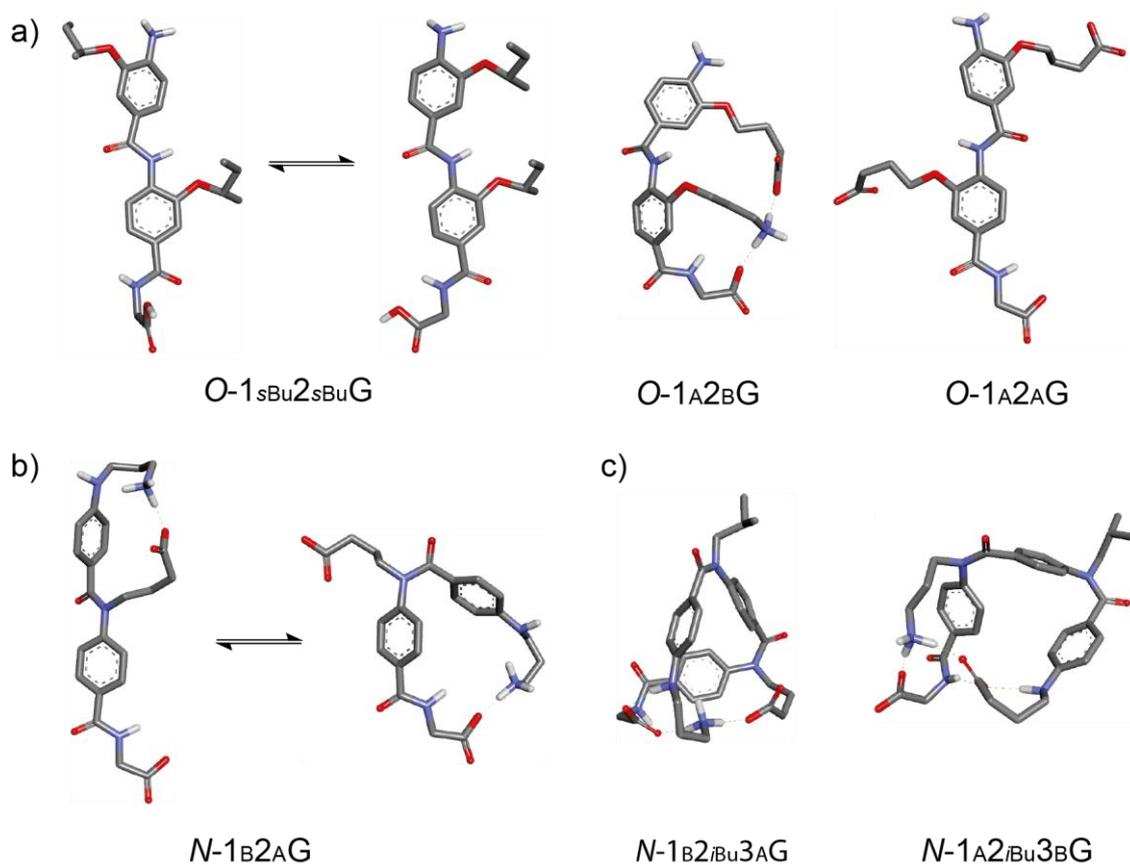
Figure 2.9: Molecular modelling predictions for a) *O*-alkylated dimers, b) *N*-alkylated dimers and c) *N*-alkylated trimers.

Self-assembly was also studied by modelling the interactions between pairs of oligoamides (Figure 2.10). In the cases of *O*-1$_A$2$_B$G (**50**) and *N*-1$_B$2$_A$G (**55**), a change of conformation happened upon dimerization: *O*-1$_A$2$_B$G (**50**) adopted an *anti* conformation, whereas *N*-1$_B$2$_A$G (**55**) adopted a *cis* conformation. The self-assembling predictions suggest the following assemblies:

<u>*O*-1$_A$2$_B$G(1) – *O*-1$_A$2$_B$G(2)</u> (Figure 2.10a): the complex was symmetrical: the amine of residue 2 on dimer 1 (*O*-1$_A$**2$_B$**G(1)) bound to the carboxylic acid of residue 1 on dimer 2 (*O*-**1$_A$**2$_B$G(2)) as well as the glycine of dimer 1 (*O*-1$_A$2$_B$**G**(1)). The carboxylic acid of residue 1 of dimer 2 (*O*-**1$_A$**2$_B$G(2)) also bound to the bottom amide of dimer 1 (*O*-1$_A$2$_B$-CON**H**-G(1)).

<u>*N*-1$_B$2$_A$G(1) – *N*-1$_B$2$_A$G(2)</u> (Figure 2.10b): the complex was symmetrical: the amine of residue 1 on dimer 1 (*N*-**1$_B$**2$_A$G(1)) bound to the carboxylic acid of residue 2 on dimer 2 (*N*-1$_B$**2$_A$**G(2)). The glycine group was not involved in the complex.

Although the trimers did not undergo a complete conformational change, some extent of self-assembly was observed (Figure 2.10c).

a)

O-1$_A$2$_B$G - *O*-1$_A$2$_B$G

b)

N-1$_B$2$_A$G - *N*-1$_B$2$_A$G

c)

N-1$_B$2*i*Bu3$_A$G - *N*-1$_B$2*i*Bu3$_A$G

N-1$_B$2*i*Bu3$_A$G (grey) - *N*-1$_A$2*i*Bu3$_B$G (black)

N-1$_A$2*i*Bu3$_B$G - *N*-1$_A$2*i*Bu3$_B$G
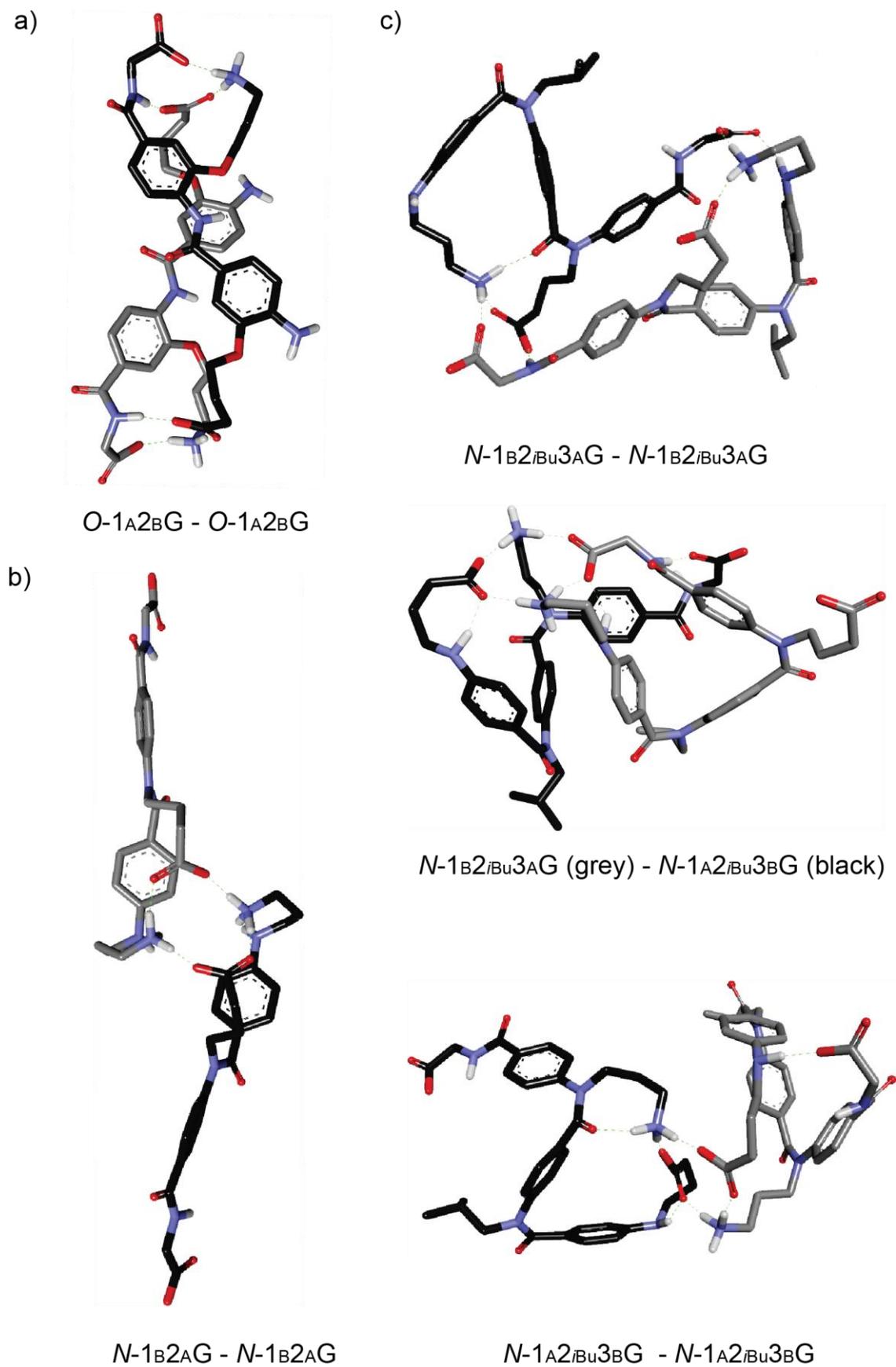
Figure 2.10: Modelling of the self-assembly of pairs of oligoamides: a) *O*-alkylated dimer *O*-1$_A$2$_B$G (**50**), b) *N*-alkylated dimer *N*-1$_B$2$_A$G (**55**) and c) *N*-alkylated trimers (**53** and **54**).

The following hydrogen bonding profiles were obtained:

N-1$_B$2$_{iBu}$3$_A$G(1) (grey) – N-1$_B$2$_{iBu}$3$_A$G(2) (black): the amine side-chain of residue 1 on trimer 1 (N-**1$_B$**2$_{iBu}$3$_A$G(1)) bound to the carboxylic acid of residue 3 from trimer 1 (N-1$_B$2$_{iBu}$**3$_A$**G(1)) as well as the glycine on trimer 2 (N-1$_B$2$_{iBu}$3$_A$G(2)). That later bond was stabilised by an interaction between that glycine (N-1$_B$2$_{iBu}$3$_A$**G**(2)) and the top amine of trimer 1 (N-N**H**-1$_B$2$_{iBu}$3$_A$G(1)). The amine side-chain from residue 1 on trimer 2 (N-**1$_B$**2$_{iBu}$3$_A$G(2)) bound to the carboxylic acid from residue 3, and amide bond from trimer 2 (N-1$_B$2$_{iBu}$-C**O**N**H**-**2$_A$**G(2)), as well as the glycine from trimer 1 (N-1$_B$2$_{iBu}$3$_A$**G**(1)). The amide bond from the glycine of trimer 1 (N-1$_B$2$_{iBu}$3$_A$-CON**H**-G(1)) also bound to carboxylic acid side-chain on residue 3 of trimer 2 (N-1$_B$2$_{iBu}$**3$_A$**G(2)).

N-1$_B$2$_{iBu}$3$_A$G(1) (black) – N-1$_A$2$_{iBu}$3$_B$G(2) (grey): The amine of residue 1 on trimer 1 (N-**1$_B$**2$_{iBu}$3$_A$G(1)) bound to the glycine from trimer 1 (N-1$_B$2$_{iBu}$3$_A$**G**(1)) and carboxylic acid from residue 1 on trimer 2 (N-**1$_A$**2$_{iBu}$3$_B$G(2)). The amine on residue 3 of trimer 2 (N-1$_A$2$_{iBu}$**3$_B$**G(2)) bound the carboxylic acid side-chain of residue 1 on trimer 2 (N-**1$_A$**2$_{iBu}$3$_B$G(2)) and the glycine trimer 1 (N-1$_B$2$_{iBu}$3$_A$**G**(1)), which bridged with the top amine on trimer 1 (N-N**H**-1$_B$2$_{iBu}$3$_A$G(1)). The carboxylic acid from residue 1 on trimer 2 (N-**1$_A$**2$_{iBu}$3$_B$G(2)) also bound to the top amine on trimer 2 (N-N**H**-1$_A$2$_{iBu}$3$_B$G(2)). The glycine on trimer 2 (N-1$_A$2$_{iBu}$3$_B$**G**(2)) bound to the bottom amide bond on trimer 2 (N-1$_B$2$_{iBu}$3$_A$-CON**H**-G(1)).

N-1$_A$2$_{iBu}$3$_B$G(1) (grey) – N-1$_A$2$_{iBu}$3$_B$G(2) (black): The carboxylic acid from the residue 1 of trimer 1 (N-**1$_A$**2$_{iBu}$3$_B$G(1)) bound to the amine groups on residues 3 of both trimers (N-1$_A$2$_{iBu}$**3$_B$**G(1) and N-1$_A$2$_{iBu}$**3$_B$**G(2)). The carboxylic acid from the residue 1 on trimer 2 (N-**1$_A$**2$_{iBu}$3$_B$G(2)) also bound to both amine groups on residues 3 (N-1$_A$2$_{iBu}$**3$_B$**G(1) and N-1$_A$2$_{iBu}$**3$_B$**G(2)) and had an extra stabilisation by binding to the nearest amide bond (N-1$_A$-CON**H**-2$_{iBu}$3$_B$G(2)). The amine group on residue 3 from trimer 2 (N-1$_A$2$_{iBu}$**3$_B$**G(2)) bound to the nearest amide (N-1$_A$2$_{iBu}$-CON**H**-3$_B$G(2)). The glycine from trimer 1 (N-1$_A$2$_{iBu}$3$_B$**G**(1)) bound to the top amine from trimer 1 (N-N**H**-1$_A$2$_{iBu}$3$_B$G(1)).

The summary of all interactions observed by molecular modelling is presented in **Table 2.2**.

Taking those calculations into account, NMR experiments were carried out in order to assess the conformation of the molecules and their potential self-assembly.

Table 2.2: Summary of all interactions observed in molecular modelling. The group involved in an interaction are denoted in bold, → symbolises a H-bond, & means "binds to both".

Complex	Interactions
O-1$_A$2$_B$G(1) – O-1$_A$2$_B$G(2)	O-1$_A$2$_B$**G**(1) → O-1$_A$**2$_B$**G(1) → O-**1$_A$**2$_B$G(2) → O-1$_A$2$_B$-**CONH**-G(1); O-1$_A$2$_B$**G**(2) → O-1$_A$**2$_B$**G(2) → O-**1$_A$**2$_B$G(1) → O-1$_A$2$_B$-**CONH**-G(2).
N-1$_B$2$_A$G(1) – N-1$_B$2$_A$G(2)	N-**1$_B$**2$_A$G(1) → N-1$_B$**2$_A$**G(2); N-**1$_B$**2$_A$G(2) → N-1$_B$**2$_A$**G(1).
N-1$_B$2$_{iBu}$3$_A$G(1) – N-1$_B$2$_{iBu}$3$_A$G(2)	N-1$_B$2$_{iBu}$**3$_A$**G(1) → N-**1$_B$**2$_{iBu}$3$_A$G(1) → N-1$_B$2$_{iBu}$3$_A$**G**(2) → N-**NH**-1$_B$2$_{iBu}$3$_A$G(1); N-1$_B$2$_{iBu}$3$_A$**G**(1) → N-**1$_B$**2$_{iBu}$3$_A$G(2) → N-1$_B$2$_{iBu}$-**CONH**-**3$_A$**G(2); N-1$_B$2$_{iBu}$3$_A$-**CONH**-G(1) → N-1$_B$2$_{iBu}$**3$_A$**G(2).
N-1$_B$2$_{iBu}$3$_A$G(1) – N-1$_A$2$_{iBu}$3$_B$G(2)	N-1$_B$2$_{iBu}$3$_A$**G**(1) → N-**NH**-1$_B$2$_{iBu}$3$_A$G(1); N-1$_B$2$_{iBu}$3$_A$**G**(1) → N-1$_A$2$_{iBu}$**3$_B$**G(2) → N-**1$_A$**2$_{iBu}$3$_B$G(2) → N-**1$_B$**2$_{iBu}$3$_A$G(1) → N-1$_B$2$_{iBu}$3$_A$**G**(1); N-**1$_A$**2$_{iBu}$3$_B$G(2) → N-**NH**-1$_A$2$_{iBu}$3$_B$G(2); N-1$_A$2$_{iBu}$3$_B$**G**(2) → N-1$_B$2$_{iBu}$3$_A$-**CONH**-G(1).
N-1$_A$2$_{iBu}$3$_B$G(1) – N-1$_A$2$_{iBu}$3$_B$G(2)	N-1$_A$2$_{iBu}$**3$_B$**G(1) & N-1$_A$-**CONH**-2$_{iBu}$3$_B$G(2) → N-1$_A$2$_{iBu}$-**CONH**-3$_B$G(2) → N-1$_A$2$_{iBu}$**3$_B$**G(2) → N-**1$_A$**2$_{iBu}$3$_B$G(1) → N-1$_A$2$_{iBu}$**3$_B$**G(1); N-1$_A$-**CONH**-2$_{iBu}$3$_B$G(2) → N-**1$_A$**2$_{iBu}$3$_B$G(2) N-1$_A$2$_{iBu}$3$_B$**G**(1) → N-**NH**-1$_A$2$_{iBu}$3$_B$G(1).

2.4. NMR studies

2.4.1. General considerations

Choice of solvent. Solvents commonly used for H-bonding studies are chloroform[184] or acetonitrile.[187] The limited solubility of the dimers and trimers in such solvents made the study impossible in those conditions. It was first decided to use pyridine-d$_5$ as it has been proven to replicate results obtained in polar solvents such as MeOD without disrupting the interactions.[188] Unfortunately the precipitation of O-alkylated compounds was observed in pyridine-d$_5$ on a few hours timescale, which meant it was not possible to acquire a good quality 2D spectra. For this reason, DMSO-d$_6$ was chosen for O-alkylated dimers as it was shown to be pertinent for H-bonding studies.[189] Although some aggregation was observed for O-1$_A$2$_B$G (**50**) at the highest concentrations (40 mM and 30 mM) even in DMSO-d$_6$, it did not affect the results (this was checked by comparing two protons spectra performed after a few minutes or after 24 hours). By extension, DMSO-d$_6$ was also the solvent of choice for N-alkylated oligomers in order to compare the behaviour of the different series of compounds.

Nomenclature of the molecules. In order to describe the NMR spectra, the same nomenclature as for characterisation is used. Any relevant hydrogen is annotated, and the full nomenclature is further explained in Chapter 5.

2.4.2. Conformational analysis

The first objective of the study was to establish a conformational analysis of the compounds. NOESY and/or ROESY NMR experiments were performed at 500 MHz (Figure 2.11). Although no correlation was observed between the two side-chains of $O\text{-}1_B2_AG$ (**50**, Figure 2.11a), the ROESY showed an interaction between 2-NH and 1-H2, but not 2-NH and 1-H6 (Figure 2.11d), showing that the conformation of the dimer was *syn*.
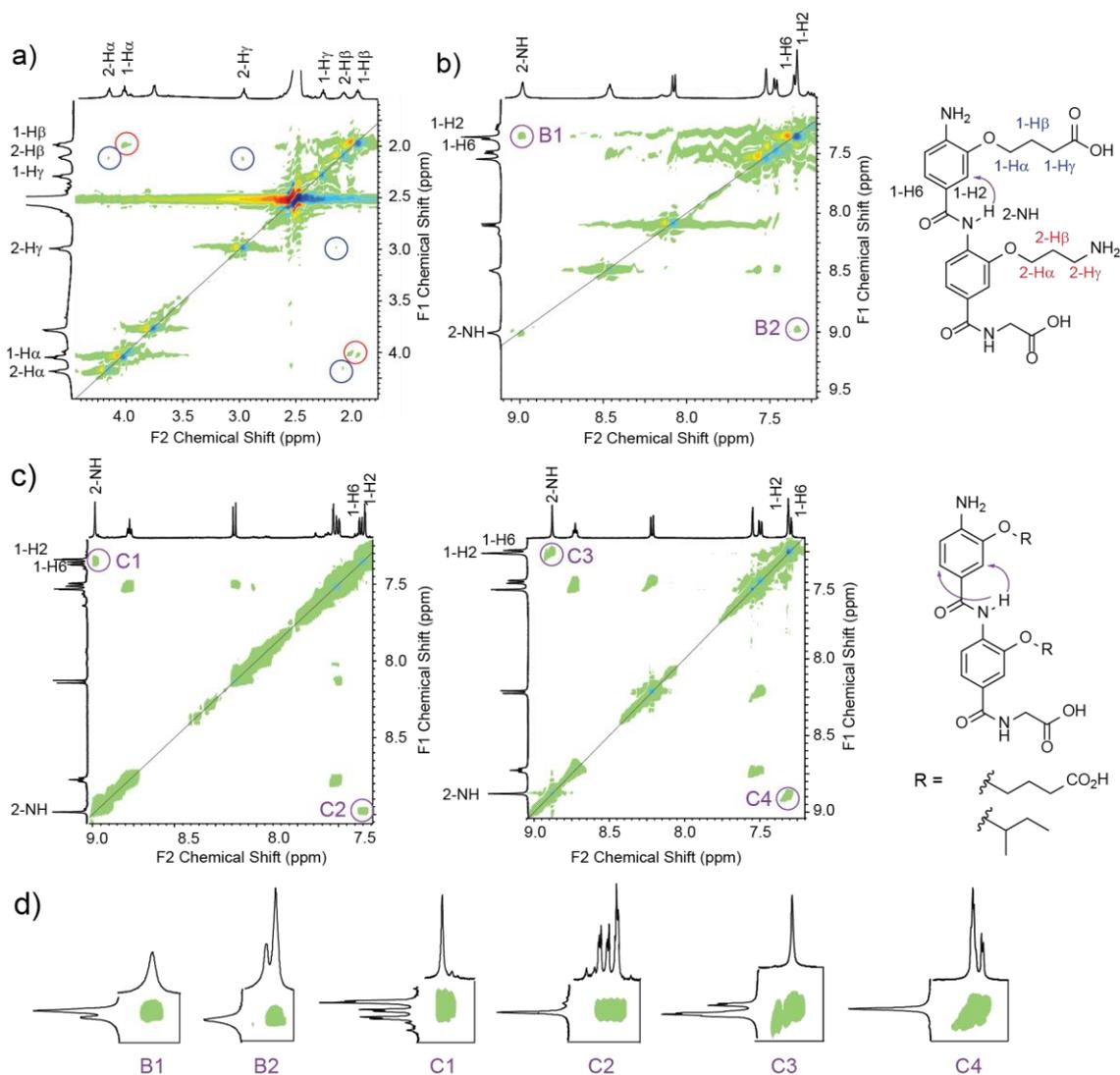
Figure 2.11: 2D NMR at 40 mM in DMSO-d₆, 500 MHz, showing a) the absence of correlation between $O\text{-}1_A2_BG$ (**50**) side-chains by ROESY, b) the correlation between 2-NH and 1-H2 only for $O\text{-}1_A2_BG$ (**50**), and c) the correlation between 2-NH and both 1-H2 and 1-H6 by NOESY for $O\text{-}1_A2_AG$ (**48**, left) and $O\text{-}1_{sBu}2_{sBu}G$ (**57**, right). d) Zoom on the relevant peaks.

Interestingly, the same correlation for O-1_A2_AG (**48**) and O-$1_{sBu}2_{sBu}G$ (**57**) is observed between 2-NH and both 1-H2 and 1-H6, showing a free rotation around the amide bond (Figure 2.11c and d).

2D NMR were also performed to establish the conformation of N-1_B2_AG (**55**, Figure 2.12). N-alkylated compounds are known to prefer a *cis* conformation,[116] and this was observed by correlation between 2-H3 and 1-H2 (Figure 2.12, purple circles), only possible in *cis*. A correlation between 1-H2 and 2-Hα was also observed (Figure 2.12, orange circles): this correlation is possible in the *trans* conformation, suggesting the dimer could be in equilibrium between the two states. Nevertheless, this signal could also be observed in a *cis* conformation, as the two aromatic rings slightly overlap. The only conclusion that can be drawn from this data is that the *cis* conformation is preferred.

In the case of N-1_B2_AG (**55**), the *cis* conformation is stabilised by the possible H-bond between the amine side-chain and the glycine: this interaction would not be possible in the case of N-1_A2_BG (**56**), and could make the *trans* conformation more stable, as suggested by the molecular modelling (Appendix I).
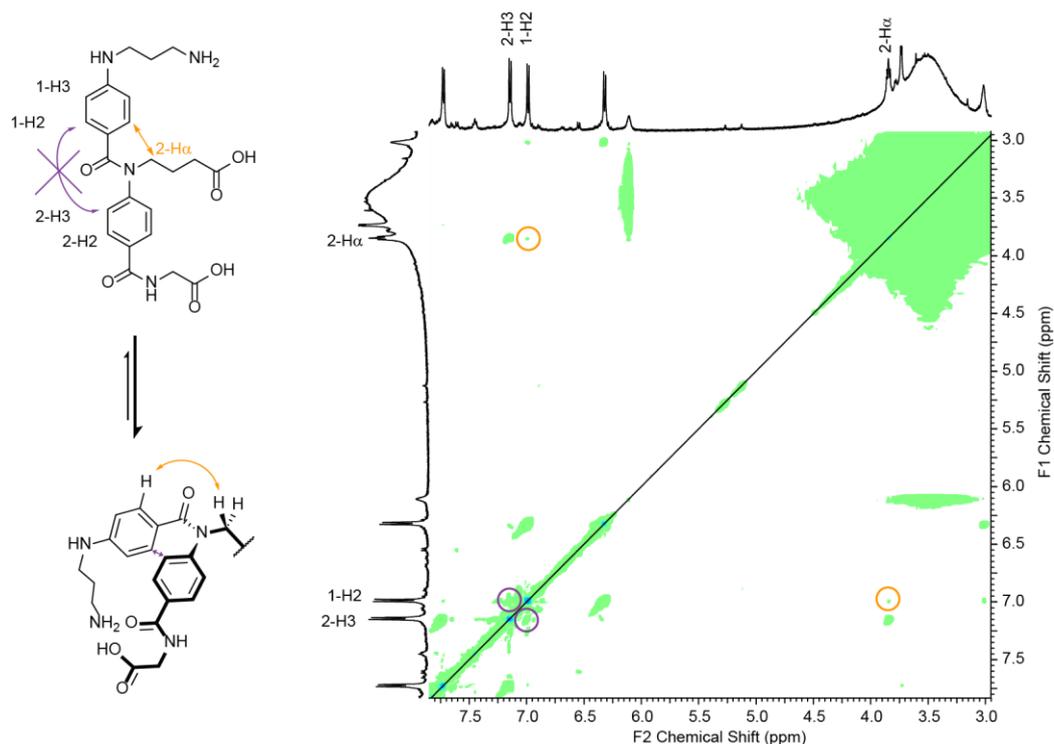
Figure 2.12: NOESY of N-1_B2_AG (**55**) 40 mM in DMSO-d_6, 500 MHz, highlighting the preference for a *cis* conformation.

For the *N*-alkylated trimers, the NOESY experiments showed a large predominance of the *cis* conformation (Figure 2.13). Indeed, although the aromatic region was quite noisy, the correlation between 3-H2 and 1-H3 was clear and out of the noise. In particular, no correlation was observed

between 3-Hα and 2-H2, or 2-Hα and 1-H2 (Appendix II), suggesting the conformation was fully *cis*.
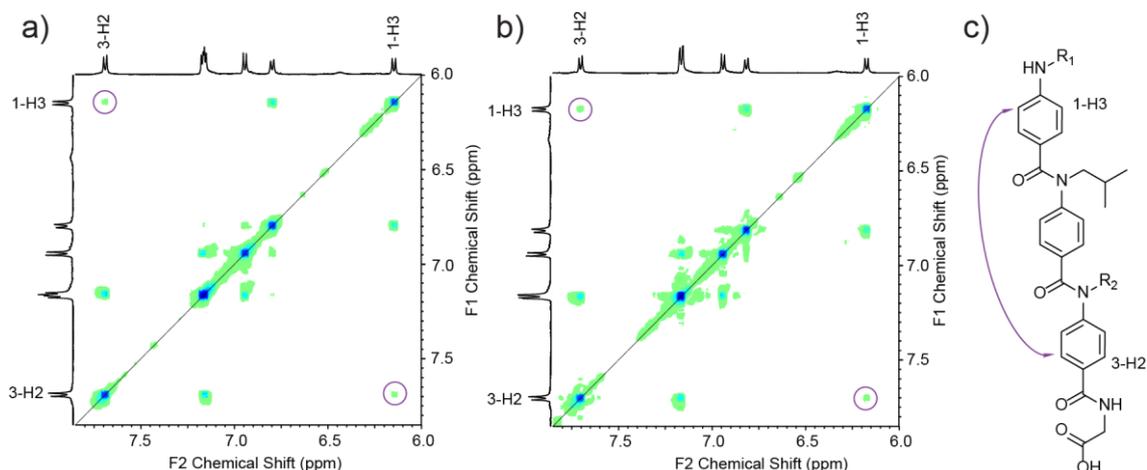
Figure 2.13: NOESY of a) N-$1_A2_{iBu}3_B$G (**54**) and b) N-$1_B2_{iBu}3_A$G (**53**), 40 mM in DMSO-d_6, 500 MHz, showing the preferred *cis* conformation (purple circle), as shown by c) the chemical structure.

2.4.3. Study of the self-assembly

2.4.3.1. Solution phase

In order to evaluate whether the designed foldamers were capable of self-assembling, NMR experiments were performed on a dilution scale from 40 mM to 1 mM. Several shifts were observed for O-1_A2_BG (**50**, Figure 2.14a), indicating changes in the electron density and therefore nuclear shielding of the compound. In particular, 1-Hγ shifted from 2.29 ppm to 2.23 ppm upon dilution from 40 mM to 1 mM (Figure 2.14a, red), showing a shielding effect therefore an increase in the local electron density, correlating with a case where the carboxylate ion is involved in a H-bonds at high concentration but not at low concentration. This matches the modelling predictions that showed the carboxylate group involved in one intramolecular H-bond on its own, and two intermolecular H-bond when self-assembled. The shift from 2-Hγ from 2.99 ppm to 2.97 ppm upon dilution (Figure 2.14a, red) also followed the prediction: in the dimeric state, the ammonium ion was surrounded by two carboxylate groups, one of which was also involved in another H-bond with 3-NH, whereas in a monomeric state, the ammonium was bound to two carboxylate groups in an exclusive manner, providing a plausible explanation for the shielding effect. The opposite shift trends observed for 3-NH, from 8.48 ppm to 8.51 ppm (Figure 2.14a, purple), was in accordance with that reasoning, as a de-shielding effect would occur upon binding with the carboxylate.
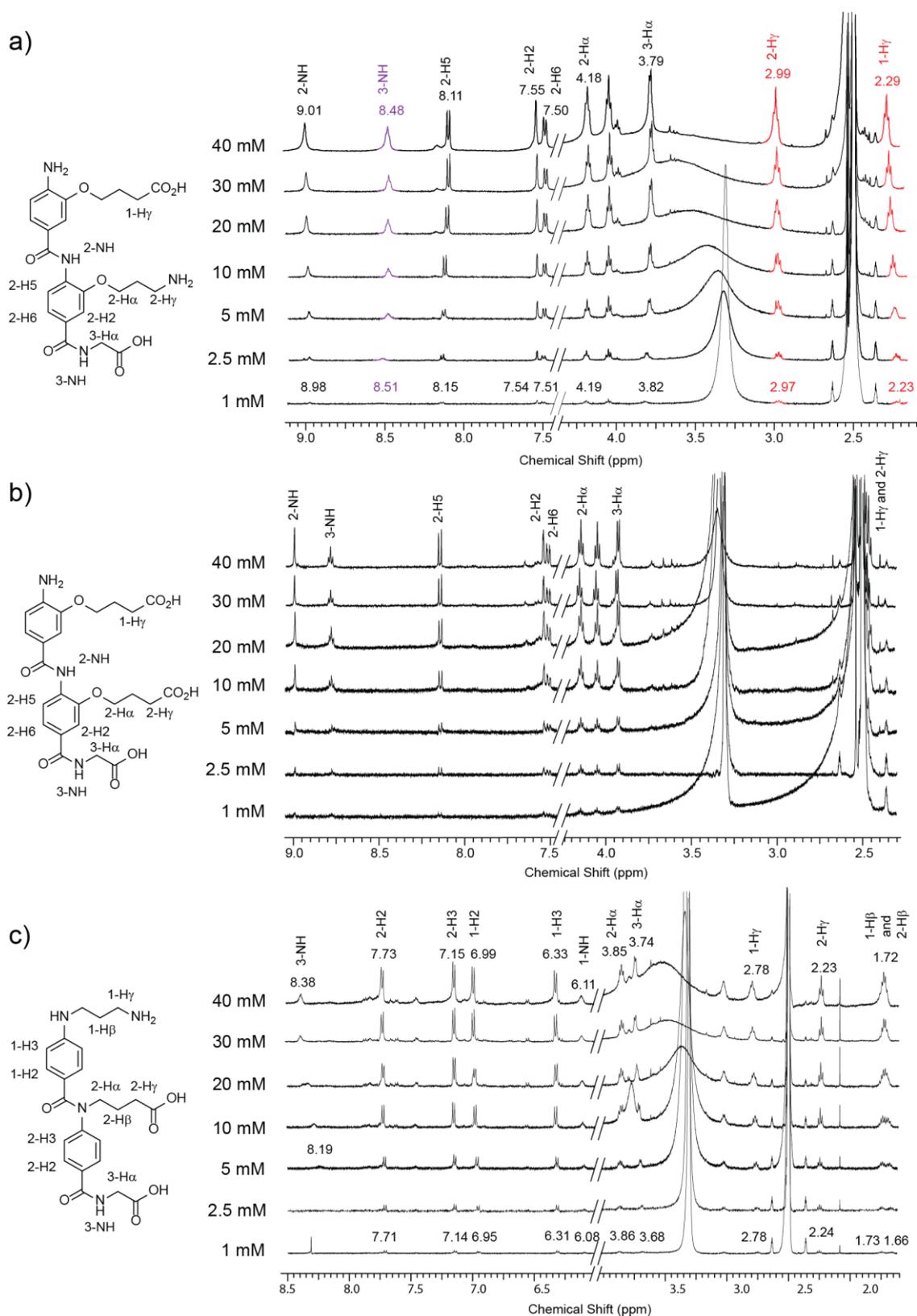
Figure 2.14: NMR dilutions in DMSO-d$_6$, at 500 MHz, of compounds a) O-1$_A$2$_B$G, b) O-1$_A$2$_A$G, and c) N-1$_B$2$_A$G, from 40 mM to 1 mM. Shifting signals are indicated.

While diluting the samples, the concentration of water in DMSO-d$_6$ remains constant, and therefore the ratio water:dimer increased. As a shift in the water peak was also observed during the dilutions, it was possible that the results could be due to interactions with water and thus not

be related to self-assembly. To assess the situation, the dilutions were also performed on O-1_A2_AG (**48**, Figure 2.14b), which cannot self-assemble, but can interact with water due to the nature of the side-chains, and O-$1_{sBu}2_{sBu}G$ (**57**, Appendix III), whose hydrophobic side-chains cannot interact with water. In both cases, no peak shifted, proving the shifts observed for O-1_A2_BG (**50**) were not due to the change of water:dimer ratio. Moreover, the water peak did not move for the controls, whereas a shift and a broadening was observed for O-1_A2_BG (**50**). A possible explanation for this change is that the self-assembly happening at higher concentration modifies the available surface for hydrogen exchange between the dimers and the water, whereas without self-assembly, this surface is constant, hence the sharp peak for O-1_A2_AG (**48**) and O-$1_{sBu}2_{sBu}G$ (**57**).

To confirm this hypothesis, DOSY NMR was performed on O-1_A2_BG (**50**) at concentrations above and below the shifts, *i.e.* 25 mM and 2.5 mM (Figure 2.15). Changes in diffusion were observed for the signals corresponding to the side-chains, and not for the aromatic peaks: this also suggested a change in behaviour in the side-chain upon dilution, although the error was too significant to extrapolate any numerical values.
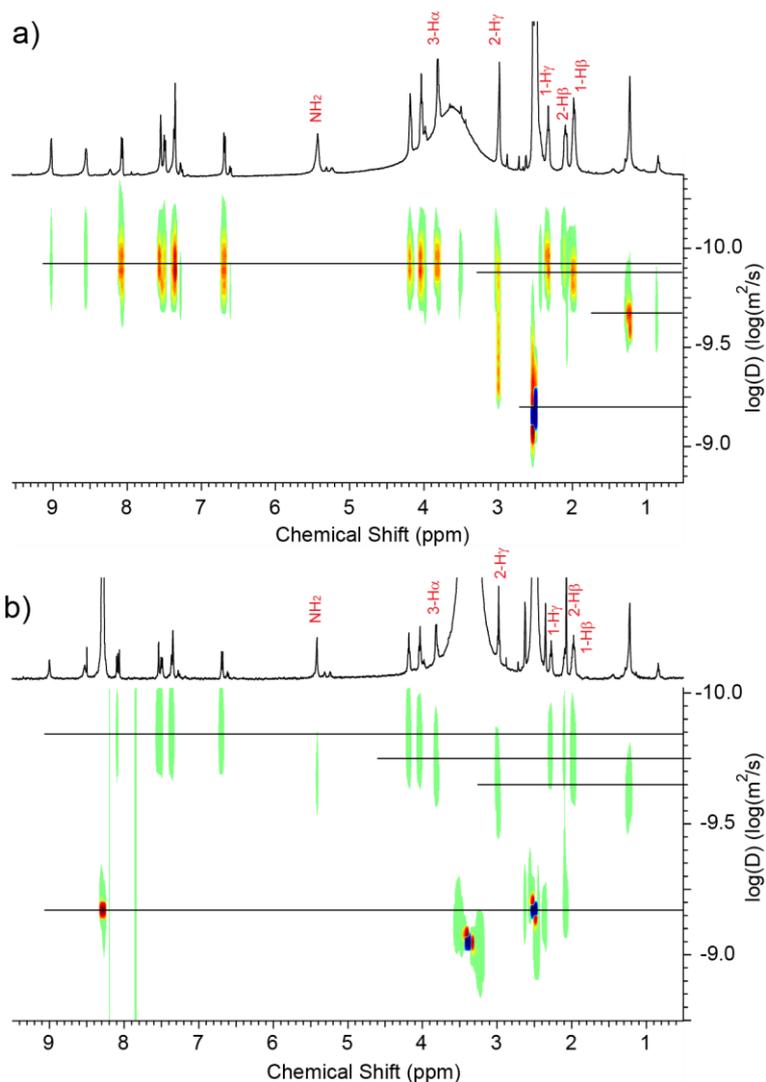
Figure 2.15: DOSY of O-1_A2_BG (**50**) at a) 25 mM and b) 2.5 mM, in DMSO-d_6. Shifting peaks are highlighted in red.

45

The dilution experiments were performed on N-1_B2_AG (**55**, Figure 2.14c). Several shifts were observed again, suggesting the same behaviour as for O-1_A2_BG (**50**).

Considering the conformational results for trimers N-$1_B2_{iBu}3_AG$ (**53**) and N-$1_A2_{iBu}3_BG$ (**54**), dilution experiments were not performed. Nevertheless, the compounds were mixed in a 1:1 ratio in order to establish a possible interaction between the trimers (Appendix II). As both trimers had very similar spectra, the mixture did not involve changes significant enough to draw any conclusion. It is possible that both trimers co-exist in solution without interacting with each other. To confirm this hypothesis, it would be necessary to synthesize different trimers, such as N-$1_A2_A3_AG$ and N-$1_B2_B3_BG$, or N-$1_A2_{iBu}3_AG$ and N-$1_B2_{iBu}3_BG$, which was not done here due to time considerations.

2.4.3.2. Solid phase-state conformation

Attempts to grow crystals were performed in order to study the solid-state conformation of the compounds. Out of all, only O-1_A2_BG (**50**), prone to precipitate even in DMSO, formed a cloudy crystal-like solid from dimethyl sulfoxide to chloroform. Several solvent systems were used, and either mixing or slow evaporation were attempted. The conditions are described on Table 2.3.

Table 2.3: Different conditions attempted for growing crystals of O-1_A2_BG (**50**).

Solvent	Method	Result
Methanol	Mixing (heating)	Nothing observed – fully soluble compound
	Slow Evaporation	Nothing observed – fully soluble compound
Ethyl Acetate	Mixing	Solidification/crystallisation on the glass– no single crystal obtained
	Slow evaporation	Solidification/crystallisation on the glass– no single crystal obtained
Diethyl Ether	Mixing	Solidification/crystallisation on the glass– no single crystal obtained
	Slow evaporation	Solidification/crystallisation on the glass – no single crystal obtained
Hexane	Mixing	Precipitate – could not be solubilised
	Slow evaporation	Solidification/crystallisation on the glass– no single crystal obtained
Chloroform	Mixing	Cloudy crystal-like solid, no single crystal
	Slow Evaporation	Small cloudy crystal-like solid

From these attempts, chloroform use was the most promising. Although all these setups were done from the NMR samples in DMSO, another solvent for the starting material can be used. Amongst all systems used, it was possible to solubilise the dimer upon gentle heating in methanol (water bath), and no precipitation or crystallisation was observed: a methanol/chloroform system will therefore be investigated for further attempts to growing crystals.

2.5. Conclusions and future work

A library of *O*- and *N*-alkylated hydrophilic oligomers was designed. The building blocks were prepared, involving the design and synthesis of a new carboxylic acid functionalised building block for each series of compounds. Because of the nature of the side-chains, the synthesis and purification was less trivial than for usual hydrophobic compounds, and therefore some cyclisation or degradation of the molecules were observed. Finally, 6 compounds were obtained: 3 *O*-alkylated dimers, 1 *N*-alkylated dimer, and 2 *N*-alkylated trimers. NMR studies lead to the following conclusions:

- The conformation of *O*-alkylated dimers can be controlled by carefully designing the side-chains: the free rotation around the amide bond can be stopped if a stabilising bridge happens between the side-chains of the compound.
- The *N*-alkylated compounds synthesised preferred the *cis* conformation, regardless of the nature of the side chains. In particular, for *N*-alkylated trimers equipped with a central hydrophobic group, the stabilising bridge cannot happen in a *trans* conformation because the distance between side-chains is smaller in the *cis* conformation. Although some combinations might generate an equilibrium, as suggested by molecular modelling, this scaffold is generally believed to be too flexible to impose a full *trans* conformation.
- Dilution experiments were performed on the dimers: only the ones designed to self-assemble exhibited shifts during the dilution, in particular the protons near the side-chain groups. The dilutions were performed on control molecules in order to assess that none of those results were due to unspecific interaction with the water present in the DMSO-d_6, proving that the molecules change state while diluting, suggesting some self-assembly happened at high concentration.

The following work will enable the completion of the results obtained in this chapter:

- Complete the initial library of oligoamides and perform the same NMR experiment in order to get more information and confirm the obtained results. In particular, the modelling for *N*-1$_A$2$_B$G (**56**) suggested a preference for a *trans* conformation, which could easily be determined by NMR and would confirm whether the *cis* conformation is always preferred or it is possible to force a *trans* conformation upon that class of compounds.
- Design and synthesise *N*-alkylated trimers equipped with the same charge, such as *N*-1$_A$2$_A$3$_A$G and *N*-1$_B$2$_B$3$_B$G, or *N*-1$_A$2$_{iBu}$3$_A$G and *N*-1$_B$2$_{iBu}$3$_B$G, which should not allow stabilising interactions in the *cis* conformation, as well as alternating side-chains to enable closer side-chain/side-chain interaction stabilising the *trans* conformation, such as *N*-1$_A$2$_B$3$_A$G or *N*-1$_B$2$_A$3$_B$G. Investigating different types of resins to have an amine at the bottom of the oligomers will also enable to avoid any kind of interaction between a side-chain and the resin.

- New crystal growth assays will be performed, changing the combinations of solvents, and varying the temperatures.
- Finally, longer molecules will be prepared in order to mimic a full coiled coil. Considering the high flexibility of *N*-alkylated series, it is more likely to achieve this goal with *O*-alkylated oligomers equipped with alternated charges on the side-chains in order to conserve a *cis* conformation.

Chapter 3. Using Affimers to identify high affinity peptide/foldamer interactions

The work reported in this chapter formed the basis of the following publication:

I. Arrata, A. Barnard, D. C. Tomlinson, A. J. Wilson, *Chem. Commun.*, **2017**, 53, 2834-2837[190]

3.1. Introduction

3.1.1. Building proteins with non-natural foldamers

Inspired by the three dimensional compact conformations adopted by proteins (helices, sheets, loops, turns and disordered domains),[1] foldamers are non-natural polymers designed to adopt well defined secondary and tertiary or quaternary structures. The ultimate aim of foldamer synthesis is to generate 3-dimensional architectures with novel, enhanced or emergent function (Chapter 1 and Figure 3.1).[18, 19, 191] This goal aligns closely with efforts to build functional proteins *de novo* in synthetic biology.[192-198]

Figure 3.1: Bottom-up design of foldamers, and examples of biologically relevant applications: building tertiary structure, PPI modulation, foldamer-protein hybrid capable of retaining a functional site, or immobilisation for use in screening protocols.

Although a major challenge in supramolecular chemical biology,[199] considerable progress has been made in the *de novo* or bottom-up design of tertiary foldamers[74, 200, 201] whilst efforts to understand and control their dynamic topology have broadened potential applications.[50, 201-210] "Protein prosthesis",[211-213] which consists in replacing segments of protein sequence with non-natural foldamer,[57, 210-215] is an alternative strategy to the design of functional foldamers, leading to bionic proteins.[120]

A third related approach would be to exploit the potential of combinatorial biology to identify natural biomacromolecule sequences (comprised of amino-acid or nucleotide building blocks) that recognise a single synthetic foldamer.[214] In identifying compatible natural and non-natural components driven by complementary molecular recognition, such an approach could be used to identify potential biological targets of a given foldamer,[214] a potentially more rapid route to ligand discovery than the painstaking construction of libraries using synthetic foldamer assembly strategies.

3.1.2. Investigating foldamer/Affimer interactions

The Wilson group extensively explored the use of proteomimetic scaffolds to inhibit the p53/*h*DM2 interaction, using libraries of 2-*O*-, 3-*O*-[113] and *N*-alkylated[114] series (*in vitro* and in cells),[118] as well as hybrid scaffolds[117] (further discussed in Chapter 1). Given the established ability of those proteomimetics to bind to *h*DM2, the basis of the work described in this chapter was to explore the extent to which biological selection methods can generate peptide sequences for future exploitation in construction of foldamer-peptide hybrids. For this purpose, a small library of *N*-alkylated trimers, including p53 mimetics, was used in phage display screening with a non-antibody-based scaffold, termed an Affimer.[122]

3.2. Phage display screening against a small foldamer library

3.2.1. Synthesis

The current work focuses on a series of six *N*-alkylated trimers **58-63** and their biotinylated equivalents Biotin-**58** to Biotin-**63** (Figure 3.2). **59** to **63** were originally designed as p53 mimetics.[118] Trimers **58**-**63** and Biotin-**58-63** were previously synthesised by Dr. Anna Barnard,[118, 178, 215] whilst the general synthesis of *N*-alkylated building blocks is described in Chapter 2. The orthogonal functionalisation of the trimers was developed in the group by Dr Anna Barnard.[215] Equipping trimers with biotin groups had previously enabled a streptavidin-biotin interaction for biological applications, such as pull down assays.[118] On this basis, the trimers were equipped with biotin tags for immobilisation on streptavidin.
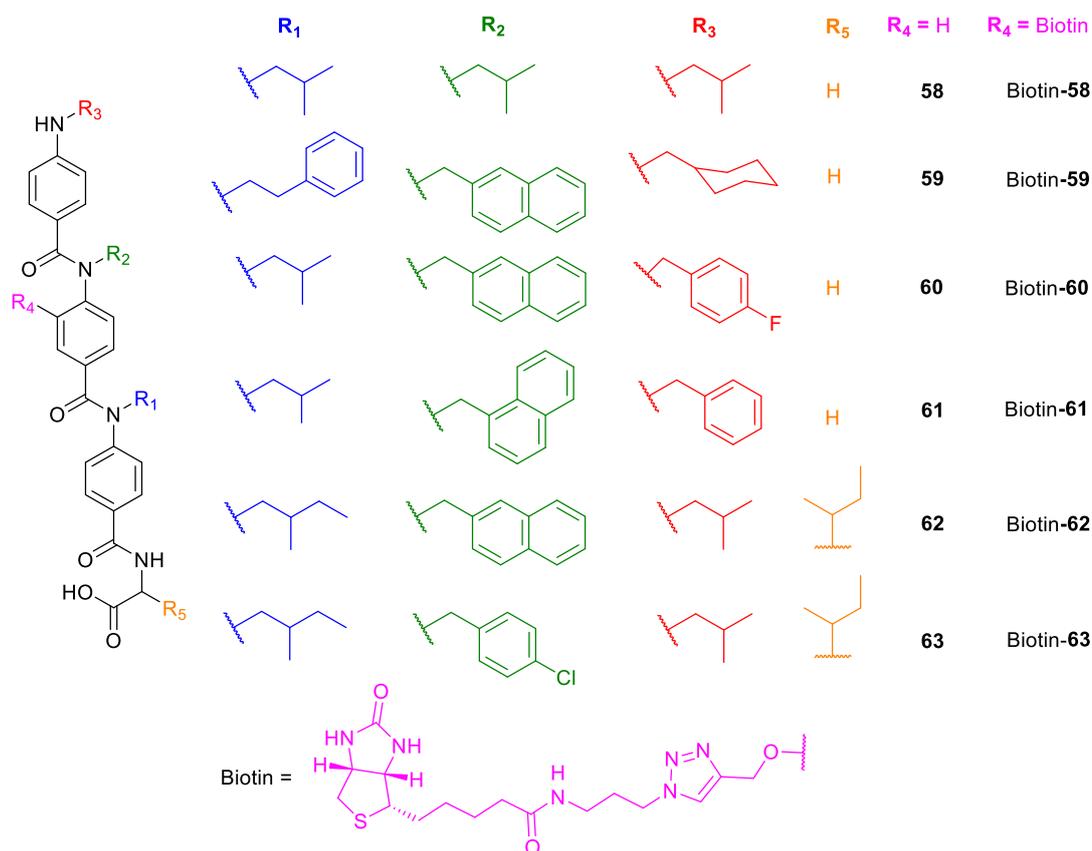
Figure 3.2: Details of the small library of *N*-alkylated trimers used for phage display.

To enable a later click-chemistry reaction, the *O*-alkyne *N*-alkylated monomer **69** (Scheme 3.1) was prepared from methyl 3-hydroxy-4-nitrobenzoate **20**. Quantitative propargylation of the phenol was followed by selective reduction of the nitro group from **65** to an amine, using tin (II) chloride. Saponification of the ester **66** led to the *O*-alkyne amino acid **67**. *N*-alkylation by reductive amination and Fmoc protection gave **69**, following the previously described procedure.[215]
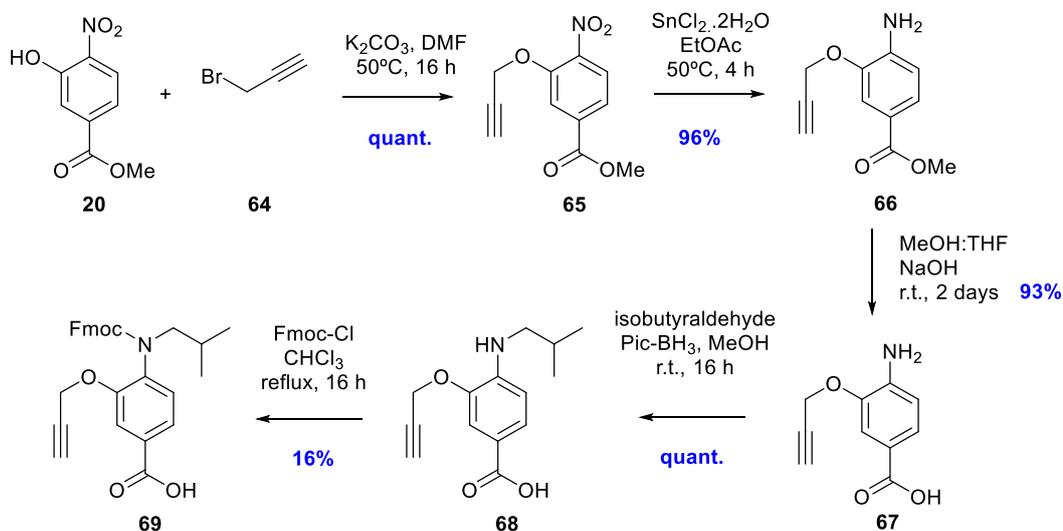
Scheme 3.1: Synthesis of the *O*-alkyne *N*-alkylated monomer **69**.

51

The trimers were prepared by SPPS on a CEM Liberty® automated microwave assisted peptide synthesizer, following the procedure previously described in the group (see Chapter 2 for discussion on trimer synthesis).[178] The samples were collected on resin after coupling: the biotin was installed on the relevant trimers by Cu(I) catalysed 'click'-chemistry using biotin azide **72**, previously synthesised in the group by Dr. Anna Barnard following the conditions described in Scheme 3.2.[118] The trimers were cleaved manually using a 50% TFA solution in dichloromethane.
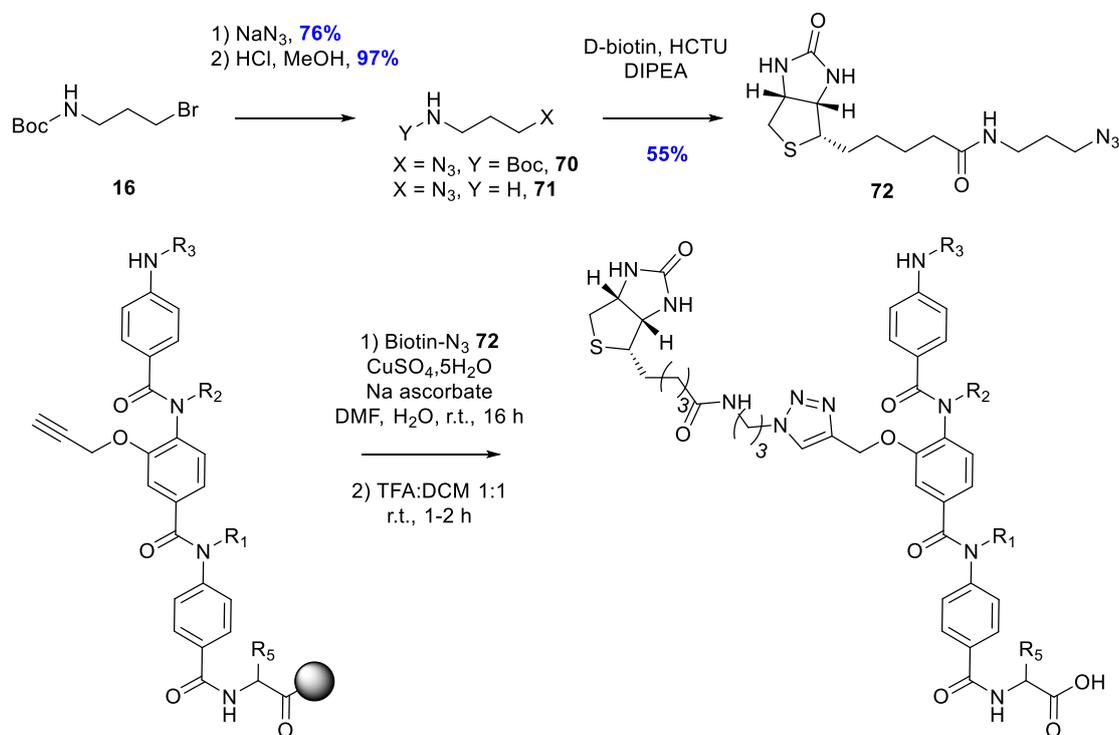
Scheme 3.2: Synthesis of the azidiobiotin **72** (top) used to introduce a biotin moiety on the trimers by click chemistry (bottom).[118]

3.2.2. Phage Display Screening using Affimers

Affimers were used in the phage display screening setting. An Affimer (Figure 3.3) is a highly thermostable artificial binding protein scaffold, based on the consensus optimisation[216, 217] of plant-derived phytocystatins,[218] comprising four β-strands, one α-helix, and two loops connecting each pairs of β-strands.[122]
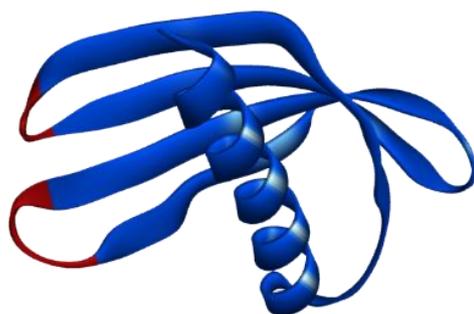
Figure 3.3: Structure of an Affimer (PDB ID: 4N6U). Insertion loops are highlighted in red.[122]

The randomisation of 9 amino acids occurs at each loop (Figure 3.3, red). Affimers have been successfully used in a number of discovery settings.[123, 219-221]

An Affimer library with a complexity of 3×10^{10} was generated and screened against the *N*-alkylated trimers Biotin-**58** to Biotin-**63**, which were immobilised on alternatively streptavidin-coated wells, neutravidin-coated wells, and streptavidin-coated magnetic beads. The phage were incubated in the presence of the foldamers, non-binding Affimers were washed away, binding Affimers were eluted and amplified. Four panning rounds were performed, and bound phage were detected by phage ELISA (Figure 3.4) using Anti-Fd-Bacteriophage-horse radish peroxidase (HRP) and SeramunBlau® fast TMB.
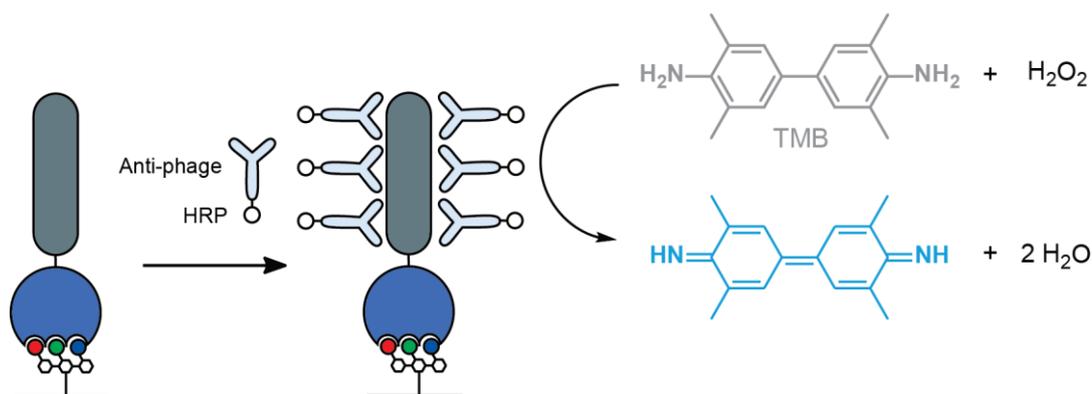
Figure 3.4: Cartoon representation of an ELISA experiment. The oxidation of TMB with hydrogen peroxide is catalysed by the HRP group covalently linked to the anti-phage.

For each trimer, 48 monoclonal Affimers were tested by ELISA against all six biotinylated trimer, in order to assess their cross-selectivity (Figure 3.5). For Biotin-**58**, there were 25 binders, 12 of which were fully selective for Biotin-**58**; for Biotin-**59**, 41 Affimers, 30 of which were selective; 7 Affimers for Biotin-**60**, 2 of which were selective; 19 Affimers for Biotin-**61**, none of which were selective; 18 Affimers for Biotin-**62**, none of which were selective; and 25 Affimers for Biotin-**63**, 4 of which were selective.

The results thus yielded high numbers of selective binders for Biotin-**58** and Biotin-**59**, while Biotin-**60** and -**61** generated fewer binders and exhibited very poor selectivity. In comparison to Biotin-**58**, Biotin-**62** and -**63** raised a good number of Affimers, nevertheless the selectivity was not as good. This is not completely unsurprising, considering the high similarities between the structures, as **59** to **63** were all designed as p53 mimetics. In particular, Biotin-**62** and Biotin-**63** only differ by one Ar to Cl replacement. The fact that all the structures are highly hydrophobic is also a possible reason for cross-binding. Although the exact reasons for the difference in results observed for each compound are unclear, it is believed that the absence of any aromatic side-chain on Biotin-**58**, and the presence of a cyclohexane side chain on Biotin-**59** may account for the higher cross-selectivity. On those considerations, it was decided to focus on the Affimers raised against Biotin-**58** and Biotin-**59** for the rest of the study.
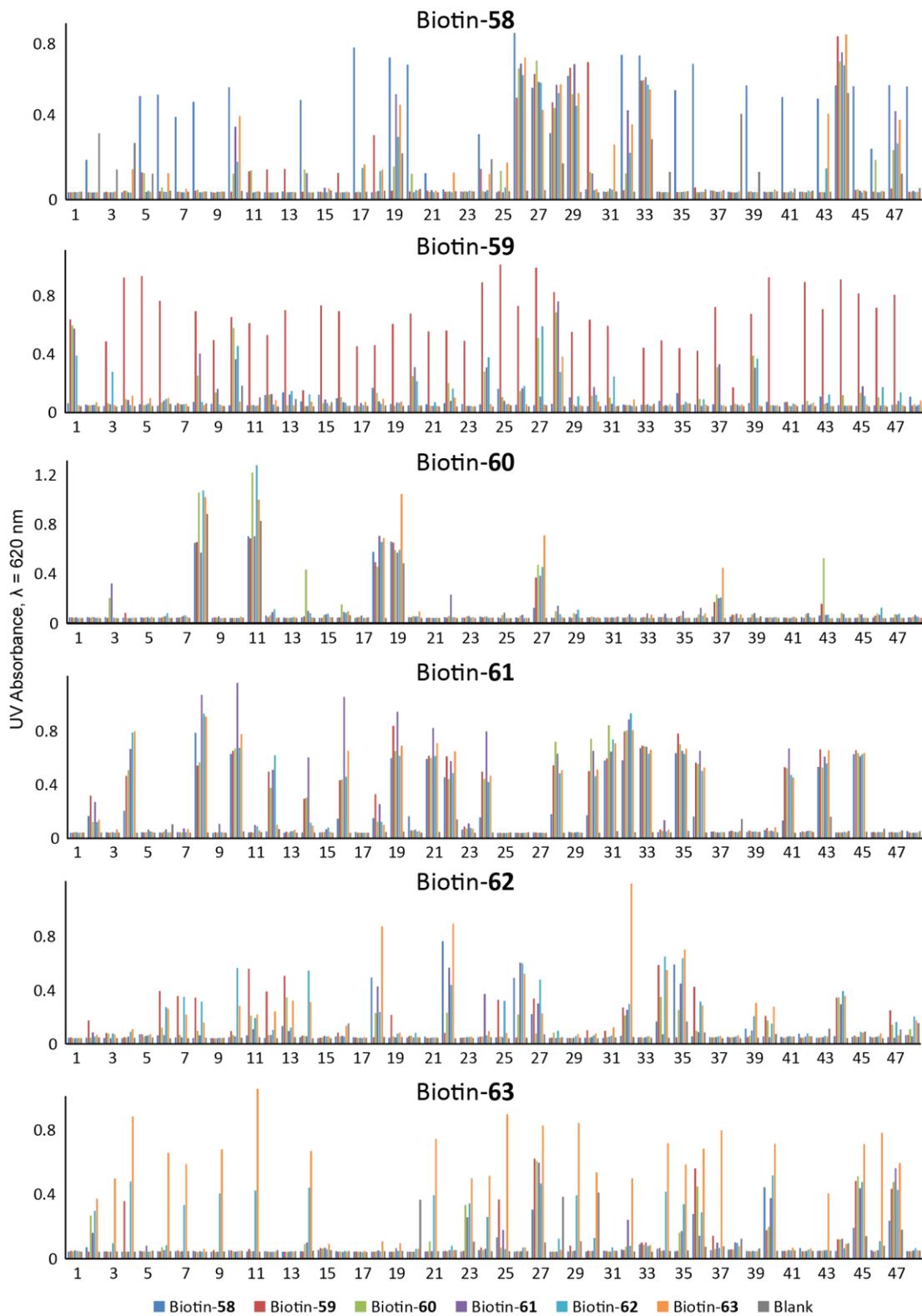
Figure 3.5: ELISA readings for 48 monoclonal Affimers for Biotin-**58** to Biotin-**63**

For Biotin-**58** and -**59**, Affimers showing differential values between test and negative control by ELISA were sequenced using a T7P primer – mixed colonies were removed from the list. Each Affimer was named using the format X-AFY, where X is the number of the trimer and Y the number of the Affimer out of 48. The data obtained is summarised in Table 3.1.

Interestingly, no pattern of amino acids stands out. Nevertheless, on average, non-polar or polar neutral amino acids tends to be predominant in the Affimer sequences, and polar positive outnumber polar negative amino acids. This may be explained by the presence of the carboxylic acid at the bottom of the foldamer, which is deprotonated in the conditions of the selection (pH 7).

Table 3.1: Summary of the Affimer sequences obtained from screening against Biotin-**58** and Biotin-**59** (red: non-selective, green: selective). The amino acids are colour-coded according to the following principle: blue: polar positive, red: polar negative, green: polar neutral, grey: non-polar aliphatic, purple: non-polar aromatic, orange: proline and glycine.

Biotin-58		
AF	**Sequence 1**	**Sequence 2**
26, 27, 29	P H R N S L V S D	Y R M A Y G F S W
28	H P V K P Q Y P T	Y K R W G I Y S W
10, 19, 32, 47	V D W Y G P V Q A	S M M T A R H N I
43	H S Q W F M Y P G	P A T V D E I M A
46	H W A H W S G D A	T Y M E E M N L N
5, 6, 39	M D Q P W W G H I	E N E G W P H L W
7	Q A W V H V H Y I	Y N G P A Y K I V
8, 14, 41	Q W V H F S G D A	Q P P Q T Y N M R
17, 36	Y V Q N E D Y F Y Q	F L R Y F P H H Y
20, 45, 48	Y Y E A W W F P I	H N E K S D H N E
35	R V Q D I Y P T Y	N A T G T T W M Q
Biotin-59		
AF	**Sequence 1**	**Sequence 2**
1	T V S T W G G D F	A N I K Q D L E M
24	S V Y T W G G P F	P I S E Q R F L P
4, 5, 25, 42	F G P R Q Y F H H	N G V P H M S M L
13, 19, 23, 29, 30, 33-35	S G Y H K D T F A	N V Y Y H N M Q I
6, 12	R G I R E W T H H	P S Q P H L H L Y
11, 45	P D K S Q P W W P	A N M K Q L Y H E
15	N G P G R H H R H	P T Y S Q W Y M M
18	W N N V G S T L M	Y K W I P W P Y I
46	R A E T W D G P W	N R S T P W D G F

For each trimer, two selective and one non-selective Affimer were selected for further study: **58**-AF8, **58**-AF17, **58**-AF26 and **59**-AF1, **59**-AF23, **59**-AF25. The choice was motivated by high occurrence (the most occurring binders were likely to be the best binders since they were selected relatively more than the others) and difference in sequences (the more varied the sequences, the more likely they were to exhibit different behaviour and deliver different results). In particular, the choice of **58**-AF17 was motivated by the fact that its first loop counts 10 amino acids instead of 9 due to an insertion, which occurs naturally but is still a relatively rare phenomenon. The

selected Affimers were subcloned and expressed as his-tagged proteins as follows:[122] from the phagemid vector, the Affimer DNA was PCR-amplified, digested with NheI and NotI restriction enzymes and ligated into a pET11a vector. The DNA was then transformed into XL1-Blue supercompetent *E.Coli* cells and the plasmid DNA was extracted and transformed into BL21 Star (DE3) *E.Coli* cells for expression. The Affimers were finally purified on a Ni-NTA resin. SDS-Page gel was performed to assess the purity of the proteins before using in assays.

To establish the affinity of interaction between foldamer and Affimer, a direct ELISA based titration of the Affimers was performed against immobilised trimer; here, anti-6X his-tag HRP and SeramunBlau® fast TMB were used for detection. The dose-response curves were fitted using a logistic model on Origin 7 to obtain EC_{50} values of: **58**-AF8: $3.8 \pm 0.9\,\mu M$; **58**-AF17: $5 \pm 1\,\mu M$; **58**-AF26: $0.19 \pm 0.02\,\mu M$; **59**-AF1: $9 \pm 1\,\mu M$; **59**-AF23: $3.3 \pm 0.8\,\mu M$; **59**-AF25: $0.98 \pm 0.08\,\mu M$. The results are summarised in Figure 3.6.
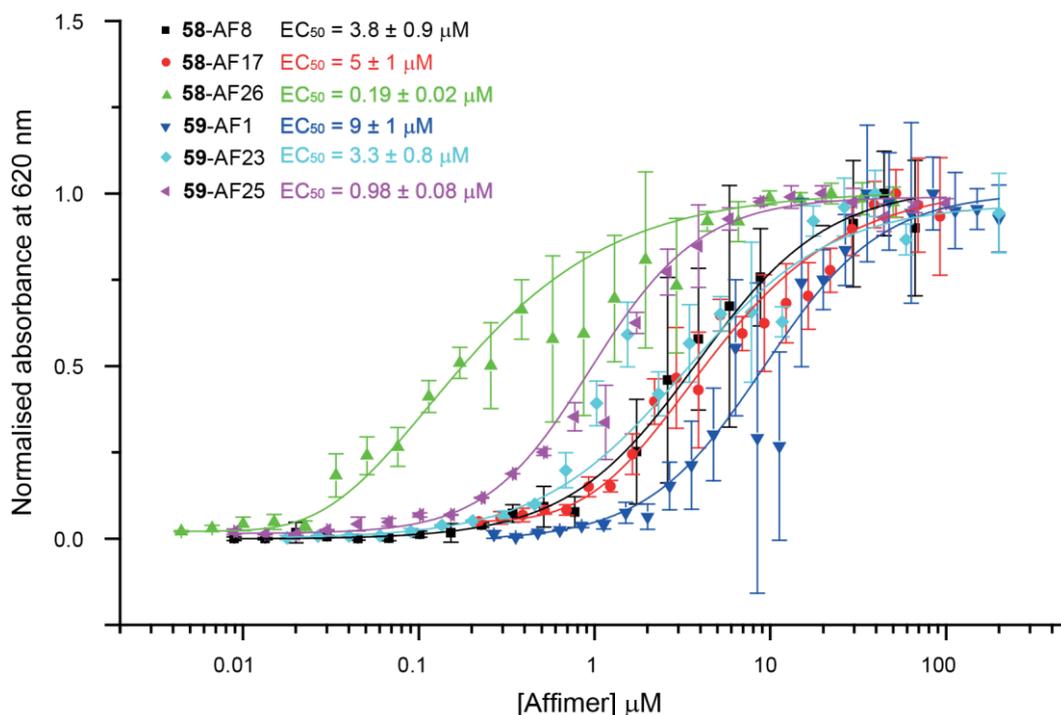
Figure 3.6: Normalised absorbance and EC_{50} values extrapolated from the ELISA direct titration for **58**-AF8 (black), **58**-AF17 (red), **58**-AF26 (green), **59**-AF1 (blue), **59**-AF23 (cyan), and **59-**AF25 (magenta), n = 3.

Based on the results of the direct ELISA, concentrations corresponding to the beginning of the top plateau were chosen for further analysis: 10 μM for **58**-AF8 and **58**-AF17, 0.3 μM for **58**-AF26, 20 μM for **59**-AF1, 2 μM for **59**-AF23 and 5 μM for **59**-AF25. First, an equimolar competition assay was realised where the competitors were the non-biotinylated trimer (**58** or **59**), the biotinylated trimer (Biotin-**58** or Biotin-**59**), streptavidin only (Str.), or a complex of streptavidin and biotinylated compound (Str-Biotin-**58** or Str-Biotin-**59**). The normalised average of four measurements (n = 3 for each) are shown in Figure 3.7.
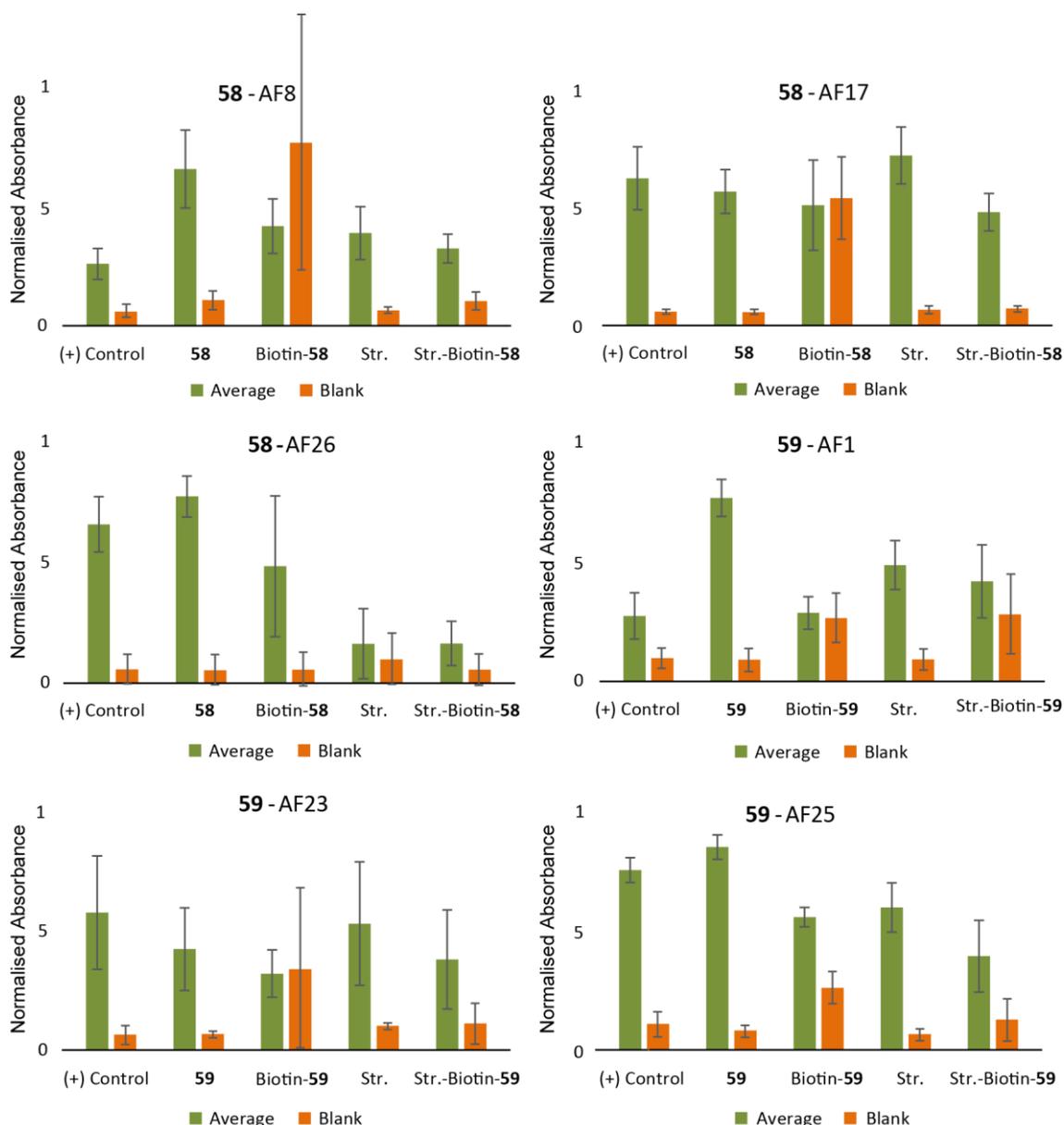
Figure 3.7: Competition ELISA on a 1:1 ratio of Affimer and competitor. The results are shown as a normalised average of four different measurements, with n = 3. Concentrations of the Affimers: [**58**-AF8] = 10 μM, [**58**-AF17] = 10 μM, [**58**-AF26] = 0.3 μM, [**59**-AF1] = 20 μM, [**59**-AF23] = 2 μM and [**59**-AF25] = 5 μM.

The general trend observed for all the Affimers was a decrease of signal from non-biotinylated competition to biotinylated competition. Even though the positive controls did not allow determination of a proper reference for **58**-AF8 and **59**-AF1 despite several repeats and for a reason that could not be identified, the presence of the biotin on the foldamer seemed to enhance the binding affinity towards the Affimer compared to a foldamer without biotin. The same observation was made in the case of **59**-AF23 and **59**-AF25. The data for **58**-AF17 and **58**-AF26 did not allow any conclusion to be drawn although streptavidin also appeared to compete for **58**-AF26 indicating strong binding between the two, which correlates well with the initial screening results where **58**-AF26 showed high binding affinity but no selectivity.

When competing with biotinylated trimer, an increase of the blank signal was observed for most Affimers: this is believed to be the result of some Affimer-trimer sticking to the plate during the incubation, despite pre-blocking the blank wells with excess of biotin.

Further analyses using 10 and 100-fold excess of competitor, focused only on **58**, **59**, Biotin-**58** and Biotin-**59** (Figure 3.8). The concentrations were adjusted in order to reduce the error bars, so that [**59**-AF23] = 5 µM and [**59**-AF25] = 10 µM.
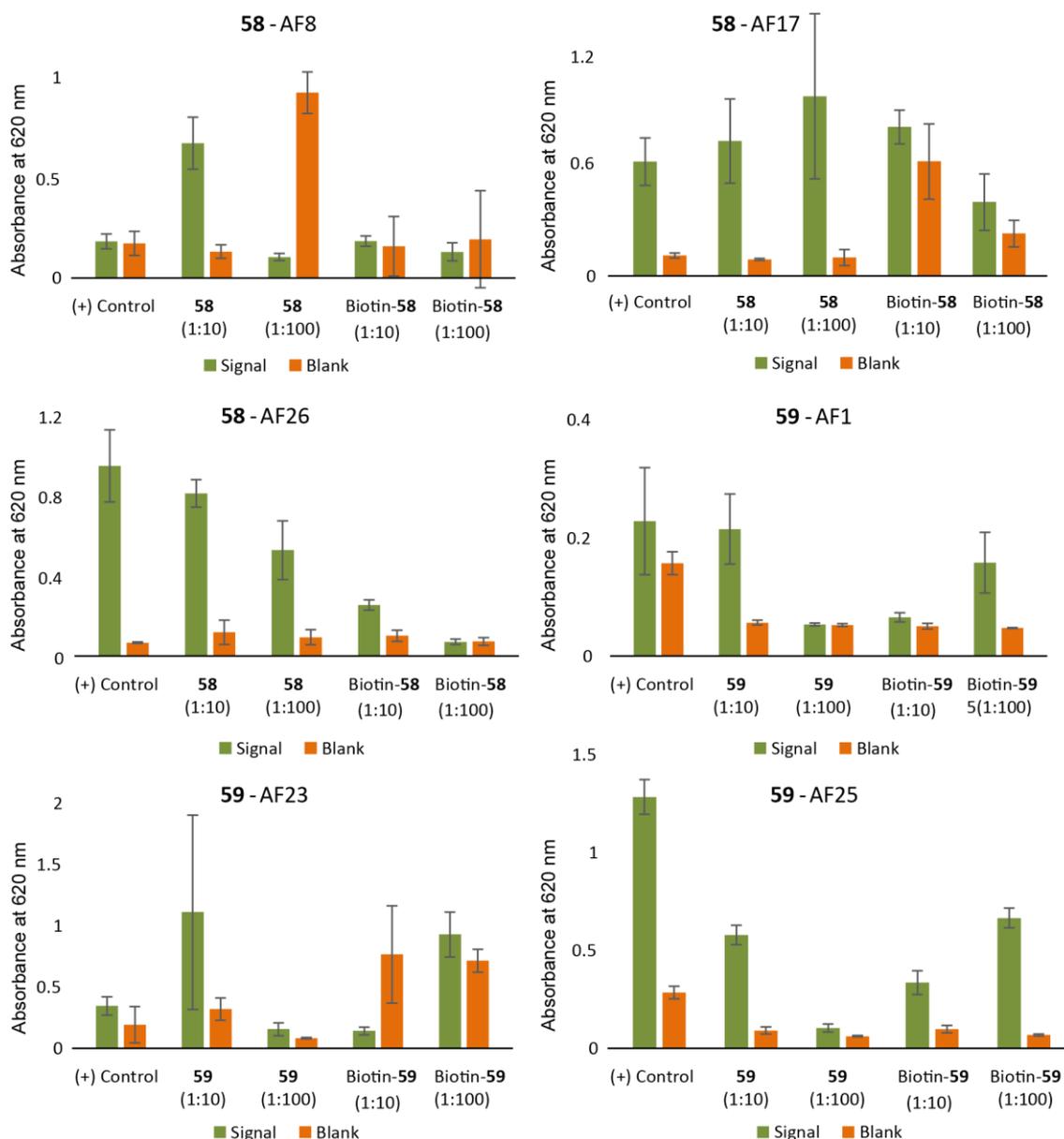
Figure 3.8: Single-point competition assay using 10- and 100-fold excess of competitor, with n = 3. Concentrations of the Affimers: [**58**-AF8] = 10 µM, [**58**-AF17] = 5 µM, [**58**-AF26] = 0.3 µM, [**59**-AF1] = 20 µM, [**59**-AF23] = 5 µM and [**59**-AF25] = 10 µM.

In this assay, the results showed that increasing the excess of competitor enhanced the competitive binding, and suggested the Affimers had more affinity for a biotinylated than a non-biotinylated compound. For all three Affimers screened against Biotin-**59**, the observed increase of signal upon

competition of Biotin-**59** at the ratio 1:100 can be explained by precipitation of the trimer in the assay – attempts to avoid this phenomenon by increasing the percentage of DMSO in the buffer were infructuous as the maximum quantity needed to be kept under 10% and ideally under 5%.

Finally, a competitive experiment was performed in a serial dilution assay, starting from 100-fold excess of competitor compared to the Affimer (Figure 3.9).
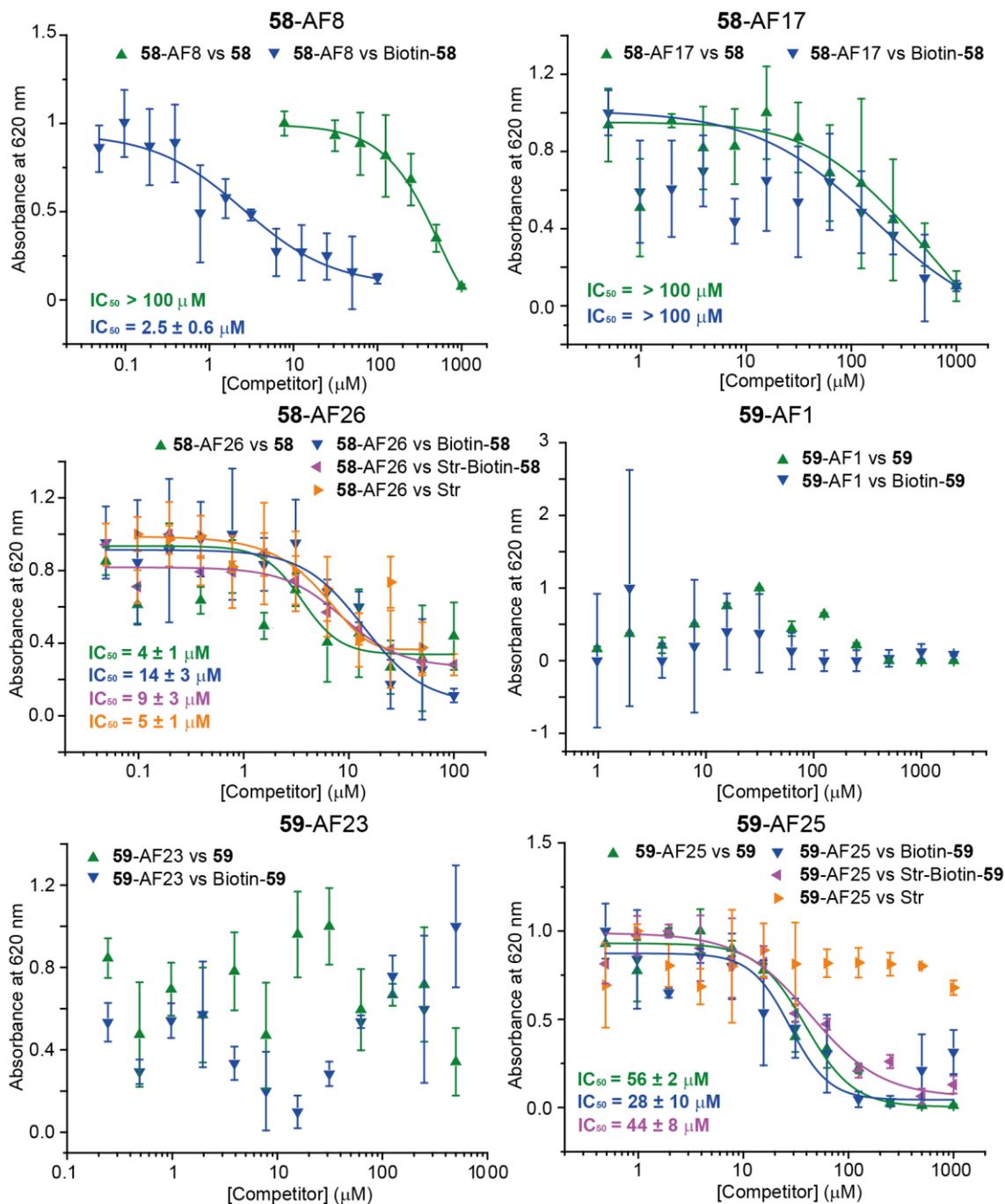
Figure 3.9: Serial dilution competitive ELISA, n = 3. The following concentrations were used: [**58**-AF8] = 10 μM, [**58**-AF17] = 10 μM, [**58**-AF26] = 1 μM, [**59**-AF1] = 20 μM, [**59**-AF23] = 5 μM and [**59**-AF25] = 10 μM.

Here again, the measurements were always prone to a high error, and data such as of **59**-AF1 and **59**-AF23 could not be interpreted. The curves corresponding to **58**-AF17 could be partially fitted

and a weak affinity for either competitor **58** or Biotin-**58** is observed ($IC_{50} > 100$ μM). In accordance with the single-concentrations experiments, streptavidin could compete for **58**-AF26 with similar IC_{50} values to **58** and Biotin-**58**. Data for **58**-AF8 and **59**-AF25 were more promising; **58**-AF8 had weak affinity for **58** ($IC_{50} > 100$ μM), but using Biotin-**58**, $IC_{50} = 2.5 \pm 0.6$ μM was obtained. This implied that trimer **58** on its own is insufficient for recognition of **58**-AF8 and that the panning process selected for an Affimer that recognises both the three helix mimicking side chains of the foldamer and the forth biotin side-chain. In contrast, for **59**-AF25 an $IC_{50} = 56 \pm 2$ μM was obtained upon competition with **59**, whereas a comparable $IC_{50} = 28 \pm 10$ μM was obtained for competition with Biotin-**59**, suggesting the biotin had no significant effect on binding and indicating Affimer selection only for the three helix mimicking side-chains of **59**.

To further confirm this binding, it was decided to use an orthogonal method Fluorescence Anisotropy was chosen (FA, Figure 3.10). FA requires the introduction of a fluorescent tag on the trimer (Scheme 3.3): unbound, the tracer is free to rotate and therefore, upon radiation with polarised light, the emitted light will not be polarised and the anisotropy will be low; upon binding, the complex formed by the tracer and the protein is much bigger and will rotate more slowly than the free tracer, therefore the emitted light will be polarised, which is observed by an increase of anisotropy.
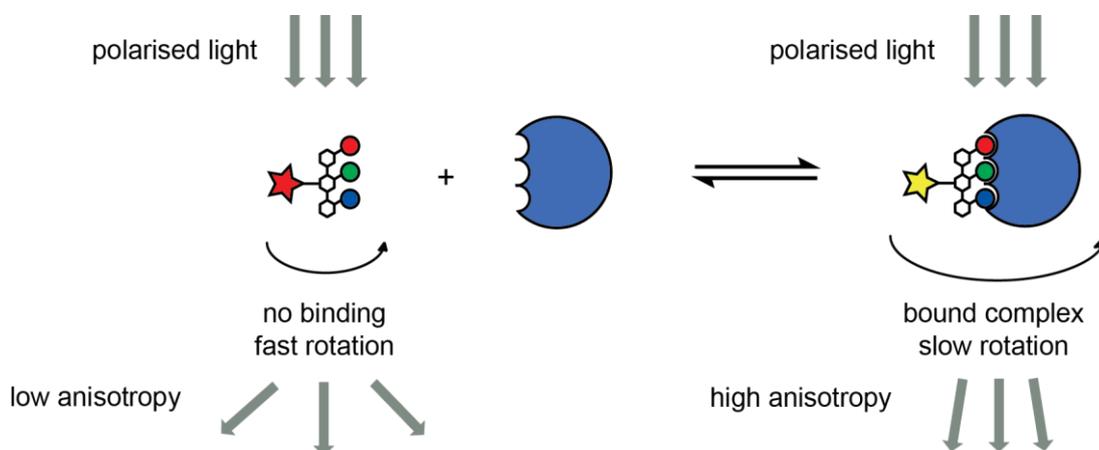
Figure 3.10: Cartoon representation of the principles of Fluorescence Anisotropy (FA).

The synthesis of Fluorescein-tagged analogues of **58** and **59** (respectively Flu-**58** and Flu-**59**, as well as Flu-**59**rev, which is the reverse sequence of building blocks of Flu-**59**) was conducted following the same procedure as for the biotinylated trimers: Flu-azide **73** was synthesized as described in Scheme 3.3 and clicked to the trimer on solid phase.[215] The yields of this coupling decreased compared to the click reaction to introduce the biotin: this was accounted for by poor solubility of **73** in water and *N,N*-dimethylformamide. Several solvents systems were tested to increase the solubility, and a mixture of water, dimethylsulfoxide and *N,N*-dimethylformamide was the only solvent system with which the product formation was observed.
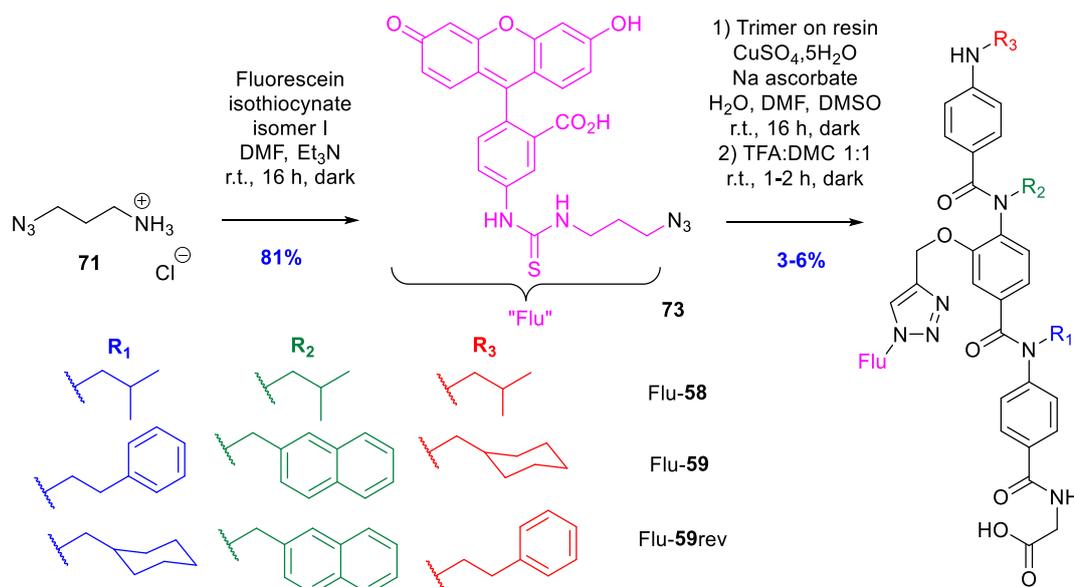
Scheme 3.3: Synthesis of Flu-azide **73** and coupling on resin to form Flu-**58**, -**59** and -**59**rev.[215]

In agreement with the data obtained during the ELISA assays, Flu-**59** was recognised by **59**-AF25, and a binding constant was measured: $K_d = 146 \pm 11$ nM (Figure 3.11a). Gratifyingly, binding was not observed for Flu-**59**rev (Figure 3.11b, blue triangles). Because the values observed gave a straight line around an anisotropy of 1, it was suggested that there could be some non-selective interactions with the plate; for this reason, repeats of the same assay were performed in a special non-stick plate (Figure 3.11b, yellow lozenges), but still no binding curve could be obtained from this assay. This result attests of the high selectivity of the Affimer for the exact sequence order and composition of side-chains. The direct titration of Flu-**58** against **58**-AF8 (Figure 3.11b, green circles) was also carried out, but did not yield a binding curve either; this confirms the previously observed recognition of the biotin as well as the proteomimetic side-chains, as the introduction of the fluorescein motif would not be recognised by the Affimer **58**-AF8.
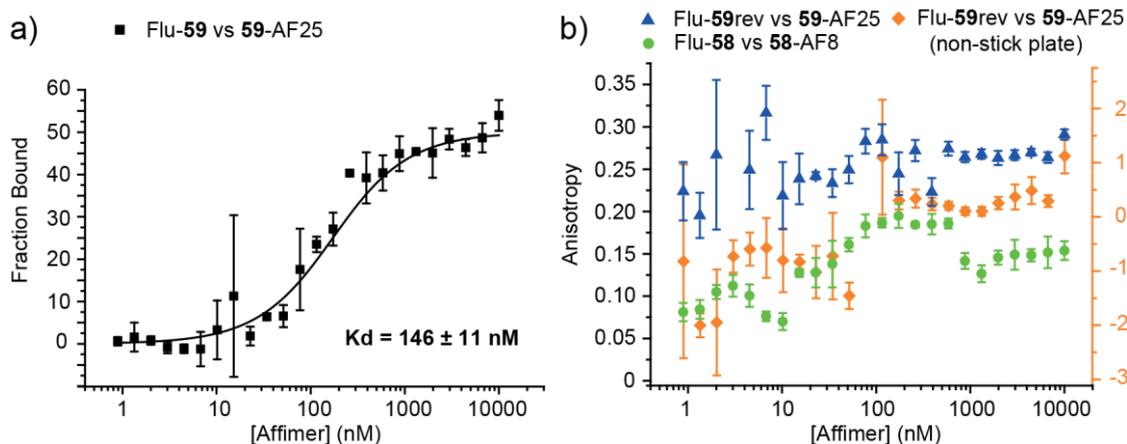
Figure 3.11: Fluorescence Anisotropy assays: a) binding curve of Flu-**59** with **59**-AF25 and b) results of Flu-**59**rev against **59**-AF25 and Flu-**58** against **58**-AF8, showing no binding occurred.

61

3.3. Synthesis of biotinylated trimers equipped with a longer linker

It was proven that the biotin group can interfere in the recognition of the Affimer, with the example of **58**-AF8 binding with micromolar affinity to Biotin-**58**, but not to **58**. Although some results indicated that the biotin is not systematically recognised by the Affimers, a way of optimising the selection by limiting this non-specific interaction may be to introduce a longer linker between the trimer and the biotin moiety. To do so, a new orthogonally functionalised scaffold involving the presence of PEG units to act as elongated linkers was designed (Figure 3.12).
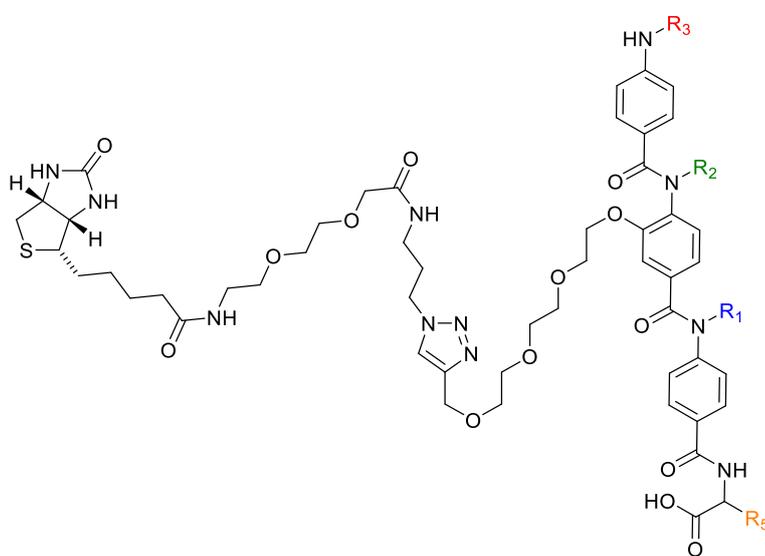
Figure 3.12: New scaffold for orthogonally-functionalised foldamers.

The synthesis of the new scaffold required the preparation of a new azide, called Biotin-PEG-Azide **76** (Scheme 3.4), as well as a new alkyne-equipped monomeric building block (Scheme 3.5). The azide synthesis started off using the acid chloride salt **71** and coupling on 8-(Fmoc-amino)-3,6-dioxaoctanoic acid. Attempts to achieve this coupling using EDC and HOBT, and HCTU were conducted but neither of them were conclusive. Finally, PyBOP was chosen as it led to **74** with 82% yield. The Fmoc group was removed using triethylamine and the biotin functionalisation was performed using HCTU and DIPEA in *N,N*-dimethylformamide.
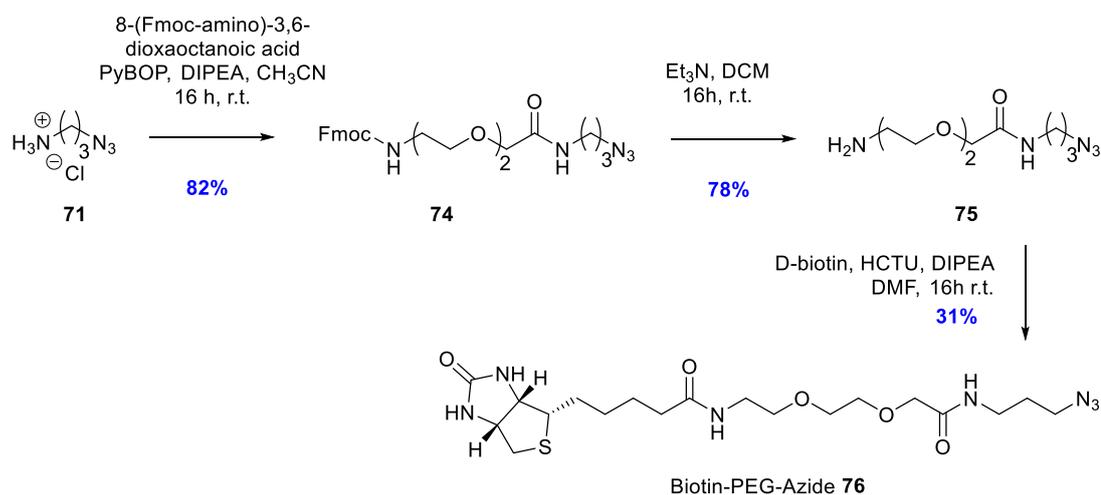
Scheme 3.4: Synthesis of Biotin-PEG-Azide **76**.

The new monomeric building block was prepared in a similar manner as previously (Scheme 3.5). 3-(2-(2-(2-chloroethoxy)ethoxy)ethoxy)prop-1-yne **77** was obtained from the Bon group at the University of Leeds, and prepared as a 4:1 mixture of Cl:Br group from a published procedure.[222] The first coupling required the presence of potassium iodide and heating to 70°C, to afford **78** with a 51% yield. Tin (II) chloride reduction followed by saponification, reductive amination with isobutyraldehyde and Fmoc protection led to the final monomer **82**.
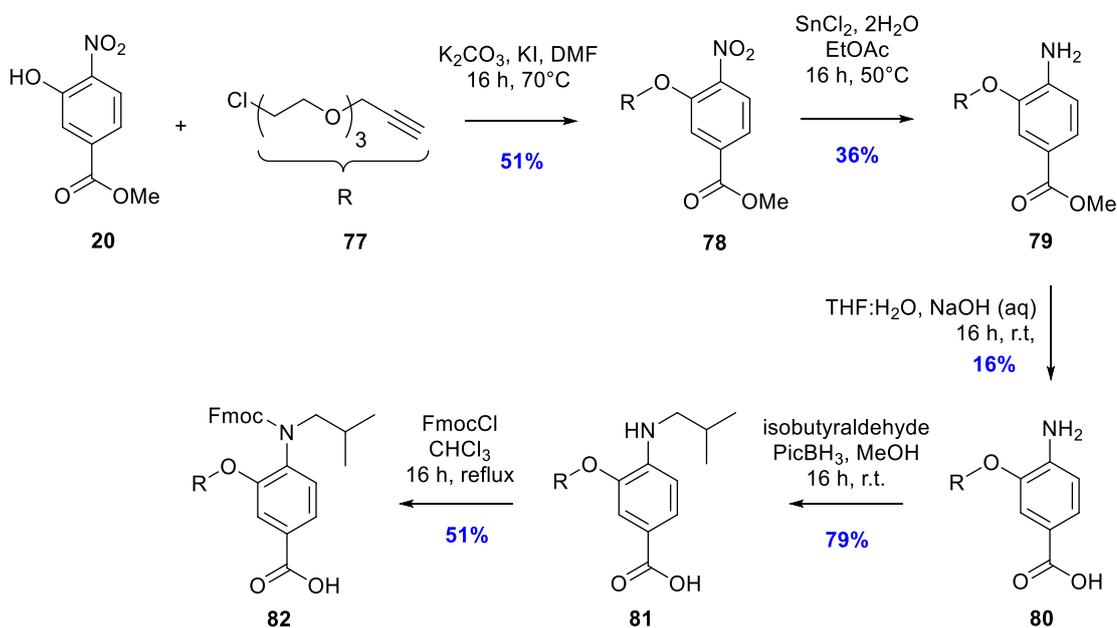
Scheme 3.5: Preparation of the PEG-equipped alkyne monomer **82**.

For time considerations and considering the final amount obtained of final monomer (40 mg) **82**, trimer synthesised was not attempted.

3.4. Conclusions and future work

A library of six biotinylated *N*-alkylated trimers was screened in phage display experiment using Affimers. Out of the six, two exhibited few binders with little selectivity (Biotin-**60** and -**61**), two exhibited a moderate number of binders and little selectivity (Biotin-**62** and -**63**), one exhibited a moderate number of binders with good selectivity (Biotin-**58**) and one exhibited a good number of binders with strong selectivity (Biotin-**59**). The Affimer hits obtained for Biotin-**58** and -**59** were sequenced, and amongst them, three were selected for each trimer (two selective and one non-selective), to be subcloned, expressed and purified. Preliminary direct titration ELISA assay were performed to assess the strength of the binding of each Affimer towards their corresponding trimer. Further single point competition assays were carried out to evaluate the degree of binding to the following competitors: non-biotinylated trimer, biotinylated trimer, streptavidin, streptavidin-biotinylated trimer. Finally, serial dilution competition assays were performed to obtain, when possible, full competition curves and IC_{50} values. In particular, three cases showed three different binding behaviours: **58**-AF8 recognises Biotin-**58** ($IC_{50} = 2.5 \pm 0.6$ μM), but its affinity for **58** without biotin is very weak ($IC_{50} > 100$ μM); **59**-AF25 recognises the side-chains of **59**, independently of the presence of a biotin or streptavidin group ($IC_{50} = 56 \pm 2$ μM for competition with **59**, $IC_{50} = 28 \pm 10$ μM for competition with Biotin-**59**, $IC_{50} = 44 \pm 8$ μM for competition with Str.-Biotin-**59**); **1**-AF26 is non-selective, and recognises all the tested competitors with similar affinities. Finally, FA assays were conducted: for this purpose, FITC tagged trimers Flu-**58**, Flu-**59** and Flu-**59**rev were prepared and tested against the Affimers in direct titrations assays. The binding curves obtained confirmed the results from the ELISA tests: Flu-**59** bound to **59**-AF25, proving its ability to recognise the side-chains independently of the orthogonal group; and no binding curve could be obtained against Flu-**59**rev, confirming the strong selectivity obtained for this Affimer and the importance of the sequence order and composition of side-chains.

On those results, a new scaffold was designed, in order to add space between the trimer and the biotin group. A new azide and a new alkyne building block were prepared, both including PEG units to lengthen the linker. The full synthesis of each compound was described and optimised.

Future work includes preparing new biotinylated trimers equipped with the new linker and screening them against Affimers. Comparing the new selection of hits with the previous ones will give important information towards the recognition during the assay. Similar ELISA and FA experiments should be conducted to assess the ability of the new hits to recognise exclusively the side-chains of the trimers.

Chapter 4.   Thesis summary and future work

Only 20 different amino acids are at the origin of a stunning array of over 18,000 proteins,[1, 2] responsible for complex and selective tasks essential for life. The complex 3-dimensional structures that enables such complexity, involving helices, sheets, loops, turns, and disordered domains, arise from a precise self-organisation of proteins, which is defined by their primary structure alone. Although great progress has been made in *de novo* design of foldamers,[50, 76, 79, 93, 200, 206, 214, 223] chemical and synthetic biologists are facing the long-term challenge to expand the genetic toolbox to generate functional bionic proteins that are not limited to sequences of α-amino acids only.[3, 199]

Considering the importance of α-helices in protein function, α-helix mimicry is a promising approach to replicating segments of proteins.[5] The previous research conducted in the Wilson Group has extensively investigated type III α-helix mimetics (proteomimetics), *i.e.* structures that match the topography of the original helix motif by mimicking the spatial orientation of its key functional residues.[11] 2-*O*-, 3-*O*- and *N*-alkylated oligoamides were designed and successfully used as PPI inhibitors[113, 114, 119] *in vitro* as well as in cells.[118] Preliminary work towards generating bionic proteins was also conducted with a bionic peptide.[120]

This project aimed at using proteomimetics to generate novel 3-dimensional assemblies and therefore advance towards engineering bionic proteins. The first part of the project attempted to design coiled coil mimetics with proteomimetics in order to establish the ability of proteomimetics to mimic α-helical supramolecular assemblies.

In that regard, *N*- and *O*-alkylated oligomers bearing hydrophilic side-chains, designed to self-assemble, were prepared (Chapter 2). The small library obtained was studied by NMR, which proved that the conformation of dimers can be, to some extent, controlled by careful design of the side-chains: for *O*-alkylated dimers, a strictly *syn* conformation was generated, while free rotation around the amide bond usually allows a *syn*/*anti* equilibrium (Figure 4.1a); for *N*-alkylated dimers, a preference for the *cis* conformation was observed (Figure 4.1b); *N*-alkylated trimers bearing central hydrophobic side-chains also proved to be too flexible to generate any *trans* conformation. The potential for self-assembly of those compounds was tested by dilution experiments. Significant shifts in the NMR were observed in the molecules designed to self-assemble, while the spectra of the control molecules were unchanged upon dilution, suggesting that self-assembly was indeed observed. Further efforts to confirm this hypothesis are nevertheless still required.

To complete this study, the synthesis of a larger library, including different patterns of side-chains, and different lengths of oligomers will be necessary. Since modelling suggested that the carboxylic acid issued from the resin was involved in different patterns of interactions, it will be

interesting to investigate the use of other resins in order to obtain an oligomer equipped with a C-terminal amine or amide group.
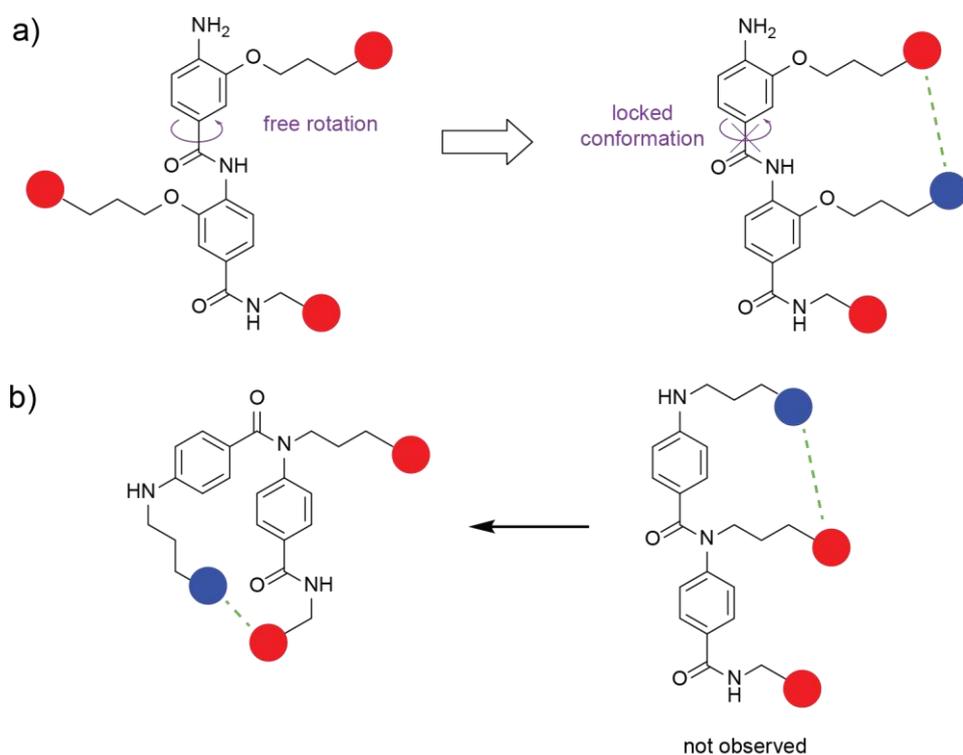
Figure 4.1: Results from chapter 2 suggest that it is possible to a) lock an *O*-alkylated dimer in the *cis* conformation, but b) no *trans* conformation was observed for the *N*-alkylated series.

Once evidence of self-assembly with dimers and trimers is observed, efforts to construct longer oligoamides will be conducted to generate a proteomimetic equivalent of a stalk. Further work will aim to diversify the types of coiled coils mimicked, and obtained a biologically functional coiled coil mimetic.

The second part of the project focused on a novel artificial binding protein scaffold developed within the Astbury Centre termed Affimer,[122] to be screened against proteomimetics in phage display so as to assess the ability for proteomimetics to act as native α-helices.

A library of 6 biotinylated *N*-alkylated trimers was screened against a library of Affimers. Two of them, Biotin-**58** and Biotin-**59**, generated an important number of hits with good selectivity. The binding affinities of the foldamers towards a small selection of Affimers were assessed by direct titration ELISA assay. Competition assays were further carried out to assess whether the Affimers recognised only the side-chains of the proteomimetics, or the full biotin-trimer assembly. Several cases were observed, where only the side-chains were recognised (**59**-AF25, Figure 4.2a); where the side-chain and the biotin were recognised (**58**-AF8, Figure 4.2b); or where the Affimer bound to any compound in solution, including the streptavidin on its own (**58**-AF26). Finally, the different patterns of selectivity observed by ELISA were confirmed by FA assays.

To improve the screening process, improvements were made to the design of the scaffold, in order to lengthen the linker between the recognition face of the foldamer and the biotin moiety. To do so, a new azide and a new alkyne building block were prepared, both including PEG units, although no trimer was prepared for time considerations. Future work should focus on preparing a new library of trimers equipped with this new linkers, and screening them again in an Affimer phage display setting in order to maximise the recognition of the side-chains only.
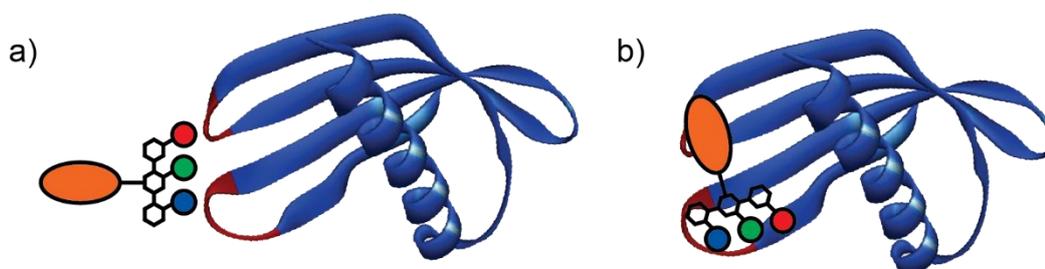
Figure 4.2: Schematic examples of an Affimer recognition a) the side-chains of a foldamer, and b) the side-chains and the streptavidin tag (orange).

Long-term efforts will focus on the binding sequence from the loops (Figure 4.3). Firstly, it is possible to express Affimers with one of the two loops mutated: this will allow to establish whether both or only one loop is required for binding. Then, excision of the binding peptide(s) will allow direct study of the binding between the trimer and the peptide, and enable to obtain more information on the 3-dimensional structure of the loop. Finally, the Affimer phage display will be used as a reverse screening method for biologically relevant targets: the binding sequences will be compared to literature databases in order to identify known sequences corresponding to a PPI, therefore allowing to discover interactions to which existing trimers have inhibitory potency.
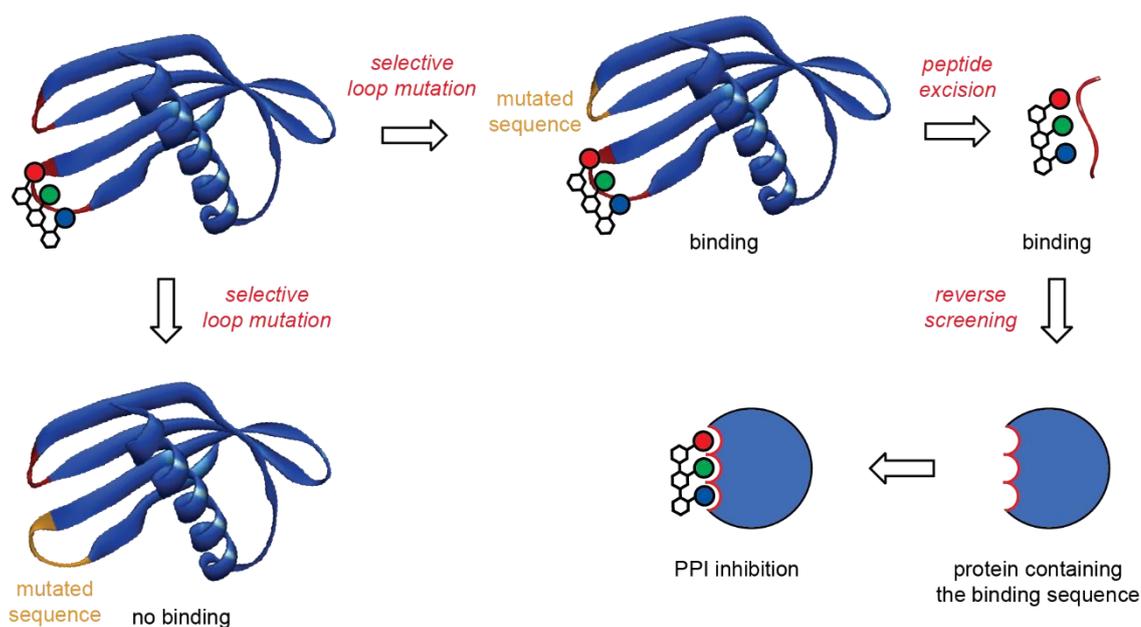
Figure 4.3: Summary of the future work concerning the Affimers.

67

Chapter 5. Experimental

5.1. Synthesis

5.1.1. General considerations

Solvents. Unless stated otherwise, solvents and reagents were used as received from major suppliers without prior purification. Anhydrous acetonitrile, chloroform, dichloromethane were obtained from the in-house solvent purification system from Innovative Technology Inc. PureSolv®. Anhydrous dimethylsulfoxide was obtained from major chemical suppliers equipped with a SureSeal or equivalent. Non-anhydrous solvents were of HPLC quality and provided by Fisher or Sigma-Aldrich. Water used for formation of aqueous solutions and quenching was deionised.

Chromatograhy. Thin Layer Chromatography (TLC) was performed on Merck Kieselgel 60 F_{24} 0.25 mm precoated aluminium plates. Product spots were visualised under UV light (λ_{max} = 254 nm) and/or staining with anisaldehyde. Purifications were performed with either silica gel 60 (0.043-0.063 mm VWR) using head bellows or by flash chromatography using an Isolera Four Biotage®. Ion-exchange columns were performed using Supleco Discovery SPE DSC-SAX columns. HPLC experiments were run on an Agilent 1290 Infinity Analytical Preparative system spectrometer.

NMR. ^{1}H NMR spectra were recorded on Bruker DPX 300 (300 MHz) or Advance 500 (500 MHz) spectrometers and referenced to residual non-deuterated solvent peaks. ^{13}C NMR were recorded on a Bruker Advance 500 (125 MHz) and referenced to the solvent peak. Chemical shifts (δ) are expressed in part per million (ppm) and coupling constants are expressed in hertz (Hz). Assignments of spectra were assisted by the results of COSY, NOESY, TOCSY, ROESY, HMQC and HMBC experiments when appropriate.

Mass. HPLC LC-MS were recorded on a Bruker HCT ultra under the conditions of electrospray ionisation (ESI). HPLC separation was performed on an Agilent 1200 series instrument equipped with a Phenomenex C18 column (50 x 2 mm) using acetonitrile/water as the eluent for positive ion spectra. HRMS were performed using a Bruker Maxis impact mass spectrometer, using ESI. Values are reported as a ratio of mass to charge.

IR. Infrared spectra were recorded on a Perkin Elmer Fourier-Transfer spectrometer. Spectra were analysed neat and only structurally important absorptions are quoted. Absorption maxima (ν_{max}) are quoted in wavenumbers (cm^{-1}).

Characterisation. For compounds previously published in the group, LC-MS and ^{1}H NMR were used to confirm the structure and purity. New compounds were either partially characterised (side chain intermediates or instable compounds: ^{1}H NMR, ^{13}C NMR) or fully characterised (^{1}H NMR, ^{13}C NMR, HRMS and IR are reported).

5.1.2. Foldamer building block side chain syntheses

tert-**Butoxycarbonylmethyl nitrate (2)**[178]
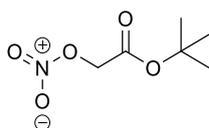
Tert-butyl bromoacetate **1** (3.0 mL, 20.3 mmol) and silver nitrate (4.5 g, 26.4 mmol) were stirred in anhydrous acetonitrile (40 mL) under a nitrogen atmosphere for 16 hours, in the dark. The reaction mixture was filtered, and the filtrate was concentrated under vacuum. The residue was washed with water (100 mL) and extracted with diethyl ether (100 mL). The organic layer was dried over magnesium sulphate, filtered and the filtrate concentrated under vacuum to afford **2** as a colourless oil, which was used without further purification (2.9 g, 81%). δ_H (500 MHz, CDCl3) 4.79 (s, 2H, C*H₂*), 1.52 (s, 9H, C(C*H₃*)₃).

di-*tert*-Butyl fumarate (8)
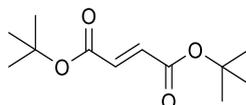
First method:[181] *tert*-Butyl acrylate **6** (4.0 mL, 27.6 mmol) was added to a solution of Grubbs second generation catalyst (24 mg, 0.028 mmol) in anhydrous dichloromethane (1 mL) under a nitrogen atmosphere. The solution was stirred and heated to reflux for 16 hours. The solution was cooled down to room temperature and the solvent was evaporated under vacuum. The residue was purified by column chromatography (eluent: ethyl acetate/hexane: 1/10) to give the target material **8** as a white crystalline solid (245 mg, 8%). **Second method:**[182] *n*-Butyl lithium (37.5 mL of 1.6 M in hexane, 60.0 mmol) was added over 10 minutes to anhydrous *tert*-butyl alcohol (10.0 mL, 105.0 mmol) at 0°C, under a nitrogen atmosphere, and the suspension was stirred at 0°C for 40 minutes. Fumaryl chloride **7** (6.5 mL, 60.0 mmol) was added over 40 minutes at room temperature and the solution was stirred at room temperature for 4 hours. The reaction was then quenched with water (50 mL) and the organic layer was washed with saturated sodium hydrogen carbonate (50 mL) and dried over magnesium sulphate, filtered and the filtrate was concentrated under vacuum. The crude product was purified by column chromatography (eluent: ethyl acetate/hexane: 1/9) to afford **8** as a white crystalline solid (5.13 g, 38%). δ_H (500 MHz, CDCl₃) 6.67 (s, 2H, C*H*=C*H*); 1.50 (s, 18H, C(C*H₃*)₃).

tert-Butyl glyoxylate (3)
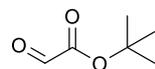
First Method:[178] *tert*-Butoxycarbonylmethyl nitrate **2** (2.9 g, 16.5 mmol) and sodium acetate (1.3 g, 16.5 mmol) were stirred in anhydrous dimethyl sulfoxide (20 mL) under a nitrogen atmosphere for 30 minutes. The mixture was poured into brine (100 mL) and extracted with dichloromethane (3 x 30 mL). The organic layer was washed with sodium hydrogen carbonate (50 mL) and brine (50 mL), then dried over magnesium sulphate, filtered and the filtrate concentrated under vacuum to give **3** as a dark yellow oil (containing some residual dimethyl sulfoxide, 560 mg, < 26%) used without further purification. **Second method:**[183] Ozone was passed through a solution of di-*tert*-butyl fumarate **8** (5.6 g, 24.6 mmol) in dichloromethane (20 mL) at -78°C over 30 minutes, after which a blue coloration of the solution was observed. Excess ozone was swept out by a stream of oxygen and further bubbled with nitrogen, until the solution became clear again. Dimethyl sulfide (9.0 mL, 123.0 mmol) was added and the solution was allowed to reach room temperature and stirred for 16 hours. The residue was stirred with magnesium sulfate for 30 minutes, filtered and the filtrate evaporated under vacuum to give **3** as a dark yellow oil (containing some residual dimethyl sulfoxide, 3.54 g, 13% estimated by NMR) used without further purification. δ_H (500 MHz, CDCl$_3$) 9.30 (s, 1H, C(O)*H*), 1.57 (s, 9H, C(C*H$_3$*)$_3$).

4-(*tert*-butoxy)-4-oxobutanoic acid (11)[224]
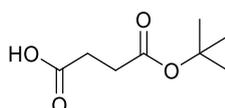
Succinic anhydride (5.0 g, 50.0 mmol) **10**, *N*-hydroxysuccinimide (1.7 g, 15.0 mmol) and 4-dimethylaminopyridine (610 mg, 5.0 mmol) were dissolved in toluene (25 mL). *tert*-Butanol (5.9 mL, 62.5 mmol) and triethylamine (2.1 mL, 15 mmol) were added sequentially and the mixture was refluxed for 16 hours. The solution was cooled down to room temperature and poured into ethyl acetate (25 mL) and washed with a 10%w citric acid solution (50 mL) and brine (50 mL). The organic layer was dried over magnesium sulfate, filtered, and the filtrate was concentrated and purified by column chromatography (eluent: dichloromethane). Product **11** was collected as a white solid (3.67 g, 42%). δ_H (CDCl$_3$, 500 Mz) 2.65 (t, 2H, J = 6.8, -C*H$_2$*CO$_2$H), 2.56 (t, 2H, J = 6.8 Hz, -C*H$_2$*CO$_2$C(CH$_3$)$_3$), 1.45 (s, 9H, C(C*H$_3$*)$_3$). δ_C (CDCl$_3$, 125 Mz): 178.0, 171.4, 81.0, 30.0, 29.1, 27.9.

tert-Butyl 4-hydroxybutanoate (12)[224]
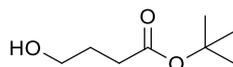
4-(*tert*-Butoxy)-4-oxobutanoic acid **11** (2.1 g, 11.8 mmol) was dissolved in anhydrous tetrahydrofuran (20 mL) and borane dimethylsulfide complex 2 M in tetrahydrofuran (6.5 mL, 13.0 mmol) was added dropwise to the mixture. The solution was stirred at room temperature for 16 hours, and poured into ethyl acetate (100 mL), then washed with water (50 mL) and brine (100 mL). The organic layer was dried over magnesium sulphate, filtered, and the filtrate was concentrated under vacuum and purified by column chromatography (eluent: dichloromethane/methanol: 9.5/0.5). **12** was collected as a colourless oil (862 mg, 46%). δ_H (CDCl$_3$, 500 Mz) 3.66 (t, 2H, J = 6.2, -CH_2OH), 2.35 (t, 2H, J = 7.1 Hz, -CH_2CO$_2$C(CH$_3$)$_3$), 2.20 (s, 1H, OH), 1.85 (quint., 2H, J = 6.2 Hz –CH_2-), 1.43 (s, 9H, C(CH_3)$_3$). δ_C (CDCl$_3$, 125 Mz): 173.4, 80.4, 62.1, 32.4, 28.0, 27.8.

tert-Butyl 4-oxobutanoate (13)
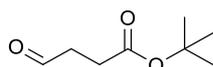
A solution of oxalyl chloride (1.1 mL, 12.7 mmol) in dichloromethane (200 mL) was cooled to -78°C. Dimethyl sulfoxide (1.0 mL, 24.1 mmol) was added and the mixture was stirred at -78°C for 1 hour. *tert*-Butyl 4-hydroxybutanoate **12** (1.13 g, 7.1 mmol) was added and the mixture was stirred at -78°C for 1 hour. Triethylamine (4.9 mL, 35.3 mmol) was added dropwise and the mixture was allowed to reach room temperature and stirred at room temperature for 16 hours. The reaction mixture was then poured into water (200 mL) and extracted with dichloromethane (2 x 150 mL). The organic layer was dried over magnesium sulphate and the solvent was removed under vacuum to give **13** as a colourless oil (326 mg, 30%). δ_H (CDCl$_3$, 500 Mz) 9.80 (s, 1H, -CHO), 2.74 (t, 2H, J = 6.6 Hz, -CH_2CHO), 2.56 (t, 2H, J = 6.6 Hz, -CH_2CO$_2$C(CH$_3$)$_3$), 1.44 (s, 9H, C(CH_3)$_3$). δ_C (CDCl$_3$, 125 Mz): 200.5, 171.7, 81.1, 38.9, 28.3, 28.1.

tert-Butyl 4-bromobutanoate (14)[225]
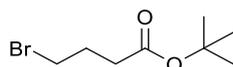
Tert-Butyl 4-hydroxybutanoate **12** (250 mg, 1.6 mmol) and carbon tetrabromide (1.0 g, 3.1 mmol) were dissolved in tetrahydrofuran (5 mL), in an ice bath. A solution of triphenylphosphine (817 mg, 3.1 mmol) in tetrahydrofuran (5 mL) was slowly added and the mixture was allowed to reach room temperature and then stirred for 16 hours. The residue was filtered off, the solvent was removed under vacuum, and the obtained oil was purified by column chromatography (eluent: ethyl acetate/hexane: 5/95) to give **14** as a yellow oil (247 mg, 71%). δ_H (CDCl$_3$, 500 Mz) 3.47

(t, 2H, J = 6.8 Hz, -CH_2Br), 2.42 (t, 2H, J = 6.8 Hz, -CH_2CO$_2$C(CH$_3$)$_3$), 2.17 (quint., 2H, J = 6.8 Hz, –CH_2-), 1.44 (s, 9H, C(CH_3)$_3$).

3-(*N*-boc-amino)-1-propanol (18)[178]
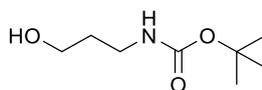
di-*tert*-Butyl dicarbonate (32.0 g, 146.5 mmol) in dichloromethane (100 mL) was added dropwise to a solution of 3-aminopropan-1-ol **17** (10.0 mL, 133.2 mmol) in dichloromethane (70 mL). The solution was stirred for 16 hours at room temperature. The reaction was quenched with saturated aqueous sodium hydrogen carbonate (100 mL) and the mixture was washed with water (50 mL) and brine (50 mL). The organic layer was dried over magnesium sulfate, filtered, and the filtrate was concentrated under vacuum to afford compound **18** as a colorless oil (18.6 g, 80%). δ_H (500 MHz, CDCl$_3$) 4.84 (s, 1H, N-*H*), 3.67 (q, 2H, J = 5.9 Hz, HOCH_2CH$_2$-), 3.29 (q, 2H, J = 5.9 Hz, CH-$_2$CH_2NH-), 3.09 (s broad, 1H, O*H*) 1.68 (qu., 2H, J = 5.9 Hz, -CH$_2$CH_2CH$_2$-), 1.44 (s, 9H, (CH_3)$_3$).

tert-Butyl (3-oxopropyl)carbamate (19)[178]
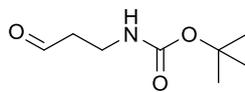
Dimethyl sulfoxide (1.6 mL, 23.4 mmol) was added at -78ºC to a solution of oxalyl chloride (1.7 mL, 21.0 mmol) in anhydrous dichloromethane (20 mL) under a nitrogen atmosphere and the solution was stirred for 1 hour at -78ºC. 3-(*N*-Boc-amino)-1-propanol **18** (1.4 mL, 8.3 mmol) was then added and the resulting mixture was further stirred for 1 hour at -78ºC. Triethylamine (7.9 mL, 57.0 mmol) was added and the mixture was allowed to reach room temperature and further stirred for 16 hours. The reaction was quenched with water (30 mL) and the mixture extracted with dichloromethane (2 x 30 mL) and concentrated under vaccum. The residue was purified by column chromatography (eluent: dichloromethane/methanol: 9.8/0.2) to afford the desired product **19** as a dark yellow viscous oil (960 mg, 67%). δ_H (500 MHz, CDCl$_3$) 9.81 (s, 1H, C(O)*H*), 4.90 (s, 1H, N-*H*), 3.44 (q, 2H, J = 5.9 Hz -CH_2CH$_2$NH-), 2.72 (t, 2H, J = 5.9 Hz, -CH$_2$CH_2CH$_2$-), 1.43 (s, 9H, (CH_3)$_3$). δ_C (125 MHz, CDCl$_3$) 200.6, 171.7, 81.1, 38.9, 28.3, 28.1.

tert-Butyl (3-bromopropyl)carbamate (16)[177]
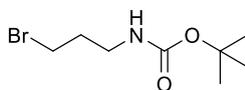
A suspension of 3-bromopropylamine **15** (18.0 g, 83.0 mmol) and di-*tert*-butyl dicarbonate (18.0 g, 83.0 mmol) in chloroform (40 mL) was stirred at 0ºC. Triethylamine (18.0 mL, 165.0 mmol)

was added dropwise over 1 hour and the mixture was stirred at room temperature for 16 hours. Chloroform (20 mL) was added, and the solution was washed with a 1M hydrochloric acid solution (2 x 20 mL) and water (2 x 20 mL). The organic layer was dried over magnesium sulfate, filtered and the filtrate was evaporated to dryness to afford the desired compound **16** as a yellow oil (14.4 g, 73%). δ_H (CDCl$_3$, 500 Mz) 4.75 (s, 1H, N*H*), 3.45 (t, 2H, *J* = 6.4 Hz, BrC*H$_2$*-), 3.29 (q, 2H, *J* = 6.4 Hz -C*H$_2$*NH-), 2.08 (quint., 2H, *J* = 6.4 Hz, -C*H$_2$*-), 1.44 (s, 9H, OC(C*H$_3$*)$_3$).

5.1.3. Monomer synthesis

5.1.3.1. General procedures

a) Procedure for *O*-alkylation

The phenol (1 eq.), alkyl halide (1.5 eq.), and potassium carbonate (3 eq.) were stirred in *N,N*-dimethylformamide (4 mL/mmol) and heated to 50°C for 16 hours. The mixture was cooled down to room temperature and poured into water (3 mL/mmol), extracted with ethyl acetate (3 mL/mmol) and the organic layer was then washed with brine (40 mL/mmol). The organic layer was dried over magnesium sulfate, filtered, and the filtrate was evaporated under vacuum and purified by column chromatography as required.

b) Procedure for nitro reduction – hydrogenation

The nitro compound (1 eq.) was dissolved in methanol:tetrahydrofuran or methanol:ethyl acetate (1:1, 5 mL/mmol) and the solution was bubbled with nitrogen for 5 minutes. Palladium 10% on carbon (10%w) was added and the solution was further bubbled with nitrogen for 5 minutes. Hydrogen was passed through the reaction mixture at room temperature for 16 hours, using a balloon. The mixture was filtered through a pad of celite, washed with methanol, and the filtrate was evaporated.

c) Procedure for nitro reduction – tin chloride

Tin (II) chloride dihydrate (5 eq.) was added to a solution of nitro compound (1 eq.) in ethyl acetate (12 mL/mmol), and the reaction mixture was stirred at 50°C for 16 hours under a nitrogen atmosphere and with a drying tube. The solution was poured into iced water (4 mL/mmol) and basified with sodium hydrogen carbonate. The product was extracted with ethyl acetate (4 mL/mmol) and the organic layer was washed with sodium hydroxide 1M (2 mL/mmol), water (2 mL/mmol) and brine (2 mL/mmol). The organic layer was dried over magnesium sulfate, filtered, and the filtrate was evaporated under vacuum.

d) Procedure for ester hydrolysis

Sodium hydroxide (6 eq., 10% aqueous solution) was added to a solution of ester (1 eq.) in a 1:1 mixture of methanol and tetrahydrofuran (4 mL/mmol), and the solution was stirred at room

temperature for 16 hours. The mixture was poured into dichloromethane (3 mL/mmol) and the organic layer was washed with water (5 mL/mmol). The aqueous layer was acidified to pH 4 with a 1M aqueous hydrochloric acid solution and extracted with dichloromethane (9 mL/mmol). The organic layer was dried over magnesium sulfate, filtered, and the filtrate was concentrated under vacuum.

e) Procedure for reductive amination

The amino benzoic acid (1 eq.), aldehyde (1 eq.) and picoline borane (1.2 eq.) were stirred in methanol (2 mL/mmol) at room temperature for 3-16 hours. The solvent was removed and the residue diluted in ethyl acetate (4 mL/mmol), and washed with a 1M aqueous HCl solution (4 mL/mmol). The organic layer was dried over magnesium sulphate, filtered, and the filtrate was concentrated under vacuum.

f) Procedure for Fmoc protection

The amine (1 eq.) and fluorenylmethyl chloroformate (1 eq.) and sodium carbonate (2 eq. for Boc-protected groups) or sodium hydrogen carbonate (1.1 eq. for *tert*-butyl esters) were refluxed for 16 hours in anhydrous chloroform (4 mL/mmol). The residue was evaporated and purified either by precipitation or column chromatography.

5.1.3.2. *O*-alkylated monomers (Chapter 2)

Methyl 3-isobutoxy-4-nitrobenzoate (21)[177]
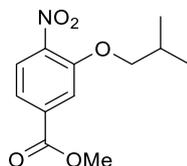
Using the general procedure for *O*-alkylation a). Methyl-4-amino-3-hydroxybenzoate **20** (3.0 g, 15.2 mmol); isobutyl bromide (2.9 mL, 26.9 mmol); potassium carbonate (7.4 g, 53.7 mmol); *N,N*-dimethylformamide (60 mL). Orange solid (2.9 g, 75%). δ_H (500 MHz, CDCl$_3$) 7.84 (d, 1H, J = 8.5 Hz, Ar-*H*), 7.74 (d, 1H, J = 1.6 Hz, Ar-*H*), 7.69 (dd, 1H, J = 8.5 Hz, 1.6 Hz, Ar-*H*), 3.99 (s, 3H, O-C*H$_3$*), 3.96 (d, 2H, J = 6.4 Hz, C*H$_2$*CH(CH$_3$)$_2$), 2.21-2.16 (m, 1H, CH$_2$C*H*(CH$_3$)$_2$), 1.09 (d, 6H, J = 6.8 Hz, CH$_2$CH(C*H$_3$*)$_2$).

Methyl 4-amino-3-isobutoxybenzoate (22)[177]

75
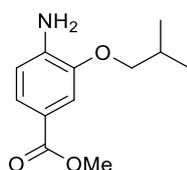
Using the general procedure for nitro reduction b). Methyl-3-isobutoxy-4-nitrobenzoate **21** (2.9 g, 11.4 mmol); methanol:tetrahydrofuran 1:1 (60 mL); palladium 10% on carbon (300 mg). Red solid (2.22 g, 86%). δ_H (500 MHz, CDCl$_3$) 7.54 (dd, 1H, J = 8.3, 0.5 Hz, Ar-*H*), 7.43 (s, 1H, Ar-*H*), 6.68 (d, 1H, J = 8.3 Hz, Ar-*H*), 4.23 (s (broad), 2H, N*H$_2$*), 3.86 (s, 3H, OC*H$_3$*), 3.83 (d, 2H, J = 6.4 Hz, C*H$_2$*CH(CH$_3$)$_2$), 2.16-2.11 (m, 1H, CH$_2$C*H*(CH$_3$)$_2$), 1.06 (d, 6H, J = 6.8 Hz, CH$_2$CH(C*H$_3$*)$_2$).

4-amino-3-isobutoxybenzoic acid (23)[177]
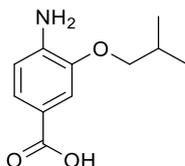
Using the general procedure for ester hydrolysis d). Methyl-4-amino-3-isobutoxybenzoate **22** (2.2 g, 9.9 mmol); sodium hydroxide (22.0 mL, 59.4 mmol); methanol:tetrahydrofuran (1:1, 40 mL). Yellow solid (1.7 g, 83%). δ_H (500 MHz, CDCl$_3$) 7.63 (d, 1H, J = 8.0 Hz, Ar-*H*), 7.48 (s, 1H, Ar-*H*), 6.70 (d, 1H, J = 8.0 Hz, Ar-*H*), 3.85 (d, 2H, J = 6.4 Hz, C*H$_2$*CH(CH$_3$)$_2$), 2.18-2.12 (m, 1H, CH$_2$C*H*(CH$_3$)$_2$), 1.07 (d, 6H, J = 6.6 Hz, CH$_2$CH(C*H$_3$*)$_2$).

4-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-isobutoxybenzoic acid (24)[177]
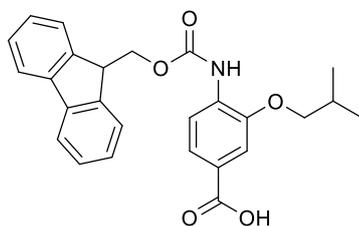
Using the general procedure for Fmoc protection f). 4-amino-3-isobutoxybenzoic acid **23** (1.6 g, 7.8 mmol); fluorenylmethyl chloroformate (2.0 g, 7.8 mmol); anhydrous chloroform (30 mL). The residue was filtered, taken up in chloroform and precipitated with hexane. The procedure was repeated three times, until a pale cream solid, **24**, was obtained (2.2 g, 65%). δ_H (500 MHz, CDCl$_3$) 8.15 (s (broad), 1H, N-*H*) 7.81 (d, 2H, J = 7.5 Hz, Ar-*H*), 7.76 (d, 1H, J = 7.3 Hz, Ar-*H*), 7.64 (d, 2H, 7.5 Hz, Ar-*H*), 7.58 (s, 1H, Ar-*H*), 7.54 (s, 1H, Ar-*H*), 7.45 (t, 2H, J = 7.5 Hz, Ar-*H*), 7.35 (t, 2H, J = 7.3 Hz, Ar-*H*), 4.54 (d, 2H, J = 6.4 Hz, C*H$_2$*(Fmoc)), 4.34 (t, 1H, J = 6.8 Hz, C*H*(Fmoc)), 3.92 (d, 2H, J = 6.4 Hz, C*H$_2$*CH(CH$_3$)$_2$), 2.26-2.18 (m, 1H, CH$_2$C*H*(CH$_3$)$_2$), 1.12 (d, 6H, J = 6.6 Hz, CH$_2$CH(C*H$_3$*)$_2$).

Methyl 3-(3-((*tert*-butoxycarbonyl)amino)propoxy)-4-nitrobenzoate (25)[177]
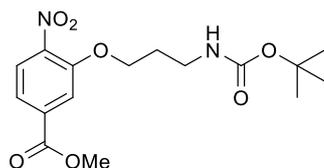
Using the general procedure for *O*-alkylation a). Methyl-4-amino-3-hydroxybenzoate **20** (5.2 g, 26.4 mmol); *tert*-butyl (3-bromopropyl)carbamate (9.4 g, 39.6 mmol); potassium carbonate (10.9 g, 79.2 mmol); *N,N*-dimethylformamide (100 mL). The residue was purified by column chromatography (eluent: hexane/ethyl acetate: 7/3) to give **25** as a yellow solid (3.03 g, 32%). δ_H (CDCl$_3$, 500 Mz) 7.86 (d, 1H, J = 8.3 Hz, Ar-H), 7.74 (s, 1H, Ar-H), 7.70 (td, 1H, J = 8.3 Hz, J = 0.7 Hz, Ar-H), 4.98 (s br., 1H, NH), 4.26 (t, 2H, J = 5.9 Hz, OCH_2), 3.96 (s, 3H, OCH_3), 3.38 (q, 2H, J = 5.9 Hz, CH_2-NH), 2.08 (quint., 2H, J = 5.9 Hz, -CH_2-), 1.44 (s, 9H, (CH_3)$_3$).

Methyl 4-amino-3-(3-((*tert*-butoxycarbonyl)amino)propoxy)benzoate (26)[177]
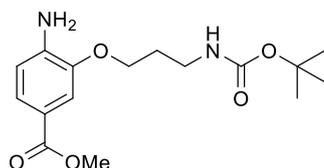
Using the general procedure for nitro reduction b). Methyl 3-(3-((*tert*-butoxycarbonyl)amino)pro-poxy)-4-nitrobenzoate **25** (3.03 g, 8.55 mmol); methanol:ethyle acetate 1:1 (40 mL); palladium 10% on carbon (300 mg). White solid (2.55 g, 93%). δ_H (MeOD, 500 Mz) 7.47 (dd, 1H, J = 8.3 Hz, J = 1.8 Hz, Ar-H), 7.41 (d, 1H, J = 1.8 Hz, Ar-H), 6.70 (d, 1H, J = 8.3 Hz, Ar-H), 4.08 (t, 2H, J = 5.9 Hz, OCH_2), 3.82 (s, 3H, OCH_3), 3.28 (t, 2H, J = 6.6 Hz, CH_2-NH), 2.01 (quint., 2H, J = 6.4 Hz, -CH_2-), 1.43 (s, 9H, (CH_3)$_3$).

4-amino-3-(3-((*tert*-butoxycarbonyl)amino)propoxy)benzoic acid (27)[177]
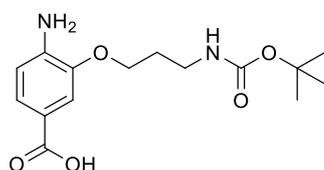
Using the general procedure for ester hydrolysis d). Methyl 4-amino-3-(3-((*tert*-butoxycarbonyl)-amino)propoxy)benzoate **26** (2.55 g, 7.9 mmol); sodium hydroxide solution (20.0 mL, 54.0 mmol); methanol: tetrahydrofuran (30 mL). Hygroscopic brown powder (1.76 g, 72%). δ_H (CDCl$_3$, 500 Mz) 7.64 (d, 1H, J = 8.3 Hz, Ar-H), 7.49 (s br., 1H, Ar-H), 6.69 (d, 1H, J = 8.3 Hz, Ar-H), 4.73 (s br., 1H, NH), 4.15 (t, 2H, J = 6.4 Hz, O-CH_2) 3.38-3.34 (m, 2H, CH_2-NH), 2.05 (t, 2H, J = 6.4 Hz, -CH_2-), 1.45 (s, 9H, C(CH_3)$_3$).

4-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(3-((*tert*-butoxycarbonyl)amino)pro-poxy)benzoic acid (28)[177]
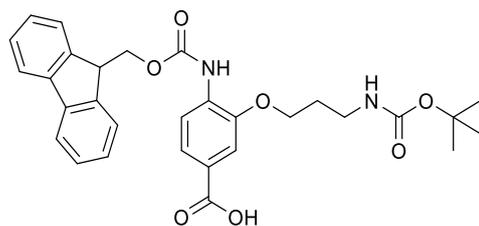
Using the general procedure for Fmoc protection f). 4-amino-3-(3-((*tert*-butoxy-carbonyl)amino)pro-poxy)benzoic acid (1.76 g, 5.7 mmol) **27**; fluorenylmethyl chloroformate (1.76 g, 6.8 mmol); sodium carbonate (1.50 g, 14.2 mmol); anhydrous chloroform (300 mL). The residue was purified by flash chromatography (eluent: dichloromethane/methanol: 10/0 to 0/10) to give the desired product **28** (1.87 g, 62%) as a white powder. δ_H (CDCl$_3$, 500 Mz) 8.14 (s br., 1H), 7.80 (d, 2H, J = 7.6 Hz, Ar-*H*), 7.75 (d, 1H, J = 6.9 Hz, Ar-*H*), 7.66 (d, 3H, J = 7.10, Ar-*H*), 7.58 (s, 1H, Ar-*H*), 7.44 (t, 2H, J = 7.6 Hz, Ar-*H*), 7.35 (t, 2H, J = 7.10 Hz, Ar-*H*), 4.70 (s br., 1H, N*H*), 4.54 (d, 2H, J = 6.9 Hz, C*H*$_2$(Fmoc)), 4.35 (t, 1H, J = 6.9 Hz, C*H*(Fmoc)), 4.19 (s br., 2H, O-C*H*$_2$), 3.39 (s br., 2H, C*H*$_2$-NH), 2.09 (t, 2H, J = 6.2 Hz, -C*H*$_2$-), 1.44 (s, 9H, C(C*H*$_3$)$_3$).

Methyl 3-(2-(*tert*-butoxy)-2-oxoethoxy)-4-nitrobenzoate (29)[177]
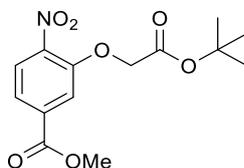
Using the general procedure for *O*-alkylation a). Methyl-4-amino-3-hydroxybenzoate **20** (3.0 g, 15.2 mmol); *tert*-butylbromoacetate (3.9 mL, 26.9 mmol); potassium carbonate (7.4 g, 53.7 mmol) *N,N*-dimethylformamide (60 mL). Brown solid (3.5 g, 73%). δ_H (500 MHz, CDCl$_3$) 7.88 (d, 1H, 8.5 Hz, Ar-*H*), 7.75 (dd, 1H, J = 8.5, 1.4 Hz, Ar-*H*), 7.63 (s, 1H, Ar-*H*), 4.73 (s, 2H, OC*H*$_2$), 3.96 (s, 3H, OC*H*$_3$), 1.45 (s, 9H, OC(C*H*$_3$)$_3$).

3-(2-(*tert*-butoxy)-2-oxoethoxy)-4-nitrobenzoic acid (30)[177]
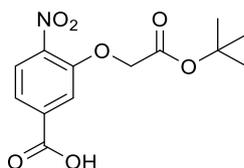
Lithium hydroxide monohydrate (418 mg, 10.0 mmol) in water (10 mL) was added to a solution of methyl 3-(2-(*tert*-butoxy)-2-oxoethoxy)-4-nitrobenzoate **29** (3.1 g, 10.0 mmol) in tetrahydrofuran (20 mL) and the solution was stirred at room temperature for 16 hours. The mixture was washed with dichloromethane (30 mL) and the organic layer was dried over

magnesium sulphate, filtered and evaporated under vacuum to recollect non-reacted starting material; the aqueous layer was adjusted to pH 4 with a 10% aqueous solution of potassium hydrogen sulphate and extracted with dichloromethane (2 x 30 mL). The organic layer was dried with magnesium sulphate, filtered, and the filtrate was evaporated under vacuum to give **30** as a yellow solid (548 mg, 19%). δ_H (500 MHz, CDCl$_3$) 7.91 (d, 1H, J = 8.3 Hz, Ar-H), 7.83 (dd, 1H, J = 8.3 Hz, 1.4 Hz, Ar-H), 7.70 (d, 1H, J = 1.4 Hz, Ar-H), 4.77 (s, 2H, OCH_2), 1.52 (s, 9H, OC(CH_3)$_3$).

4-amino-3-(2-(*tert*-butoxy)-2-oxoethoxy)benzoic acid (31)[177]
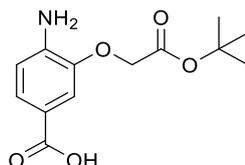
Using the general procedure for nitro reduction b). 3-(2-(tert-butoxy)-2-oxoethoxy)-4-nitrobenzoic acid **30** (649 mg, 2.2 mol); methanol (15 mL); palladium 10% on carbon (65 mg). Beige solid (553 mg, 95%). δ_H (500 MHz, CDCl$_3$) 7.67 (dd, 1H, J = 8.3 Hz, 1.4 Hz, Ar-H), 7.44 (d, J = 1.4 Hz, Ar-H), 6.71 (d, 1H, J = 8.3 Hz), 4.40 (s, 2H, OCH_2), 1.51 (s, 9H, OC(CH_3)$_3$).

4-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(2-(*tert*-butoxy)-2-oxoethoxy)benzoic acid (32)[177]
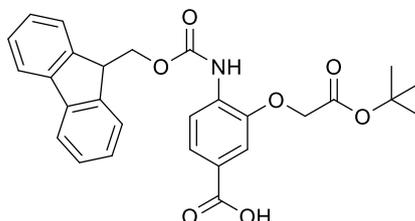
Using the general procedure for Fmoc protection f). 4-amino-3-(2-(*tert*-butoxy)-2-oxoethoxy)-benzoic acid **31** (533 mg, 2.1 mmol); fluorenylmethyl chloroformate (534 mg, 2.1 mmol); sodium hydrogen carbonate (174 mg, 2.1 mmol); anhydrous chloroform (10 mL). The residue was taken up with chloroform and precipitated with hexane three times and finally purified by flash chromatography (eluent: dichloromethane/methanol: 9.5/0.5) to give the target material **32** as a white solid (54 mg, 5%). δ_H (500 MHz, MeOD) 8.05 (s broad, 1H, Ar-H), 7.86 (d, 2H, J = 7.6 Hz, Ar-H), 7.76 (d, 2H, J = 7.6 Hz, Ar-H), 7.70 (d, 1H, J = 7.8 Hz, Ar-H), 7.57 (d, 1H, J = 1.8 Hz, Ar-H), 7.45 (t, 2H, J = 7.6 Hz, Ar-H), 7.38 (td, 2H, J = 7.6 Hz, 1.2 Hz, Ar-H), 4.76 (s, 2H, OCH_2), 4.57 (d, 2H, J = 6.6 Hz, CH_2(Fmoc)), 4.36 (t, 1H, J = 6.6 Hz, CH(Fmoc)), 1.53 (s, 9H, C(CH_3)$_3$).

**Methyl 3-(4-(*tert*-butoxy)-4-oxobutoxy)-4-nitrobenzoate (33)**
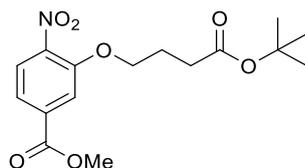
Using the general procedure for *O*-alkylation a). Methyl-4-amino-3-hydroxybenzoate **20** (180 mg, 0.9 mmol); *tert*-butyl 4-bromobutanoate **14** (247 mg, 1.1 mmol); potassium carbonate (380 mg, 2.76 mmol); *N*,*N*-dimethylformamide. White solid (104 mg, 33%). δ_H (500 MHz, CDCl$_3$) 7.84 (d, 1H, *J* = 8.3 Hz, Ar-*H*), 7.74 (d, 2H, *J* = 1.5 Hz, Ar-*H*), 7.70 (d, 1H, *J* = 1.5 Hz, Ar-*H*), 4.24 (t, 2H, *J* = 6.0, -OC*H*$_2$-), 3.97 (s, 3H, -OC*H*$_3$), 2.49 (t, 2H, *J* = 7.3 Hz, -C*H*$_2$CO2(CH$_3$)$_3$), 2.16 (quint., 2H, *J* = 6.3 Hz -C*H*$_2$-), 1.46 (s, 9H, -C(C*H*$_3$)$_3$). δ_C (125 MHz, CDCl$_3$) 172.3, 165.2, 151.7, 142.5, 134.8, 125.3, 121.4, 115.6, 80.6, 68.8, 52.8, 31.4, 28.1, 24.3. ν_{max}/cm^{-1} (solid state) = 2985, 2952, 1720 (COOMe, COO^{t}Bu), 1530, 1293 (NO$_2$). HRMS: Calcd. [M+Na]$^+$ (C$_{16}$H$_{21}$NO$_7$) *m/z* = 362.1216. Found [M+Na]$^+$ *m/z* = 362.1217.

3-(4-(*tert*-butoxy)-4-oxobutoxy)-4-nitrobenzoic acid (34)
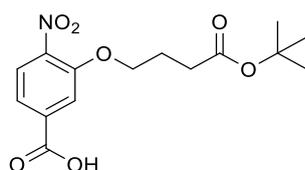
Methyl 3-(4-(*tert*-butoxy)-4-oxobutoxy)-4-nitrobenzoate **33** (650 mg, 1.9 mmol) was dissolved in tetrahydrofuran (25 mL) and water (25 mL), in an ice bath. A 0.25M aqueous solution of sodium hydroxide was added and the mixture was stirred for 4 hours at 0°C. The mixture was then diluted with water (50 mL) and extracted with ethyl acetate (2 x 50 mL). The organic layer was washed with brine (50 mL) and the solvent was evaporated. The residue was purified by column chromatography (eluent: dichloromethane/methanol 10/0 to 9/1) to give the product **34** as a pale yellow solid (406 mg, 66%). δ_H (500 MHz, MeOD) 7.82 (d, 2H, *J* = 8.3 Hz, Ar-*H*), 7.69 (dd, 1H, *J* = 1.3 Hz, 8.3 Hz, Ar-*H*), 4.23 (t, 2H, *J* = 6.0, -OC*H*$_2$-), 2.47 (t, 2H, *J* = 7.3 Hz, -C*H*$_2$CO$_2$(CH$_3$)$_3$), 2.10 (quint., 2H, *J* = 6.4 Hz, -C*H*$_2$-), 1.44 (s, 9H, -C(C*H*$_3$)$_3$). δ_C (125 MHz, MeOD) 199.3, 174.1, 152.6, 143.6, 125.6, 122.6, 116.9, 81.6, 69.9, 32.5, 28.2, 25.6 (1 x C missing). ν_{max}/cm^{-1} (solid state) = 2981, 2931, 1735 (COOH, COO^{t}Bu), 1520, 1287 (NO$_2$). HRMS: Calcd. [M+Na]$^+$ (C$_{15}$H$_{19}$NO$_7$) *m/z* = 348.1059. Found [M+Na]$^+$ *m/z* = 348.1059.

4-amino-3-(4-(*tert*-butoxy)-4-oxobutoxy)benzoic acid (35)
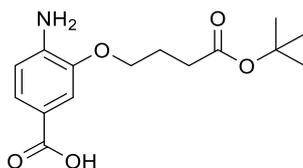
Using the general procedure for nitro reduction b). 3-(4-(*tert*-butoxy)-4-oxobutoxy)-4-nitrobenzoic acid **34** (782 mg, 2.4 mmol); methanol (20 mL); palladium 10% on carbon (78 mg). Brown powder (651 mg, 92%). δ_H (500 MHz, MeOD) 7.50 (d, 2H, J = 11.4 Hz, Ar-H), 6.70 (d, 1H, J = 8.0), 4.07 (t, 2H, J = 6.2 Hz, -OCH_2-), 2.48 (t, 2H, J = 7.3 Hz, -CH_2CO$_2$C(CH$_3$)$_3$), 2.09 (quint., 2H, J = 6.4 Hz, -CH_2-), 1.45 (s, 9H, -C(CH_3)$_3$). δ_C (125 MHz, DMSO-d$_6$): 172.6, 144.6, 124.3, 112.5, 80.1, 67.3, 31.9, 28.3, 24.9 (4 x C missing). ν_{max}/cm^{-1} (solid state) = 3362 (NH) 2925, 1716 (COOH, COO^{t}Bu), 1146, 751 (NH). HRMS: Calcd. [M+Na]$^+$ (C$_{15}$H$_{21}$NO$_5$) m/z = 318.1317. Found [M+Na]$^+$ m/z = 318.1318.

4-(((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-3-(4-(*tert*-butoxy)-4-oxobutoxy)benzoic acid (36)
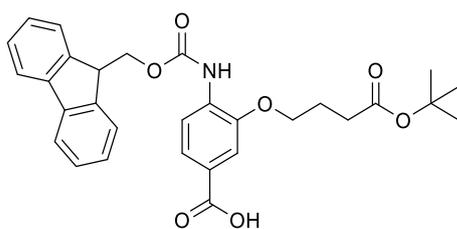
Using the general procedure for Fmoc protection f). 4-amino-3-(4-(*tert*-butoxy)-4-oxobutoxy)-benzoic acid **35** (1.6 g, 5.4 mmol); fluorenylmethyl chloroformate (1.4 g, 5.4 mmol), sodium hydrogen carbonate (498 mg, 5.9 mmol); anhydrous chloroform (50 mL). The residue was purified by flash chromatography (eluent: dichloromethane/methanol: 10/0 to 9/1) and the product **36** was collected as a white powder (1.73 g, 62%). δ_H (500 MHz, CDCl$_3$) 8.18 (s broad, 1H, NH), 7.81 (d, 2H, J = 7.5 Hz, Ar-H), 7.78-7.76 (m, 1H, Ar-H(Fmoc)), 7.69 (d, 1H, J = 7.5 Hz, Ar-H), 7.65-7.62 (m, 2H, Ar-H (Fmoc)), 7.59 (d, 1H, J = 1.5 Hz, Ar-H(Fmoc)), 7.45 (t, 2H, J = 7.6 Hz, Ar-H(Fmoc)), 7.37 (t, 2H, J = 7.6 Hz Ar-H(Fmoc)), 4.55 (d, 2H, J = 7.1 Hz, CH_2(Fmoc)), 4.36 (t, 1H, J = 7.1 Hz, CH(Fmoc)), 4.20 (t, 2H, J = 6.2 Hz, -OCH_2-), 2.50 (t, 2H, J = 7.1 Hz, -CH_2CO$_2$(CH$_3$)$_3$), 2.23 (quint., 2H, J = 6.6 Hz, -CH_2-), 1.48 (s, 9H, CO$_2$(CH_3)$_3$). δ_C (125 MHz, CDCl$_3$) 172.2, 171.0, 153.1, 146.3, 143.7, 141.4, 133.0, 127.9, 127.2, 125.1, 124.5, 123.3, 120.1, 117.4, 112.1, 80.8, 68.3, 67.6, 47.1, 32.2, 28.1, 24.5. ν_{max}/cm^{-1} (solid state) = 3426 (NH amide), 2971, 1715 (COOH, COO^{t}Bu), 1681 (C=O amide) 1195. HRMS: Calcd. [M+Na]$^+$ (C$_{30}$H$_{31}$NO$_7$) m/z = 540.1998. Found [M+]$^+$ m/z = 540.1996.

5.1.3.3. *N*-alkylated monomers (Chapter 2)

4-(isobutylamino)benzoic acid (38)[178]
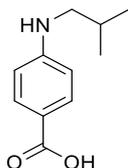
Using the general procedure for reductive amination e). 4-aminobenzoic acid **37** (2.0 g, 14.6 mmol); isobutyraldehyde (1.34 mL, 14.6 mmol); picoline borane (2.8 g, 26.3 mmol); methanol (25 mL). The residue was then taken up with ethyl acetate and precipitated with hexane to obtain the target material **38** as a white solid (2.6 g, 91%). δ_H (500 MHz, CDCl$_3$), 7.92 (d, 2H, J = 8.7 Hz, Ar-*H*), 6.56 (d, 2H, J = 8.7 Hz, Ar-*H*), 3.01 (d, 2H, J = 6.8 Hz, C*H$_2$*CH(CH$_3$)$_2$), 2.76 (s, 1H, N-*H*), 1.94-1.88 (m, 1H, CH$_2$C*H*(CH$_3$)$_2$), 1.03 (d, 6H, J = 6.8 Hz, CH$_2$CH(C*H$_3$*)$_2$).

4-(((((9H-fluoren-9-yl)methoxy)carbonyl)(isobutyl)amino)benzoic acid (39)[178]
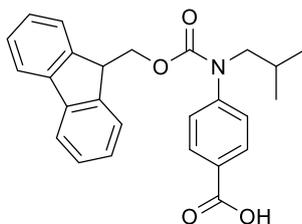
Using the general procedure for Fmoc protection f). 4-(isobutylamino)benzoic acid **38** (2.5 g, 13.2 mol); fluorenylmethyl chloroformate (4.1 g, 15.8 mmol); anhydrous chloroform (70 mL). The solvent was removed and the residue was taken up in chloroform and precipitated with hexane, and the precipitate was isolated by centrifugation, dissolved in chloroform (20 mL) and evaporated under vacuum to get the target **39** as an off-white foam (3.21 g, 59%). δ_H (500 MHz, CDCl$_3$) 8.04 (d, 2H, J = 8.3 Hz, Ar-*H*), 7.72 (d, 2H, J = 7.3 Hz, Ar-*H*), 7.38 (t, 2H, J = 7.3 Hz, Ar-*H*), 7.33 (d, 2H, J = 8.3 Hz, Ar-*H*), 7.24 (t, 2H, J = 7.3 Hz, Ar-*H*), 7.16 (d, 2H, J = 8.3 Hz, Ar-*H*), 4.55 (d, 2H, J = 5.9 Hz, C*H$_2$*(Fmoc)), 4.13 (t, 1H, J = 5.9 Hz, C*H*(Fmoc)), 3.47 (d, 2H, J = 7.3 Hz, C*H$_2$*CH(CH$_3$)$_2$), 1.70-1.62 (m, 1H, CH$_2$C*H*(CH$_3$)$_2$), 0.79 (d, 6H, J = 6.8 Hz, CH$_2$CH(C*H$_3$*)$_2$).

4-((2-(*tert*-butoxy)-2-oxoethyl)amino)benzoic acid (40)[178]
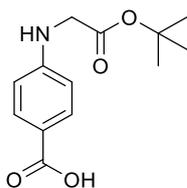
Using the general procedure for reductive amination e). 4-aminobenzoic acid **37** (589 mg, 4.3 mmol); *tert*-butyl glyoxylate **3** (559 mg, crude, ≈3.1 mmol); picoline borane (556 mg, 5.2

mmol); methanol (20 mL); The residue was purified by flash chromatography (eluent: ethyl acetate/hexane: 0/10 to 5/5), yielding target compound **40** as a white solid (335 mg, 44%). δ_H (500 MHz, CDCl$_3$) 7.95 (d, 2H, J = 8.7 Hz, Ar-*H*), 6.58 (d, 2H, J = 8.7 Hz, Ar-*H*), 4.80 (s, 1H, N-*H*), 3.87 (s, 2H, N-C*H$_2$*), 1.51 (s, 9H, C(CH$_3$)$_3$).

4-(((((9H-fluoren-9-yl)methoxy)carbonyl)(2-(*tert*-butoxy)-2-oxoethyl)amino)benzoic acid (41)[178]
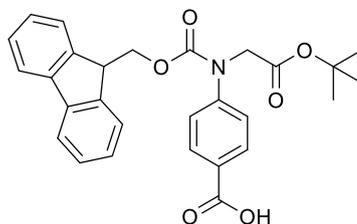
Using the general procedure for Fmoc protection f). 4-((2-(*tert*-butoxy)-2-oxoethyl)amino)-benzoic acid **40** (268 mg, 1.1 mmol); fluorenylmethyl chloroformate (276 mg, 1.1 mmol); sodium hydrogen carbonate (99 mg, 1.2 mmol); anhydrous chloroform (10 mL). The solvent was removed and the residue was taken up in chloroform and precipitated with hexane, and the precipitate was isolated by centrifugation. The solid was dissolved in chloroform and evaporated under vacuum to give the desired product **41** as an off-white foam (263 mg, 52%). δ_H (500 MHz, CDCl$_3$): 8.08 (d, 2H, J = 8.3 Hz, Ar-*H*), 7.74 (d, 2H, J = 7.6 Hz, Ar-*H*), 7.39 (t broad, 3H, J = 7.3 Hz, Ar-*H*) 7.33 (m broad, 3H, Ar-*H*), 7.24 (m broad, 2H, Ar-*H*), 4.51 (d, 2H, J = 6.6 Hz, C*H$_2$*(Fmoc)), 4.27 (s, 2H, N-C*H$_2$*), 4.16 (s broad, 1H, C*H*(Fmoc)), 1.46 (s, 9H, C(C*H$_3$*)$_3$).

4-((4-(*tert*-butoxy)-4-oxobutyl)amino)benzoic acid (42)
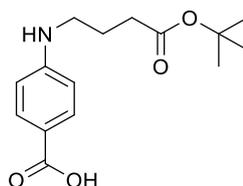
Using the general procedure for reductive amination e). 4-aminobenzoic acid **37** (2.5 g, 18.2 mmol); *tert*-butyl-4-oxobutanoate **13** (2.87 g, 18.2 mmol); picoline borane (2.3 g, 21.8 mmol); methanol (100 mL). The obtained pink residue was filtered and washed with dichloromethane and the product **42** was collected as a white powder (2.58 g, 51%). δ_H (500 MHz, MeOD) 7.78 (d, 2H, J = 9.0 Hz, Ar-*H*), 6.59 (d, 2H, J = 9.0 Hz, Ar-*H*), 3.18 (t, 2H, J = 7.0 Hz, CO$_2$-C*H$_2$*), 2.34 (t, 2H, J = 7.0 Hz, NH-C*H$_2$*), 1.87 (quint., 2H, J = 7.0 Hz, CH$_2$-C*H$_2$*-CH$_2$), 1.45 (s, 9H, -C(C*H$_3$*)$_3$). δ_C (125 MHz, MeOD) 174.8, 170.9, 154.6, 132.9, 118.3, 112.2, 81.7, 43.4, 33.9, 28.5, 25.7. ν_{max}/cm^{-1} (solid state) = 3382 (NH), 2959, 2933, 2520 (COOH, COO^{t}Bu), 1705 (COO^{t}Bu) 1659, 1414, 1159 (COOH). HRMS: Calcd. [M+H]$^+$ (C$_{15}$H$_2$NO$_4$) m/z = 280.1504. Found [M+H]$^+$ m/z = 280.1547.

4-(((((9H-fluoren-9-yl)methoxy)carbonyl)(4-(*tert*-butoxy)-4-oxobutyl)amino)benzoic acid (43)
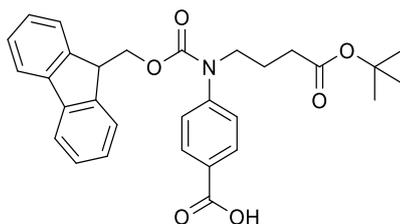
Using the general procedure for Fmoc protection f). 4-((4-(*tert*-butoxy)-4-oxobutyl)-amino)benzoic acid **42** (2.5 g, 8.6 mol); fluorenylmethyl chloroformate (2.2 g, 8.6 mmol); sodium hydrogen carbonate (798 mg, 9.5 mmol); anhydrous chloroform (50 mL) The residue was purified by flash chromatography (eluent: dichloromethane/methanol: 10/0 to 8.5/1.5) and product **43** was collected as a white powder (2.6 g, 59%). δ_H (500 MHz, CDCl$_3$): 8.07 (d, 2H, J = 8.2 Hz Ar-H), 7.73 (d, 2H, J = 7.6 Hz, Ar-H), 7.39-7.33 (m, 4H, Ar-H), 7.24 (t, 2H, J = 7.3 Hz, Ar-H), 7.19 (d, 2H, J = 8.2 Hz, Ar-H), 4.56 (d, 2H, J = 5.8 Hz, CH_2(Fmoc)), 4.14 (t, 1H, J = 5.5 Hz, CH(Fmoc)), 3.65 (t, 2H, J = 7.0 Hz, CO$_2$CH_2), 2.15 (t, 2H, J = 7.0 Hz, NH-CH_2), 1.74 (quint., 2H, J = 7.0 Hz, CH$_2$-CH_2-CH$_2$), 1.42 (s, 9H, -C(CH_3)$_3$). δ_C (125 MHz, CDCl$_3$): 172.1, 170.9, 154.8, 146.5, 143.6, 141.4, 131.1 127.7, 127.0, 126.7, 124.7, 119.9, 80.5, 67.3, 49.4, 47.2, 32.4, 28.1, 23.7 (3 x C missing). ν_{max}/cm^{-1} (solid state) = 2975, 2656 (COOH, COO^{t}Bu), 1685, 1418, 1146 (COOH). HRMS: Calcd. [M+Na]$^+$ (C$_{30}$H$_{31}$NO$_6$) m/z = 524.2083. Found [M+Na]$^+$ m/z = 524.2050.

4-((3-((*tert*-butoxycarbonyl)amino)propyl)amino)benzoic acid (44)[178]
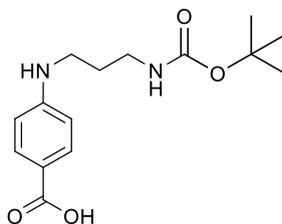
Using the general procedure for reductive amination e). 4-aminobenzoic acid **37** (792 mg, 5.8 mmol); *tert*-butyl (3-oxopropyl)carbamate **19** (1.0 g, 5.8 mmol); picoline borane (74 mg, 6.9 mmol); methanol (40 mL). The residue was precipitated from diethyl ether and water to give **44** as a white powder (943 mg, 55%). δ_H (500 MHz, CDCl$_3$) 7.92 (d, 2H, J = 8.7 Hz, Ar-H), 6.59 (d, 2H, J = 8.7 Hz, Ar-H), 4.64 (s, 1H, N-H), 3.28-3.24 (m, 4H, NCH_2CH$_2$CH_2N), 1.80 (t, 2H, J = 5.9 Hz, NCH$_2$CH_2CH$_2$N), 1.46 (s, 9H, C(CH_3)$_3$).

4-((((9H-fluoren-9-yl)methoxy)carbonyl)(3-((*tert*-butoxycarbonyl)amino)propyl)-amino)benzoic acid (45)[178]
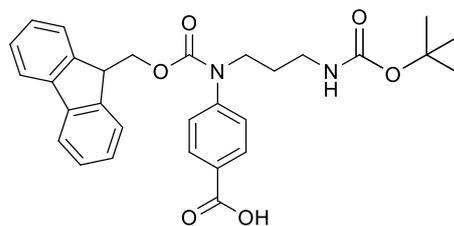
Using the general procedure for Fmoc protection f). 4-((3-((*tert*-butoxycarbonyl)amino)-propyl)amino)benzoic acid **44** (202 mg, 0.7 mmol); fluorenylmethyl chloroformate (248 mg, 0.9 mmol); sodium hydrogen carbonate (0.5g, 5.9 mmol); anhydrous chloroform (10 mL). The residue was taken up in chloroform and precipitated with hexane, and the precipitate was isolated by centrifugation, dissolved in chloroform and evaporated under vacuum to get the target **45** as a white foam (123 mg, 35%). δ_H (500 MHz, CDCl$_3$) 8.01 (d, 2H, J = 8.3 Hz, Ar-*H*), 7.72 (d, 2H, J = 7.3 Hz, Ar-*H*), 7.39 (t, 2H, J = 7.3 Hz, Ar-*H*), 7.31 (d broad, 2H, J = 5.5 Hz, Ar-*H*), 7.27 (t, 2H, J = 7.8 Hz, Ar-*H*), 7.12 (d, 2H, J = 7.8 Hz, Ar-*H*), 4.85 (s, 1H, N-*H*), 4.58 (d, 2H, J = 5.5 Hz, C*H$_2$*(Fmoc)), 4.12 (t, 1H, J = 5.5 Hz, C*H*(Fmoc)), 3.66 (t, 2H, J = 5.5 Hz C*H$_2$*-N), 3.02 (t broad, 2H, J = 6.4 Hz, C*H$_2$*-N), 1.57 (m broad, 2H, NCH$_2$C*H$_2$*CH$_2$N), 1.47 (s, 9H, C(C*H$_3$*)$_3$).

5.1.3.4. Alkene functionalised *N*-alkylated monomers (Chapter 3)

Methyl 4-nitro-3-(prop-2-yn-1-yloxy)benzoate (65)[215]
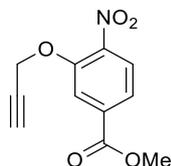
Using the general procedure for *O*-alkylation a). Methyl 3-hydroxy-4-nitrobenzoate **20** (5.0 g, 25.4 mmol); propargyl bromide **64** (4.26 ml, 38.2 mmol); potassium carbonate (18.9 g, 136.5 mmol); *N,N*-dimethylformamide (100 ml). Pale yellow solid (6.0 g, quant.). δ_H (500 MHz, CDCl$_3$) 7.94 (d, 1H, J = 1.3 Hz, Ar-*H*), 7.89 (d, 1H, J = 8.7 Hz, Ar-*H*), 7.79 (dd, 1H, J = 8.7 Hz, 1.3 Hz, Ar-*H*), 4.93 (s, 2H, C*H$_2$*C≡CH), 4.01 (s, 3H, OC*H$_3$*), 2.64 (t, 1H, J = 2.3 Hz, CH$_2$C≡C*H*).

Methyl 4-amino-3-(prop-2-ynyloxy)benzoate (66) [215]
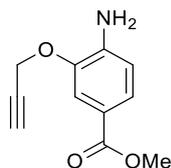
Using the general procedure for nitro reduction c). Tin (II) chloride dihydrate (28.6 g, 127.0 mmol); methyl 4-nitro-3-(prop-2-ynyloxy)benzoate **65** (6.0 g, 25.4 mmol); ethyl acetate (300

mL). Orange solid (5.0 g, 96%). δ_H (500 MHz, CDCl$_3$) 7.62 (dd, 1H, J = 8.3, 1.8 Hz, Ar-H), 7.59 (d, 1H, J = 1.8 Hz, Ar-H), 6.75 (d, 1H, J = 8.3 Hz, Ar-H), 4.81 (d, 2H, J = 2.3 Hz, CH_2C≡CH), 3.89 (s, 3H, CH_3), 2.57 (t, 1H, J = 2.3 Hz, CH$_2$C≡CH).

4-amino-3-(prop-2-yn-1-yloxy)benzoic acid (67)[215]
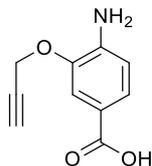
Using the general procedure for ester hydrolysis d). Methyl 4-amino-3-(prop-2-ynyloxy)benzoate **66** (2.9 g, 14.1 mmol); sodium hydroxide 20% (21 mL, 42.3 mmol); methanol:tetrahydrofuran (100 mL). Beige solid (2.5 g, 93%). δ_H (500 MHz, MeOD) 7.61 (d, 1H, J = 1.8 Hz, Ar-H), 7.56 (dd, 1H, J = 8.3, 1.8 Hz Hz, Ar-H), 6.76 (d, 1H, J = 8.3 Hz, Ar-H), 4.83 (d, 2H, J = 2.5 Hz CH_2C≡CH), 3.01 (t, 1H, J = 2.5 Hz, CH$_2$C≡CH).

4-[(2-methylpropyl)amino]-3-(prop-2-yn-1-yloxy)benzoic acid (68)
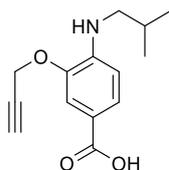
Using the general procedure for reductive amination e). 4-Amino-3-(prop-2-yn-1-yloxy) benzoic acid **67** (3.92 g, 20.5 mmol); isobutyraldehyde (1.87 mL, 20.5 mmol); picoline borane (2.63 g, 24.6 mmol); methanol (100 mL). Pale yellow solid (5.08 g, quant.). δ_H (500 MHz, DMSO-d$_6$) 12.08 (s, 1H, OH) 7.47 (dd, 1H, J = 8.1 Hz, 1.7 Hz, Ar-H), 7.41 (d, 1H, J = 1.7 Hz, Ar-H), 6.56 (d, 1H, J = 8.1 Hz, Ar-H), 5.59 (t, 1H, J = 5.9 Hz, N-H) 4.83 (d, 2H, J = 2.1 Hz, CH_2C≡CH), 3.57 (t, 1H, J = 2.1 Hz, CH$_2$C≡CH), 2.9 (t, 2H, J = 6.4 Hz, CH_2CH(CH$_3$)$_2$), 1.81-1.90 (m, 1H, CH$_2$CH(CH$_3$)$_2$), 0.90 (d, 6H, J = 6.4 Hz, CH$_2$CH(CH_3)$_2$). δ_C (125 MHz, DMSO-d$_6$) 167.4, 142.9, 142.8, 125.0, 116.2, 112.2, 107.8, 79.2, 78.4, 55.9, 49.8, 27.1, 20.2, 20.0. v_{max}/cm^{-1} (solid state) = 3401 (NH), 3278 (CH alkyne), 2961 (COOH), 2122 (C≡C), 1659, 1410, 1280 (COOH). HRMS: Calcd. [M+H]$^+$ (C$_{14}$H$_{17}$NO$_3$) m/z = 248.1208. Found [M+H]$^+$ m/z = 248.1283.

4-((((9H-fluoren-9-yl)methoxy)carbonyl)(isobutyl)amino)-3-(prop-2-yn-1-yloxy)benzoic acid (69)
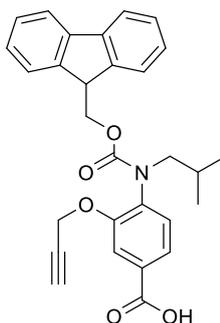
Using the general procedure for Fmoc protection f). 4-[(2-methylpropyl)amino]-3-(prop-2-yn-1-yloxy)benzoic acid **68** (4.97 g, 20.1 mmol); fluorenylmethyl chloroformate (6.24 g, 24.1 mmol); anhydrous chloroform (100 ml). The residue was purified by flash chromatography (eluent: ethyl acetate/hexane: 0/10 to 5/5). The collected oil was taken up in chloroform and precipitated with hexane, and the precipitate was isolated by centrifugation, dissolved in chloroform and evaporated under vacuum to get **69** as an off-white foam (1.55 g, 16%). δ_H (500 MHz, DMSO-d$_6$) 13.09 (s, 1H, C(O)O*H*) 7.78 (s broad, 2H, Ar-*H*), 7.68 (s, 1H, Ar-*H*), 7.56 (d, 1H, *J* = 7.7 Hz, Ar-*H*), 7.34 (s broad, 2H, Ar-*H*), 7.21 (d, 2H, *J* = 8.1 Hz, Ar-*H*), 7.15 (s, 3H, Ar-*H*), 4.83 (s, 2H, C*H$_2$*C≡CH), 4.22 (s broad, 2H, C*H$_2$*(Fmoc)), 4.02 (s broad, 1H, C*H*(Fmoc)), 3.54 (s, 2H, C*H$_2$*CH(CH$_3$)$_2$), 1.61 (s, 1H, CH$_2$C*H*(CH$_3$)$_2$), 0.81 (s, 6H, CH$_2$CH(C*H$_3$*)$_2$). δ_c (125 MHz, DMSO-d$_6$) 166.6, 154.7, 152.7, 143.5, 140.6, 134.4, 130.7, 129.6, 127.4, 126.8, 124.9, 122.3, 119.9, 114.1, 79.2, 78.6, 66.7, 56.1, 55.9, 46.4, 26.8, 19.9. ν_{max}/cm^{-1} (solid state) = 2957, 1685, 1406, 1285 (COOH). HRMS: Calcd. [M+Na]$^+$ (C$_{14}$H$_{17}$NO$_3$) *m/z* = 492.1709. Found [M+Na]$^+$ *m/z* = 492.1787.

Methyl 4-nitro-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoate (78)
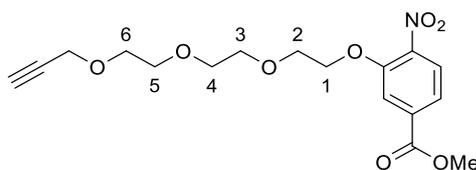
Using the general procedure for *O*-alkylation a). Methyl 3-hydroxy-4-nitrobenzoate **20** (1.0 g, 5.1 mmol); 3-(2-(2-(2-chloroethoxy)ethoxy)ethoxy)prop-1-yne **77** (4:1 Cl:Br, 1.1 g, 5.1 mmol); potassium carbonate (2.1 g, 15.3 mmol); *N,N*-dimethylformamide (50 ml). 70°C. The residue was purified by flash chromatography (eluent: dichloromethane/methanol 10/0 to 9/1) and **78** was collected as a pale oil (957 mg, 51%). δ_H (500 MHz, CDCl$_3$) 7.82 (d, 1H, *J* = 8.4 Hz, Ar-*H*), 7.78 (d, 1H, *J* = 1.6 Hz, Ar-*H*), 7.70 (dd, 1H, *J* = 8.4, 1.6 Hz, Ar-*H*), 4.33 (t, 2H, *J* = 4.6 Hz, *H$_{5 or 6}$*), 4.19 (d, 2H, *J* = 2.3 Hz, C*H$_2$*C≡CH), 3.95 (s, 3H, OC*H$_3$*), 3.92 (t, 2H, *J* = 4.6 Hz, *H$_{5 or 6}$*), 3.75-3.73 (m, 2H, *H$_{1-4}$*), 3.70-3.66 (m, 6H, *H$_{1-4}$*), 2.43 (t, 1H, *J* = 2.3 Hz, CH$_2$C≡C*H*). δ_C (125 MHz, CDCl$_3$) 165.2, 151.8, 142.8, 34.8, 125.3, 121.7, 116.0, 79.7, 74.5, 71.1, 70.7, 70.4, 69.9, 39.2,

69.1, 58.4, 52.8. ν_{max}/cm^{-1} (solid state) = 3283 (CH alkyne), 2870 (C≡C), 1293 (COOMe), 1241 (C-O ether). HRMS: Calcd. [M+Na]$^+$ (C$_{17}$H$_{21}$NO$_8$) m/z = 390.1165 Found [M+Na]$^+$ m/z = 390.1163.

Methyl 4-amino-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoate (79)
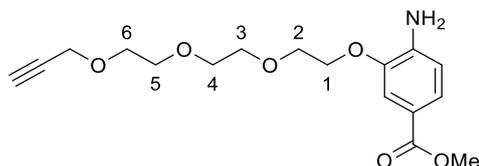
Using the general procedure for nitro reduction c). Tin (II) chloride dihydrate (4.0 g, 17.9 mmol); methyl 4-nitro-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoate **78** (1.3 g, 3.6 mmol); ethyl acetate (50 mL). The residue was purified by flash chromatography (eluent: dichloromethane/methanol: 9.5/0.5) to get the desired product **79** as a yellow oil (436 mg, 36%). δ_H (500 MHz, CDCl$_3$) 7.54 (dd, 1H, J = 8.1, 1.7 Hz, Ar-H), 7.49 (d, 1H, J = 1.7 Hz, Ar-H), 6.68 (d, 1H, J = 8.1 Hz, Ar-H), 4.42 (s broad, 2H, NH_2), 4.23 (q, 4H, J = 2.3 Hz, H_{5-6}), 3.87-3.84 (m, 2H, CH_2C≡CH), 3.86 (s, 3H, OCH_3), 3.75-3.67 (m, 8H, H_{1-4}), 2.45 (t, 1H, J = 2.3 Hz, CH$_2$C≡CH). δ_C (125 MHz, CDCl$_3$) 166.8, 144.6, 141.8, 124.3, 118.8, 113.7, 113.0, 79.2, 74.2, 70.2, 70.0, 69.2, 68.7, 68.2, 60.0, 51.2 (1 x C missing). ν_{max}/cm^{-1} (solid state) = 3478 (NH), 3361 (CH alkyne), 3252, 2872 (CH), 2113 (C≡C), 1291 (COOMe), 1261 (C-O ether). HRMS: Calcd. [M+H]$^+$ (C$_{17}$H$_{23}$NO$_6$) m/z = 338.1559. Found [M+H]$^+$ m/z = 338.1622.

4-amino-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoic acid (80)
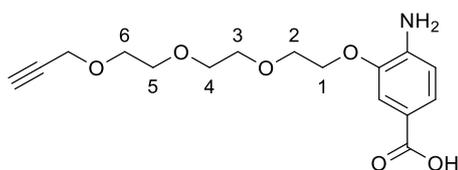
Using the general procedure for ester hydrolysis d). Methyl 4-amino-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoate **79** (400 mg, 1.2 mmol); sodium hydroxide 10% (10 mL, 11.8 mmol); tetrahydrofuran (10 mL). Pale yellow oil (62 mg, 16%). δ_H (500 MHz, CDCl$_3$) 7.65 (dd, 1H, J = 8.1, 1.7 Hz, Ar-H), 7.53 (d, 1H, J = 1.7 Hz, Ar-H), 6.69 (d, 1H, J = 8.1 Hz, Ar-H), 4.23 (m, 4H, H_{5-6}), 3.88 (m, 2H, CH_2C≡CH), 3.70 (m, 8H, H_{1-4}), 2.46 (t, 1H, J = 2.3 Hz, CH$_2$C≡CH). δ_c (125 MHz, CDCl$_3$) 171.7, 144.9, 143.1, 125.7, 118.0, 114.5, 113.3, 79.6, 74.6, 70.6, 70.6, 70.4, 69.6, 69.1, 68.6, 58.4. ν_{max}/cm^{-1} (solid state) = 3478 (NH), 3359 (CH alkyne), 3286, 3012 (CO2H) 2871 (CH), 2119 (C≡C), 1291 (CO2H), 1243 (C-O ether). HRMS: Calcd. [M+Na]$^+$ (C$_{16}$H$_{21}$NO$_7$) m/z = 346.1267. Found [M+Na]$^+$ m/z = 346.1280.

4-(isobutylamino)-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoic acid (81)
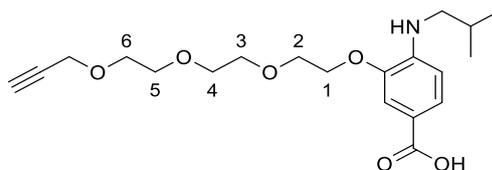
Using the general procedure for reductive amination e). 4-amino-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoic acid **80** (55 mg, 0.17 mmol); isobutyraldehyde (16 μL, 0.17 mmol); picoline borane (22 mg, 0.20 mmol); methanol (2 mL). The residue was purified by flash chromatography (eluent: dichloromethane) to give **81** as a pale yellow oil (51 mg, 79%). δ_H (500 MHz, CDCl$_3$) 7.72 (dd, 1H, J = 8.5, 1.6 Hz, Ar-H), 7.47 (d, 1H, J = 1.6 Hz, Ar-H), 6.55 (d, 1H, J = 8.5 Hz, Ar-H), 4.23 (t, 2H, J = 4.6 Hz, H_5 or H_6), 4.20 (d, 2H, J = 2.3 Hz, CH_2C≡CH), 3.90 (t, 2H, J = 4.6 Hz, H_5 or H_6), 3.74-3.68 (m, 8H, H_{1-4}), 3.03 (d, 2H, 6.8 Hz, CH_2CH(CH$_3$)$_2$), 2.44 (t, 1H, J = 2.3 Hz, CH$_2$C≡CH), 1.97 (quint., 1H, CH$_2$CH(CH$_3$)$_2$), 1.00 (d, 6H, J = 6.6 Hz, CH$_2$CH(CH_3)$_2$). δ_c (125 MHz, CDCl$_3$) 172.2, 144.4, 143.9, 126.2, 115.5, 112.4, 107.8, 79.6, 74.5, 70.6, 70.6, 70.4, 69.6, 69.0, 68.2, 58.3, 50.6, 27.9, 20.4. ν_{max}/cm^{-1} (solid state) = 3440 (NH), 3312 (CH alkyne), 2953 (CO2H), 2874 (CH), 2061 (C≡C), 1279 (CO2H), 1245 (C-O ether). HRMS: Calcd. [M+Na]$^+$ (C$_{20}$H$_{29}$NO$_6$) m/z = 402.1926. Found [M+Na]$^+$ m/z = 402.1908.

4-((((9H-fluoren-9-yl)methoxy)carbonyl)(isobutyl)amino)-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoic acid (82)
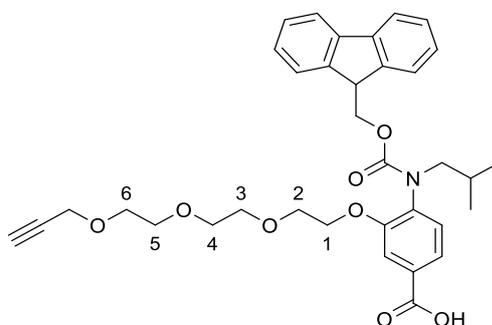
Using the general procedure for Fmoc protection f). 4-(isobutylamino)-3-(2-(2-(2-(prop-2-yn-1-yloxy)ethoxy)ethoxy)ethoxy)benzoic acid **81** (51 mg, 0.13 mmol); fluorenylmethyl chloroformate (35 mg, 0.13 mmol); anhydrous chloroform (1 mL). The residue was purified by flash chromatography (eluent: dichloromthane/methanol: 10/0 to 9.5/0.5). Product **82** was collected as a pale brown oil (40 mg, 51%). δ_H (500 MHz, CDCl$_3$) 7.72 (dd, 1H, J = 8.0, 1.6 Hz, Ar-H), 7.69-7.62 (m, 3H, Ar-H), 7.35 (t broad, J = 6.6 Hz, 2H, Ar-H), 7.16 (m broad, 4H, Ar-H), 7.09 (s broad, 1H, Ar-H), 4.46 (s broad, 1H, H_5 or H_6), 4.29 (s broad, 1H, H_5 or H_6), 4.18 (d, J = 2.3 Hz, 2H, CH_2C≡CH), 4.16-4.11 (m, 2H, H_5 or H_6), 4.02 (m, 2H, H_5 or H_6), 3.98 (s broad, 1H, CH(Fmoc)) 3.75 (s broad, 2H, CH_2(Fmoc)), 3.68-3.56 (m, 10H, CH_2CH(CH$_3$)$_2$ and H_{1-4}), 2.42 (t, J = 2.4 Hz, 1H, CH$_2$C≡CH), 1.75 (quint. broad, 1H, CH$_2$CH(CH$_3$)$_2$), 0.93 (d, 6H, J = 5.5 Hz,

CH$_2$CH(C*H$_3$*)$_2$). δ$_c$ (125 MHz, CDCl$_3$) 170.7, 155.8, 154.5, 143.9, 141.3, 129.7, 129.3, 127.5, 126.9, 124.9, 123.0, 120.1, 119.8, 114.1, 79.7, 74.6, 70.8, 70.6, 70.4, 69.5, 69.1, 68.1, 67.3, 58.4, 56.8, 47.1, 27.5, 20.1. ν$_{max}$/cm^{-1} (solid state) = 3247 (CH alkyne), 2924 (CO$_2$H), 2870 (CH), 2116 (C≡C), 1284 (CO$_2$H). HRMS: Calcd. [M+H]$^+$ (C$_{35}$H$_{39}$NO$_8$) *m/z* = 602.2676. Found [M+H]$^+$ *m/z* = 602.2668.

5.1.4. Azide synthesis (Chapter 3)

tert-**Butyl (3-azidopropyl)carbamate (70)**[215]
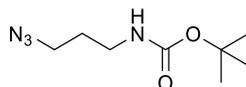
A solution of *tert*-butyl (3-bromopropyl)carbamate **16** (5.0 g, 21.0 mmol) and sodium azide (2.73 g, 42.0 mmol) in water:dioxane 1:1 (30 mL) was refluxed for 2 hours, behind a safety screen. The solution was allowed to cool down to room temperature, and stirred at room temperature for 16 hours. The mixture was diluted with dichloromethane (30 mL) and washed with water (2 x 25 mL). The organic layer was dried over magnesium sulphate, filtered and evaporated to give product **70** as a pale yellow oil (2.94 g, 70%), used without further purification. δ$_H$ (500 MHz, CDCl$_3$) 4.71 (s, 1H, N*H*), 3.36 (t, 2H, *J* = 6.64 Hz, N$_3$C*H$_2$*-), 3.21 (q, 2H, *J* = 6.64 Hz, -C*H$_2$*NH-), 1.79 (quint., 2H, *J* = 6.64 Hz, -C*H$_2$*-), 1.43 (s, 9H, OC(C*H$_3$*)$_3$).

3-chloropropylamine hydrochloride (71) [215]

tert-butyl (3-azidopropyl)carbamate **70** (1.0 g, 5.0 mmol) was dissolved in methanolic hydrochloric acid 3M and the solution was stirred at room temperature for 2 hours, behind a safety screen. The solvent was then evaporated to give the hydrochloric acid salt **71** as a white solid (663 mg, 98%). δ$_H$ (500 MHz, MeOD): 3.35 (t, 2H, *J* = 6.42 Hz, N$_3$CH$_2$-), 3.07 (t, 2H, *J* = 7.33 Hz, -C*H$_2$*NH$_2$), 1.97 (quint., 2H, *J* = 7.33, -C*H$_2$*-).

5-(3-(3-azidopropyl)thioureido)-2-(6-hydroxy-3-oxo-3H-xanthen-9-yl)benzoic acid (73, Flu-azide)[215]
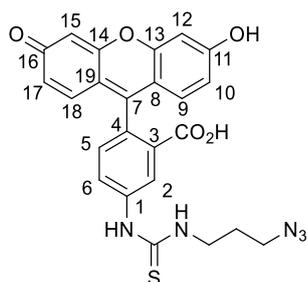
Fluorescein isocyanate isomer I (700 mg, 1.8 mmol) was dissolved in anhydrous *N,N*-dimethylformamide (60 mL). Triethylamine (5.6 mL, 4.0 mmol) and 3-chloropropylamine hydrochloride **71** (272 mg, 2.0 mmol) were added and the mixture was stirred for 16 hours at room temperature, in the dark. The solution was diluted with ethyl acetate (200 mL) and washed with hydrochloric acid 1M (2 x 200 mL) and brine (2 x 200 mL). The organic layer was dried with magnesium sulphate, filtered and the filtrate was dried under reduced pressure to afford **73** as an orange solid (710 mg, 81%). δ_H (500 MHz, MeOD) 8.12 (d, 1H, $J = 1.8$ Hz, Ar-H_2), 7.77 (dd, 1H, $J = 8.3$ Hz, 2.0 Hz, Ar-H_6), 7.16 (d, 1H, $J = 8.25$ Hz, Ar-H_5), 6.68 (t, 4H, $J = 2.5$ Hz, Ar-$H_{9, 10, 17, 18}$), 6.55 (dd, 2H, $J = 8.7$, 2.3 Hz, Ar-$H_{12, 15}$), 3.71 (s br, 2H, CH_2N$_3$), 3.45 (t, 2H, $J = 6.6$ Hz, CH_2NH), 1.94 (quint., 2H, $J = 6.6$ Hz, -CH_2-). δ_C (250 MHz, MeOD) 182.9, 172.8, 170.9, 161.9, 154.0, 142.2, 132.0, 130.1, 128.8, 125.5, 119.9, 13.4, 111.3, 103.3, 50.4, 42.9, 24.0, 20.6.

(9H-fluoren-9-yl)methyl (2-(2-(2-((3-azidopropyl)amino)-2-oxoethoxy)ethoxy)ethyl)carba-mate (74)
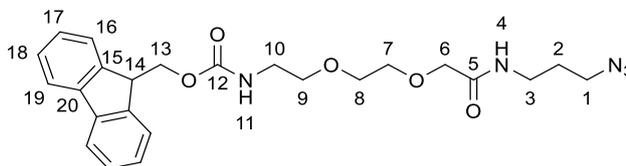
To a solution of 3-azidopropylamine hydrochloride **71** (1.1 g, 7.8 mmol) and 8-(Fmoc-amino)-3,6-dioxaoctanoic acid (3.0 g, 7.8 mmol) in anhydrous acetonitrile (40 mL) were added PyBOP (4.9 g, 9.4 mmol) and *N,N*-diisopropylethylamine (4.1 mL, 23.4 mmol) and the mixture was stirred at room temperature for 16 hours. The mixture was diluted with ethyl acetate (2x 50 mL), washed with a 10%w solution of citric acid (2x 50 mL), saturated sodium hydrogen carbonate (50 mL) and brine (50 mL). The organic layer was dried over magnesium sulphate, and the solvent was removed under vacuum and the desired product **74** was obtained as a yellow oil (3.0 g, 82%). δ_H (500 MHz, CDCl$_3$): 7.77 (d, 2H, $J = 7.5$ Hz, Ar-H_{19}), 7.60 (d, 2H, $J = 7.5$ Hz, Ar-H_{16}), 7.41 (t, 2H, $J = 7.5$ Hz, Ar-H_{17}), 7.32 (td, 2H, $J = 7.5$ Hz, $J = 0.8$ Hz, Ar-H_{18}), 6.99 (s, 1H, NH_{11}), 5.31 (s, 1H, NH_4), 4.42 (d, 2H, $J = 6.6$ Hz, H_{13}), 4.23 (t, 1H, $J = 6.6$ Hz, H_{14}), 3.98 (s, 2H, H_6), 3.69-3.60 (m, 4H, $H_{9,10}$), 3.57 (d, 2H, $J = 4.7$ Hz, H_1), 3.41 (d, 2H, $J = 4.7$ Hz, H_3), 3.37-3.31 (m, 4H, $H_{7,8}$), 1.81 (quint., 2H, $J = 6.6$ Hz, H_2). δ_D (250 MHz, CDCl$_3$): 169.8, 156.4, 143.8, 141.2, 127.6, 126.9, 124.8, 119.9, 70.8, 70.4, 70.1, 69.9, 66.5, 49.1, 47.1, 40.7, 36.2, 28.6. ν_{max}/cm^{-1} (solid state) = 3041 (NH), 2885 (CH), 2093 (N$_3$), 1663 (C=O amide), 1261 (C-O ether). HRMS: Calcd. [M+H]$^+$ (C$_{24}$H$_{29}$N$_5$O$_5$) $m/z = 468.2241$. Found [M+H]$^+$ $m/z = 468.2272$.

2-(2-(2-aminoethoxy)ethoxy)-*N*-(3-azidopropyl)acetamide (75)
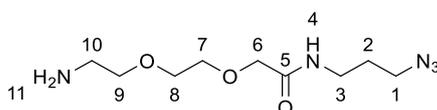
A solution of diethylamine in dichloromethane (5.4 mL, 50% volume, 26 mmol) was added to a solution of (9H-fluoren-9-yl)methyl (2-(2-(2-((3-azidopropyl)amino)-2-oxoethoxy)ethoxy)-ethyl)carbamate **74** (1.2 mg, 2.6 mmol) in dichloromethane (50 mL). The reaction mixture was stirred for 16 hours, then the solvent was removed under vacuum and the residue was purified by column chromatography (eluent: dichloromethane/methanol: 8.5/1.5). Compound **75** was obtained as a pale yellow oil (500 mg, 78%). δ_H (500 MHz, CDCl$_3$) 7.18 (s, 1H, NH_4), 4.83 (s br., 2H, NH_{11}) 4.06 (s, 2H, H_6), 3.73-3.68 (m, 6H, $H_{1,9,10}$), 3.41-3.37 (m, 4H, $H_{7,8}$), 3.11 (t br., 2H, H_3), 1.84 (quint., 2H, J = 6.6 Hz, H_2). δ_C (125 MHz, CDCl$_3$) 170.3, 70.9, 70.7, 70.4, 69.6, 49.3, 40.4, 36.5, 28.8.

N-(2-(2-(2-((3-azidopropyl)amino)-2-oxoethoxy)ethoxy)ethyl)-5-((4S)-2-oxohexahydro-1H-thieno[3,4-d]imidazol-4-yl)pentanamide (76, Biotin-PEG-azide**)**
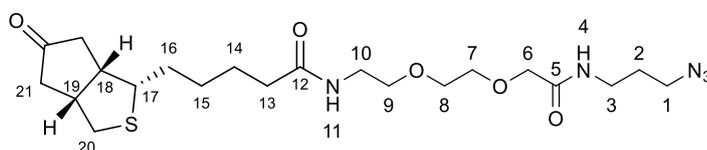
N,N-Diisopropylethylamine (185 μl, 1.06 mmol) and HCTU (328 mg, 0.79 mmol) were added to a solution of D-biotin (129 mg, 0.53 mmol) in dimethylformamide (4 mL). 2-(2-(2-aminoethoxy)ethoxy)-N-(3-azidopropyl)acetamide **75** (130 mg, 0.53 mmol) in dimethyl-formamide (1 mL) was added to the mixture and the solution was stirred for 16 hours at room temperature. The solvent was removed under vacuum and the crude product purified on a column (eluent: dichloromethane/methanol 10/0 to 9/1) and Biotin-PEG-Azide **76** was collected as a brown oil (78 mg, 31%). δ_H (500 MHz, CDCl$_3$) 7.11 (t, 1H, J = 5.6 Hz, H_4), 6.71 (t, 1H, J = 5.6 Hz, H_{11}), 6.48 (s, 1H, H_{23}), 5.69 (s, 1H, H_{21}), 4.51-4.48 (m, 1H, H_{19}), 4.32-4.29 (m, 1H, H_{18}), 4.00 (s, 2H, H_6), 3.68-3.62 (m, 4H, $H_{7,8}$), 3.58 (t, 2H, J = 5.3 Hz, H_9), 3.45-3.42 (m, 2H, H_{10}), 3.40-3.36 (m, 4H, $H_{1,3}$), 3.15-3.12 (m, 1H, H_{17}), 2.92-2.88 (dd, 1H, J = 12.8 Hz, H_{20}), 2.74 (d, 1H, J = 12.8 Hz, H_{20}), 2.24 (t, 2H, J = 7.5 Hz, H_{13}), 1.85 (quint., 2H, J = 6.6 Hz, H_2), 1.75-1.62 (m, 4H, $H_{15,16}$), 1.49-1.37 (m, 4H, H_{14}). δ_C (250 MHz, CDCl$_3$) 173.6, 170.2, 164.1, 70.9, 70.6, 70.0, 61.9, 60.2, 55.6, 49.3, 40.6, 39.1, 36.5, 35.9, 28.8, 28.2, 28.1, 25.6. ν_{max}/cm^{-1} (solid state) = 3278, 3079 (NH), 2927, 2886 (CH), 2096 (N$_3$), 1695, 1654 (C=O amide), 1261 (C-O ether). HRMS: Calcd. [M+Na]$^+$ (C$_{19}$H$_{33}$N$_7$O$_5$S) m/z = 494.2126. Found [M+Na]$^+$ m/z = 494.2156.

5.1.5. Trimer and dimer synthesis

5.1.5.1. General procedures

g) SPPS procedure

Glycine loaded Wang resin (0.79 mmol.g^{-1}, 100-200 mesh; carrier: polystyrene, crosslinked with 1% DVB; 127 mg, 0.1 mmol) was swelled in anhydrous N-methyl-2-pyrrolidone (5 mL) 15

minutes prior to reaction. The appropriate monomers were dissolved in anhydrous *N*-methyl-2-pyrrolidone and preactivated for coupling with Ghosez's reagent (20% in chloroform) for 1 hour at 50°C. The reactions were carried out on a CEM Liberty Blue® automated microwave assisted peptide synthesizer, under microwave heating, following the conditions detailed in Table 5.1. Before each coupling, standard washing and deprotection (25% piperidine solution in *N*-Methyl-2-pyrrolidone) cycles were carried out on the synthesizer. The samples on resin were washed with dichloromethane and ether. When required, cleavage off the resin was performed manually, using a 50% solution of TFA in dichloromethane (2 mL). TFA was removed by bubbling the solution with nitrogen and the solvent was evaporated under vacuum.

Table 5.1: Conditions for SPPS depending on the type of scaffold.

Scaffold	Monomer (mmol)	Ghosez (mmol)	NMP (mL)	Method	Coupling time (mins)	Temperature (°C)
N-alkylated	0.25-0.5	0.25-0.5	5	Double	20	60
O-alkylated	0.2	0.2	4	Single	30	50

h) General click-chemistry procedure

The trimer on resin was swelled in tetrahydrofuran (1 mL). Azidobiotin **72** (0.1 mmol), water (1 mL), copper (II) sulphate (0.01 mmol) and sodium ascorbate (0.02 mmol) were added in that order. The reaction was carried out on a Stuart SB2 rotator for 16 hours. Addition of more equivalents of reagents was sometimes required to observe the completion of the reaction. The resin was then washed with water, dichloromethane and ether, then the trimer was cleaved off the resin using a 50% solution of TFA in dichloromethane (2 mL). TFA was removed by bubbling the solution with nitrogen and the remaining solvent was evaporated under vacuum.

5.1.5.2. Numbering system

To simplify the NMR assignment of the trimers, the following nomenclature is used. The monomers constituting the dimers/trimers are considered separately, numbered 1 to 2 or 3 starting from the *N*-terminus, and the glycine is numbered 3 or 4. All the monomers are numbered following the same standard system: the carbon bearing the carboxylic acid is C1 and the one bearing the amine is C4. The carbon attached to the nitrogen is Cα, and the numbering of the aliphatic part of the chain continues with Cβ, etc. The numbering of the protons corresponds to the numbering of the carbons (Figure 5.1). For clarity purposes, the monomer number is added as a prefix to the proton number.

Figure 5.1: Examples of numbering of an *N*-alkylated trimer 1$_B$2$_{iBu}$3$_A$G (left) and an *O*-alkylated dimer 1$_B$2$_A$G (right).

To avoid confusion, in the case of *O*-substituted *N*-alkylated compounds, a first number indicates the monomer bearing the side chain. A second number followed by ^{O} specifies the position of the functionalised oxygen. The aliphatic protons are numbered Hα, Hβ, etc. Triazole protons are indicated as HT (Figure 5.2).

Figure 5.2: Numbering system for biotin (left) and FITC (right) orthogonal groups.

5.1.5.3. Synthesis of dimer and trimers (Chapter 2)

4-(2-amino-5-((2-(3-aminopropoxy)-4-((carboxymethyl)carbamoyl)phenyl)carbamoyl)-phenoxy)butanoic acid (50)

Using the general SPPS procedure g). **24** (107 mg, 0.2 mmol); **36** (104 mg, 0.2 mmol). The residue was purified using semi-preparative HPLC (eluent: acetonitrile/water: 0.5/9.5 to 9.5/0.5) and the product was collected as a pale cream powder (9.3 mg, 19%). δ$_H$ (500 MHz, DMSO-d$_6$) 9.01 (s, 1H, 2-NH), 8.48 (s, 1H, 3-NH), 8.11 (d, 1H, *J* = 8.3 Hz, 2-H5), 7.55 (s, 1H, 2-H2), 7.50 (d, 1H, *J* = 8.3 Hz, 2-H6), 7.38 (s, 1H, 1-H6), 7.36 (s, 1H, 1-H2), 6.71 (d, 1H, *J* = 8.1 Hz, 1-H5), 5.43 (s broad, 2H, NH$_2$), 4.18 (t, 2H, *J* = 5.5 Hz, 2-Hα), 4.05 (t, 2H, *J* = 6.5 Hz, 1-Hα), 3.79 (d, 2H, *J* = 4.9 Hz, 3-Hα), 2.99 (t, 2H, *J* = 6.6 Hz, 2-Hγ), 2.29 (t, 2H, *J* = 6.6 Hz, 1-Hγ), 2.08 - 2.15 (m, 2H, 2-Hβ), 1.96 - 2.01 (m, 2H, 1-Hβ). δ$_C$ (125 MHz, DMSO-d$_6$) 175.5, 175.0,

171.6, 165.6, 165.4, 151.1, 148.9, 144.6, 142.2, 131.2, 130.8, 129.5, 129.4, 122.7, 120.9, 120.0, 68.2, 67.3, 32.4, 27.7, 25.0, 24.8, 24.5. HRMS: Calcd. $[M+H]^+$ ($C_{23}H_{28}N_4O_8$) $m/z = 489.1907$. Found $[M+H]^+$ $m/z = 489.1902$.
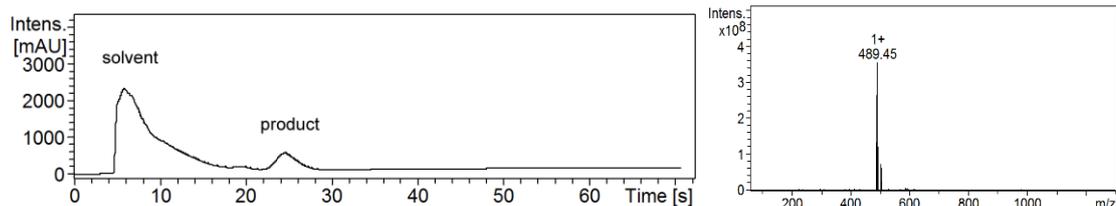
4-(2-(4-amino-3-(3-carboxypropoxy)benzamido)-5-((carboxymethyl)carbamoyl)phenoxy)-butanoic acid (48)
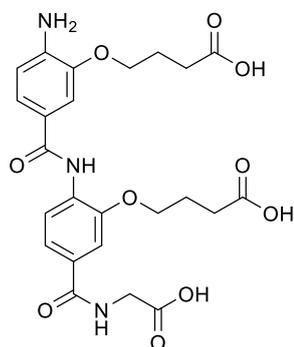
Using the general SPPS procedure g). **36** (208 mg, 0.4 mmol). The residue was purified using semi-preparative HPLC (eluent: acetonitrile/water: 0.5/9.5 to 9.5/0.5) and the product was collected as a pale cream powder (4.6 mg, 9%). δ_H (500 MHz, DMSO-d_6) ^{1}H NMR (500 MHz, DMSO-d_6) δ ppm 8.98 (s, 1H, 2-NH), 8.77 (t, 1H, $J = 5.9$ Hz, 3-NH), 8.14 (d, 1H, $J = 8.4$ Hz, 2-H5), 7.53 (d, 1H, $J = 1.7$ Hz, 2-H2), 7.50 (dd, 1H, $J = 8.5, 1.7$ Hz, 2-H6), 7.36 (dd, 1H, $J = 8.1, 1.9$

Hz, 1-H6), 7.34 (d, 1H, $J = 1.9$ Hz, 1-H2), 6.68 (d, 1H, $J = 8.1$ Hz, 1-H5), 4.14 (t, 2H, $J = 6.2$ Hz, 2-Hα), 4.04 (t, 2H, $J = 6.2$ Hz, 1-Hα), 3.92 (d, 2H, $J = 5.8$ Hz, 3-Hα), 2.44 - 2.47 (m, 4H, 1-Hγ and 2-Hγ), 2.05 (quint., 2H, $J = 6.7$ Hz, 2-Hβ), 1.99 (quint, 2H, $J = 6.7$ Hz, 1-Hβ). δ_C (125 MHz, DMSO-d_6) 175.0, 171.8, 166.2, 165.1, 164.0, 148.8, 148.4, 144.4, 141.9, 130.6, 129.2, 129.1, 121.2, 120.8, 119.0, 110.7, 110.3, 92.2, 79.4, 78.7, 24.9, 24.1, 16.8 (1 x C missing). HRMS: Calcd. $[M+H]^+$ ($C_{24}H_{27}N_3O_{10}$) $m/z = 518.1696$. Found $[M+H]^+$ $m/z = 518.1769$.
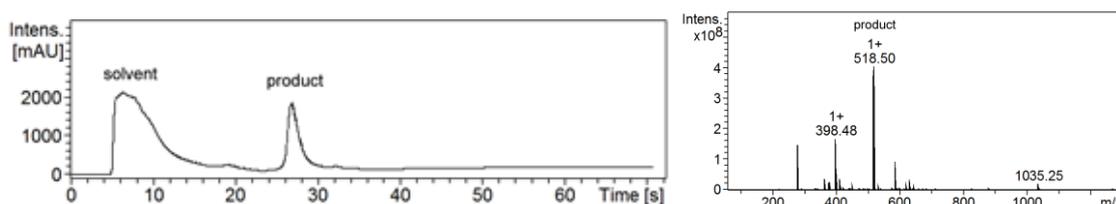
(4-(4-amino-3-(sec-butoxy)benzamido)-3-(sec-butoxy)benzoyl)glycine (57)
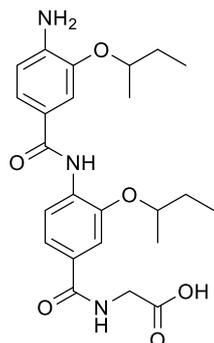
Dark yellow solid, prepared by Dr Sílvia Rodriguez-Marin. δ_H (500 MHz, DMSO-d_6) 8.89 (s, 1H, 2-NH), 8.74 (t, 1H, 3-NH) 8.24 (d, 1H, $J = 8.3$ Hz, 2-H5), 7.56 (d, 1H, $J = 1.7$ Hz, 2-H2), 7.52 (dd, 1H, $J = 8.4, 1.7$ Hz, 2-H6), 7.32 (t, 1H, $J = 1.5$ Hz, 1-H2), 7.30 (d, 1H, $J = 8.2, 1.9$ Hz, 1-H6), 6.72 (d, 1H, $J = 8.1$ Hz, 1-H5), 5.41 (s, 2H, 1-NH$_2$), 4.56 (quint., 1H, $J = 5.9$ Hz, 2-Hα), 4.40 (quint., 1H, $J = 59$ Hz, 1-Hα), 3.92 (d, 2H, $J = 5.9$ Hz, 3-Hα), 1.79-1.61 (m, 4H, 1-Hβ, 2-Hβ), 1.32 (dd, 3H, $J = 6.1$ Hz, 1.3 Hz, 2-

CHα(C*H₃*)), 1.28 (dd, 3H, *J* = 6.1 Hz, 2.3 Hz, 1-CHα(C*H₃*)), 0.99-0.94 (m, 6H, 1-Hγ, 2-Hγ). δ$_C$ (125 MHz, DMSO-d₆) 171.4, 165.8, 164.4, 147.0, 143.5, 143.2, 131.7, 128.9, 121.1, 120.9, 119.9, 112.8, 112.2, 76.0, 75.3, 41.3, 40.4, 28.6, 28.5, 19.1, 19.0 18.9, 9.6, 9.4. HRMS: Calcd. [M+H]⁺ (C₂₄H₃₂N₃O₆) *m/z* = 458.2286. Found [M+H]⁺ *m/z* = 458.2298.
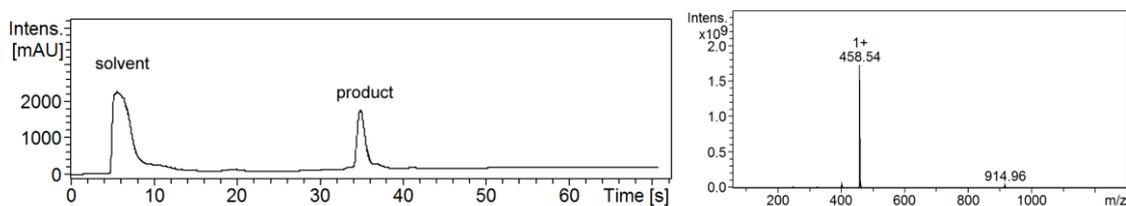
4-(4-((3-aminopropyl)amino)-*N*-(4-((carboxymethyl)carbamoyl)phenyl)benzamido)buta-noic acid (55)
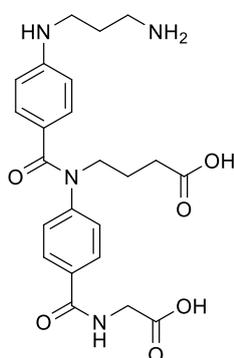
Using the general SPPS procedure g). **43** (126 mg, 0.25 mmol); **45** (129 mg, 0.25 mmol). The residue was purified using semi-preparative HPLC (eluent: acetonitrile/water: 0.5/9.5 to 9.5/0.5) and the product was collected as a pale cream powder (2.8 mg, 6%). δ$_H$ (500 MHz, DMSO-d₆) 8.38 (t broad, 1H, 3-NH), 7.73 (d, 2H, *J* = 8.2 Hz, 2-H2), 7.15 (d, 2H, *J* = 8.1 Hz, 2-H3), 6.99 (d, 2H, *J* = 8.6 Hz, 1-H2), 6.33 (d, 2H, *J* = 8.4 Hz, 1-H3), 6.11 (s broad, 1H, 1-NH), 3.85 (d, 2H, *J* = 7.4 Hz, 2-Hα), 3.74 (d, 2H, *J* = 5.0 Hz, 3-Hα), 3.02 (t broad, 2H, 1-Hα), 2.78 (t broad, 2H, 1-Hγ), 2.23 (t, *J* = 7.4 Hz, 2-Hγ), 1.65 - 1.77 (m, 4H, 1-Hβ and 2-Hβ). δ$_C$ (125 MHz, DMSO-d₆) 199.6, 174.5, 169.6, 168.2, 146.8, 131.6, 131.5, 130.0, 128.0, 121.8, 118.2, 85.9, 54.7, 51.9, 28.1, 27.6, 23.1, 21.5, 12.5. HRMS: Calcd. [M+H]⁺ (C₂₃H₂₈N₄O₆) *m/z* = 457.2009. Found [M+H]⁺ *m/z* = 457.2092.
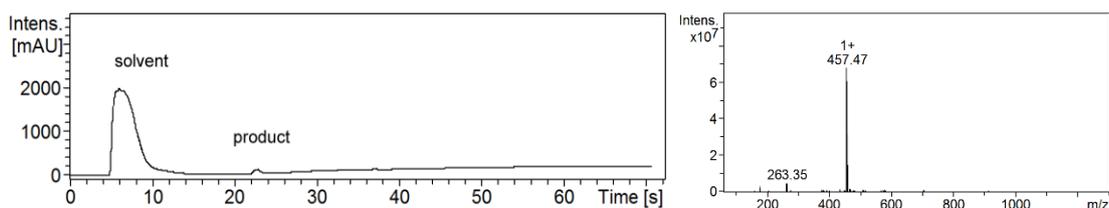
4-(4-(4-((3-aminopropyl)-amino)-*N*-isobutylbenzamido)-*N*-(4-((carboxymethyl)carba-moyl)phenyl)benzamido)buta-noic acid (53)

Using the general SPPS procedure g). **43** (251 mg, 0.5 mmol); **39** (208 mg, 0.5 mmol); **45** (258 mg, 0.5 mmol). The residue was purified using semi-preparative HPLC (eluent: acetonitrile/water: 0.5/9.5 to 9.5/0.5) and the product was collected as a pale cream powder (9.3 mg, 15%). δ$_H$ (500 MHz, pyr.-d₅) 9.27 (s broad, 1H, 4-NH), 8.21 (d, 2H, *J* = 8.3 Hz, 3-H2), 7.47
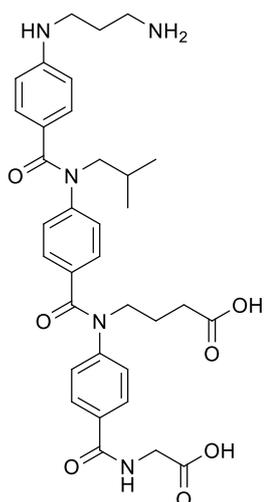
(d, 2H, *J* = 8.2 Hz, 2-H2), 7.35 (d, 2H, *J* = 8.8 Hz, 1-H2), 7.26 (d, 2H, *J* = 8.3 Hz, 3-H3), 7.13 (d, 2H, *J* = 8.3 Hz, 2-H3), 6.61 (d, 2H, *J* = 8.8 Hz, 1-H3), 4.56 (d, 2H, *J* = 4.8 Hz, 4-Hα), 4.22 (t, 2H, *J* = 7.2 Hz, 3-Hα), 3.86 (d, 2H, *J* = 7.3 Hz, 2-Hα), 3.41 (t, 2H, *J* = 6.6 Hz, 1-Hα), 3.34 (t, 2H, *J* = 5.7 Hz, 1-Hγ), 2.67 (t, 2H, *J* = 7.4 Hz, 3-Hγ), 2.18 - 2.27 (m, 4H, 1-Hβ and 3-Hβ), 1.91 - 1.99 (m, 1H, 2-Hβ), 0.90 (d, 6H, *J* = 6.7 Hz, 2-Hγ). δ$_H$ (500 MHz, DMSO-d$_6$) 8.47 (t, 1H, *J* = 5.3 Hz, 4-NH), 7.72 (d, 2H, *J* = 8.5 Hz, 3-H2), 7.18-7.16 (m, 4H, 2-H2, 3-H3), 6.96 (d, 2H, *J* = 8.4 Hz, 2-H3), 6.83 (d, 2H, *J* = 8.5 Hz, 1-H2), 6.19 (d, 2H, *J* = 8.8 Hz, 1-H3), 3.87 (t, 2H, *J* = 7.4 Hz, 3-Hα), 3.70 (d, 2H, *J* = 5.3 Hz, 4-Hα),

3.63 (d, 2H, *J* = 7.4 Hz, 3-Hγ), 3.06 (t br., 2H, 1-Hα), 2.87 (t, 2H, *J* = 6.9 Hz, 1-Hγ), 2.27 (t, 2H, *J* = 7.5 Hz, 2-Hα), 1.83-1.78 (m, 2H, 1-Hβ), 1.75-1.67 (m, 3H, 2-Hβ, 3-Hβ), 0.80 (d, 6H, *J* = 6.6 Hz, 2-Hγ). δ$_C$ (125 MHz, DMSO-d$_6$) 172.7, 169.4, 168.8, 165.2, 150.1, 145.3, 145.1, 133.9, 132.2, 130.9, 128.7, 127.8, 127.3, 126.2, 121.2, 110.0, 58.3, 56.1, 48.5, 43.2, 36.2, 31.1, 26.6, 26.0, 22.8, 20.1 (1 x C missing). HRMS: Calcd. [M+H]$^+$ (C$_{34}$H$_{41}$N$_5$O$_7$) *m/z* = 632.3006. Found [M+H]$^+$ *m/z* = 632.2992.
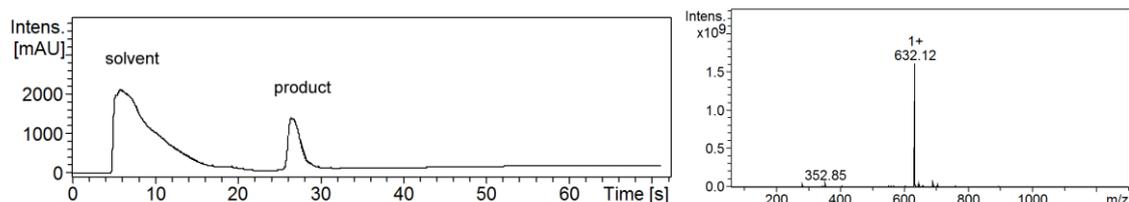
4-((4-((4-((3-aminopropyl)(4-((carboxymethyl)carbamoyl)phenyl)carbamoyl)phenyl)(iso-butyl)carbamoyl)phenyl)amino)butanoic acid (54)
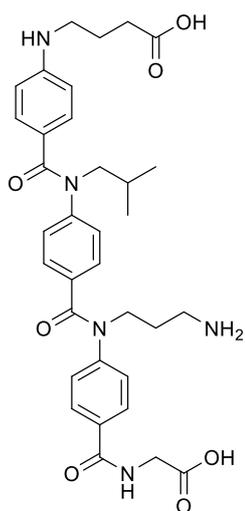
Using the general SPPS procedure g). **45** (258 mg, 0.5 mmol); **39** (208 mg, 0.5 mmol); **43** (251 mg, 0.5 mmol). The residue was purified using semi-preparative HPLC (eluent: acetonitrile/water: 0.5/9.5 to 9.5/0.5) and the product was collected as a pale cream powder (10.2 mg, 16%). δ$_H$ (500 MHz, pyridin-d$_5$) 9.13 (s broad, 1H, 4-NH), 8.20 (d, 2H, *J* = 8.4 Hz, 3-H2), 7.48 (d, 2H, *J* = 8.4 Hz, 2-H2), 7.35 (d, 2H, *J* = 8.7 Hz, 1-H2), 7.25 (d, 2H, *J* = 8.6 Hz, 3-H3), 7.14 (d, 2H, *J* = 8.6 Hz, 2-H3), 6.60 (d, 2H, *J* = 8.8 Hz, 1-H3), 4.55 (d, 2H, *J* = 4.9 Hz, 4-Hα), 4.22 (t, 2H, *J* = 7.2 Hz, 3-Hα), 3.86 (d, 2H, *J* = 7.5 Hz, 2-Hα), 3.44-3.26 (m, 4H, 1-Hα and 1-Hγ), 2.68 (t, 2H, *J* = 7.3 Hz, 3-Hγ), 2.26-2.15 (m, 4H, 1-Hβ, 3-Hβ), 2.01-1.90 (m, 1H, 2-Hβ), 0.90 (d, 6H, *J* = 6.6 Hz, 2-Hγ). δ$_H$ (500 MHz, DMSO-d$_6$)

8.38 (t, 1H, *J* = 5.5 Hz, 4-NH), 7.71 (d, 2H, *J* = 8.5 Hz, 3-H2), 7.19-7.16 (m, 4H, 2-H2 and 3-H3), 6.96 (d, 2H, *J* = 8.4 Hz, 2-H3), 6.82 (d, 2H, *J* = 8.4 Hz, 1-H2), 6.17 (d, 2H, *J* = 8.7 Hz, 1-H3), 3.86 (t, 2H, *J* = 7.0 Hz, 3-Hα), 3.64 (d, 2H, *J* = 6.5 Hz, 4-Hα and 3-Hγ), 3.06 (t br., 2H, 1-

Hα), 2.86 (t, 2H, J = 7.0 Hz, 1-Hγ), 2.26 (t, 2H, J = 7.5 Hz, 2-Hα), 1.84-1.79 (m, 2H, 1-Hβ), 1.75-1.68 (m, 3H, 2-Hβ, 3-Hβ), 0.81 (d, 6H, J = 6.6 Hz, 2-Hγ). δ_C (125 MHz, DMSO-d$_6$) 174.1, 173.1, 169.5, 168.6, 165.1, 150.1, 145.3, 145.0, 134.0, 132.4, 131.0, 128.7, 127.8, 127.3, 126.2, 120.9, 109.9, 56.1, 48.5, 43.8, 36.0, 31.2, 26.6, 25.8, 22.8, 20.1 (1 x C missing). HRMS: Calcd. [M+H]$^+$ (C$_{34}$H$_{41}$N$_5$O$_7$) m/z = 632.3006. Found [M+H]$^+$ m/z = 632.2984.
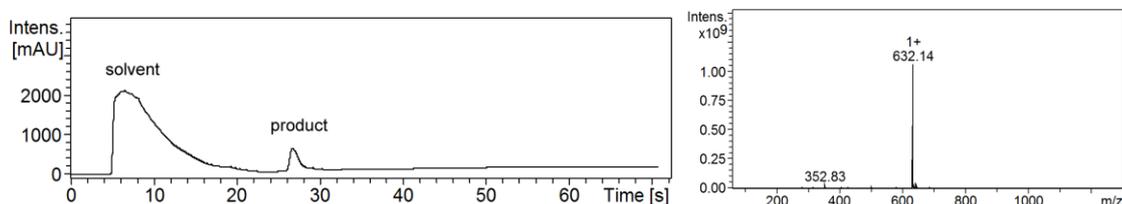
5.1.5.4. Synthesis of trimers (Chapter 3)

(4-(*N*-isobutyl-4-(*N*-isobutyl-4 (isobutylamino)benzamido)benzamido)benzoyl)glycine (58)
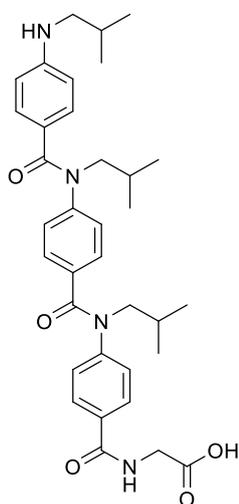
Using the general SPPS procedure g). **39** (624 mg, 1.5 mmol). The product was collected as a brown solid (57.9 mg, 96%). δ_H (500 MHz, DMSO-d$_6$) 8.82 (t, 1H, J = 5.9 Hz, 4-NH), 7.74 (d, 2H, J = 8.7 Hz, 3-H2), 7.19 (d, 2H, J = 8.5 Hz, 3-H3), 7.12 (d, 2H, J = 8.3 Hz, 2-H2), 6.93 (d, 2H, J = 8.5 Hz, 2-H3), 6.86 (d, 2H, J = 8.7 Hz, 1-H2), 6.27 (d, 2H, J = 8.7 Hz, 1-H3), 3.90 (d, 2H, J = 5.9 Hz, 4-Hα), 3.73 (d, 2H, J = 7.3 Hz, 3-Hα), 3.59 (d, 2H, J = 7.3 Hz, 2-Hα), 2.79 (d, 2H, J = 6.6 Hz, 1-Hα), 1.79-1.71 (m, 2H, 1-Hβ and 3-Hβ), 1.67-1.62 (m, 1H, 2-Hβ) 0.89 (d, 6H, J = 6.6 Hz, 1-Hγ), 0.85 (d, 6H, J = 6.6 Hz, 3-Hγ), 0.77 (d, 6H, J = 6.6 Hz, 2-Hγ). δ_C (125 MHz, DMSO-d$_6$) 171.7, 169.8, 169.7, 166.1, 150.8, 146.2, 145.9, 133.8, 131.7, 131.1, 129.2, 128.4, 127.7, 126.7, 122.0, 110.6, 56.4, 50.7, 41.6, 31.1, 27.8, 27.2, 27.1, 20.8, 20.5, 20.4. HRMS: Calcd. [M+H]$^+$ (C$_{35}$H$_{44}$N$_4$O$_5$) m/z = 601.3312. Found [M+H]$^+$ m/z = 601.3416.
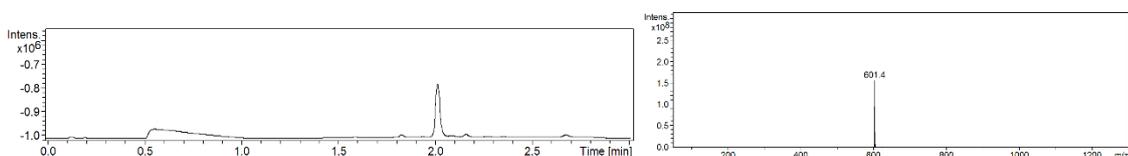
(4-(*N*-isobutyl-4-(*N*-isobutyl-4-(isobutylamino)benzamido)benzamido)-3-(prop-2-yn-1-yloxy)benz-amido)benzoyl)glycine (Alk-**58**)
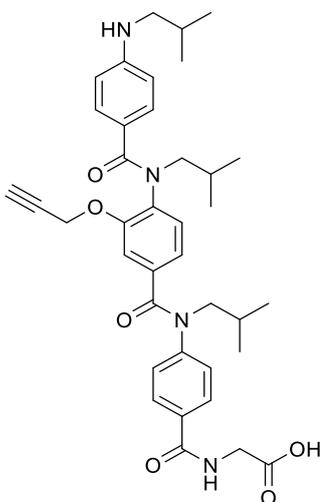
Using the general SPPS procedure g). **39** (415 mg, 1.0 mmol); **68** (234 mg, 0.5 mmol). A small amount of the sample was cleaved off the resin with TFA (50% in dichloromethane) to confirm the structure. HRMS: Calcd. [M+H]$^+$ (C$_{38}$H$_{46}$N$_4$O$_6$) m/z = 655.3417. Found [M+H]$^+$ m/z = 655.3502.
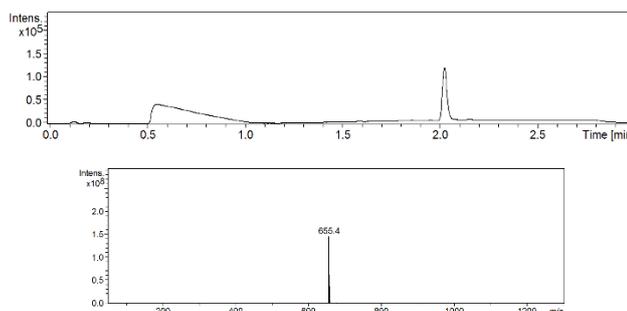
(4-(*N*-isobutyl-4-(*N*-isobutyl-4-(isobutylamino)benzamido)benzamido)-3-((1-(3-(4-(2-oxohexahy-dro-1H-thieno[3,4-d]imidazol-4-yl)butanamido)propyl)-1H-1,2,3-triazol-5-yl)methoxy)-benzamido)benzoyl)glycine (Biotin-**58**)
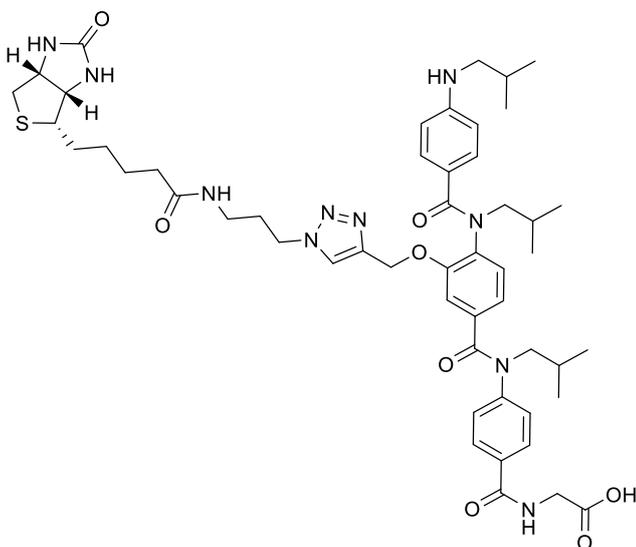
Using the general click-chemistry procedure h) on Alk-**58**. The product was collected as a brown solid (59.9 mg, 60% over 2 steps). δ_H (500 MHz, DMSO-d_6) 8.82 (t, 1H, J = 5.3 Hz, 4-NH), 7.98 (s, 1H, 2-3°T), 7.87 (s br., 1H, 4-CO$_2$H), 7.70 (d, 2H, J = 8.1 Hz, 3-H2), 7.17 (d, 2H, J = 7.9 Hz, 3-H3), 7.06 (s, 1H, 2-H2), 6.90 (d, 1H, J = 7.7 Hz, 2-H5), 6.84 (d, 2H, J = 8.1 Hz, 1-H2), 6.67 (d, 1H, J = 7.7 Hz, 2-H6), 6.36 (d, 2H, J = 7.9 Hz, 1-H3), 4.93 (s broad, 1H, 2-3°NH), 4.83 (s broad, 1H, 2-3°NH), 4.36 (t broad, 2H, J = 6.2 Hz, 2-3°Hβ), 4.29 (t broad, 1H, J = 6.2 Hz, 2-3°Hλ), 4.11 (t broad, 1H, J = 6.2 Hz, 2-3°Hλ), 3.89 (d, 2H, J = 5.3 Hz, 4-Hα), 3.69 (d, 2H, J = 6.6 Hz, 3-Hα), 3.55 (s broad, 2H, 2-3°Hδ) 3.15 (s, 2H, 2-3°Hα), 3.08 (d, 1H, J = 4.6 Hz, 2-3°Hι), 3.05 (d, 2H, J = 5.3 Hz, 2-Hα), 2.81 (d, 2H, J = 6.6 Hz, 1-Hα), 2.77 (dd (under previous peak), 1H, J = 4.5 Hz, , 2-3°Hκ), 2.53 (d, 1H, J = 4.9 Hz, 2-3°Hκ), 2.07-2.05 (m, 2H, 2-3°Hε), 1.95-1.91 (m, 2H, 2-3°Hγ), 1.81-1.74 (m, 1H, 1-Hβ), 1.73-1.66 (m, 1H, 3-Hβ), 1.63-1.56 (m, 2H, 2-3°Hθ) 1.55-1.47 (m, 4H, 2-Hβ and 2-3°Hζ), 1.34-1.24 (m, 2H, 2-3°Hη), 0.89 (d, 6H, J = 6.6 Hz, 1-Hγ), 0.84 (d, 6H, J = 6.4 Hz, 3-Hγ), 0.71 (d, 6H, J = 6.4 Hz, 2-Hγ). δ_c (125 MHz, DMSO-d_6) 172.2, 171.1, 170.2, 169.8, 169.0, 165.8,

152.2, 148.1, 145.6, 142.0, 135.9, 133.6, 131.2, 129.4, 128.9, 127.8, 127.1, 124.4, 120.5, 113.5, 111.5, 61.5, 61.1, 59.6, 59.3, 55.3, 54.7, 51.2, 48.4, 47.3, 41.1, 35.6, 35.1, 30.0, 28.1, 27.9, 27.0, 26.7, 26.5, 25.1, 20.6, 20.1, 19.9. HRMS: Calcd. $[M+H]^+$ ($C_{51}H_{68}N_{10}O_8S$) m/z = 981.4942. Found $[M+H]^+$ m/z = 981.5024.
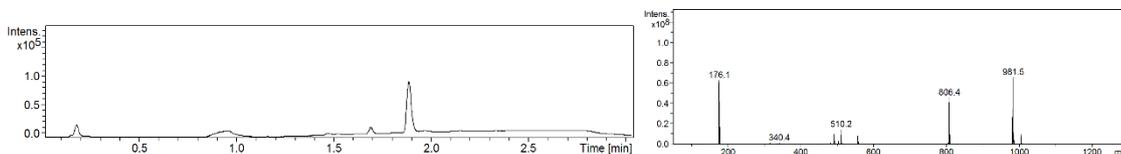
5-(3-(3-(4-((5-((4-((carboxymethyl)carbamoyl)phenyl)(isobutyl)carbamoyl)-2-(N-isobutyl-4-(isobutylamino)benzamido)phenoxy)methyl)-1H-1,2,3-triazol-1-yl)propyl)ureido)-2-(6-hydroxy-3-oxo-3H-xanthen-9-yl)benzoic acid (Flu-58)
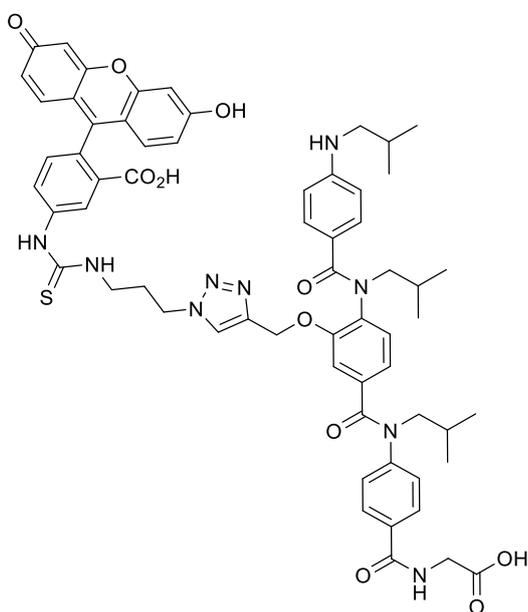
Using the general click-chemistry procedure h) on Alk-**58**. The crude product was purified by semi-preparative HPLC (eluent: acetonitrile/ water: 0.5/9.5 to 9.5/0.5), the fractions were evaporated to afford Flu-**58** as a dark red solid (3.5mg, 6%). δ_H (500 MHz, DMSO-d6) 8.81 (t, 1H, J = 5.8 Hz, 4-NH), 8.21 (s broad, 2-3°H2), 7.97 (s, 1H, , 2-3°OH), 7.75 (d, 1H (under previous peak), 2-3°H6), 7.73 (d, 2H, J = 8.5 Hz, 3-H2), 7.21 (d, 2H, J = 8.5 Hz, 3-H3), 7.16 (d, 1H, J = 8.3 Hz, 2-3°H5), 7.09 (s broad, 1H, 2-H2), 6.91 (d, 1H, J = 8.1 Hz, 2-H5), 6.81 (d, 2H, J = 8.7 Hz, 1-H2), 6.68 (d, 1H, J = 8.1 Hz, 2-H6), 6.65 (d, 2H, J = 2.3 Hz, 2-3°H10), 6.59 (t, 2H, J = 8.7 Hz, 2-3°H9), 6.54 (dd, 2H, J = 8.7 Hz, 2.3 Hz, 2-3°H11), 6.22 (d, 2H, J = 8.7 Hz, 1-H3), 5.82 (s broad, 1H, 1-NH), 4.93 (s broad, 1H, 2-3°NH), 4.87 (s broad, 1H, 2-3°NH), 4.47 (t, 2H, J = 7.0 Hz, 2-3°Hδ), 3.89 (d, 2H, J = 6.0 Hz, 4-Hα), 3.71 (d, 2H, J = 7.0 Hz, 3-Hα), 3.55 (d, 2H, J = 5.3 Hz, 2-3°Hβ), 2.77 (d, 2H, J = 6.6 Hz, 1-Hα), 2.15 (quint., 2H, J = 7.0 Hz, 2-3°Hγ) 1.77-1.68 (m, 2H, 3-Hβ and 1-Hβ), 1.54-1.49 (m, 1H, 2-Hβ), 0.87 (d, 6H, J = 6.6 Hz, 1-Hγ), 0.84 (d, 6H, J = 6.8 Hz, 3-Hγ), 0.71 (d, 6H, J = 6.8 Hz, 2-Hγ). ^{13}C NMR could not be acquired (not enough material). HRMS: Calcd. $[M+H]^+$ ($C_{62}H_{65}N_9O_{11}S$) m/z = 1144.4558. Found $[M+H]^+$ m/z = 1144.4598.
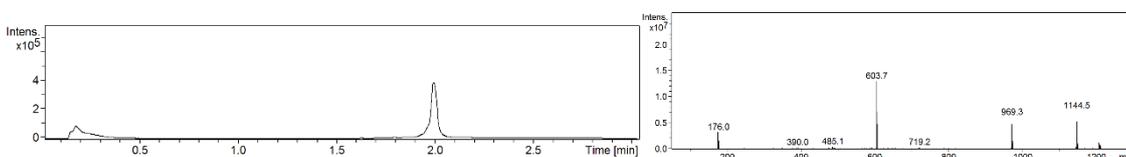
(4-(*N*-(cyclohexylmethyl)-4-(*N*-(naphthalen-2-ylmethyl)-4-(phenethylamino)benzamido)-3-(prop-2-yn-1-yloxy)benzamido)benzoyl)glycine (Alk-59rev)
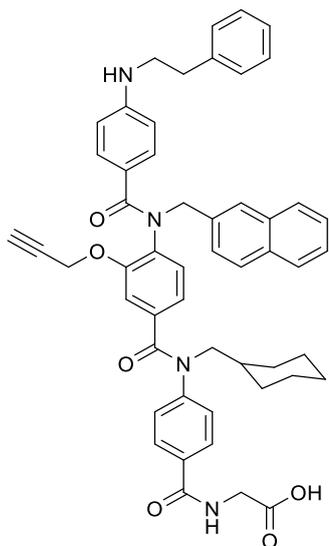
Alk-**59**rev was prepared following the general SPPS procedure. The readily available monomers "cyclohexane" **83** (232 mg, 0.5 mmol) and "2-naphtyl" **84** (277 mg, 0.5 mmol) and "ethyl-Ph" **85** (228 mg, 0.5 mmol) were used. A small amount of the sample was cleaved off the resin with TFA (50% in dichloromethane) to confirm the structure. HRMS: Calcd. $[M+H]^+$ ($C_{52}H_{50}N_4O_6$) m/z = 827.3764. Found $[M+H]^+$ m/z = 827.3769.
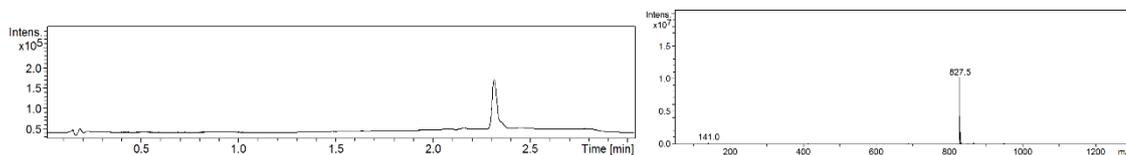
5-(3-(3-(4-((5-((4-((carboxymethyl)carbamoyl)phenyl)(cyclohexylmethyl)carbamoyl)-2-(N-(naphthalen-2-ylmethyl)-4-(phenethylamino)benzamido)phenoxy)methyl)-1H-1,2,3-triazol-1-yl)propyl)thioureido)-2-(6-hydroxy-3-oxo-3H-xanthen-9-yl)benzoic acid (Flu-59rev)
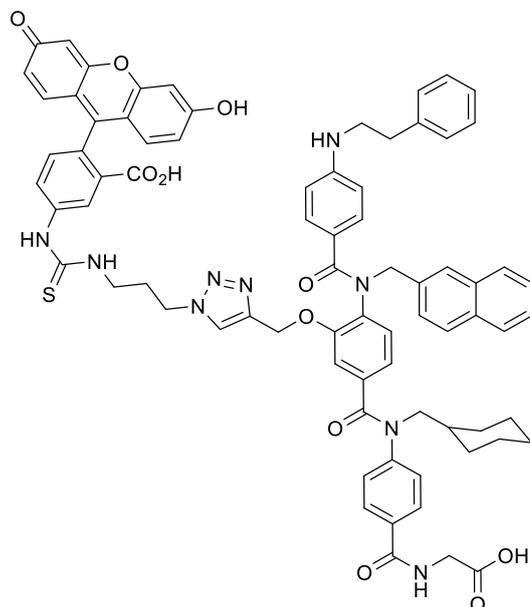
Using the general click-chemistry procedure h) on Alk-**59**rev. The crude product was purified by semi-preparative HPLC (eluent: acetonitrile/water: 0.5/9.5 to 9.5/0.5), the fractions were evaporated to afford a dark red solid (1.7 mg, 3%). δ_H (500 MHz, DMSO-d_6) 10.1 (s, 2H, 2-3°NH), 8.17 (s, 1H, 2-3°H2), 7.81-7.80 (m, 1H, 2-3°ArH), 7.77 (d, 3H, J = 8.3 Hz, 2-3°ArH), 7.71 (d, 2H, J = 8.3 Hz, 2-3°ArH), 7.64 (s, 1H, 1-NH), 7.45-7.41 (m, 2H, 1-Ph), 7.25 (d, 4H, J = 4.3 Hz, ArH), 7.23 (d broad, 1H, J = 8.1 Hz, 2-H5), 7.17-7.15 (m, 3H, 1-Ph), 7.09 (d, 1H, J = 8.3 Hz, 2-3°H6), 6.93 (d broad, 2H, J = 8.1 Hz, 1-H2), 6.70 (d, 1H, J = 7.1 Hz, 2-H$_{napht.}$), 6.66 (d, 2H, J = 2.3 Hz, 2-H$_{napht.}$), 6.58 (d, 2H, J = 8.6 Hz, 2-H$_{napht.}$), 6.55 (dd, 2H, J = 8.6 Hz, 2.3 Hz, 2-H$_{napht.}$), 6.32 (d, 2H, J = 8.6 Hz, 1-H3), 5.98 (s broad, 1H, 1-NH), 4.43 (t, 2H, J = 6.4 Hz, 2-3°Hγ), 3.93 (s broad, 2H), 3.71 (s broad, 2H), 3.25 (t broad, 2H, J = 7.3 Hz, 1-Hα), 3.09 (s,

broad, 2H), 3.02 (t broad, 1H, J = 5.6), 2.81 (t, 2H, J = 7.3 Hz, 1-Hβ), 2.01 (t broad, 2H, J = 6.4 Hz, 2-3⁰Hα), 1.67-1.63 (m, 3H), 1.63-1.58 (m, 3H, 3-Hβ,γ) 1.56-1.53 (m, 2H, 3-Hδ), 1.43 (s broad, 2H), 1.24 (s, 2H), 1.06 (s broad, 4H, 3-Hγ,δ), 0.90 (s broad, 1H), 0.87 (t, 2H, J = 7.3 Hz, 3-Hε). ^{13}C NMR could not be acquired (not enough material).
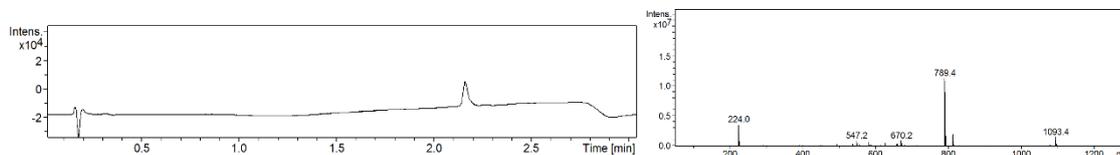
5.2. NMR experiments

The NMR experiments performed to determine the conformation of the oligomers described in chapter 2 were acquired on a Bruker Advance 500, at 500 MHz and 26°C, using the following parameters:

NMR	Number of scan (ns)	Relaxation time (d1)
^{1}H	64	2s
COSY	8	1s
NOESY	64	1s
ROESY	200	1s

5.3. Selection of the Affimers

5.3.1. Phage Display

The Affimer library was prepared by the BioScreening Technology Group in the School of Molecular and Cellular Biology.[122]

5.3.1.1. Screening

For the selection against Biotin-**58** to Biotin-**63**, four panning rounds were performed, described as follows:

First panning round

A colony of ER2738 *E. coli* cells was picked into 5 ml of 2TY media with 12 μg/mL tetracycline and incubated overnight at 37°C, 230 rpm. Streptavidin Coated (High Binding Capacity, Thermo Scientific) 8-well strips (4 wells per target) were pre-blocked with 300 μL 2x Blocking Buffer (BB) and incubated overnight at 37°C. The wells were washed 3 times with 300 μL phosphate buffered saline with 0.1% tween (PBST) and 100 μL 2x BB per well was added. 5 μL of phage library (10^{12} colony-forming units/mL) were pre-panned three times, and each time incubated at

room temperature for 40 minutes on a platform shaker (Heidolph VIBRAMAX 100, speed 3). 5 μL of the biotinylated target (1 mM in DMSO) was added to the panning well and incubated for 1 hour on the platform shaker. The phage was added to the target and incubated for 2 hours at room temperature on the platform shaker, while the overnight ER2738 culture was diluted 15 times with 2TY media with 12 μg/mL tetracycline and incubated for 1 hour at 37°C, 230 rpm. The panning wells were washed 6 times with 300 μL of PBST and eluted with 100 μL of glycine (0.2 mM, pH 2.2) for 10 min, neutralised with 15 μL of Tris-HCl (1M, pH 9.1), further eluted with 100 μL of triethylamine (100 mM in PBS) for 6 minutes at room temperature, and neutralised with 50 μL of Tris–HCl (1 M, pH 7). The eluted phage were incubated with ER2738 for 1 hour at 37°C and 90 rpm and plated onto LB agar plates with 100 μg/mL carbenicillin (carb) and grown overnight. Colonies were scraped with 2TY media containing 100 μg/mL carbenicillin (2TYcarb) (5 mL then 2 mL), diluted to get a 8 mL solution with an absorbance at 600 nm of 0.2, and incubated for 1 hour at 37°C, 230 rpm. 0.32 μL of M13K07 helper phage (10^{14}/mL) was added and the cells were incubated for 30 minutes at 37°C, 90 rpm. 16 μL of kanamycin (25 mg/mL) was added and the cells were incubated overnight at 25°C, 170 rpm. The phage were precipitated with 2 mL of polyethylene glycol (20% (w/v) PEG 8000, 2.5 M NaCl), incubated overnight at 4°C, and re-suspended with 320 μl of TE buffer (10 mM Tris, pH 8.0, 1 mM EDTA).

Second panning round

A colony of ER2738 *E. coli* cells was picked into 5 mL of 2TY media with 12 μg/mL tetracycline and incubated overnight at 37°C, 230 rpm. 20 μL of streptavidin beads (Dynabeads® MyOne™ Streptavidin T1, 10 mg/mL) were pre-blocked in 100 μL of 2x BB and incubated overnight at room temperature and 20 rpm.

KingFisher Flex (Thermo Scientific) plate were pre-blocked:

1 panning deep well 96 plate: 1 mL per well of 2x BB

2 elution well 96 plates: 300 μL per well of 2x BB

4 washing deep well 96 plates: 950 μL per well of 2x BB

After 2 hours at 37°C, the buffer was removed from the panning and the elution plates. 100 μL per well of glycine (0.2 M, pH 2.2) was added to one elution plate and 100 μL per well of triethylamine (100 mM in PBS) was added to the other. The Streptavidin beads were centrifuged for 1 min at 800 x*g*, the buffer was removed, and the beads were re-suspended in 100 μL of fresh 2x BB. 5 μL of the phage from the first panning round were pre-panned with 245 μL of 2x BB and 25 μL of pre-blocked Streptavidin beads. After incubating for 1 hour at room temperature at 20 rpm, the phage were centrifuged at 800 x*g* for 1 minute and the supernatant was collected and pre-panned as previously. 15 μL of the biotinylated target (1 mM in DMSO) was bound to 50 μL of pre-blocked Streptavidin beads in 200 μL of 2x BB and incubated for 1 hour at room

temperature on the rotator (20 rpm). The biotinylated target was centrifuged for 1 minute at 800 x*g*, and washed 3 times with 500 μL of 2x BB. The pre-panned phage were centrifuged for 1 minute at 800 x*g* and the supernatant was transferred to the biotinylated targets, re-suspended, and transferred to the pre-blocked deep well plate. The beads were washed 4 times with 950 μL of PBST and eluted as previously on an automated KingFisher robotic platform (ThermoFisher). The eluted phage were neutralised, amplified and precipitated following the same procedure as for the first panning round.

Third panning round

A colony of ER2738 *E. coli* cells was picked into 5 ml of 2TY media with 12 μg/mL tetracycline and incubated overnight at 37°C, 230 rpm. Neutravidin Coated (High Binding Capacity, Thermo Scientific) 8-well strips (5 wells per target) were pre-blocked with 300 μL of 2x BB and incubated overnight at 37°C. The wells were washed 3 times with 300 μL of PBST and 100 μl of 2x BB was added. 8 μL of the phage library from the previous round was pre-panned four times, and each time incubated at room temperature for 1 hour on a platform shaker. 10 μL of the biotinylated target (1 mM in DMSO) was added to the panning well, incubated for 1 hour on the platform shaker, and washed three times with PBST. The phage was then incubated with the target for 45 minutes at room temperature on the platform shaker. The panning wells were washed 27 times with 300 μL of PBST and eluted with 100 μL of glycine (0.2 mM, pH 2.2) for 10 min, neutralised with 15 μL of Tris-HCl (1M, pH 9.1), further eluted with 100 μL of triethylamine (100 mM in PBS) for 6 minutes at room temperature, and neutralised with 50 μL Tris–HCl (1 M, pH 7). The phage were eluted, neutralised, amplified and precipitated following the same procedure as for the first panning round.

Fourth panning round

A colony of ER2738 *E. coli* cells was picked into 5 mL of 2TY media with 12 μg/ml tetracycline and incubated overnight in an orbital incubator at 37°C, 230 rpm. Streptavidin Coated (High Binding Capacity, Thermo Scientific) 8-well strips (6 wells per target) were pre-blocked were filled with 300 μl of 2x BB and incubated overnight at 37°C. The wells were washed 3 times with 300 μL of PBST and 100 μl of 2x BB was added. 8 μL of the phage library from the previous round was pre-panned four times, and each time incubated at room temperature for 1 hour on a platform shaker. 10 μL of the biotinylated target (1 mM in DMSO) was added to the panning well, incubated for 1 hour on the platform shaker, and washed three times with 300 μL of PBST. The last well (blank) was washed with 300 μl of PBST. Half of the phage (100 μL) was transferred to the target and the other half (100 μL) transferred to the control well, and both were incubated for 45 minutes at room temperature on the platform shaker. The panning wells were washed 27 times with 300 μL of PBST and the phage were eluted, neutralised, and amplified following the same procedure as for the first panning round.

5.3.1.2. Phage ELISA

50 μL of Streptavidin (5 μg/ml in PBS) was placed in a F96 Maxisorp Nunc-Immuno Plate (Thermo Scientific) and incubated for 4 hours at room temperature. The plates were blocked with 200 μL of 2x BB, incubated overnight at 37°C, and washed with 300 μL of PBST. 50 μL of the biotinylated targets (1 μM in of 2x BB) was placed in the 6 fist columns of the streptavidin-coated wells, and negative control wells were prepared in the 6 last columns with 50 μL of 2x BB. The plate was incubated for 1 hour at room temperature on a platform shaker and washed with 300 μL of PBST.

48 individual colonies from the final panning round of the phage display were picked, inoculated into 96 well V bottom deep filled with 200 μL of 2TYcarb, and incubated for six hours at 37°C, 1050 rpm. 25 μL of the cultures were transferred into a 96 well V bottom deep well plate with 200 μL of 2TY carb and incubated for 1 hour at 37°C, 1050 rpm. 10 μL of a 10^{11}/mL solution of M13K07 helper phage in 2TY carb was added to the freshly grown cultures and incubated for 30 minutes at room temperature. 10 μL of a kanamycin solution in 2TY carb (1.25 mg/mL) was added to the cultures and incubated overnight at room temperature, 750 rpm. The cultures were centrifuged at 3,500 x*g* for 10 minutes and 40 μL of the phage-containing supernatant was transferred to the previously prepared streptavidin-coated 96-well plates with 10 μL of 10x BB. The plate was incubated for 1 hour at room temperature on a platform shaker and washed with 300 μL of PBST. 50 μl Anti-Fd-Bacteriophage-HRP (Seramun, diluted 1000 times in 2x BB from the commercial bottle) was added to the wells, the plate was incubated for 1 hour at room temperature on a vibrating platform shaker and washed 10 times with 300 μL of PBST. 50 μL of TMB (SeramunBlau® fast TMB/substrate solution) was added to the wells and allowed to develop for 3 minutes and the absorbance was measured at 620 nm on a Thermo Multiskan Ascent plate reader. The reading of the phage ELISA plates are shown below.

5.3.1.3. Sequencing

When sequencing was required, the plasmid DNA was extracted from selected hits using a QIAprep Spin Miniprep Kit (QIAGEN), the plasmids were eluted with water (50 μL) and an aliquot of DNA (15 μL, 100 ng/μl) was sent to Beckman Coulter Genomics for sequencing using a T7P primer.

5.3.2. Subcloning of the Affimers

Preparation of the vector DNA

20 μg of pET11a plasmid, 50 μL of CutSmart™ Buffer, 20 μg of NheI-HF™ and 20 μg of NotI-HF™ were mixed together and sterile deionised water was added to a total volume of 500 μL. The mixture was incubated overnight at 37°C, and divided into 4 tubes of 125 μL each. 14 μL of

Antartic Phosphatase Reaction Buffer (10x) and 1 µL of Antartic Phosphatase were added to each tube which was then mixed and incubated for 15 minutes at 37°C. The Antartic Phosphatase was heat inactivated by incubating for 5 minutes at 65°C, and the resulting digested vector was purified on a 0.7% agarose gel (80 V), using 14 µL of 10x Orange G Loading Dye for each tube. The purified vector was extracted from the gel using a Gel Extraction kit (QIAGEN) and stored at -20°C for later use.

PCR Amplification of the Affimer DNA sequence from the Phagemid Vector

For each Affimer, one 0.2 mL PCR tube was set up, containing: 13.8 µL of sterile water, 5 µL of 5X Phusion HF Buffer, 0.2 µL of dNTP Mix (25 nM), 0.75 µL of DMSO, 2 µL of forward primer (5' – ATGGCTAGCAACTCCCTGGAAATCGAAG, 10 µM), 2 µL of reverse primer (5' – TACCCTAGTGGTGATGATGGTGATGC, 10 µM), 0.25 µL of Phusion DNA Polymerase, and 1 µL of Template DNA.

The following PCR cycle were conducted:

Initial denaturation: 98°C, 30 seconds; 1 cycle

Denaturation: 98°C, 20 seconds; Annealing: 54°C, 20 seconds; Extension: 72°C, 20 seconds; 30 cycles

Final extension: 72°C, 10 minutes, 1 cycle

0.5 µL of DpnI was added to the PCR mixture and incubated for 1 hour at 37°C. The amplified DNA was purified on a 2% agarose gel with 1 µL of 10x Orange G Loading Dye, and the purified DNA was extracted from the gel using a Gel Extraction kit (QIAGEN), and eluted with 15 µL of sterile water.

Digestion of the PCR-amplified Affimer sequence with NheI and NotI Restriction Enzymes

3 µL of sterile water, 6 µL of CutSmart™ Buffer, 0.5 µL of NheI-HF™ and 0.5 µL of NotI-HF™ were added to each PCR product, mixed and incubated overnight at 37°C. The digested products were purified using a PCR Clean-up kit (QIAGEN) and eluted with 50 µL of sterile water.

Ligation of the NheI-NotI digested inserts into the pET11a vector

2 µL of 10X T4 DNA Ligase Buffer, 75 ng (0.4 µL) of Vector DNA, 25 ng (0.5 µL) of Insert DNA and 0.5 µL of T4 DNA Ligase were mixed, with 3 µL of sterile water. The mixture was incubated overnight at room temperature.

Transformation of ligations into *E.coli*

1 µL of the ligation mix was added to 10 µL of Agilent Technologies XL1-Blue supercompetent cells and the resulting mixture was incubated on ice for 30 minutes, heat shocked in at 42°C for

45 seconds, and incubated on ice for 2 further minutes. 190 μL of Recovery Media was added and the mixture was incubated at 37°C for 1 hour with shaking at 225 – 250 rpm. 100 μL of this transformation mixture was plated onto LB agar plates containing 100 μg/mL of carenicilin (LB carb plates) and incubated overnight at 37°C.

Purification of the subcloned plasmid DNA

Two single colonies were picked for each subcloning, and grew overnight in 4 mL of 2TY carb at 37°C, 230 rpm. The plasmid DNA was then extracted using a QIAprep Spin Miniprep Kit (QIAGEN), eluted with water (50 μl) and an aliquot of DNA (15 μL, 100 ng/μL) was sent to Beckman Coulter Genomics for sequencing using a T7P primer.

5.3.3. Expression and purification of the Affimers

Transformation of Adhion-pET11a Plasmid into BL21Star (DE3) E.Coli

0.5 μL of DNA was aliquoted into 50 μL of BL21 Star cells on ice, and incubated for 30 minutes on ice. The tubes were heat shocked in a 42°C water bath for 45 secods and further incubated 2 minutes on ice. 200 μL of Recovery Media was added and the mixture was incubated at 37°C for 1 hour with shaking (230 rpm). 10 μL of the transformation mixture was plated with 100 μL of Recovery Media on LB carb plates and incubated overnight at 37°C.

IPTG-induced Expression of the Affimer

One transformant was picked from the overnight culture and grown in 2 mL of 2TY carb + 1% glucose overnight at 37°C, 230 rpm. 1 mL of this culture was added to 50 mL of LB carb and grown to an OD of 0.8 (3 hours). 25 μL of 1M IPTG was added, and the culture was incubated for an additional 6 hours at 30°C, 150 rpm. The cells were harvested by centrifugation at 4,816 *xg* for 15 minutes and the supernatant was poured off.

Extraction and purification of the Affimer

The pellets were resuspended in the following buffer: 100 μL of BugBuster 10X Protein Extraction Reagent (Novagen), 0.4 μL of Benzonase Nuclease 25 U/μl (Novagen), 10 μL of Halt Protease Inhibitor Cocktail 100X (ThermoScientific), and 889.6 μL of Lysis Buffer. The mixture was incubated for 20 minutes at room temperature on a rotator (Stuart SB2), then incubated for 20 minutes in a 50°C water bath, and centrifuged at 16,000 *xg* for 20 minutes.

300 μL of Amintra Ni-NTA resin (Expedeon) was resuspended in 1 mL of Lysis Buffer and centrifuged at 1,000 *xg* for 1 minute and the buffer was aspirated off. The previous supernatant containing soluble proteins were added to the resin and incubated at room temperature for 2 hours on a rotator (Stuart SB2). The resin was centrifuged at 1,000 *xg* for 1 minute and the supernatant containing unbound proteins was collected and stored at -20°C. The resin was transferred to a

Pierce Centrifuge Column (ThermoScientific) and washed with Wash Buffer (50 mM NaH$_2$PO$_4$, 500 mM NaCl, 20 mM imidazole, pH 7.4) by gravity filtration until the concentration of the exiting Wash Buffer was consistently below 0.09 mg/mL (approx. 20 mL). The resin was then eluted by incubating for 10 minutes with 500 μL of Elution Buffer (50 mM NaH$_2$PO$_4$, 500 mM NaCl, 300 mM imidazole, 20% glycerol, pH 7.4) and the eluted Affimer was collected.

5.3.4. Purity assessment of the Affimers

A 15% SDS-PAGE gel was performed to check the purity of the proteins:
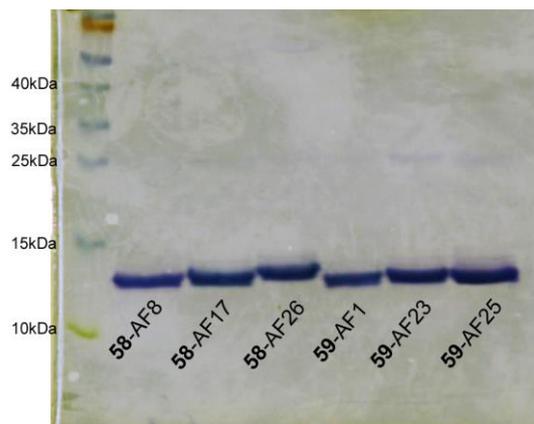
HRMS traces were recorded:

58-AF8 Calcd. 12259.78. Found: 12259.0982.
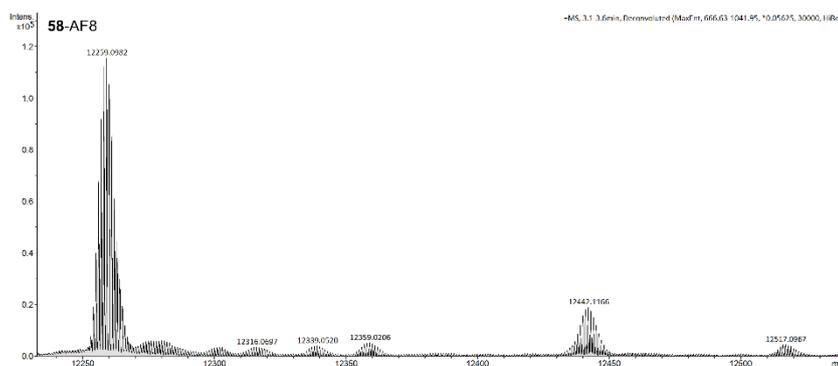
58-AF17: Calcd. 12727.33 Found. 12726.3642.
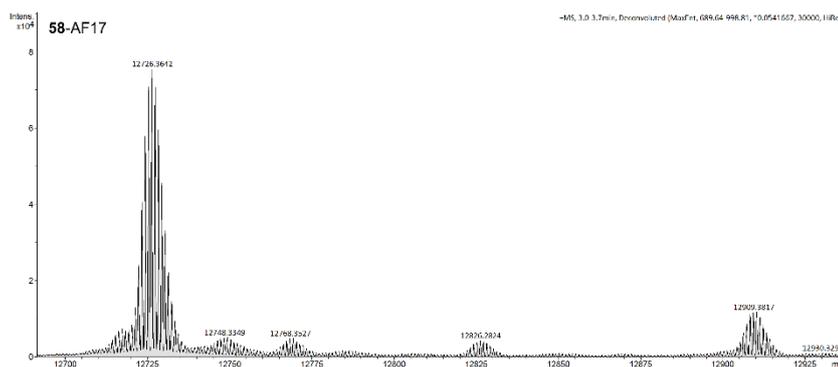
58-AF26: Calcd. 12283.89. Found. 12283.2020.
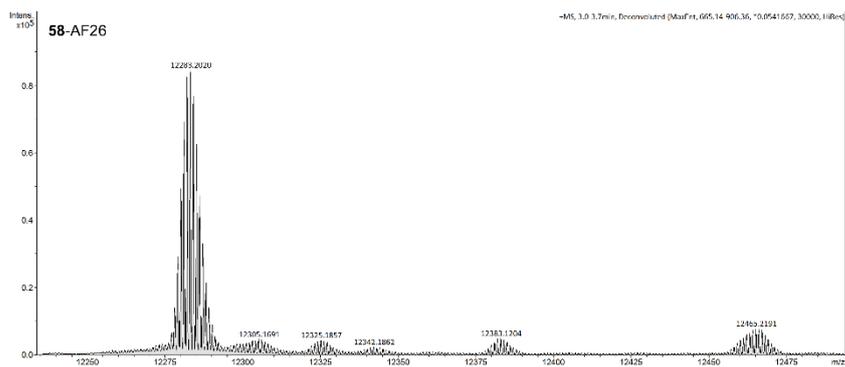
59-AF1: Calcd. 12109.68. Found. 12109.1284.
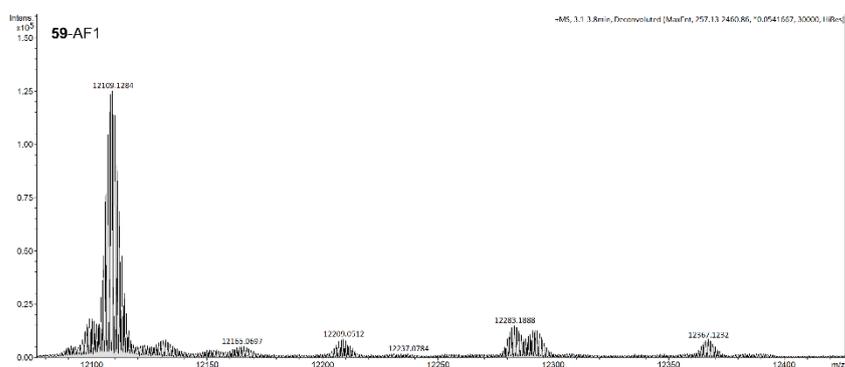
59-AF23: Calcd. 12285.86. Found. 12285.1838.
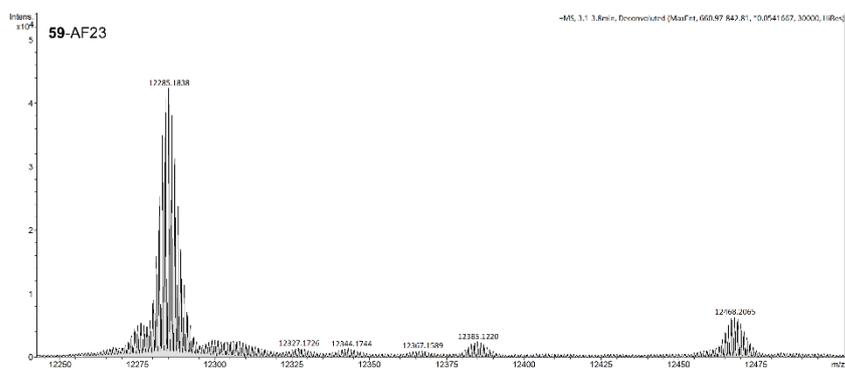
59-AF25: Calcd. 12252.94. Found. 12252.2021.
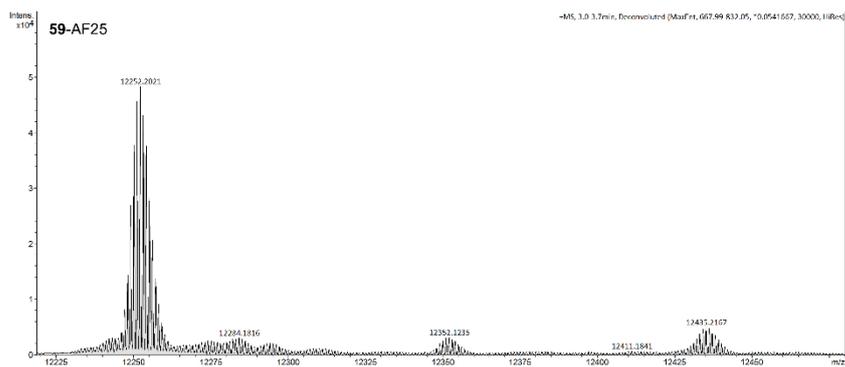
References

1. C.-I. Brändén and J. Tooze, *Introduction to Protein Structure, Second Edition*, Garland Science, 1998.
2. M. Wilhelm, J. Schlegl, H. Hahne, A. M. Gholami, M. Lieberenz, M. M. Savitski, E. Ziegler, L. Butzmann, S. Gessulat, H. Marx, T. Mathieson, S. Lemeer, K. Schnatbaum, U. Reimer, H. Wenschuh, M. Mollenhauer, J. Slotta-Huspenina, J.-H. Boese, M. Bantscheff, A. Gerstmair, F. Faerber and B. Kuster, *Nature*, 2014, **509**, 582-587.
3. K.-H. Altmann, B. Johannes, H. Kessler, F. Diederich, B. Krautler, S. Lippard, R. Liskamp, K. Muller, E. M. Nolan, B. Samori, G. Schneider, S. L. Schreiber, H. Schwalbe, C. Toniolo, C. A. A. v. Boeckel, H. Waldmann and C. T. Walsh, *ChemBioChem*, 2009, **10**, 16-29.
4. K. Lang and J. W. Chin, *Chem. Rev.*, 2014, **114**, 4764-4806.
5. V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161-173.
6. M. N. Fodje and S. Al-Karadaghi, *Protein Eng.*, 2002, **15**, 353-358.
7. T. A. Edwards and A. J. Wilson, *Amino Acids*, 2011, **41**, 743-754.
8. J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001-1100.
9. A. S. Ripka and D. H. Rich, *Cuss. Opin. Chem. Biol.*, 1998, **2**, 441.
10. L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi and E. A. Liu, *Science*, 2004, **303**, 844-848.
11. B. P. Orner, J. T. Ernst and A. D. Hamilton, *J. Am. Chem. Soc.*, 2001, **123**, 5382-5383.
12. S. Jones and J. M. Thornton, *Proc. Natl. Acad. Sci. USA.*, 1996, **93**, 13-20.
13. L. L. Conte, C. Chothia and J. Janin, *J. Mol. Biol.*, 1999, **285**, 2177-2198.
14. T. Clackson and J. A. Wells, *Science*, 1995, **267**, 383-386.
15. A. S. Voisin-Chiret and S. Rault, *Pure Appl. Chem.*, 2012, **84**, 2467-2478.
16. A. L. Jochim and P. S. Arora, *ACS Chem. Biol.*, 2010, **5**, 919-923.
17. S. Fahs, Y. Patil-Sen and T. J. Snape, *ChemBioChem*, 2015, **16**, 1840-1853.
18. S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173-180.
19. G. Guichard and I. Huc, *Chem. Commun.*, 2011, **47**, 5933–5941.
20. J. P. Schneider and J. W. Kelly, *Chem. Rev.* , 1995, **95**, 2169-2187.
21. M. J. I. Andrews and A. B. Tabor, *Tetrahedron*, 1999, **55**, 11711-11743.
22. C. Bracken, J. Gulyas, J. W. Taylor and J. Baum, *J. Am. Chem. Soc.*, 1994, **116**, 6431-6432.
23. J. C. Phelan, N. J. Skelton, A. C. Braisted and R. S. McDowell, 1997, **119**, 455-460.
24. J. H. B. Pease, R. W. Storrs and D. E. Wemmer, *Proc. Natl. Acad. Sci. U.S.A.*, 1990, **87**, 5643-5647.
25. D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh and P. G. Schultz, *J. Am. Chem. Soc.*, 1991, **113**, 9391-9392.
26. M. Pellegrini, M. Royo, M. Chorev and D. F. Mierke, *J. Pept. Res.* , 1997, **49**, 404-414.
27. A. K. Galande, K. S. Bramlett, J. O. Trent, T. P. Burris, J. L. Wittliff and A. F. Spatola, *ChemBioChem*, 2005, **6**, 1991-1998.
28. M. Chorev, E. Roubini, R. L. McKee, S. W. Gibbons, M. E. Goldman, M. P. Caulfield and M. Rosenblatt, *Biochemistry* 1991, **30**, 5968–5974.
29. T. R. Geistlinger and R. K. Guy, *J. Am. Chem. Soc.*, 2003, **125**, 6852–6853
30. S. K. Sia, P. A. Carr, A. G. Cochran, V. N. Malashkevich and P. S. Kim, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 14664-14669.
31. H. E. Blackwell and R. H. Grubbs, *Angew. Chem., Int. Ed.* , 1998, **37**, 3281-3284.
32. C. E. Schafmeister, J. Po and G. L. Verdine, *J. Am. Chem. Soc.*, 2000, **122**, 5891-5892.
33. D. J. Yeo, S. L. Warriner and A. J. Wilson, *Chem. Commun.*, 2013, **49**, 9131-9133.
34. J. A. Miles, D. J. Yeo, P. Rowell, S. Rodriguez-Marin, C. M. Pask, S. L. Warriner, T. A. Edwards and A. J. Wilson, *Chemical Science*, 2016.
35. A. Patgiri, A. M. Jochim and P. S. Arora, *Acc. Chem. Res*, 2008, **41**, 1289-1300.
36. D. Wang, M. Lu and P. S. Arora, *Angew. Chem. Int. Ed.*, 2008, **47**, 1879-1882.
37. A. Skerra, *Curr. Opin. Chem. Biol.* , 2003, **7**, 683-693.
38. J. Pande, M. M. Szewczyk and A. K. Grover, *Biotechnol. Adv.*, 2010, **28**, 849–858.
39. K. Hertveldt, T. Belien and G. Volckaert, *Methods Mol Biol.* , 2009, **502**, 321-339.
40. A. Kiewitz and H. Wolfes, *FEBS Letters*, 1997, **415** 258-262.
41. M. Voss, M. Lettau and O. Janssen, *BMC Immunology*, 2009, **10**, 53-64.

42. K. J. James, M. A. Hancock, J.-N. Gagnon and J. W. Coulton, *Biochemistry*, 2009, **48**, 9212–9220.
43. G. P. Smith and V. A. Petrenko, *Chem. Rev.*, 1997, **97**, 391-410.
44. S. F. Parmley and G. P. Smith, *Gene*, 1988, **73** 305-318.
45. V. A. Petrenko, G. P. Smith, M. M. Mazooji and T. Quinn, *Protein Eng.*, 2002, **15**, 943-950.
46. P. Diderich, D. Bertoldo, P. Dessen, M. M. Khan, I. Pizzitola, W. Held, J. Huelsken and C. Heinis, *ACS Chem. Biol.*, 2016.
47. G. P. Dado and S. H. Gellman, *J. Am. Chem. Soc*, 1994, **116**, 1054-1062.
48. D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell and S. H. Gellman, *J. Am. Chem. Soc.*, 1996, **118**, 13071-13072.
49. D. H. Appella, L. A. Christianson, D. A. Klein, D. R. Powell, X. Huangt, J. J. J. Barchi and G. S. H., *Nature*, 1997, **387**, 381-384.
50. J. W. Checco, E. F. Lee, M. Evangelista, N. J. Sleebs, K. Rogers, A. Pettikiriarachchi, N. J. Kershaw, G. A. Eddinger, D. G. Belair, J. L. Wilson, C. H. Eller, R. T. Raines, W. L. Murphy, B. J. Smith, S. H. Gellman and W. D. Fairlie, *J. Am. Chem. Soc.*, 2015, **137**, 11365-11375.
51. J. A. Kritzer, J. D. Lear, M. E. Hodsdon and A. Schepartz, *J. Am. Chem. Soc*, 2004, **126**, 9468-9469.
52. O. M. Stephens, S. Kim, B. D. Welch, M. E. Hodsdon, M. S. Kay and A. Schepartz, *J. Am. Chem. Soc.*, 2005, **127**, 13126-13127.
53. A. D. Bautista, O. M. Stephens, L. Wang, R. A. Domaoal, A. K. S. and A. Schepartz, *Bioorg. Med. Chem. Lett*, 2009, **19**, 3736–3738.
54. C. M. Grison, S. Robin and D. J. Aitken, *Chem. Commun.*, 2015, **51**, 16233-16236.
55. B. Baptiste, F. Godde and I. Huc, *ChemBioChem*, 2009, **10**, 1765-1767.
56. R. David, R. Günther, L. Baumann, T. Lühmann, D. Seebach, H.-J. Hofmann and A. G. Beck-Sickinger, *J. Am. Chem. Soc.*, 2008, **130**, 15311–15317.
57. Z. E. Reinert, G. A. Lengyel and W. S. Horne, *J. Am. Chem. Soc.*, 2013, **135**, 12528-12531.
58. T. Etezady-Esfarjani, C. Hilty, K. Wuthrich, M. Rueping, J. Schreiber and D. Seebach, *HelV. Chim. Acta*, 2002, **85**, 1197-1209.
59. S. A. Hart, A. B. Bahadoor, E. E. Matthews, X. J. Qiu and A. Schepartz, *J. Am. Chem. Soc.*, 2003, **125**, 4022-4023.
60. J. X. Qiu, E. J. Petersson, E. E. Matthews and A. Schepartz, *J. Am. Chem. Soc.*, 2006, **128**, 11338-11339.
61. D. S. Daniels, E. J. Petersson, J. X. Qiu and A. Schepartz, *J. Am. Chem. Soc.*, 2007, **129**, 1532-1533.
62. I. Huc, *Eur. J. Org. Chem.*, 2004, 17-29.
63. B. Gong, *Chem. Eur. J.*, 2001, **7**, 4336-4342.
64. I. Huc, *Polym. Prep.*, 2003, **44**, 587-588.
65. K. Yamato, M. Kline and B. Gong, *Chem. Comm.*, 2012, **48**, 12142-12158.
66. T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Comm.*, 2012, **48**, 6337–6339.
67. M. Kudo, V. Maurizot, B. Kauffmann, A. Tanatani and I. Huc, *J. Am. Chem. Soc.*, 2013, **135**, 9628-9631.
68. E. R. Gillies, C. Dolain, J.-M. Léger and I. Huc, *J. Org. Chem.*, 2006, **71**, 7931-7939.
69. X. Hu, S. J. Dawson, Y. Nagaoka, A. Tanatani and I. Huc, *The Journal of Organic Chemistry*, 2016, **81**, 1137-1150.
70. S. Shirude, E. R. Gillies, S. Ladame, F. Godde, K. Shin-ya, I. Huc and S. Balasubramanian, *J. Am. Chem. Soc.*, 2007, **129**.
71. L. Delaurière, Z. Dong, K. Laxmi-Reddy, F. Godde, J.-J. Toulmé and I. Huc, *Angew. Chem.*, 2012, **51**, 473-477.
72. P. K. Mandal, B. Baptiste, B. Langlois d'Estaintot, B. Kauffmann and I. Huc, *ChemBioChem*, 2016, **17**, 1911-1914.
73. J. Garric, J.-M. Léger and I. Huc, *Chem. Eur. J.*, 2007, **13**, 8454-8462.
74. N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil and I. Huc, *Nat. Chem.*, 2015, **7**, 334-341.

75. N. Chandramouli, M. F. El-Behairy, G. Lautrette, Y. Ferrand and I. Huc, *Org. Biomol. Chem.*, 2016, **14**, 2466-2472.
76. N. Chandramouli, Y. Ferrand, B. Kauffmann and I. Huc, *Chem. Commun.*, 2016, **52**, 3939-3942.
77. J. T. Ernst, J. Becerril, H. S. Park, H. Yin and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2003, **42**, 535-539.
78. G. N. Tew, D. Liu, B. Chen, R. J. Doerksen, J. Kaplan, P. J. Carroll, M. L. Klein and W. F. DeGrado, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 5110-5114.
79. J. T. Ernst, O. Kutzki, A. K. Debnath, S. Jiang, H. Lu and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2002, **41**.
80. H. Yin, G.-I. Lee, K. A. Sedey, O. Kutzki, H. S. Park, B. P. Orner, J. T. Ernst, H.-G. Wang, S. M. Sebti and A. D. Hamilton, *J. Am. Chem. Soc.*, 2005, **127**, 10191-10196.
81. J. M. Rodriguez and A. D. Hamilton, *Tetrahedron Lett.*, 2006, **47**, 7443-7446.
82. M. J. Adler and A. D. Hamilton, *J. Org. Chem.*, 2011, **76**, 7040-7047.
83. J. Plante, F. Campbell, B. Malkova, C. Kilner, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2008, **6**, 138–146.
84. J. L. Yap, X. Cao, K. Vanommeslaeghe, K.-Y. Jung, C. Peddaboina, P. T. Wilder, A. Nan, A. D. MacKerell Jr., W. R. Smythe and S. Fletcher, *Org. Biomol. Chem.*, 2012, **10**, 2928-2933.
85. S. Marimganti, M. N. Cheemala and J.-M. Ahn, *Org. Lett.*, 2009, **11**, 4418-4421.
86. S. Thompson, R. Vallinayagam, M. J. Adler, R. T. W. Scott and A. D. Hamilton, *Tetrahedron*, 2012, **68**, 4501-4505.
87. B. N. Bullock, A. L. Jochim and P. S. Arora, *J. Am. Chem. Soc.*, 2011, **133**, 14220–14223.
88. S. Thompson and A. D. Hamilton, *Org. Biomol. Chem.*, 2012, **10**, 5780–5782.
89. J. M. Rodriguez and A. D. Hamilton, *Angew. Chem.*, 2007, **119**, 8768 –8771.
90. K.-Y. Jung, K. Vanommeslaeghe, M. E. Lanning, J. L. Yap, C. Gordon, P. T. Wilder, A. D. MacKerell Jr. and S. Fletcher, *Org. Lett.*, 2013, **15**, 3234-3237.
91. C. G. Cummings and A. D. Hamilton, *Tetrahedron*, 2013, **69**, 1663-1668.
92. S. S. and Y. W., *Cell Signal.*, 2008, **20**, 1927-1934.
93. O. V. Kulikov, S. Thompson, H. Xu, C. D. Incarvito, R. T. W. Scott, I. Saraogi, L. Nevola and A. D. Hamilton, *Eur. J. Org. Chem.*, 2013, 3433–3445.
94. S. B. H. Kent, *Annu. Rev. Biochem.*, 1988, **57**, 957-989.
95. S. B. H. Kent, *Chem. Soc. Rev.*, 2009, **38**, 338-351.
96. B. L. Nilsson, M. B. Soellner and R. T. Raines, *Annu. Rev. Biophys. Biomol. Struct.*, 2005, **34**, 91-118.
97. L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498-500.
98. L. Wang, J. Xie, A. A. Deniz and P. G. Schultz, *J. Org. Chem.*, 2003, **68**, 174-176.
99. L. Wang, Z. Zhang, A. Brock and P. G. Schultz, *Proc. Natl. Acad. Sci. USA.*, 2003, **100**, 56-61.
100. V. W. Cornish, D. Mendel and P. G. Schultz, *Angew. Chem. Int. Ed.*, 1995, **34**, 621-633.
101. C. A. Hutchinson, S. Phillips, M. H. Edgell, S. Gillam, P. Janhke and M. Smith, *J. Biol. Chem*, 1978, **253**, 6551-6559.
102. T. Clackson, M. H. Ultsch, J. A. Wells and A. M. de Vos, *J. Mol. Biol.*, 1998, **277**, 1111-1128.
103. C. J. Noren, S. J. Anthony-Cahill, M. C. Griffith and P. G. Schultz, *Science*, 1989, **244**, 182-188.
104. A. Rinaldi, *EMBO reports*, 2004, **5**, 336-339.
105. C. C. Liu and P. G. Schultz, *Annu. Rev. Biochem.*, 2010, **79**, 413-444.
106. L. Wang, A. Brock, B. Herberich and P. G. Schultz, *Science*, 2001, **292**, 498-500.
107. J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang and P. G. Schultz, *Science*, 2003, **301**, 964-967.
108. M. Taki, T. Hohsaka, H. Murakami, K. Taira and M. Sisido, *FEBS Letters*, 2001, **507**, 35-38.
109. C. H. Kim, J. Y. Axup and P. G. Schultz, *Curr. Op. Chem. Biol.*, 2013, **17**, 412-419.
110. H. Xiao, A. Chatterjee, S.-H. Choi, K. M. Bajjuri, S. C. Sinha and P. G. Schultz, *Angew. Chem.*, 2013, **125**, 1-5.

111. T. Hohsaka, Y. Ashizuka, H. Taira, H. Murakami and M. Sisido, *Biochemistry*, 2001, **40**, 11060-11064.
112. T. Hohsaka and M. Sisido, *Curr. Op. Chem. Biol.*, 2002, **6**, 809-815.
113. V. Azzarito, P. Prabhakaran, A. I. Bartlett, N. S. Murphy, M. J. Hardie, C. A. Kilner, T. A. Edwards, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2012, **10**, 6469-6472.
114. F. Campbell, J. Plante, T. A. Edwards, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2010, **8**, 2344-2351.
115. P. Prabhakaran, A. Barnard, N. S. Murphy, C. Kilner, T. A. Edwards and A. J. Wilson, *Eur. J. Org. Chem.*, 2013, 3504–3512.
116. S. Saito, Y. Toriumi, N. Tomioka and A. Itai, *J. Org. Chem.*, 1995, **60**, 4715-4720.
117. V. Azzarito, J. A. Miles, J. Fisher, T. A. Edwards, S. L. Warriner and A. J. Wilson, *Chem. Sci.*, 2015, **6**, 2434–2443.
118. A. Barnard, K. Long, H. L. Martin, J. A. Miles, T. A. Edwards, D. C. Tomlinson, A. Macdonald and A. J. Wilson, *Angew. Chem. Int. Ed.*, 2015, **54**, 2960-2965.
119. G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *ChemBioChem*, 2014, **15**, 1083-1087.
120. G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *Chem. Commun.*, 2016, **52**, 5421-5424.
121. S. Rodriguez-Marin, N. S. Murphy, H. J. Shepherd and A. J. Wilson, *RSC Adv.*, 2015, **5**, 104187-104192.
122. C. Tiede, A. Tang, S. Deacon, U. Mandal, J. Nettleship, R. Owen, R. Owens, D. Tomlinson and M. McPherson, *Protein Eng. Des. Sel.*, 2014, **27**, 145–155.
123. H. F. Kyle, K. F. Wickson, J. Stott, G. M. Burslem, A. L. Breeze, C. Tiede, D. C. Tomlinson, S. L. Warriner, A. Nelson, A. J. Wilson and T. A. Edwards, *Mol. BioSys.*, 2015, **11**, 2738-2749.
124. A. Lupas, *Trends Biochem. Sci.*, 1996, **21**, 375-382.
125. E. Moutevelis and D. N. Woolfson, *J. Mol. Biol.*, 2009, **385**, 726-732.
126. F. H. C. Crick, *Acta. Crystallogr.*, 1953, **6**, 689–697.
127. J. Walshaw and D. N. Woolfson, *J. Struct. Biol.*, 2003, **144**, 349-361.
128. W. Landschulz, P. Johnson and S. McKnight, *Science*, 1988, **240**, 1759-1764.
129. J. N. M. Glover and S. C. Harrison, *Nature*, 1995, **373**, 257-261.
130. K. Arndt and G. R. Fink, *Proc. Natl. Acad. Sci. U.S.A.*, 1986, **83**, 8516-8520.
131. M. R. Gartenberg, C. Ampe, T. A. Steitz and D. M. Crothers, *Proc. Natl. Acad. Sci. U.S.A.*, 1990, **87**, 6034-6038.
132. L. Carlier, A. S. Haase, M. Y. Burgos Zepeda, E. Dassa and O. Lequin, *J. Struct. Biol.*, 2012, **180**, 577-584.
133. M. Y. B. Zepeda, K. Alessandri, D. Murat, C. El Amri and E. Dassa, *Biochim. Biophys. Acta*, 2010, **1804**, 755-761.
134. M. Wiener, D. Freymann, P. Ghosh and R. M. Stroud, *Nature*, 1997, **385**, 461-464.
135. P. R. Tsuruda, D. Julius and D. L. Minor, Jr., *Neuron*, **51**, 201-212.
136. J. Maupin-Furlow, *Nat Rev Micro*, 2012, **10**, 100-111.
137. J. Martin, M. Gruber and A. N. Lupas, *Trends Biochem. Sci.*, 2004, **29**, 455-458.
138. R. Siegert, M. R. Leroux, C. Scheufler, F. U. Hartl and I. Moarefi, *Cell*, **103**, 621-632.
139. S. Djuranovic, M. D. Hartmann, M. Habeck, A. Ursinus, P. Zwickl, J. Martin, A. N. Lupas and K. Zeth, *Mol. Cell*, 2009, **34**, 580-590.
140. L. C. Welsh, M. F. Symmons and D. A. Marvin, *Acta Cryst.*, 2000, **56**, 137-150.
141. L. Sun, L. N. Young, X. Zhang, S. P. Boudko, A. Fokine, E. Zbornik, A. P. Roznowski, I. J. Molineux, M. G. Rossmann and B. A. Fane, *Nature*, 2014, **505**, 432-435.
142. D. Linke, T. Riess, I. B. Autenrieth, A. Lupas and V. A. J. Kempf, *Trends Microbiol.*, 2006, **14**, 264-270.
143. H. U. Ferris, K. Zeth, M. Hulko, S. Dunin-Horkawicz and A. N. Lupas, *J. Struct. Biol.*, 2014, **186**, 349-356.
144. M. Rabe, C. Schwieger, H. R. Zope, F. Versluis and A. Kros, *Langmuir*, 2014, **30**, 7724-7735.
145. J. J. Skehel and D. C. Wiley, *Cell*, **95**, 871-874.
146. T. Zheng, J. Voskuhl, F. Versluis, H. R. Zope, I. Tomatsu, H. R. Marsden and A. Kros, *Chem. Commun.*, 2013, **49**, 3649-3651.

147. I. A. Wilson, J. J. Skehel and D. C. Wiley, *Nature*, 1981, **289**, 366-373.
148. Y. Yu, D. King and Y. Shin, *Science*, 1994, **266**, 274-276.
149. L. J. Calder and P. B. Rosenthal, *Nat. Struct. Mol. Biol.*, 2016, **23**, 853-858.
150. J. S. Blijleven, S. Boonstra, P. R. Onck, E. van der Giessen and A. M. van Oijen, *Semin. Cell Dev. Biol.*, 2016, **60**, 78-88.
151. V. N. Malashkevich, B. J. Schneider, M. L. McNally, M. A. Milhollen, J. X. Pang and P. S. Kim, *Proc. Natl. Acad. Sci. U.S.A.*, 1999, **96**, 2662-2667.
152. S. Watanabe, A. Takada, T. Watanabe, H. Ito, H. Kida and Y. Kawaoka, *J. Virol.*, 2000, **74**, 10194-10201.
153. W. Weissenhorn, A. Carfí, K.-H. Lee, J. J. Skehel and D. C. Wiley, *Mol. Cell*, 1998, **2**, 605-616.
154. W. Weissenhorn, A. Dessen, S. C. Harrison, J. J. Skehel and D. C. Wiley, *Nature*, 1997, **387**, 426-430.
155. D. C. Chan, C. T. Chutkowski and P. S. Kim, *Proc. Natl. Acad. Sci. U.S.A.*, 1998, **95**, 15613-15617.
156. C. A. Bewley, J. M. Louis, R. Ghirlando and G. M. Clore, *J. Biol. Chem.*, 2002, **277**, 14238-14245.
157. D. M. Eckert, V. N. Malashkevich, L. H. Hong, P. A. Carr and P. S. Kim, *Cell*, 1999, **99**, 103-115.
158. D. M. E. and and P. S. Kim, *Annu. Rev. Biochem.*, 2001, **70**, 777-810.
159. E. K. O'Shea, K. J. Lumb and P. S. Kim, *Curr. Biol.*, 1993, **3**, 658-667.
160. R. A. Kammerer and M. O. Steinmetz, *J. Struct. Biol.*, 2006, **155**, 146-153.
161. D. N. Woolfson, in *Advances in Protein Chemistry*, Academic Press, 2005, vol. Volume 70, pp. 79-112.
162. D. N. Woolfson, G. J. Bartlett, A. J. Burton, J. W. Heal, A. Niitsu, A. R. Thomson and C. W. Wood, *Curr. Op. Struct. Biol.*, 2015, **33**, 16-26.
163. A. Lomander, W. Hwang and S. Zhang, *Nano Lett.*, 2005, **5**, 1255-1260.
164. G. R. Dieckmann, A. B. Dalton, P. A. Johnson, J. Razal, J. Chen, G. M. Giordano, E. Muñoz, I. H. Musselman, R. H. Baughman and R. K. Draper, *J. Am. Chem. Soc.*, 2003, **125**, 1770-1777.
165. F. Thomas, N. C. Burgess, A. R. Thomson and D. N. Woolfson, *Angew. Chem. Int. Ed.*, 2016, **55**, 987-991.
166. H. Dong, S. E. Paramonov and J. D. Hartgerink, *J. Am. Chem. Soc.*, 2008, **130**, 13691-13695.
167. D. N. Woolfson and M. G. Ryadnov, *Curr. Op. Chem. Biol.*, 2006, **10**, 559-567.
168. D. N. Woolfson, G. J. Bartlett, M. Bruning and A. R. Thomson, *Curr. Op. Struct. Biol.*, 2012, **22**, 432-441.
169. K. R. Mahendran, A. Niitsu, L. Kong, A. R. Thomson, R. B. Sessions, D. N. Woolfson and H. Bayley, *Nat Chem*, 2016.
170. W. S. Horne, J. L. Price, J. L. Keck and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 4178-4180.
171. T. Sawada and S. H. Gellman, *J. Am. Chem. Soc.*, 2011, **133**, 7336-7339.
172. E. K. Nyakatura, J. Mortier, V. S. Radtke, S. Wieczorek, R. Rezaei Araghi, C. Baldauf, G. Wolber and B. Koksch, *ACS Med. Chem. Lett.*, 2014, **5**, 1300-1303.
173. W. S. Horne, M. K. Yadav, C. D. Stout and M. R. Ghadiri, *J. Am. Chem. Soc.*, 2004, **126**, 15366-15367.
174. M. G. Wuo, A. B. Mahon and P. S. Arora, *J. Am. Chem. Soc.*, 2015, **137**, 11618-11621.
175. L. A. Estroff, C. D. Incarvito and A. D. Hamilton, *J. Am. Chem. Soc.*, 2004, **126**, 2-3.
176. J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards and A. J. Wilson, *Chem. Commun.*, 2009, 5091-5093.
177. N. S. Murphy, P. Prabhakaran, V. Azzarito, J. P. Plante, M. J. Hardie, C. A. Kilner, S. L. Warriner and A. J. Wilson, *Chem. Eur. J.* , 2013, **19**, 5546 – 5550.
178. K. Long, T. A. Edwards and A. J. Wilson, *Bioorg. Med. Chem.*, 2013, **21**, 4034–4040.
179. B. Neises and W. Steglich, *Angew. Chem. Int. Ed. Engl.*, 1978, **17**, 522-524.
180. J. Våbenø, M. Brisander, T. Lejon and K. Luthman, *J. Org. Chem.*, 2002, **67**, 9186-9191.
181. T.-L. L. Choi, C. W.; Chatterjee, A. K.; Grubbs, R. H., *J. Am. Chem. Soc.*, 2001, **123**, 10417-10418.

182. D. J. Ritson, R. J. Cox and J. Berge, *Org. Biomol. Chem.*, 2004, **2**, 1921-1933.
183. J. E. Bishop, J. F. O'Connell and H. J. Rapoport, *J. Org. Chem.* , 1991, **56**, 5079-5091.
184. P. Prabhakaran, V. Azzarito, T. Jacobs, M. J. Hardie, C. A. Kilner, T. A. Edwards, S. L. Warriner and A. J. Wilson, *Tetrahedron*, 2012, **68**, 4485-4491.
185. O. V. Kulikov, C. Incarvito and A. D. Hamilton, *Tetrahedron Letters*, 2011, **52**, 3705-3709.
186. K. Long, *Development of Synthetic α-Helix Mimetics as Potent Anticancer Agents*, University of Leeds, 2013.
187. S. H. Gellman, G. P. Dado, G. B. Liang and B. R. Adams, *J. Am. Chem. Soc.*, 1991, **113**, 1164-1173.
188. A. Altmayer-Henzien, V. Declerck, J. Farjon, D. Merlet, R. Guillot and D. J. Aitken, *Angew. Chem. Int. Ed.*, 2015, **54**, 10807-10810.
189. B. R. Linton, M. S. Goodman, E. Fan, S. A. van Arman and A. D. Hamilton, *J. Org. Chem.*, 2001, **66**, 7313-7319.
190. I. Arrata, A. Barnard, D. C. Tomlinson and A. J. Wilson, *Chem. Commun.*, 2017.
191. B. A. F. Le Bailly and J. Clayden, *Chem. Commun.*, 2016, **52**, 4852-4863.
192. A. R. Thomson, C. W. Wood, A. J. Burton, G. J. Bartlett, R. B. Sessions, R. L. Brady and D. N. Woolfson, *Science*, 2014, **346**, 485-488.
193. J. M. Fletcher, R. L. Harniman, F. R. H. Barnes, A. L. Boyle, A. Collins, J. Mantell, T. H. Sharp, M. Antognozzi, P. J. Booth, N. Linden, M. J. Miles, R. B. Sessions, P. Verkade and D. N. Woolfson, *Science*, 2013, **340**, 595-599.
194. A. J. Burton, A. R. Thomson, W. M. Dawson, R. L. Brady and D. N. Woolfson, *Nat Chem*, 2016, **8**, 837-844.
195. E. H. C. Bromley, K. Channon, E. Moutevelis and D. N. Woolfson, *ACS Chem. Biol.*, 2008, **3**, 38-50.
196. S. E. Boyken, Z. Chen, B. Groves, R. A. Langan, G. Oberdorfer, A. Ford, J. M. Gilmore, C. Xu, F. DiMaio, J. H. Pereira, B. Sankaran, G. Seelig, P. H. Zwart and D. Baker, *Science*, 2016, **352**, 680-687.
197. S. Gonen, F. DiMaio, T. Gonen and D. Baker, *Science*, 2015, **348**, 1365-1368.
198. P.-S. Huang, G. Oberdorfer, C. Xu, X. Y. Pei, B. L. Nannenga, J. M. Rogers, F. DiMaio, T. Gonen, B. Luisi and D. Baker, *Science*, 2014, **346**, 481-485.
199. D. A. Uhlenheuer, K. Petkau and L. Brunsveld, *Chem. Soc. Rev.*, 2010, **39**, 2817-2826.
200. C. S. Wood, T. K. Ronson, A. M. Belenguer, J. J. Holstein and J. R. Nitschke, *Nat Chem*, 2015, **7**, 354-358.
201. P. S. P. Wang, J. B. Nguyen and A. Schepartz, *J. Am. Chem. Soc.*, 2014, **136**, 6810-6813.
202. M. De Poli, W. Zawodny, O. Quinonero, M. Lorch, S. J. Webb and J. Clayden, *Science*, 2016, **352**, 575-580.
203. D. Mazzier, M. Crisma, M. De Poli, G. Marafon, C. Peggion, J. Clayden and A. Moretto, *J. Am. Chem. Soc.*, 2016, **138**, 8007-8018.
204. C. Mayer, M. M. Müller, S. H. Gellman and D. Hilvert, *Angew. Chem.*, 2014, **126**, 7098-7101.
205. J. W. Checco, D. F. Kreitler, N. C. Thomas, D. G. Belair, N. J. Rettko, W. L. Murphy, K. T. Forest and S. H. Gellman, *Proc. Natl. Acad. Sci. U.S.A.*, 2015, **112**, 4552-4557.
206. C. M. Grison, J. A. Miles, S. Robin, A. J. Wilson and D. J. Aitken, *Angew. Chem.*, 2016, **128**, 11262-11266.
207. B. B. Lao, I. Grishagin, H. Mesallati, T. F. Brewer, B. Z. Olenyuk and P. S. Arora, *Proc. Natl. Acad. Sci. USA.*, 2014, **111**, 7531-7536.
208. S. Kushal, B. B. Lao, L. K. Henchey, R. Dubey, H. Mesallati, N. J. Traaseth, B. Z. Olenyuk and P. S. Arora, *Proc. Natl. Acad. Sci. USA.*, 2013, **110**, 15602-15607.
209. Q. Gan, Y. Ferrand, C. Bao, B. Kauffmann, A. Grélard, H. Jiang and I. Huc, *Science*, 2011, **331**, 1172-1175.
210. W. H. Jeong, H. Lee, D. H. Song, J.-H. Eom, S. C. Kim, H.-S. Lee, H. Lee and J.-O. Lee, *Nat. Commun.*, 2016, **7**, 11031.
211. U. Arnold, M. P. Hinderaker, B. L. Nilsson, B. R. Huck, S. H. Gellman and R. T. Raines, *J. Am. Chem. Soc.*, 2002, **124**, 8522-8523.
212. U. Arnold, M. P. Hinderaker, J. Köditz, R. Golbik, R. Ulbrich-Hofmann and R. T. Raines, *J. Am. Chem. Soc.*, 2003, **125**, 7500-7501.

213. A. Tam, U. Arnold, M. B. Soellner and R. T. Raines, *J. Am. Chem. Soc.*, 2007, **129**, 12670-12671.
214. L. Delaurière, Z. Dong, K. Laxmi-Reddy, F. Godde, J.-J. Toulmé and I. Huc, *Angew. Chem. Int. Ed.*, 2012, **51**, 473-477.
215. A. Barnard, K. Long, D. J. Yeo, J. A. Miles, V. Azzarito, G. M. Burslem, P. Prabhakaran, T. A. Edwards and A. J. Wilson, *Org. Biomol. Chem.*, 2014, **12**, 6794-6799.
216. B. Steipe, B. Schiller, A. Plückthun and S. Steinbacher, *J. Mol. Biol.*, 1994, **240**, 188-192.
217. L. K. Mosavi, D. L. Minor and Z.-y. Peng, *Proc. Natl. Acad. Sci. USA.*, 2002, **99**, 16029-16034.
218. H. Kondo, K. Abe, Y. Emori and S. Arai, *FEBS Letters*, 1991, **278**, 87-90.
219. M. Raina, R. Sharma, S. E. Deacon, C. Tiede, D. Tomlinson, A. G. Davies, M. J. McPherson and C. Walti, *Analyst*, 2015, **140**, 803-810.
220. R. Sharma, S. E. Deacon, D. Nowak, S. E. George, M. P. Szymonik, A. A. S. Tang, D. C. Tomlinson, A. G. Davies, M. J. McPherson and C. Wälti, *Biosens. Bioelectron.*, 2016, **80**, 607-613.
221. A. E. Rawlings, J. P. Bramble, A. A. S. Tang, L. A. Somner, A. E. Monnington, D. J. Cooke, M. J. McPherson, D. C. Tomlinson and S. S. Staniland, *Chem. Sci.*, 2015, **6**, 5586-5594.
222. M. Scobie, M. F. Mahon and M. D. Threadgill, *J. Chem. Soc., Perkin Trans. 1*, 1994, 203-210.
223. H. Yin, G.-i. Lee, H. S. Park, G. A. Payne, J. M. Rodriguez, S. M. Sebti and A. D. Hamilton, *Angew. Chem.*, 2005, **117**, 2764-2767.
224. A. Nortcliffe, I. N. Fleming, N. P. Bottinga and D. O'Hagana, *Tetrahedron*, 2014, **70**, 8343-8347.
225. S. Machida, N. Kato, K. Harada and J. Ohkanda, *J. Am. Chem. Soc.*, 2011, **133**, 958–963.

Appendix I.   Molecular modelling

AI.1. Conformational analyses

The conformational analyses by molecular modelling of $O\text{-}1_B2_BG$ (**47**), $O\text{-}1_B2_AG$ (**49**) and $N\text{-}1_A2_BG$ (**56**) are shown in Figure A1. The *anti* conformation is observed for $O\text{-}1_B2_BG$ (**47**), which was stabilised by H-bond between the amine group and the carboxylic acid from the glycine at the bottom, and the oxygen from the *O*-alkylation at the top (Figure A1a). For $O\text{-}1_B2_AG$ (**49**), a *syn* conformation was preferred, stabilised by side-chain/side-chain H-bonding (Figure A1b). Finally, $N\text{-}1_A2_BG$ (**56**) adopted a *trans* conformation, thanks to side-chain/side-chain interactions (Figure A1c).
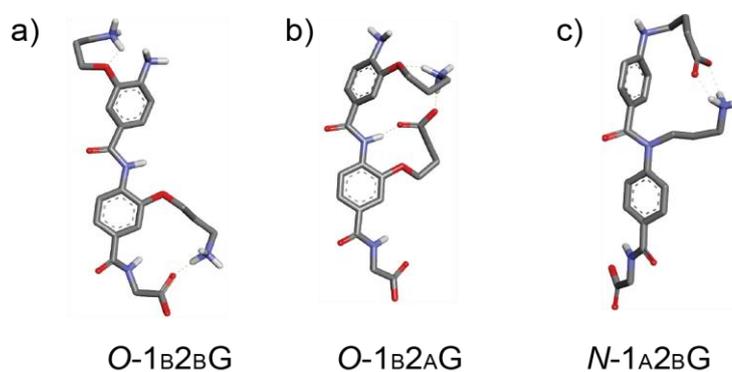

Figure A1: Conformational analyses of a) $O\text{-}1_B2_BG$ (**47**), b) $O\text{-}1_B2_AG$ (**49**) and c) $N\text{-}1_A2_BG$ (**56**).

AI.2. Self-assembly

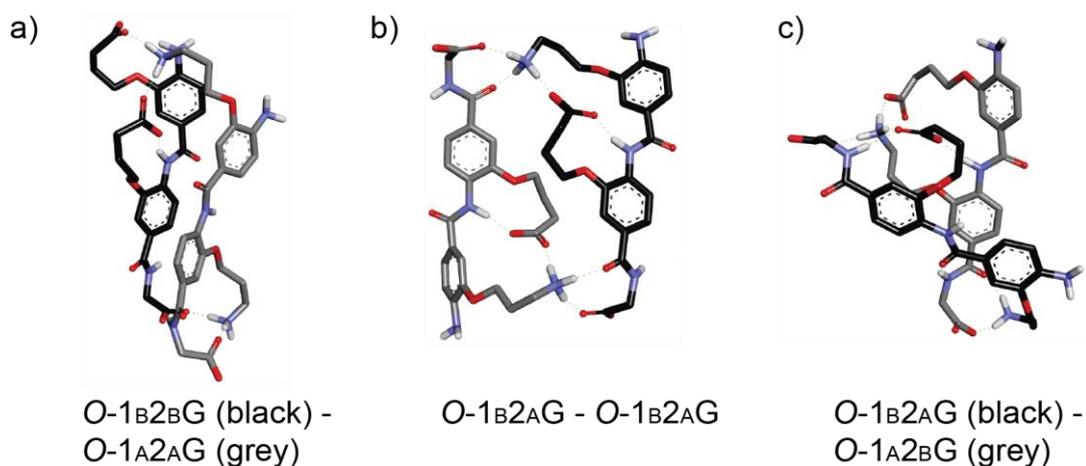The self-assembly was also studied by modelling the interactions between pairs of oligoamides:

Figure A2: Self-assembly of *O*-alkylated dimers: a) $O\text{-}1_B2_BG$ (**47**) with $O\text{-}1_A2_AG$ (**48**), b) $O\text{-}1_B2_AG$ (**49**) with itself, and $O\text{-}1_B2_AG$ (**49**) with $O\text{-}1_A2_BG$ (**50**).

O-1$_B$2$_B$G(1) – *O*-1$_A$2$_A$G(2) (**Figure A2**a): amine of residue 1 of dimer 1 (*O*-**1$_B$**2$_B$G(1)) bound to both carboxylic acids of dimer 2 (*O*-1$_A$**2$_A$**G(2)). The amine on residue 2 of dimer 1 (*O*-1$_B$**2$_B$**G(1)) bound to the carboxylic acids from the glycine of both dimers (*O*-1$_B$2$_B$**G**(1) and *O*-1$_A$2$_A$**G**(2)). The carboxylic acid of residue 2 on dimer 2 (*O*-1$_A$**2$_A$**G(2)) acid also bound to the middle amide on dimer 1 (*O*-1$_B$-CON**H**-2$_B$G(1)).

O-1$_B$2$_A$G(1) – *O*-1$_B$2$_A$G(2) (**Figure A2**b): the complex was symmetrical: the amine on residue 1 of dimer 1 (*O*-**1$_B$**2$_A$G(1)) bound to the carboxylic acid side chain from the same dimer, on residue 2 (*O*-1$_B$**2$_A$**G(1)) and the middle amide group of dimer 1 (*O*-1$_B$-CON**H**-2$_A$G(1)), as well as both the glycine and the bottom amide of dimer 2 (*O*-1$_B$2$_A$-CON**H**-**G**(2)).

O-1$_B$2$_A$G(1) – *O*-1$_A$2$_B$G(2) (**Figure A2**c): the carboxylic acid on residue 1 of dimer 2 (*O*-**1$_A$**2$_B$G(2)) bound to the amine on residue 2 of the same dimer (*O*-1$_A$**2$_B$**G(2)), which bound to both the carboxylic acid of residue 2 and the glycine of dimer 1 (*O*-1$_B$**2$_A$G**(1)). The carboxylic acid of residue 2 on dimer 1 also bound to the middle amide of dimer 2 (*O*-1$_A$-CON**H**-2$_B$G(2)). The amine of residue 1 on dimer 1 (*O*-**1$_B$**2$_A$G(1)) bound to the glycine and bottom amide of dimer 2 (*O*-1$_A$2$_B$-CON**H**-**G**(2)).
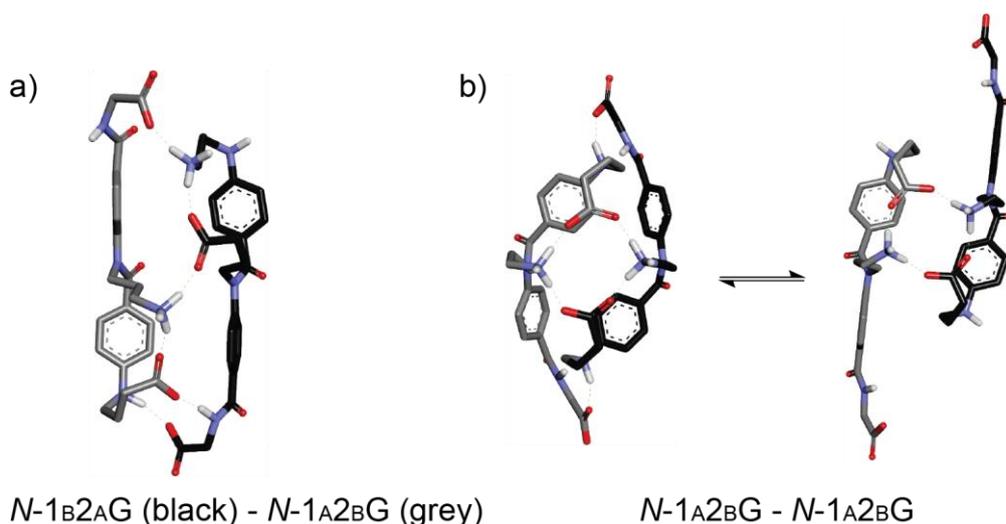

Figure A3: Self-assembly of *N*-alkylated dimers: a) *N*-1$_B$2$_A$G (**55**) with *N*-1$_A$2$_B$G (**56**) and b) *N*-1$_A$2$_B$G (**56**) with itself.

N-1$_B$2$_A$G(1) – *N*-1$_A$2$_B$G(2) (**Figure A3**a): the amine of residue 1 on dimer 1 (*N*-**1$_B$**2$_A$G(1)) bridged between the carboxylic acid of residue 2 on dimer 1 (*N*-1$_B$**2$_A$**G(1)) and glycine on dimer 2 (*N*-1$_A$2$_B$**G**(2)). The amine of residue 2 on dimer 2 (*N*-1$_A$**2$_B$**G(2)) bridged between the carboxylic acid on residue 1 of dimer 2 (*N*-**1$_A$**2$_B$G(2)) and the carboxylic acid of residue 2 on dimer 1 (*N*-1$_B$**2$_A$**G(1)). Finally, the carboxylic acid of residue 1 on dimer 2 (*N*-**1$_A$**2$_B$G(2)) also bound to the bottom glycine amide bond on dimer 1 (*N*-1$_B$2$_A$-CONH-G(1)).

N-1$_A$2$_B$G(1) – *N*-1$_A$2$_B$G(2) (**Figure A3**b): two conformation were observed, both symmetrical. In both cases, the carboxylic acid of residue 1 on dimer 1 (*N*-**1$_A$**2$_B$G(1)) bound to the amine of residue

2 on both dimers (*N*-1$_A$**2$_B$**G(1) and *N*-1$_A$**2$_B$**G(2)). In one case, the glycine was not involved; otherwise, the glycine of dimer 1 (*N*-1$_A$2$_B$**G**(1)) bound to the top aniline of dimer 2 (*N*-1-N**H**-$_A$2$_B$G(2)).

The summary of all interactions observed by molecular modelling is presented in **Table A1**.

Table A1: Summary of all interactions observed in molecular modelling. The group involved in an interactions are denoted in bold, → symbolises a H-bond, & means "binds to both".

Complex	Interactions
O-1$_B$2$_B$G(1) – *O*-1$_A$2$_A$G(2)	*O*-**1$_B$**2$_B$G(1) → *O*-**1$_A$2$_A$**G(2); *O*-1$_B$2$_B$**G**(1) → *O*-1$_B$**2$_B$**G(1) → *O*-1$_A$2$_A$**G**(2); *O*-1$_A$**2$_A$**G(2) → *O*-1$_B$-CON**H-**2$_B$G(1).
O-1$_B$2$_A$G(1) – *O*-1$_B$2$_A$G(2)	*O*-**1$_B$**2$_A$G(1) → *O*-1$_B$**2$_A$**G(1) → *O*-1$_B$-CON**H**-2$_A$G(1) & *O*-1$_B$2$_A$-**C**ONH-**G**(2); *O*-**1$_B$**2$_A$G(2) → *O*-1$_B$**2$_A$**G(2) → *O*-1$_B$-CON**H**-2$_A$G(2) & *O*-1$_B$2$_A$-**C**ONH-**G**(1).
O-1$_B$2$_A$G(1) – *O*-1$_A$2$_B$G(2)	*O*-**1$_A$**2$_B$G(2) → *O*-1$_A$**2$_B$**G(2) → *O*-1$_B$**2$_A$G**(1); *O*-1$_B$**2$_A$**G(1) → *O*-1$_A$-CON**H**-2$_B$G(2); *O*-**1$_B$**2$_A$G(1) → *O*-1$_A$2$_B$-**C**ONH-**G**(2).
N-1$_B$2$_A$G(1) – *N*-1$_A$2$_B$G(2)	*N*-1$_B$2$_A$-CONH-G(1) → *N*-**1$_A$**2$_B$G(2) → *N*-1$_A$**2$_B$**G(2) → *N*-1$_B$**2$_A$**G(1) → *N*-**1$_B$**2$_A$G(1) → *N*-1$_A$2$_B$**G**(2)).
N-1$_A$2$_B$G(1) – *N*-1$_A$2$_B$G(2) **1**	*N*-1$_A$**2$_B$**G(1) → *N*-**1$_A$**2$_B$G(1) → *N*-1$_A$**2$_B$**G(2) → *N*-**1$_A$**2$_B$G(2).
N-1$_A$2$_B$G(1) – *N*-1$_A$2$_B$G(2) **2**	*N*-1$_A$**2$_B$**G(1) → *N*-**1$_A$**2$_B$G(1) → *N*-1$_A$**2$_B$**G(2) → *N*-**1$_A$**2$_B$G(2); *N*-1$_A$2$_B$**G**(1) → *N*-1-N**H**-$_A$2$_B$G(2)

Appendix II. NMR of *N*-alkylated trimers

AII.1. NOESY

The full NOESY spectra of *N*-1$_A$2$_{iBu}$3$_B$G (**54**) and *N*-1$_B$2$_{iBu}$2$_A$G (**53**) in pyridine are shown in Figure A4 and Figure A6. The observed correlations suggest the *cis* conformation (purple circle), and no indication of the presence of *trans* conformation is observed (orange), showing the compounds are fully *cis*. The full NOESY spectra of *N*-1$_A$2$_{iBu}$3$_B$G (**54**) and *N*-1$_B$2$_{iBu}$2$_A$G (**53**) in DMSO are shown in Figure A5 and Figure A7, confirming the *cis* conformation (purple circle) – the correlation or absence of correlation between 1-H2 and 2-Hα, or 2-H2 and 3-Hα could not be observed due to the superimposition of 2-H2 and 3-H3 in the ^{1}H spectra.
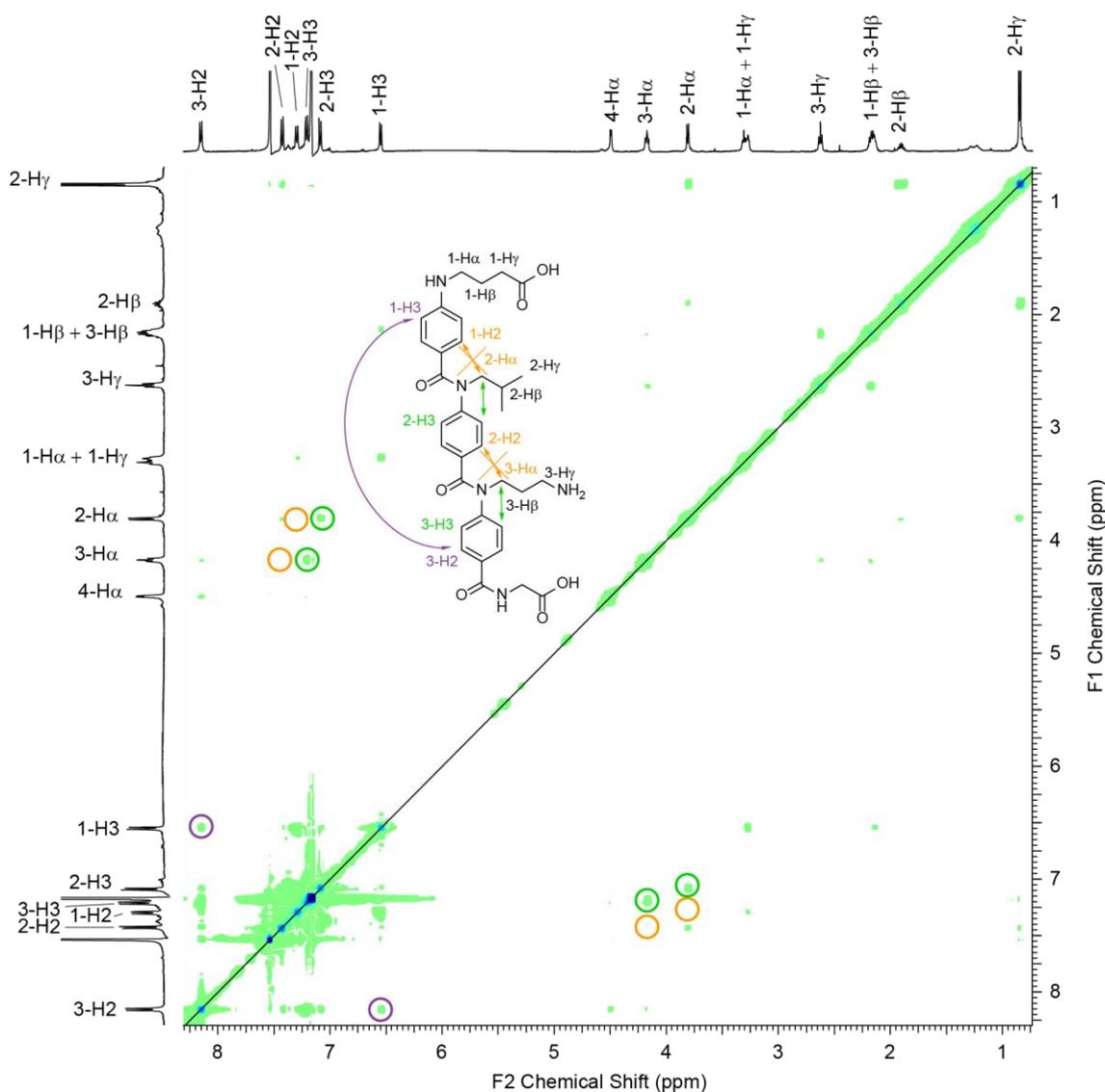
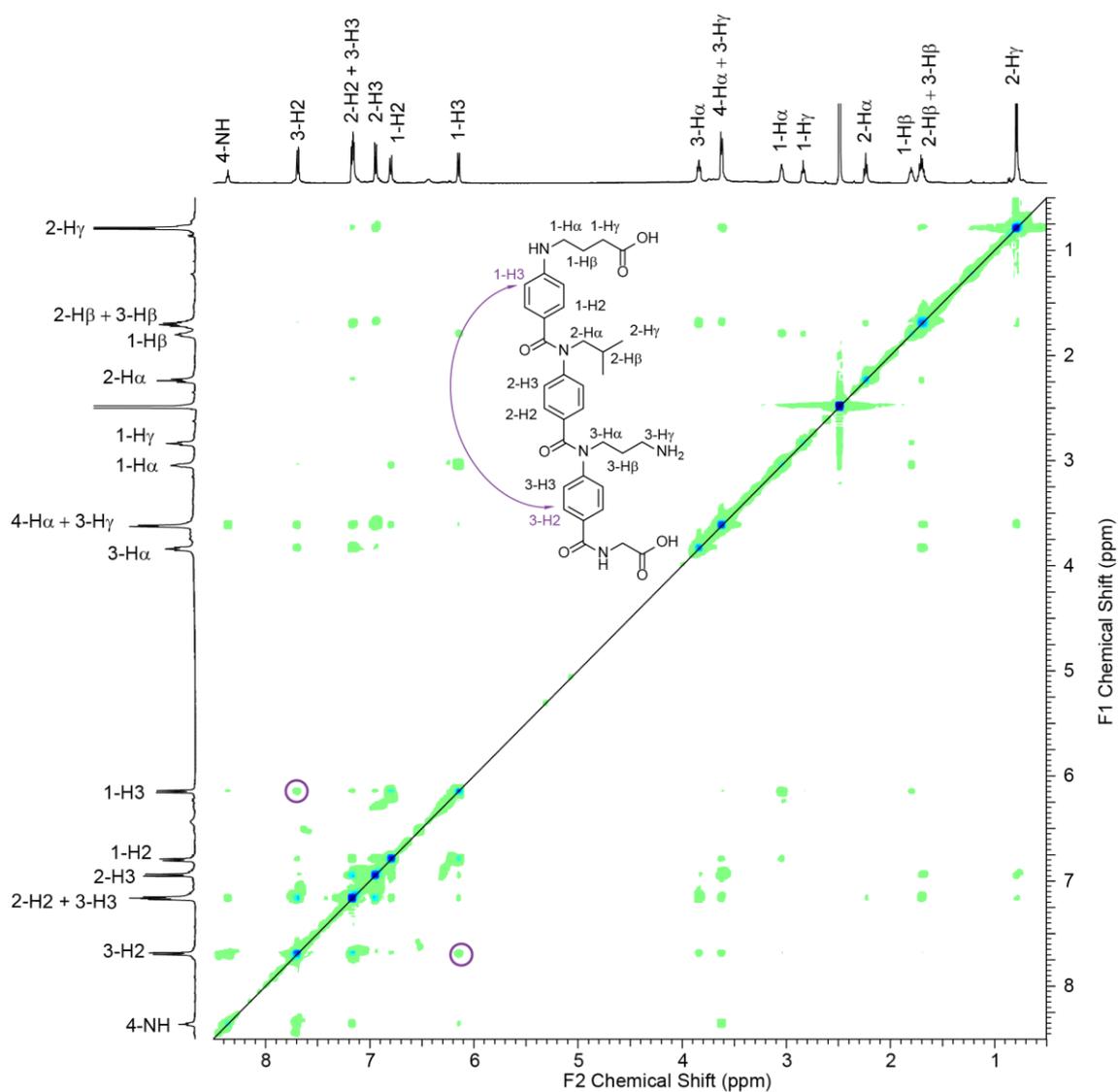**Figure A4:** NOESY of *N*-1$_A$2$_{iBu}$3$_B$G (**54**) in pyr-d$_5$, 16 mM, 500 MHz.

Figure A5: NOESY of *N*-1$_A$2$_{iBu}$3$_B$G (**54**) in DMSO-d$_6$, 40 mM, 500 MHz.
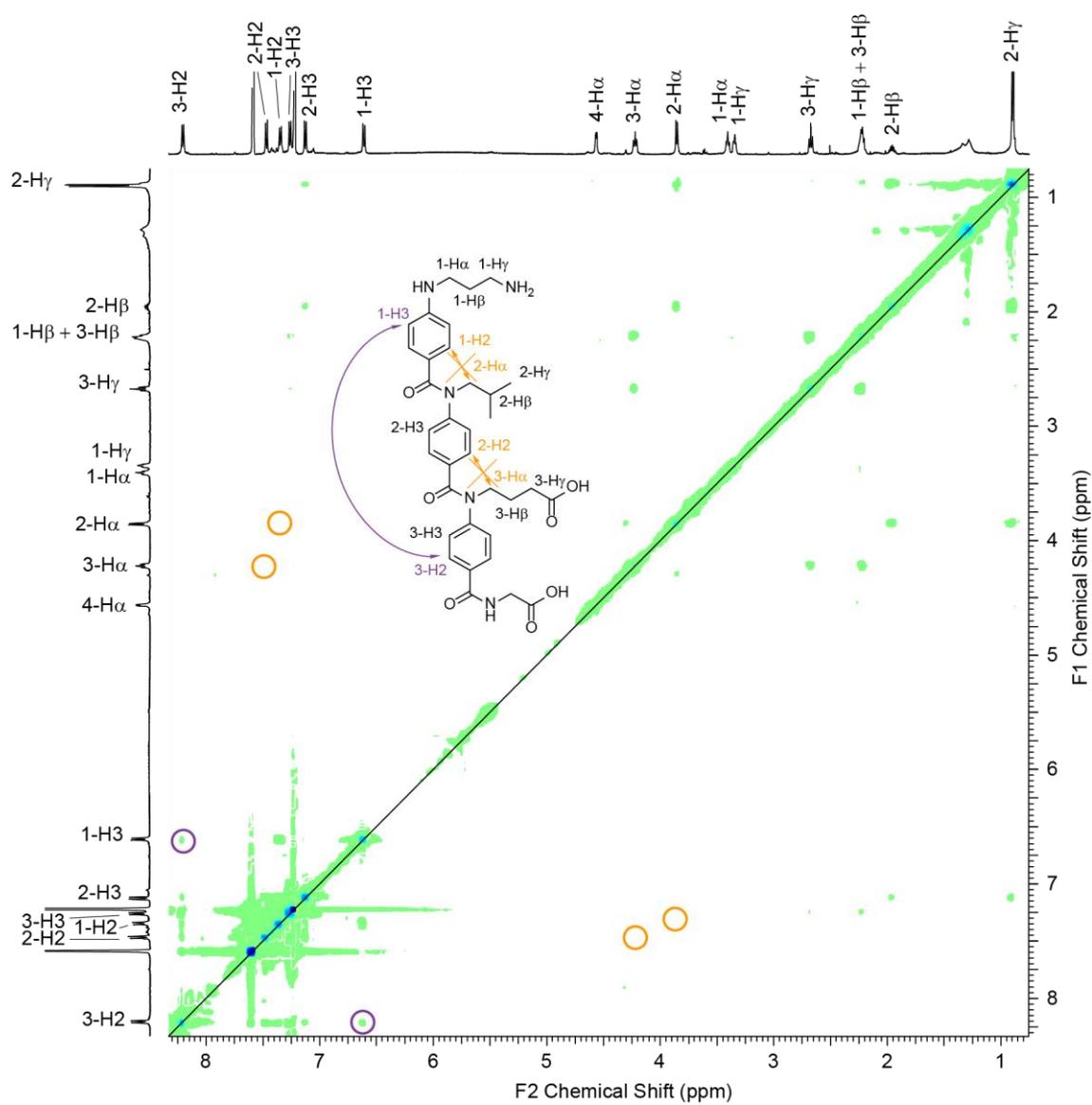
Figure A6: NOESY of *N*-1$_B$2$_{iBu}$3$_A$G (**53**) in pyr-d5, 16 mM, 500 MHz.
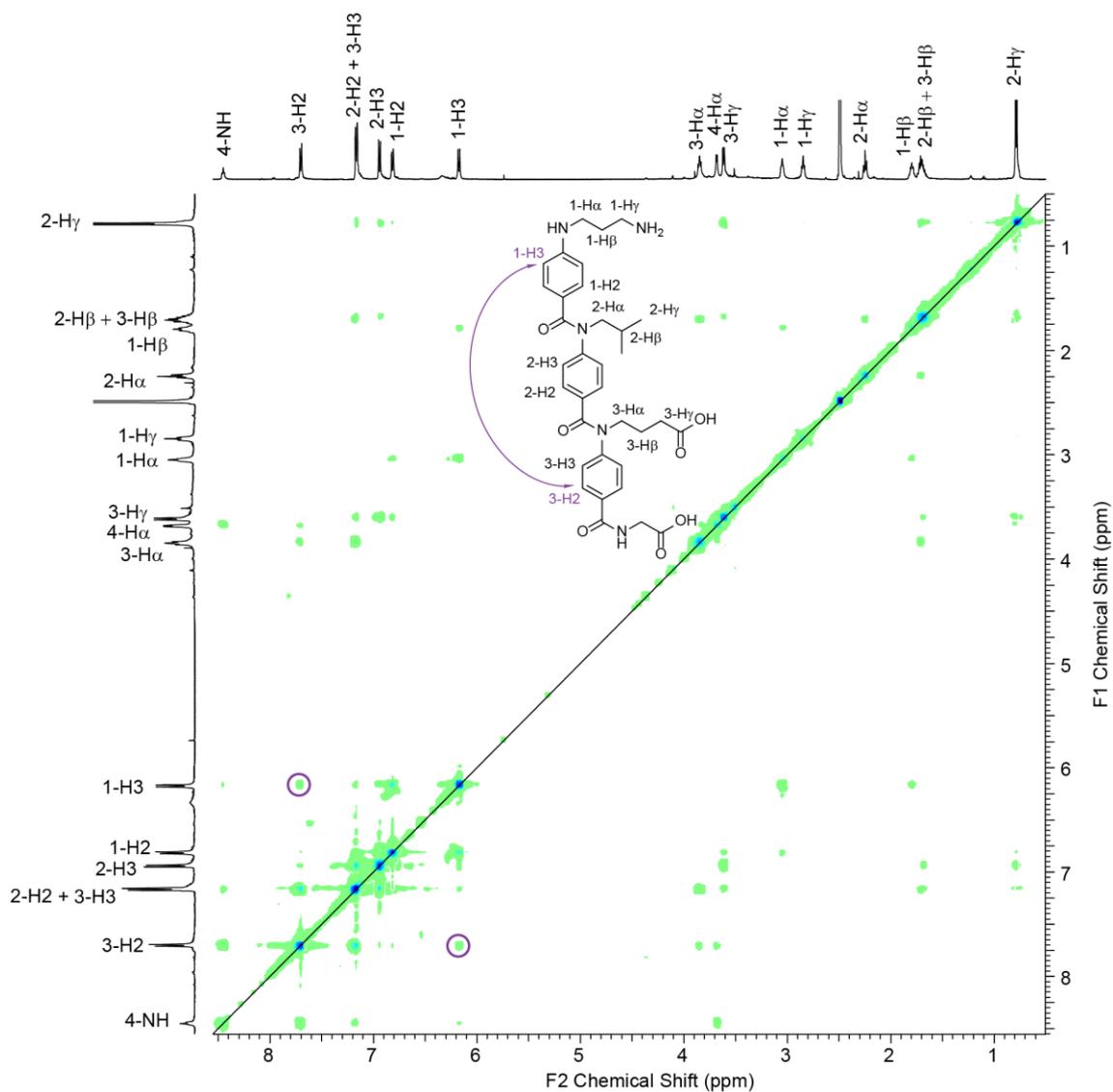
Figure A7: NOESY of *N*-1$_B$2$_{iBu}$2$_A$G (**53**) in DMSO-d$_6$, 40 mM, 500 MHz.

AII.2. ^{1}H NMR

Bottom: *N*-1$_A$2$_{iBu}$2$_B$G (**54**), top: *N*-1$_B$2$_{iBu}$2$_A$G (**53**), middle: 1:1 mixture of both; 16 mM, 500 MHz, pyr-d5.
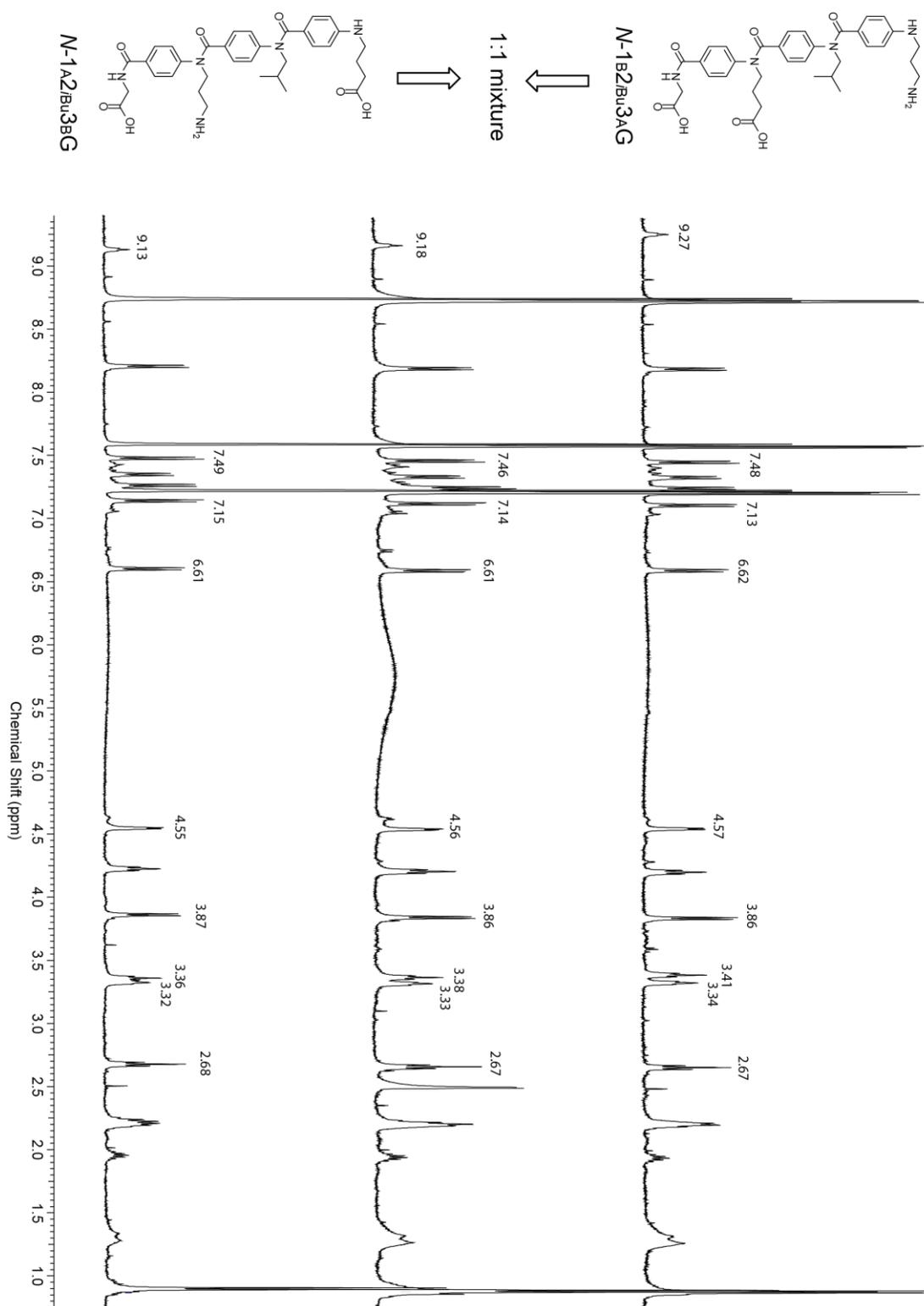
Figure A8: 1H NMR of *N*-1$_A$2$_{iBu}$2$_B$G (**53**, top), *N*-1$_B$2$_{iBu}$2$_A$G (**54**, bottom) and their 1:1 mixture (middle) in pyr-d5, 16 mM, 500 MHz.

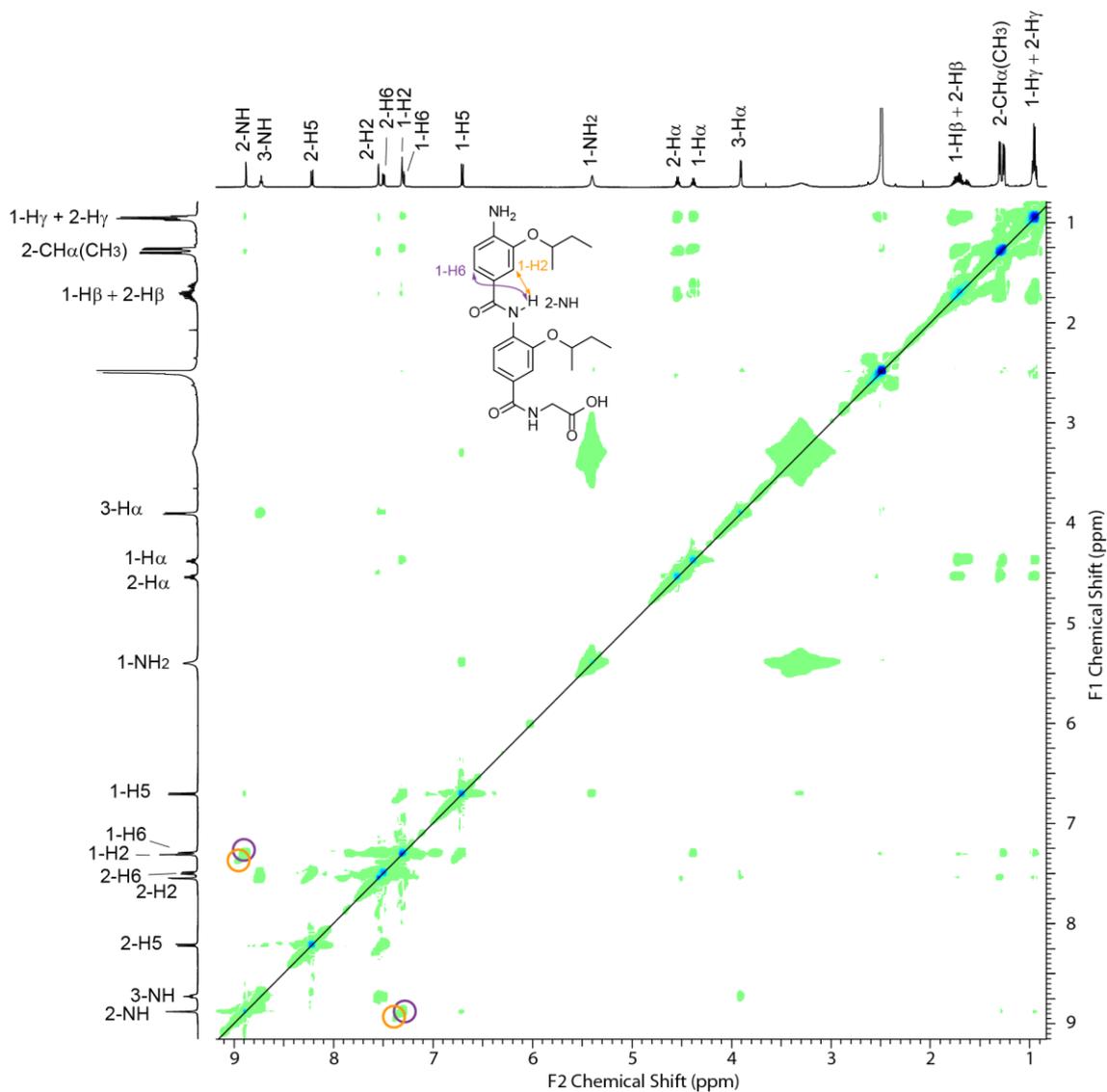# Appendix III. NMR analyses of *O*-1*s*Bu2*s*BuG (57)

AIII.1. NOESY

Figure A9: NOESY NMR of *O*-1*s*Bu2*s*BuG (**57**), highlighting the two conformation *syn* (orange circle) and *anti* (purple circle).
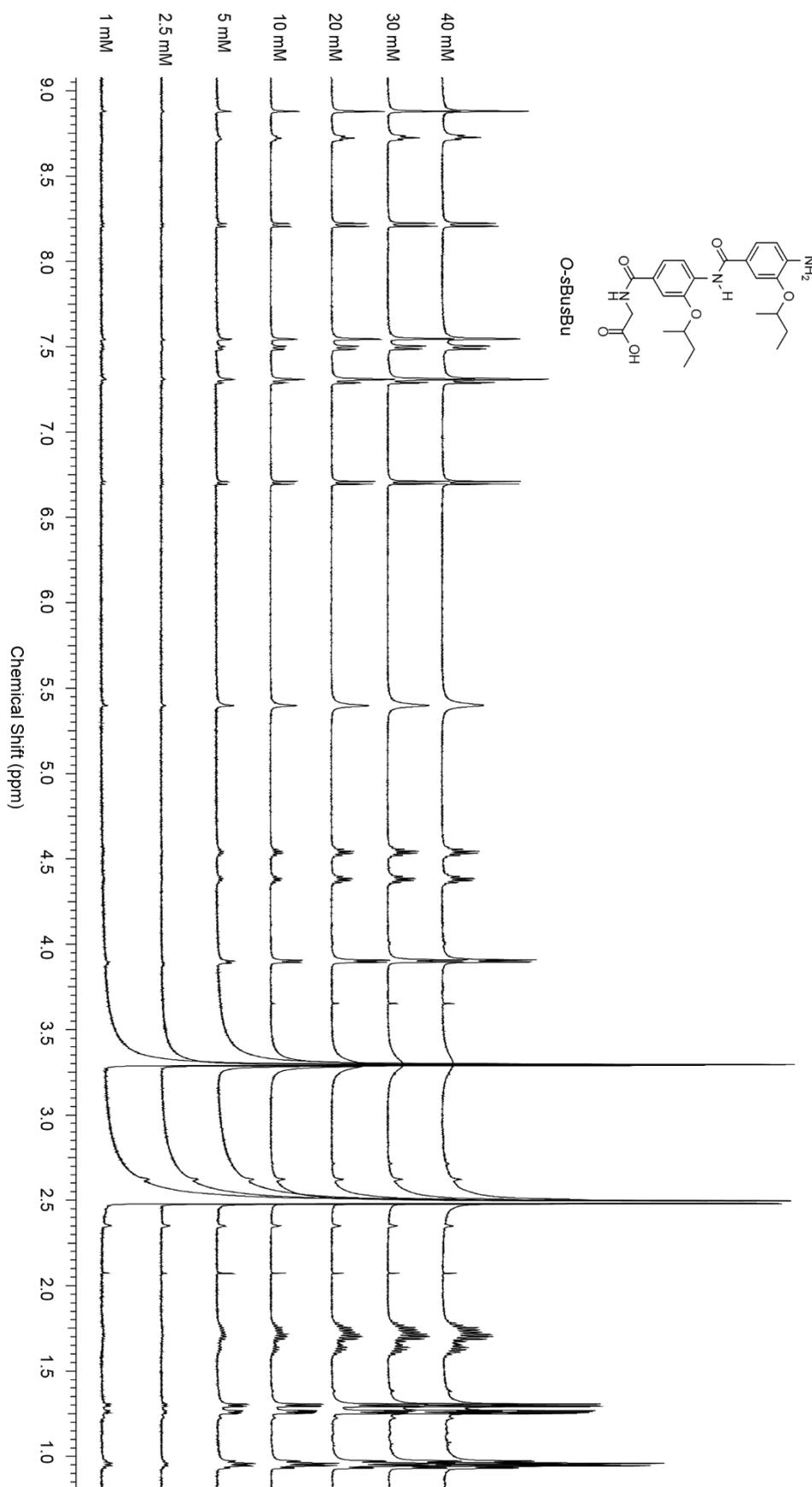
AIII.2. ¹H NMR dilutions

Figure A10: NMR dilutions of O-$1_{sBu}2_{sBu}$G (**57**) from 40 to 1 mM, DMSO-d6, 500 MHz.