Multifunctional Scaffolds for Selective Protein-Protein Inhibition

Silvia Rodriguez Marin

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

University of Leeds

School of Chemistry

October 2016

Intellectual Property and Publication Statements

The candidate confirms that the work submitted is his/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 described in *Chapter 2* constituted the basis for the research article: *'Design, Synthesis and Conformational Analyses of Bifacial Benzamide Based Foldamers',* S. Rodriguez-Marin, N. S. Murphy, H. J. Shepherd, A. J. Wilson, *RSC Advances,* **2015**, *5*, 104187-104192. The contributions of the authors are as follows: SRM (candidate) was the lead author on this piece of work and drafted the original manuscript. AJW edited the manuscript into its present form. N. S. Murphy started the synthesis of the first generation of inhibitors, performed the preliminary molecular modelling and docking for them and also obtained the reported crystal structure. H. J. Shepherd solved the reported crystal structure.

Acknowledgments

First of all, I would like to thank my supervisor Prof. Andy J. Wilson for his constant support and guidance thorough all these years and for never stop challenging me. I will keep with me all the Christmas meals, annual BBQ and Wilson group days out memories. Most importantly, for giving me the opportunity four years ago of starting this adventure, which has completely changed my life. Thanks to the European Research Council (ERC) for funding my PhD and to the people who have supported the project: Simon Barret for his assistance with NMR; Dr Chris Empson for the support with the plate-reader and robot; and specially Martin Huscroft for his continuous help with the HPLC. I would also like to thank other members of staff who helped me during this journey: Dr Stewart Warriner, for his endless positive energy and precious technical support; Dr Sri Shridaran, for simply being "Sri", for always telling me that I was working hard and for driving me to the airport on Eid day; and Dr Julie Fisher for giving me the most comforting words when I just landed at the University of Leeds.

My deepest gratitude goes to the past and current members of the Wilson group. It has been a real pleasure to be part of this little family. I could not have imagined a better lot to share the joys and miseries of the PhD. To the initial members of the team, who made me feel at home since the moment I arrived in this country: Kérya, I wish we had more time together; Tasha for always being genuine; my European mate Valeria, for being a huge support and helping me to get used to the English way; Dave, for being the most manly man in the lab; Anna, my lab mum and salsa introducer; Kelly, thanks for cleaning that desk *for me* and for everything since; George, I will always have another question for you; and Hannah, for making me not the shortest in the group. To the new bunch: Jenny, my hamster buddy and protein provider; Irene, do not need to say: I am very happy you are here!; Sarah, for always being willing to help; Ludwig, for bringing the German touch to the lab; Jayapal, for the countless shared laughs and songs sang out of tune; Raquel, my Catalonian ally; Phil, for always keeping positive; Sam, ¿qué? (Spanish impression); Claire, it is a shame you did not join before; Heather, for her never-ending good mood; Kat, for being the best acquisition that the group could have had. To my lovely "students" Kris and Kate for their great enthusiasm and for making me realise all the things I do know instead than the ones I still need to learn. To others like Steven, who showed me you can share an awful lot of good moments even when you do not understand a word each other is saying. Finally, thank you to all the people that pass through the department and that I do not give a mention to, my PhD would not have been the same without you.

Furthermore, I would like to thank you Dan. From the day you let me in the LC-MS room you have been my main support, thank you for putting up with me and being there at all times (even

in the mornings). You are one of the best things I have got from this PhD, every knock in that office glass brightened my day, and still does.

I would like to give a special recognition to those people whose passion for science and for teaching left a deep mark on me and who led me to do this PhD in Organic Chemistry: David Ballesteros, Marisa Salgado, Dr Francesc Rabanal and Dr Karim Cassimjee.

Als meus amics de casa: Marta, Maria, Vidal, Jordi después de un largo camino lo logramos! simepre interconectados a distancia, i a la Eli, my sister from another mother. Saber que us tinc allà tot i la distancia m'ha ajudat a seguir endavant aquests quatre anys, moltíssimes gràcies!

Finalment, a tota la meva família que és el meu gran orgull i recolzament. En especial als meus avis, als que hi són i als que ja no hi són, perquè sempre esteu amb mi donant-me forces per continuar. Però sobretot, aquesta tesis està dedicada als meus pares, pel seu suport infinit al llarg de tota la meva vida. No hauria pogut arribar aquí si no fos per vosaltres papa i mama, mai us podré agrair tot el que heu fet i seguiu fent per mi. Moltes gràcies, us estimo!

Abstract

Protein-protein interactions (PPIs) play an important role in numerous biological processes. Consequently, modulating PPIs is fundamental for understanding and manipulating mechanisms that govern many diseases. Among the wide range of topographies that PPIs display, the α-helix is the most common secondary structure in nature and thus represents a good generic template for inhibitor design.[1] Some of the most relevant approaches in this field are the proteomimetic approach, which recapitulate the key binding residues of an α-helix on a non-peptidic scaffold; and the constrained peptides, which aim to reproduce the helical structure by stabilising a helical peptide. Both approaches have generated potent inhibitors of a great diversity of α-helix mediated PPIs. However, developing a better understanding of the key features that govern the modulation of protein recognition is necessary to further advance the field and fully exploit each class of foldamer.

In that context, we developed functionalised aromatic oligoamide backbones to mimic residues located on multiple faces of an α-helix to target the ER/co-activator PPI. The novel scaffolds are based on bis-benzamide and *N*-(4-aminophenyl)terephthalamidic acid backbones functionalised with *iso*butyl groups to reproduce the key side chains of the co-activator α-helix. Conformational studies in combination with molecular modeling and docking analysis provide evidence that the new oligomers can adopt conformations that mimic the residues at *i*, *i*+3 and *i*+4 positions of the native co-activator α-helix.

In addition, the rules that govern molecular recognition of protein surfaces were further investigated through the optimisation of the oligobenzamide hybrid scaffold using a structure-activity relationship (SAR) study. A library of compound analogues has been synthesised incorporating five variable sites. The modifications focused on size, polarity and stereochemistry to obtain more potent and selective proteomimetic inhibitors of the p53/*h*DM2 and Mcl-1/NOXA B PPIs.

Finally, using existing methodologies a 3-*O*-alkylated proteomimetic scaffold and hydrocarbon stapling peptide strategy, have been used to design inhibitors of the Asf1/H3 interaction. The application of both approaches allowed the different inhibitor designs to be directly compared when targeting the same PPI.

Table of contents

List of Tables

List of Figures

List of Schemes

List of Abbreviations

aa	Amino acid
AF	Activation Function
AR	Androgen Receptor
Asf1	Anti-silencing function 1 protein
AsLOV2	LOV2 domain of Avena Sativa phototropin 1
Bak	Bcl-2 homologous Antagonist Killer protein
Bax	Bcl-2 Associated X-Protein
Bcl-2	B-cell lymphoma 2
Bcl-x_L	B-cell lymphoma extralarge
BID	BH3-Interacting Domain death agonist
BIM	Bcl-2 Interacting Mediator of cell death
BSA	Bovin Serum Albumin
$^tBoc_2O$	*tertiary*-Butoxycarbonyl anhydride
CAF-1	Chromatin Assembly Factor 1
CaM	Calmodulin
CBP	CREB-binding protein
CD	Circular Dichroism
Cdc42	Cell division cycle 42
CDI	1,1'-Carbonyldiimidazole
CoA	Co-Activator
CPP	Cell Penetrating Peptides
DBD	DNA Binding Domain
Dbl	Diffuse B-cell Lymphoma
Dbs	Dbl's big sister
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCM	Dichloromethane
DHT	Dihydrotestosterone
DIPEA	*N*,*N*′-Diisopropylethylamine
DMF	*N*,*N*′-Dimethylformamide
DMSO	Dimethylsulfoxide
DPPA	Diphenyl Phosphoryl Azide
DTT	Dithiothreitol
E2	17-β-estradiol
EDT	1,2-Ethanedithiol
ER	Estrogen Receptor

ERE	Estrogen Responsive Element
ESI	Electrospray Ionization
FA	Fluorescence Anisotropy
FITC	Fluorescein-isothiocyanate
Fmoc	Fluorenylmethyloxycarbonyl
GLP-1	Glucagon-like peptide-1
GLP-1R	Glucagon-like peptide-1 receptor
gp41	Glycoprotein 41
HATU	1-[Bis(dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-*b*]pyridinium 3-oxid hexafluorophosphate
HBS	Hydrogen Bond Surrogate
HCTU	2-(6-Chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate
*h*DM2	Human Double Minute two
*h*DMX	Human Double Minute X
*h*GH	Human Growth Hormone
HIF-1α	Hypoxia-Inducible Factor 1 alpha
HIRA	Histone Regulator A
HIV-1	Human Immunodeficiency Virus type 1
HMQC	Heteronuclear Multiple Quantum Coherence
HPLC	High-Pressure Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HTS	High-Throughput Screening
IAPP	Islet Amyloid Polypeptide
IC_{50}	Half maximal Inhibitory Concentration
iNOS	Inducible Nitric Oxide Synthase
ITC	Isothermal Titration Calorimetry
K_d	Dissociation constant
K_i	Inhibition constant
LBD	Ligand Binding Domain
LC-MS	Liquid Chromatography–Mass Spectrometry
MBHA	4-Methylbenzhydrylamine
Mcl-1	Induced Myeloid Leukemia cell differentiation protein
*m*DM2	Murine Double Minute two
*m*DMX	Murine Double Minute X
MMFFs	Merck Molecular Force Fields
MRE	Mean Residual Ellipticity

NBS	N-Bromosuccinimide
NMR	Nuclear Magnetic Resonance
NMP	*N*-Methylpyrrolidinone
NOESY	Nuclear Overhauser Effect Spectroscopy
NOP	Nociceptin opioid receptor
NOXA B	Phorbol-12-myristate-13-acetate-induced protein 1
NR	Nuclear Receptor
NTD	N-Terminal Domain
p53	Tumor Protein 53
p300	E1A binding Protein 300
Pbf	2,2,4,6,7-pentamethyldihydrobenzofurane
PDB	Protein Data Bank
PEG	Polyethyleneglycol
PPI	Protein-protein interaction
PUMA	p53 Upregulated Modulator of Apoptosis
QSAR	Quantitative Structure Activity Relationship
RAS	Rat sarcoma
RCM	Ring-Closing Metathesis
RMSD	Root-Mean-Square Deviation
RXR	Receptor X Retinoide
SAHBs	Stabilized alpha-helices of Bcl-2 domains
SAR	Structure Activity Relationship
smMLCK	Smooth Muscle Myosin Light Chain Kinase
SOS	Son of Sevenless
SPPS	Solid Phase Peptide Synthesis
SRC	Steroid receptor coactivator
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
THF	Tetrahydrofuran
TIS	Triisopropylsilane
TLC	Thin Layer Chromatography
UV	Ultraviolet
WT	Wild Type

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
Naphtylalanine	Nal	-
Ornithine	Orn	O
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter 1. Inhibition of protein-protein interactions

1.1. Inhibition of protein-protein interactions

Protein-protein interactions (PPIs) are probably the most complex and diverse biological macromolecules, and have the highest regulatory impact among the class of macromolecular interactions.[2] They mediate a large number of important regulatory pathways and control essential cellular processes involved in signalling, growth and survival.[3, 4] Thus, there is much interest in targeting the interfaces between interacting proteins for therapeutic purposes.[2] However, they display greater structural and chemical diversity than the classical targets, such as protein kinases and proteases. Therefore, the use of small-molecules to modulate PPIs is important for the aforementioned key therapeutic benefits, as well as to gain further insight complex biological signalling pathways.[1]

1.2. Structural features of PPIs

Protein-protein interactions are generally not considered attractive targets for small molecule drug design. As a result of the unique characteristics of each of these interfaces, it is difficult to establish general guidelines for effective and selective inhibition of PPIs using small molecules.[5] The contact surfaces involved in PPIs are large (~1500-3000 $Å^2$) compared with those involved in protein-small-molecule interactions (~300-1000 $Å^2$).[2] In addition, PPIs present flat or moderately convex surfaces with fewer well-defined concave binding sites, such as grooves and pockets, than classical enzymes (Figure 1.1b a).[4] The interacting regions present both hydrophobic and charged recognition elements with a poorly defined spatial relationship. Furthermore, PPIs may contain discontinuous epitopes formed by peptidic strands from different regions of the protein sequence (Figure 1.1b).[3, 4]

However, most contact surfaces are dynamic. They display adaptability and flexibility, involving motions of side chains and small perturbations of loops. The reorganization of the surface residues occasionally creates available ligand-binding cavities that are not seen in structures of either the free-protein or the protein-protein complex.[2] These transient pockets and temporary structures represent promising targets for small molecule PPI inhibition.[6]

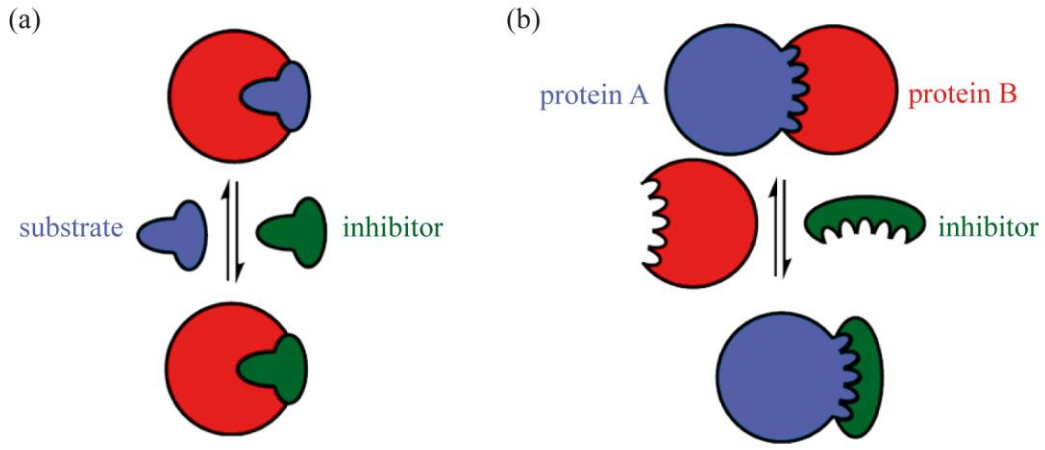


Figure 1.1 (a) Recognition and inhibition of enzymes; (b) Recognition and inhibition of PPIs.[2]

1.2.1. Binding features of PPIs

Numerous studies have focused on protein–protein complex formation. However, the principles governing PPIs are still not fully understood.[7] The most important factor that has obstructed their elucidation is the role of plasticity in protein–protein interfaces, including protein flexibility, presence of disordered regions, protein promiscuity and cooperativity in PPIs.[8,9]

Nevertheless, there are tools that allow the study of protein-protein association. In 1995 Clackson and Wells used the alanine scanning mutagenesis technique to explore the energetic contributions of individual side-chains in protein binding. This study showed that a PPI usually involves a few key residues that contribute the majority of the binding affinity to the interaction. These residues are known as "hot spots"[10] and are usually densely packed in clusters where they form a network of interactions (Figure 1.2).[11] Bogan and Thorn made a further contribution in 1998 and found that certain amino acid residues, particularly tryptophan (21%) and tyrosine (12%), appear more frequently in hot spots. These residues can perform aromatic-π interactions and hydrogen bonds through the indole nitrogen on the tryptophan and the phenolic hydroxyl on the tyrosine. Furthermore, their large hydrophobic surfaces presumably protect these hydrogen bonding interactions from water molecules. Arginine (13%) residues are also important, as they can form a similar range of favourable interactions in addition to ion pairs. Importantly, an energetically less important ring of residues, known as an O-ring, often surrounds the hot spots and seems to occlude bulk-solvent access.[4,11,12]

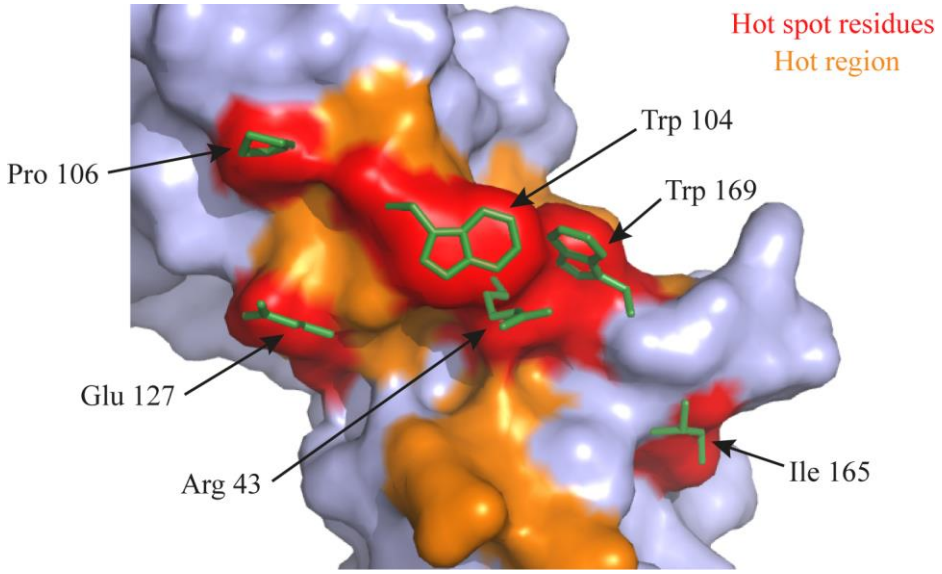


Figure 1.2 Binding site for human growth hormone (hGH) on the human growth hormone receptor (hGHbp). The "hot spot" residues identified by alanine scanning mutagenesis are shown in green (PDB ID: 3HHR).[10, 13]

1.3. Inhibition of α-helix mediated protein-protein interactions

The α-helix is the most common secondary structure in nature; over 30% of protein structure is helical. Therefore, α-helices represent a good generic template for inhibitor design given the high likelihood of PPIs involving this structural motif.[5, 14] Nevertheless, α-helix mediated PPIs still exhibit considerable diversity. They can vary in the number of proteins involved in the interaction, as well as in the number of helical faces found at the interface.

A typical α-helix has 3.4 amino acid residues per turn, is defined by backbone dihedral angles close to Φ= -60º and Ψ= -45º and has a rise of 1.5 Å/residue or 5.4 Å/turn (Figure 1.3).[14] The helix can be considered to have three distinct faces; side chains placed at a distance of 3-4 residues in the peptide sequence are located above one another and, therefore, are projected from the same face (Figure 1.3c). This structural characteristic plays a major role in its molecular function; the residues located on the central polypeptide backbone of an α-helix structure are projected along individual faces of the scaffold, which allows selective and specific molecular recognition.[1]

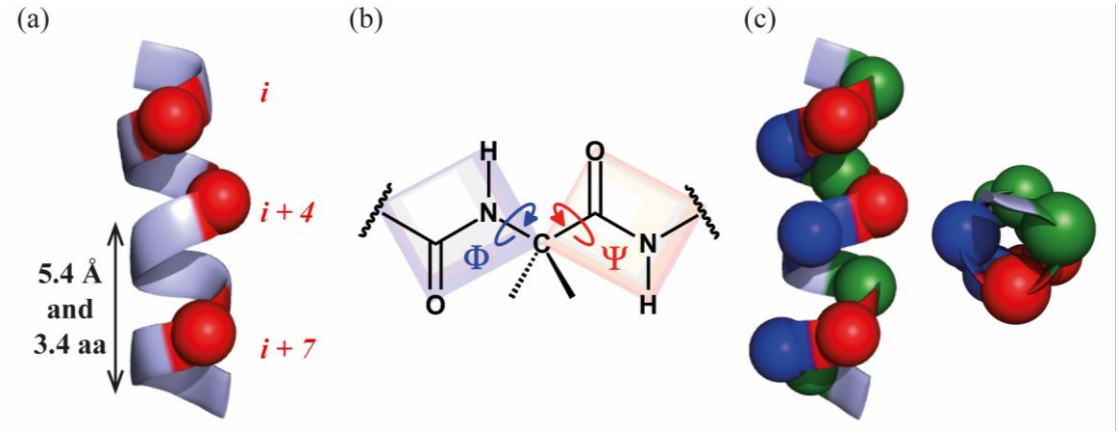


Figure 1.3 (a) α-Helix with the residues at *i*, *i+4* and *i+7* positions shown in red; (b) peptide bond showing the dihedral angles Φ and Ψ, and planar segments in purple and pink rectangles; (c) α-Helix with the three distinct faces highlighted in red, green and blue (side and top views are given).

1.3.1. Lead discovery in α-helix mediated PPIs

Different approaches have been established with the final aim of developing molecules that effectively and selectively inhibit α-helix mediated PPIs. The most important strategies can be classified according to the backbone that they utilise to connect the binding functionalities:[15]

- Type I mimetics: They mimic the topography of the original structural α-helix backbone at the atomic level.

- Type II mimetics: They mimic the function rather than the structure of the original α-helix; they are generally small non-peptidic molecules that bind to the corresponding target protein.

- Type III mimetics: They mimic the side chain projection of the key amino acid residues of the original α-helix; they use non-peptidic scaffolds.

1.3.1.1. Type I mimetics

Peptides are attractive candidates for stabilizing or disrupting PPIs. However, they present some severe drawbacks for therapeutic purposes, such as i) limited secondary structure as isolated sequences;[16] ii) poor cell permeability and transport properties;[17] iii) low stability due to proteolysis.[18] Type I mimetics consist of short peptidic oligomers that reproduce the local topography of an α-helical structural motif and focus on maximizing helicity whilst enhancing the proteolytic stability and the pharmacokinetic properties. The different strategies developed

within this approach can be classified in two main groups: helical foldamers and constrained peptides.[1]

Helical foldamers

Helical foldamers are structures that adopt well-defined conformations reminiscent of protein secondary structures.[19] Their synthesis is based on the oligomerization of building blocks and their structure is stabilized by intramolecular non-covalent interactions between non-consecutive residues along the peptide. The most extensively studied examples are β-peptides and α/β-peptides.[20]

β-peptides

The use of β-peptides in substitution of their natural α-counterparts has been extensively studied due to some of their attractive features for therapeutic use. The addition of an α-methylene group in the β-peptides provides them with an increased degree of freedom compared to the α-peptides. Consequently, β-peptides were expected to be entropically disfavoured from adopting well defined folded states in solution. However, this minor backbone modification resulted in higher propensity to adopt helical conformations, which permits the presence of defined structures within relatively short sequences. Moreover, β-peptides presented an increased number of accessible helical secondary structures (Figure 1.4b) when compared with the natural α-peptides (Figure 1.4a). In addition, this class of compounds is characterized by an enhanced resistance to proteolysis and thus a more favorable pharmacodynamic profile.[21] Different types of β-amino acids have been described: β^2- or β^3-amino acids bearing a single side chain either at C2 or C3, and $\beta^{2,3}$-building blocks with both carbon atoms substituted.

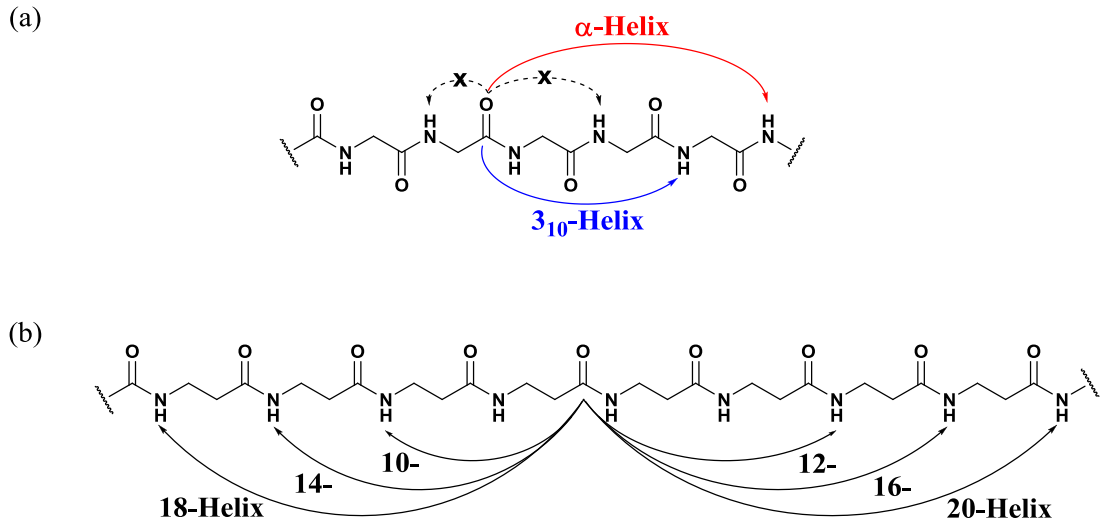


Figure 1.4 Intramolecular hydrogen bonding interactions occurring in different helical secondary structures of: (a) α-peptides; (b) β-peptides.

Several β-peptidic scaffolds have been used to inhibit PPIs due to the above described desirable therapeutic properties. Seebach and co-workers were the first to report the synthesis and complete characterisation of a β-peptide sequence.[22] They used circular dichroism (CD) and X-ray analysis to elucidate the intramolecular hydrogen bonding network that allows the β-peptides to adopt the desired helical conformation. Continuing with this research, they also reported short-chain amphipathic β-peptides that mimicked lipoproteins and inhibited intestinal cholesterol absorption.[23] Despite exhibiting subtle structural differences, the synthetic peptides exploited their increased resistance against pancreatic proteases compared to the corresponding proteins and natural α-peptide based inhibitors of lipid absorption.

The biological relevance of β-peptides was further explored by the Schepartz group, which reported a series of β^3-decapeptides that effectively inhibited the human double minute 2/tumour protein 53 (*h*DM2/p53) PPI.[20] Their 14-helical conformation was stabilized by intramolecular salt bridges and an electrostatic macrodipole interaction. Subsequent structure-activity relationship (SAR) studies exploited the introduction of non-natural side chains to increase the affinity of the β^3-decapeptides to the *h*DM2 protein. In particular, the substitution of the key tryptophan residue of p53 by a 6-chlorotryptophan analogue resulted in a 10-fold increase in activity. The β^3-peptides has also been used to successfully target other PPIs of interest, such as such as glycoprotein 41/human immunodeficiency virus type 1 (gp41/HIV)[24, 25] and glucagon-like peptide-1/glucagon-like peptide-1 receptor (GLP-1/GLP-1R).[26] The poor cell permeability of these types of molecules made deeper understanding of the correlation between affinity and cell uptake necessary, in order to obtain derivatives with improved pharmacokinetic properties.[27, 28]

α/β-peptides

The synthesis of foldamers containing α- and β-amino acids generated a wide range of accessible heterogeneous combinations. Importantly, conformational control and predictable folding patterns can be achieved by modifications of the peptide residue arrangement.[29-31] The main purposes of the α/β-peptides were to increase α-helix mimicry whilst retaining resistance to proteolysis.[32] Consequently, this type of foldamer contains an epitope formed by the α-amino acids responsible for surface recognition, whereas the β-amino acids increase the helical secondary structure through intramolecular salt bridges or by introducing structural constraints.

The Gellman group has extensively studied the use of α/β-peptidic scaffolds as ligands for the BH3-recognition cleft of the B-cell lymphoma 2 (Bcl-2) protein family. Early efforts focused on the structure-based design of pure α/β-foldamer backbones, which proved ineffective.[31, 33] These

studies led to the chimeric (α/β + α)-peptide family, which was formed by a 1:1 α/β-residue alternation in the *N*-terminal segment and exclusively α-residues at the *C*-termini.[34] This class of foldamers inhibited the Bcl extralarge/Bak (a homologous antagonist killer peptide) (Bcl-x_L/Bak) interaction with IC_{50} values in the low nM range and showed high proteolytic stability. Subsequent studies following the chimeric approach identified weak inhibitors of the Bcl-x_L/BIM (a Bcl-2 interacting mediator of cell death peptide) PPI.[35] The first crystal structure of the chimeric foldamers bound to Bcl-x_L provided fundamental insight into the peptide binding mode, highlighting the importance of the subtle changes on the α-helix for side chain matching, and the relevance of the solvent exposed β-amino acids for backbone helicity.[36]

Building on these results, the Gellman group adapted their chimeric approach to a novel sequence-based design approach.[37] This approach involved replacing subsets of regularly spaced α-residues with the corresponding β^3-residues, creating patterns such as "ααβαααβ" or "αααβ". This strategy was applied to the design of hybrid peptide mimetics of Puma (another pro-apoptotic member of the Bcl-2 family) (Figure 1.5), which led to the identification of potent inhibitors of Bcl-xL and induced myeloid leukemia cell differentiation protein (Mcl-1) and demonstrated that the affinity and selectivity of the α/β-peptides were dependent on the position of the β-residues along the sequence.[38]

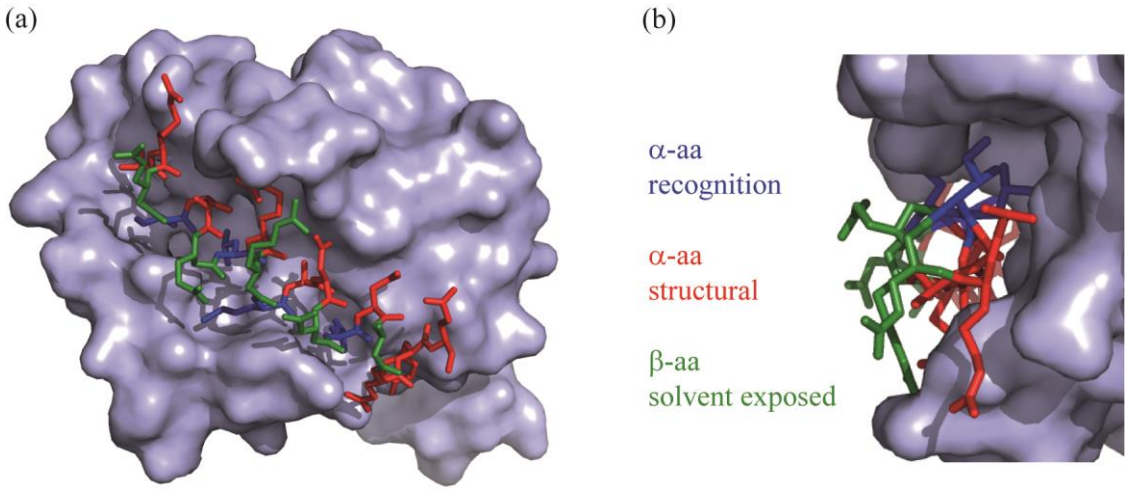


Figure 1.5 (a) X-Ray structure of the complex between the Bcl-x_L protein and the ααβαααβ-peptide derivative of the native PUMA-BH3 domain (PDB ID: 2YJ1); (b) top view illustrating the alignment of the α- and β-amino acids.

The scope of the sequence-based strategy has expanded in recent years with introduction of new β-residues, achievement of selectivity between target molecules, and enhancement of potency and proteolytic stability.[39, 40] Moreover, this approach has been used to successfully inhibit other PPIs of therapeutic interest, such as gp41/HIV[41] and GLP-1/GLP-1R.[42] However, the use of α/β-

peptides as drug candidates is still limited by their poor cell membrane permeability properties.[43]

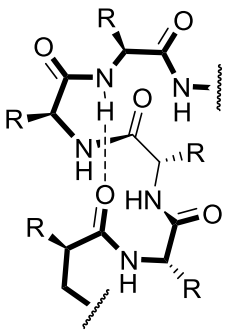
Constrained peptides

In principle, peptides retain excellent surface recognition properties whilst presenting reduced toxicity. However, in most cases they suffer from proteolytic instability and low cell permeability. These limitations are related to the unstructured conformation that short peptides adopt in solution. This fact causes entropic penalties when the peptides transition to more restricted conformational states upon binding, which ultimately has an effect on the target affinity.[44]

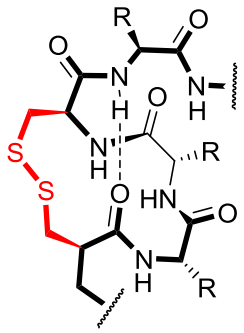
Therefore, major efforts have focused on the introduction of conformational constraints into peptides in order to stabilize bioactive conformations. This would presumably reduce the entropic penalty upon binding and thus engender more drug-like properties whilst increasing target affinity.[45] The structural stabilization of helical peptides by covalent linkages between residues suitably positioned in space is one of the most important approaches in this area.[46] Some of the most relevant methodologies for these purposes are described below.

Thiol-based crosslinks

One of the first methods used to stabilise the helical conformation of peptides was the introduction of thiol-based crosslinks (Figure 1.6). Spatola and co-workers used a simple disulfide bond between cysteine residues to constrain a nonapeptide inhibitor of the estrogen receptor (ER)/co-activator PPI. Interestingly, the X-ray structure of the complex revealed the helical conformation adopted by the constrained peptide when bound to the protein surface, which contrasted with the minimal helicity shown in solution.[47]



α-Helix



Cys-Cys disulfide bridge

Figure 1.6 Generic structure of a peptide constrained with a Cys-Cys disulfide bridge.

One of the main disadvantages of the disulfide cross-links is their lability under the reductive conditions found in most eukaryotic cells. For this reason, chemically more stable thioether moieties have been designed; most of them including the use of electrophiles that selectively react with the thiol side chain. A variety of biselectrophilic molecules have been used to cross-link two properly aligned cysteine residues of peptides thus stabilizing their helical conformation. In particular, aryl and bis-aryl methylene bromides were introduced by Lin and co-workers (Figure 1.7a).[48, 49] They were successfully utilised to crosslink cysteine containing peptides providing inhibitors for the p53/murine double minute 2 (*m*DM2) and Mcl-1/NOXA B (another pro-apoptotic member of the Bcl-2 family) (Figure 1.7) PPIs. Likewise, the Pentelute group has reported a new class of α-helix-induced peptides which utilised perfluorinated aryl linkers to mildly functionalize cysteine containing peptides (Figure 1.7a).[50] Recently, Chou and co-workers expanded the variety of thiol-based crosslinks with the introduction of a robust and versatile thiol-ene coupling approach (Figure 1.7a), which has provided p53 constrained derivatives that bind to the *h*DM2 protein partner and retain activity in cells.[51]

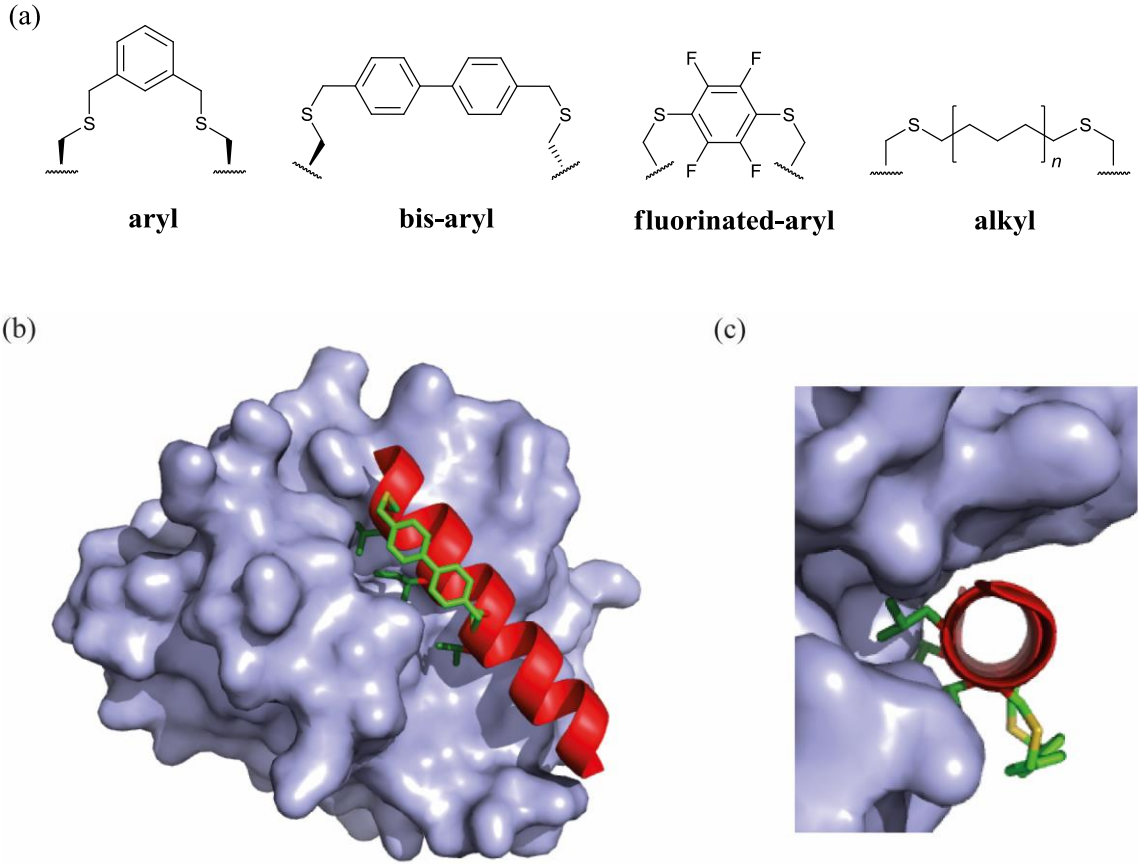


Figure 1.7 (a) Different cross-linking systems between cysteine residues resulting from the reaction with: an aryl and a bis-aryl methylene bromides, a fluorinated aryl and an alkyl thiol-ene linkers, respectively from left to right. (b) X-Ray structure of the complex between the Mcl-1 protein and a cysteine bisaryl methylene bridged peptide derivative of the native NOXA-BH3

domain (PDB ID: 4G35); (c) top view illustrating the alignment of the binding α-residues in green and the solvent exposed bisaryl methylene linkage between cysteine residues.

Photo-controlled helices or azobenzene photo-switches

A slightly different approach introduced by Woolley and co-workers was based on an azobenzene molecule that crosslinked suitably positioned cysteine residues of a peptide.[52] Photo-isomerization of the crosslinker can be used to switch peptides between the α-helical and random coil-like conformations. Building on that work, the Allemann group developed a family of photo-controllable peptide-based switches based on the BH3 region of Bak and BIM proteins (Figure 1.8).[53] Interestingly, the resulting helix-stabilized peptides bound to the Bcl-x_L target protein with greater affinities than the helix-de-stabilized forms. Recently, this group have fused a BID modified peptide to the LOV2 domain of Avena Sativa phototropin 1(AsLOV2) protein to create optically controlled intracellular modulators of the Bcl-x_L protein.[54]

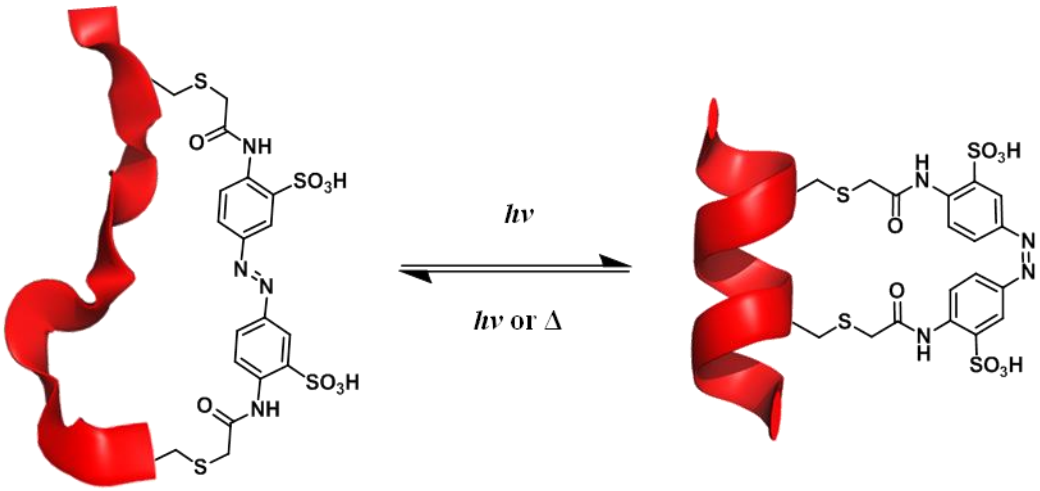
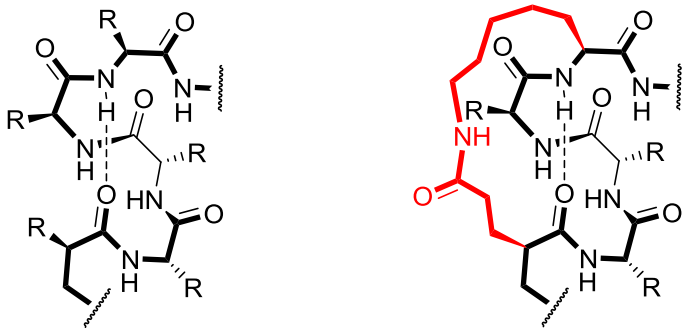


Figure 1.8 Peptide conformational change controlled by photo-isomerisation of the azobenzene cross-linker. The example illustrates helical conformation stabilised in the *cis* configuration with *i*, *i*+4 and *i*, *i*+7 Cys linkages.

Lactam bridge

The lactam linkage was another of the earliest approaches for constraining peptides (Figure 1.9). This strategy was introduced by the Rosenblatt group, who for the first time stabilised the helical secondary structure of a peptide by forming a lactam bridge between a lysine and an aspartic acid residue located at *i* and *i*+4 positions of the peptide sequence.[55]



α-Helix **Lactam bridge**

Figure 1.9 Generic structure of a peptide constrained with a Lys-Glu lactame bridge.

This work was followed up by Kim and co-workers, who stabilised a 14-residue *C*-terminal peptide of gp41 by crosslinking two glutamic acid residues at *i* and *i*+7 positions with an α,ω diaminoalkane group (Figure 1.10). The introduction of this constraint resulted in a potent inhibitor of the HIV-1/gp41 PPI.[56] Additionally, the extensive work of the Fairlie group in this area focused on downsizing protein helical epitopes by strategically locking them in highly α-helical structures through the introduction of two adjacent lactam cross-links.[57] Following this approach, they managed to successfully constrain a wide variety of peptides from viral, bacterial, or human proteins. One of the most remarkable examples is the nociceptin peptide, which was constrained to produce a pM agonist of the nociceptin receptor (NOP), the most potent known to date.[58] Likewise, the constrained 11-residue analogue of the *N*-terminus GLP-1 was recently identified as a potent agonist of GLP-1R.[59]

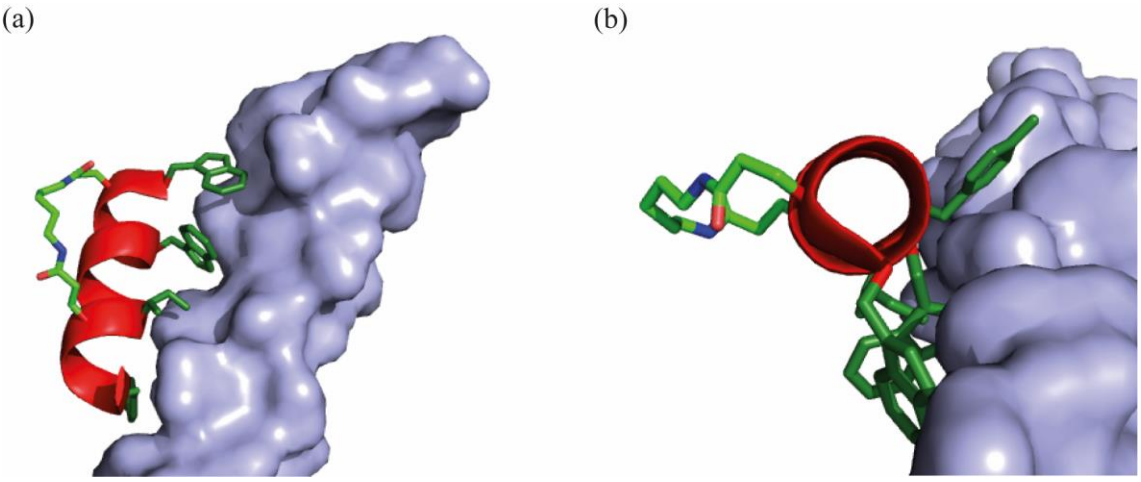


Figure 1.10 (a) X-Ray structure of one of the domains from the trimeric coil-coil complex between the HIV-1 protein and a glutamic acid α,ω diaminoalkane bridged peptide derivative of the native gp41 (PDB ID: 1GZL); (b) top view illustrating the alignment of the binding α-residues in green and the solvent exposed α,ω diaminoalkane linkage between glutamic acid residues.

Hydrocarbon staple

Inspired by the initial work of Grubbs with cross-linked *O*-allyl serine residues,[60] Verdine and co-workers established the hydrocarbon stapling methodology by introducing unnatural α,α'-disubstituted amino acids with olefin tethers into the peptide sequence and cross-linking them *via* ring-closing metathesis (RCM) (Figure 1.11).[61] Initial studies focused on the role of residue positioning, stereochemistry and linker length, in order to obtain higher binding affinities and enhanced proteolytic stability. The use of this strategy by Korsmeyer and co-workers generated ligands based on the BID BH3 sequence, which represented one of the main breakthroughs in the field.[46] It generated peptides with a significant enhancement in peptide α-helicity, protease resistance and *in vitro* and *in vivo* biological activity, showing tumour suppression and regression in leukaemia xenografts.[62]

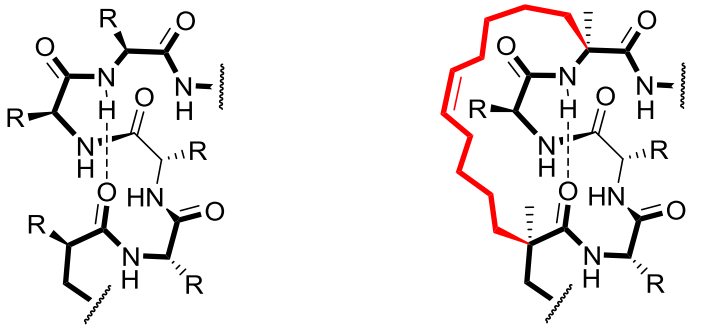


Figure 1.11 Generic structure of a hydrocarbon stapled peptide.

The Walensky group has extensively expanded the use of hydrocarbon stapled peptides to study in depth the interactions between members of the Bcl-2 family and identify potent inhibitors for those PPIs (Figure 1.12).[63-65] In addition, they also reported a hydrocarbon double-stapled peptide, which achieved helical conformations in long peptides whilst maintaining resistance to proteases both *in vitro* and *in vivo* and enhanced inhibitory activity against the HIV-1/gp41 interaction.[66]

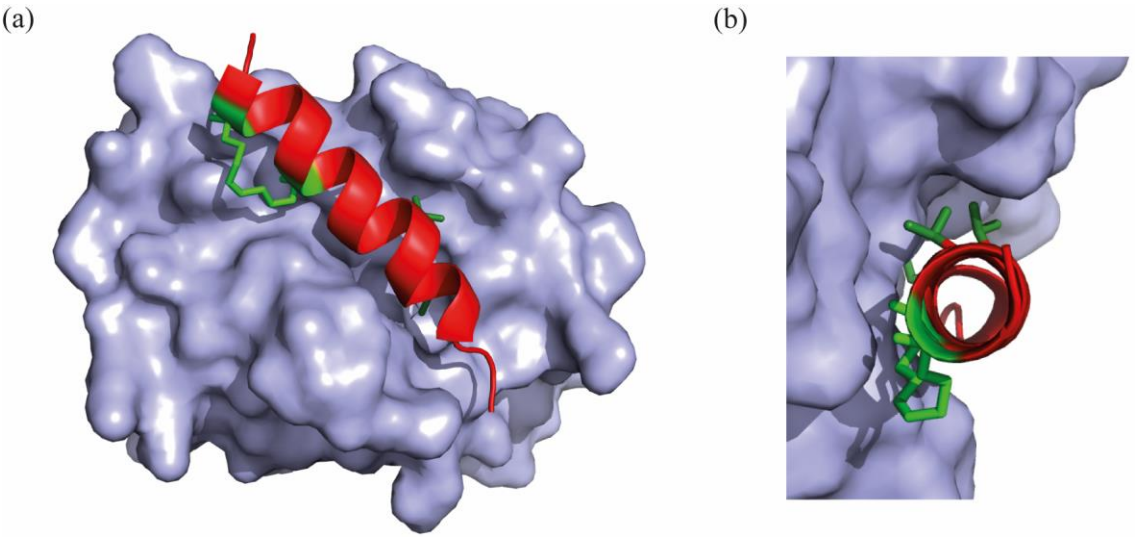


Figure 1.12 (a) X-Ray structure of the complex between the Mcl-1 protein and a stapled peptide derivative of a stabilized α-helix of a Bcl-2 domain (SAHBs) (PDB ID: 3MK8); (b) side view illustrating the alignment of the binding α-residues in green and the solvent exposed hydrocarbon linkage.

Furthermore, it is important to highlight the crystal structures of the stapled peptides targeting the p53/*h*DM2[67] (Figure 1.13a) and ERα/co-activator (Figure 1.13b).[68] Both structures showed the hydrocarbon staple chain actively participating in the interaction with the binding cleft from the corresponding protein partners. These discoveries highlighted the need for careful case by case analysis of these kinds of molecules, in order to avoid miss-interpretation of the structure-potency relationships.

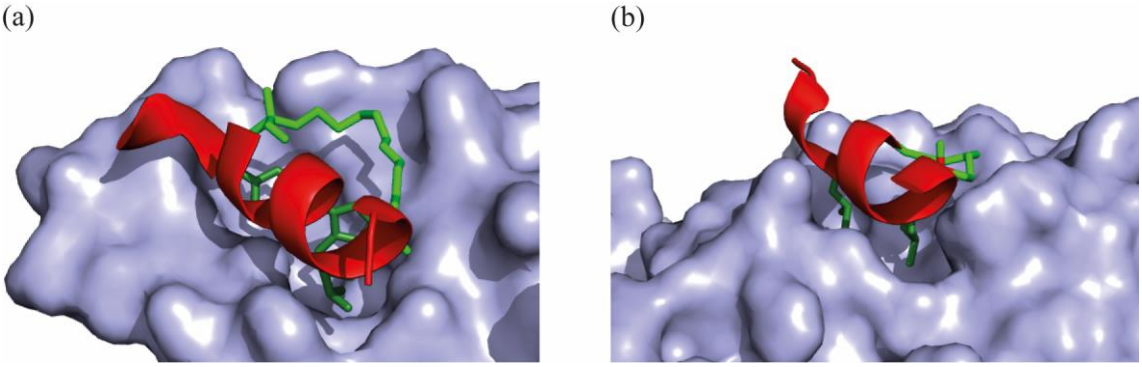


Figure 1.13 (a) X-Ray structure of the complex between the *h*DM2 protein and a stapled peptide derivative of p53 (PDB ID: 3V3B); (b) X-Ray structure of the complex between the ERα protein and a co-activator stapled peptide derivative (PDB ID: 2YJD). Both illustrating the participation of the hydrocarbon staple in binding to the protein surface.

In that context, the Wilson group has recently reported an α-alkenyl mono-substituted amino acid which has been incorporated in the BID peptide sequence.[69] Significantly, it was demonstrated to increase peptide helicity, enhance proteolytic stability and provide similar inhibitory activity towards Bcl-x_L/Bak whilst involving a simpler synthetic route than the α,α'-disubstituted amino acid. Furthermore, an extensive investigation on the mode of action of these mono-substituted stapled peptides, in particular BID and BIM derivatives, has recently shown evidence for induced-fit binding and enthalpy-entropy compensation.[70] Those findings join the concerns raised by Czabotar and co-workers on the need for more in-depth understanding of the effects of pre-organisation in protein-ligand binding processes and its ultimate role in binding potency.[71, 72]

Nevertheless, hydrocarbon stapled peptides might represent the most successful example of therapeutic PPI inhibitors, resulting in the creation of AILERON Therapeutics who have brought this class of compounds into Phase I clinical trials.[73]

Hydrogen bond surrogate

Inspired by the original hydrazone linker from Cabezas and Satterthwait,[74, 75] the hydrogen bond surrogate (HBS) methodology was established by Arora and co-workers (Figure 1.14).[76] It substitutes the natural intramolecular hydrogen bond between residues at *i* and *i*+4 positions of a peptide employing a non-natural carbon-carbon bond formed by ring-closing metathesis (RCM). This approach has the advantage of not adding steric constraints to the natural helix; however, its use is limited to the *N*-terminal position of a peptide. It has been extensively exploited to inhibit PPIs of interest,[77] such as Bcl-x_L/Bak,[78] p53/*m*DM2,[79] hypoxia-inducible factor 1 alpha/E1A binding protein 300 (HIF-1α/p300)[80] and more recently *in cellulo* inhibition has been reported for the RAS/son of sevenless (SOS) interaction.[81] The Arora group also developed a reversible disulfide and thioether-linked versions of the HBS peptide, which disrupted the p53/*m*DM2 PPI.[82, 83] Another remarkable HBS is the covalent ethylene bridge from Alewood, which can be positioned anywhere on the peptide sequence.[84]

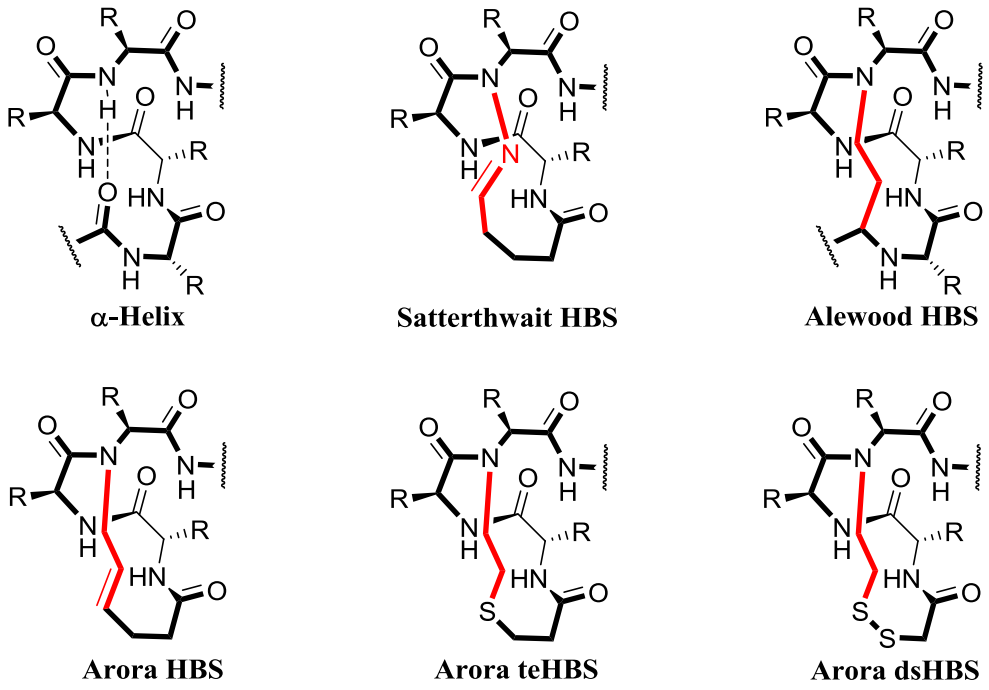


Figure 1.14 Generic structures of the most relevant HBS covalent constraints developed to date.

1.3.1.2. Type II mimetics

Type II mimetics are based on small non-peptidic molecules that inhibit α-helix mediated PPIs by binding to the target receptor.[85] Distinctively, they inhibit PPIs without necessarily mimicking the original helix.[1] Several examples of potent inhibitors identified using this strategy have been reported, such as the tetra-substituted imidazoles (Nutlins)[86] and the ABT-737-based compounds,[87] which represent some of the more relevant and ground breaking examples in the utilisation of small molecules to effectively inhibit PPIs.

Nutlin-3

F. Hoffmann-La Roche identified by a high throughput screen (HTS) a series of tetra-substituted cis-imidazoline analogues known as Nutlins.[86] These compounds were used as lead structures for the development of p53/*h*DM2 inhibitors. Several rounds of chemical optimization finally yielded Nutlin-3, a potent and selective inhibitor of the p53/*h*DM2 interaction with an IC_{50} of 90 nM. An X-ray crystal structure verified the mode of binding of this family of compounds and provided the first structural information of a non-peptidic small-molecule inhibitor bound to the *m*DM2 protein. The Nutlins display analogous interactions to the natural p53 peptide, with the chlorophenyl moieties and the isopropyl substituent occupying the Trp, Leu and Phe pockets from the *m*DM2 binding site respectively.

Nutlin-3 was the first *m*DM2 inhibitor to enter Phase I clinical trials and proved its activity *in vitro* and in tumour xenografts *in vivo*, providing the first *in vivo* proof-of-concept of this approach to cancer therapy (Figure 1.15a). Additional efforts from F. Hoffmann-La Roche around this compound series has led to candidates such as RG7112[88, 89] and RG7388[90] with increased binding affinity (Figure 1.15b), cellular potency, pharmacokinetic properties and chemical stability, and which are currently in Phase I/II clinical trials.

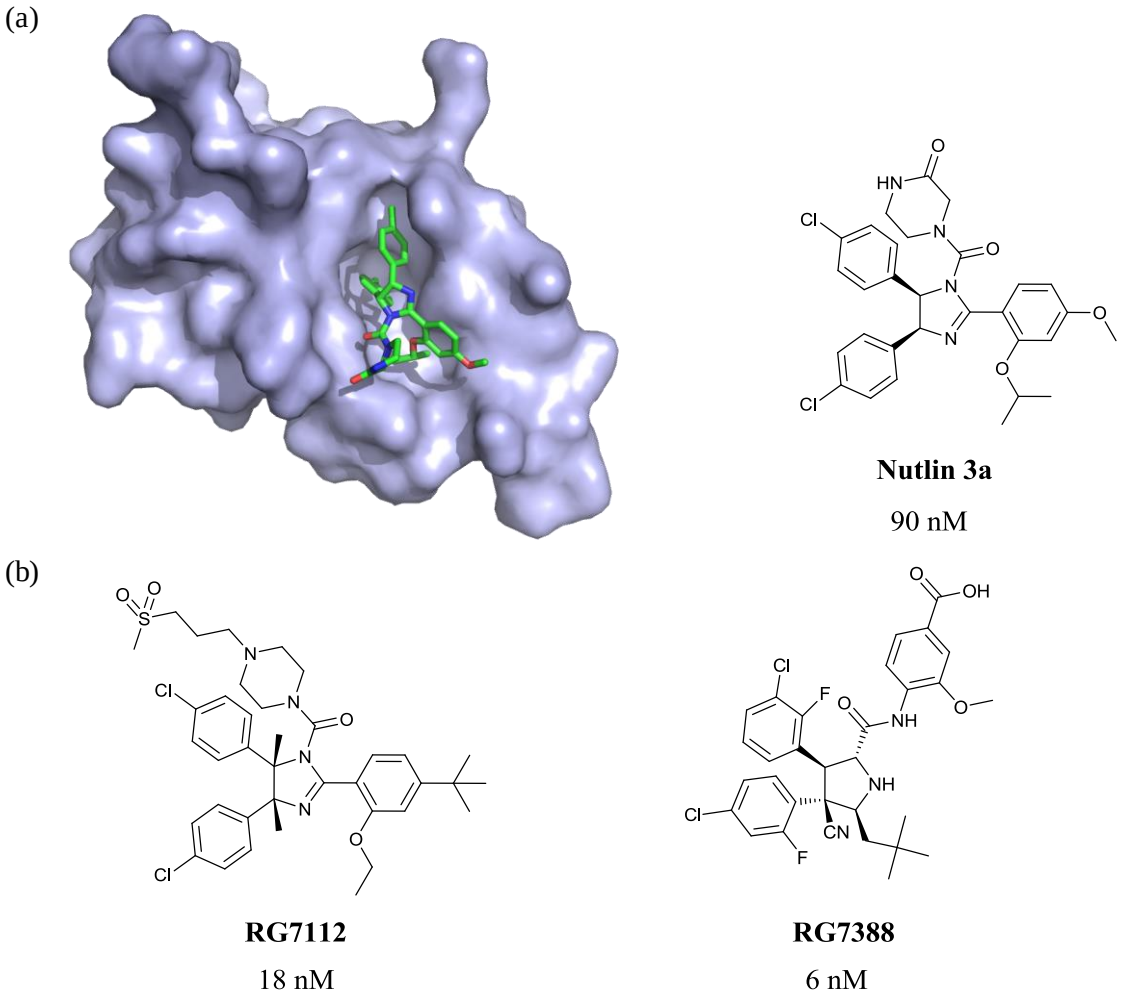


Figure 1.15 Small molecule inhibitors of PPIs identified via HTP screening (a) X-Ray structure of the Nutlin 3a/*m*DM2 complex (PDB ID: 4J3E), Nutlin 3a chemical structure and corresponding *m*DM2 binding affinity; (b) Chemical structures of RG7112 and RG7388 and corresponding *m*DM2 binding affinities.

ABT-737

Abbott Laboratories discovered, by an NMR-based high throughput fragment screen, a group of high-affinity small molecules that bind to some members of the Bcl-2 family, such as Bcl-x_L

and Bcl-2, through their helical binding site.[91] The most potent inhibitor was ABT-737,[62] which binds to the BH3 α-helix binding grove of Bcl-x_L with a K_i of 0.6 nM (Figure 1.16a). However, ABT-737 ultimately failed in clinical trials due to its poor bioavailability. Further structural optimization resulted in derivative ABT-263 (Navitoclax) (Figure 1.16b),[87] which displayed improved oral bioavailability and similar affinity (*Ki* < 1 nM) for the Bcl-2 family proteins. ABT-263 failed phase II clinical trials for small-cell lung carcinoma treatment due to its hematologic toxicity; however, it led to the selective Bcl-2 candidate ABT-199 (Venetoclax) (*Ki* < 0.01 nM), which has been approved for some forms of chronic lymphocytic leukaemia (Figure 1.16b),[92] being one of the first success stories of PPI inhibition originate from fragment based screening.[93]

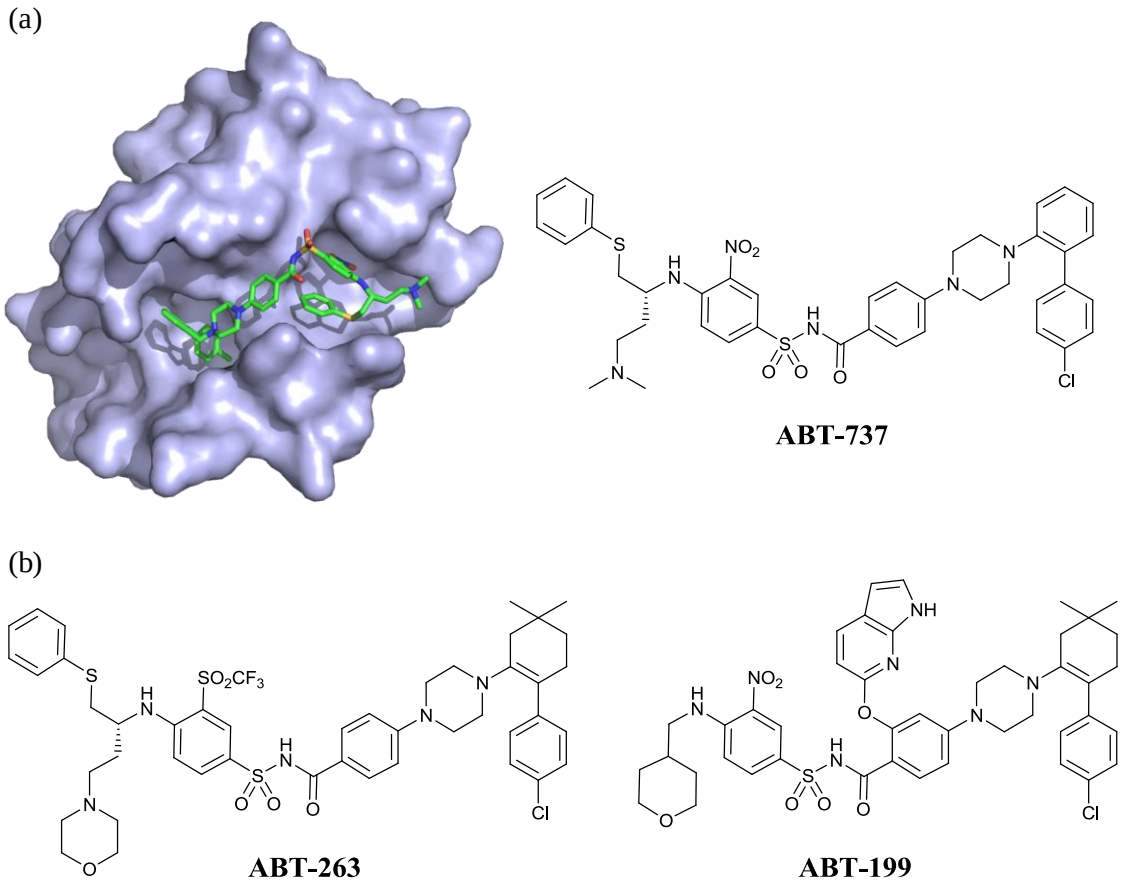


Figure 1.16 Small molecule inhibitors of PPIs identified via structure-based screening (a) X-Ray structure of the ABT-737/Bcl-x_L complex (PDB ID: 2YXJ) and ABT-737 chemical structure;[94] (b) Chemical structures of ABT-263 and ABT-199.

Despite the increasing number of reported small molecules as potent inhibitors of PPI, the significant developments in computational design and docking algorithms and the assembly of more focused screening libraries, the use of traditional drug discovery approaches to target PPI is still challenging. Furthermore, these approaches usually focus on the development of

inhibitors that specifically target a particular PPI, which may not directly contribute to the elucidation of the general guidelines for PPI inhibition. Therefore, this type of inhibitors will not further be discussed.

1.3.1.3. Type III mimetics (or proteomimetics)

The proteomimetic approach is based on non-peptidic scaffolds that mimic the spatial orientation of the key recognition residues on the native α-helix surface (Figure 1.17). The method focuses on simplifying the pharmacophore to a rod-shaped object which projects the side chains in a similar orientation to the original α-helix. The ligands potentially present more accessible syntheses and improved drug-like properties than the original peptides. Furthermore, due to its modular synthesis, this approach could be easily adapted to different PPIs by changing the side chains according to the target of interest.[1]

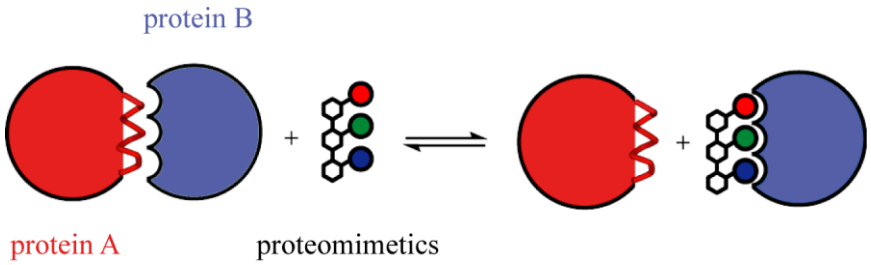


Figure 1.17 Schematic illustrating the proteomimetic approach.

Early scaffolds

The first small molecules designed to inhibit PPIs by mimicking the side chain residues of an α-helix were the trisubstituted indanes reported by Horwell *et al.* which mimic the residues at *i*-1, *i* and *i*+1 positions (Figure 1.18a). However, due to its small size and limited mimicry potential of just two consecutive residues, they cannot be considered effective inhibitors for α-helix mediated PPIs.[95, 96] Inspired by that work, the Hamilton group established the field of proteomimetics by reporting the terphenyl as the first real α-helix mimetic scaffold (Figure 1.18b).[97] In the tris-functionalised 3,2',2''-terphenyl derivatives the aryl core adopts a staggered conformation which projects the *ortho*-substituents in a spatial orientation that mimic the *i*, *i*+3, and *i*+7 residues of an α-helix.

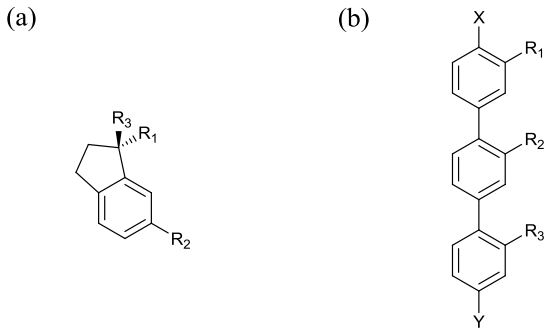


Figure 1.18 (a) Indane scaffold: helix mimetics proof-of-concept; (b) Terphenyl scaffold: first true proteomimetics.

Since its development, the terphenyl scaffold was an attractive template due to the simplicity of the structure and the synthetic potential. Thus, terphenyl derivatives incorporating a wide variety of functionalized side chains were developed and used to successfully target many PPIs of interest, such as calmodulin/smooth muscle myosin light chain kinase (CaM/smMLCK),[97] gp41/HIV-1,[98] Bcl-x_L/Bak[98] and p53/*h*DM2.[99] This family of mimetics has achieved inhibition of PPIs with high potency, good selectivity and proven activity in cells.[100, 101]

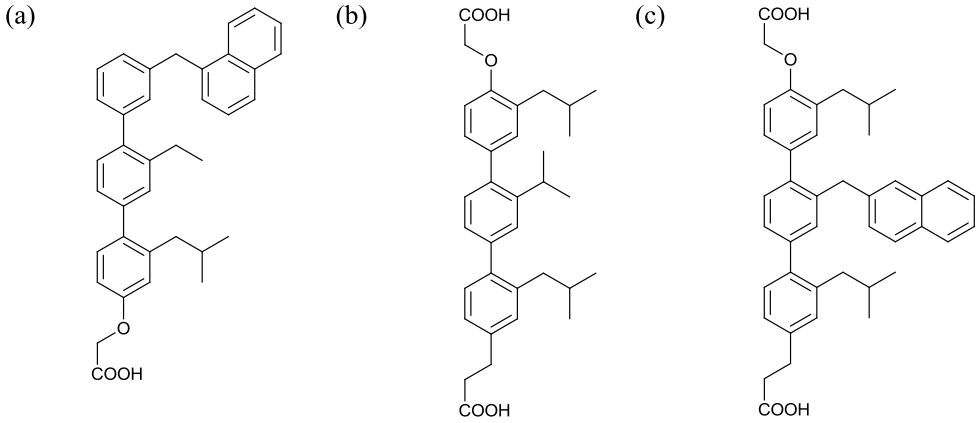


Figure 1.19 Terphenyl derivatives inhibitors of (a) CaM/smMLCK; (b) Bcl-x_L/Bak; (c) p53/*h*DM2.

Next generation of scaffolds

Despite the success of the terphenyl scaffold, some of its disadvantages include difficulties in the synthesis and most importantly its highly hydrophobic character.[102] Considerable effort has therefore been focused on the development of scaffolds with more versatile synthetic strategies leading to molecules with enhanced drug-like properties. Some of the most remarkable scaffolds developed by Hamilton and co-workers include the terephtalamide[103] (Figure 1.20a) and 4,4'-

dicarboxamide[104] (Figure 1.20b) templates, which benefit from accessible syntheses, hold rigidity due to intramolecular hydrogen bonding and present good water solubility due to the their higher polarity. These scaffolds inhibited Bcl-x_L/Bak; however, with lower affinity than the original terphenyl derivatives. Importantly, the mimicry of an additional fourth hot spot in the 4,4'-dicarboxamide template compared with the terephtalamide did not result in an increase of the binding affinity. Thus, highlighting the complexity of drug design for disruption of protein surface recognition processes.

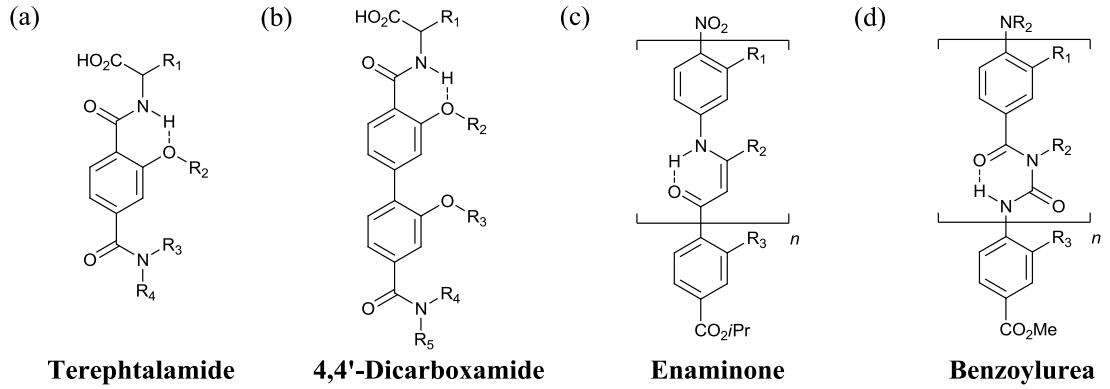


Figure 1.20 Second generation of Hamilton's scaffolds: (a) Terephtalamide; (b) 4,4'-dicarboxamide; (c) Enaminone; (d) Benzoylurea.

Further investigation on extended α-helix mimetics was accomplished with the introduction of the enaminone scaffold (Figure 1.20c), a development of the previously reported terphenyl.[105] In this template the central aromatic ring is substituted for a six-membered isostere formed *via* an intramolecular hydrogen bond, which locks the molecule in the desired conformation. Similarly, the benzoylurea template (Figure 1.20d) contains a central core formed by six membered hydrogen-bonded acylurea structures.[106] These scaffolds gave access to longer oligomers with improved water solubility properties.[107] In addition, the benzoylurea inhibited the Bcl-x_L/Bak interaction with lower binding affinity than the terphenyl and the 4,4'-dicarboxamide scaffolds,[108] further stressing the importance of the balance between the number of hot spots mimicked and the molecular size of the inhibitor.

Continuing the search towards more soluble scaffolds, Rebek and co-workers developed new amphiphilic α-helix mimetics with a series of scaffold based on oxazole-pyridazine-piperazine rings as backbones (Figure 1.21a).[109-111] The design included a hydrophobic face for protein surface recognition and another face rich in hydrogen bonding groups, also known as the "wet edge", which was anticipated to be directed towards the solvent thus increasing water solubility. However, the compounds obtained from this scaffold presented low inhibition towards the Bcl-x_L/Bak interaction. Similarly, the Hamilton group introduced the 5-6-5-imidazole–phenyl–

thiazole scaffold (Figure 1.21b),[112, 113] in which the terminal monomers of the terphenyl backbone are replaced with more hydrophilic five-membered heterocycles. The resulting compound also had limited success, inhibiting the Dbl's big sister/cell division control protein 42 (Dbs/Cdc42) interaction with an IC_{50} value of 67.0 μM. In addition, Lim and co-workers reported a pyrrolopyrimidine-based scaffold[114] (Figure 1.21c) containing a "wet edge", which showed potent activity against p53/*h*DM2 and improved solubility and cell permeability properties. Importantly, this scaffold is accessed by a facile solid-phase synthetic route, amenable to large library generation. Recently, the Fletcher group has reported a similar scaffold based on a purine derivative.[115]

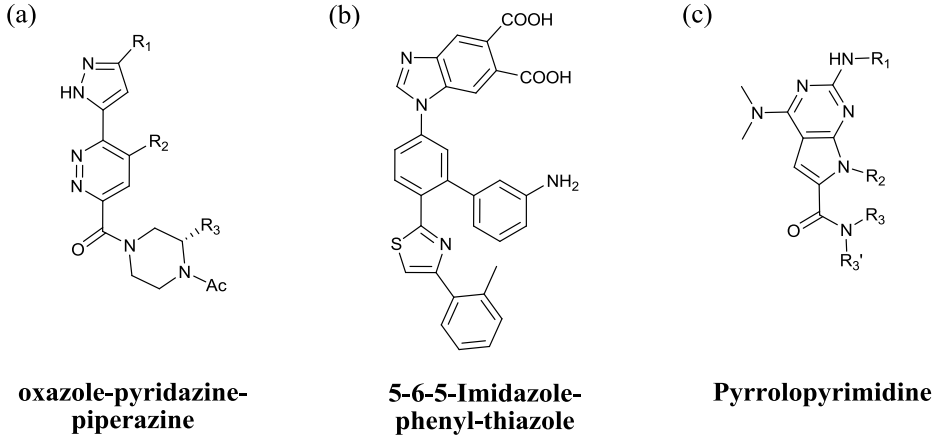
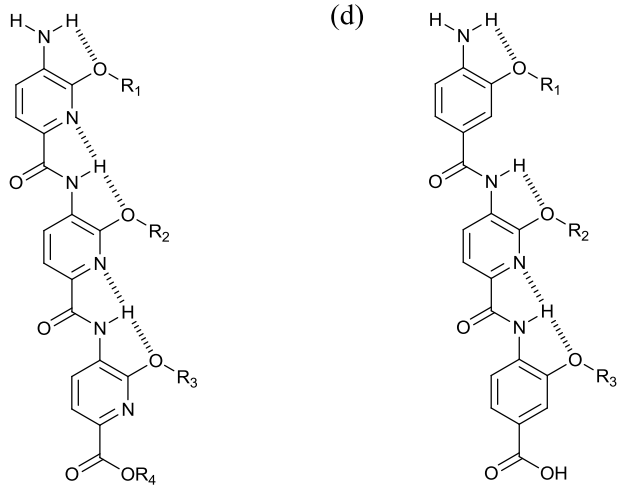


Figure 1.21 Amphipilic α-helix mimetics scaffolds: (a) Oxazole-pyridazine-piperazine; (b) 5-6-5-Imidazole-phenyl-thiazole; (c) Pyrrolopyrimidine.

The continuous efforts of Hamilton and co-workers in the field yielded the trispyridylamide scaffold (Figure 1.22a).[116, 117] Importantly, the preferred conformation adopted by this template is controlled by intramolecular H-bonds between the amide NH proton, the *ortho* alkoxy functionalities and the pyridyl nitrogen. The resulting geometry is almost planar and projects the three side chains on the same face of the backbone and in a similar orientation as the *i*, *i*+3/4 and *i*+7 residues of an α-helix. Moreover, the modular synthetic route to the trispyridylamide scaffold permitted the production of small libraries of compounds as inhibitors of the Bcl-x_L/Bak PPI and the islet amyloid polypeptide (IAPP) aggregation.[118, 119] Since then, other groups have designed derived scaffolds where some of the pyridyl monomers were substituted by benzene rings, thus increasing the flexibility of the backbone and favouring an "induce-fit" interaction. The Fletcher group introduced a series of mixed scaffolds mimicking the BH3 sequence (Figure 1.22b),[120] which resulted in nM inhibitors of the Bcl-x_L/Bak interaction and induced apoptosis on multiple cell lines.[121]

(c) (d)



Trispyridylamide **Trispyridylamide mixed**

Figure 1.22 (a) Trispyridylamide scaffold; (b) Trispyridylamide mixed scaffold.

Oligobenzamides

The Wilson group introduced the 3-*O*-alkylated oligobenzamide scaffold (Figure 1.23a),[122] which presented increased flexibility in the backbone and thus was anticipated to maximise interactions with the target protein surfaces. This template has proven effective to generate low μM inhibitors against the p53/*h*DM2, Bcl-x_L/Bak, Mcl-1/NOXA B and HIF-1α/p300 interactions.[123, 124] The extensive studies performed in this class of analogues resulted in the development of a microwave-assisted solid phase synthetic route, which permitted the assembly of larger libraries of compounds.[125] In order to improve the drug-like properties of these derivatives, a "wet edge" approach was designed with the introduction of a poly ethylene glycol (PEG) chain in one side of the helix mimetic (Figure 1.23b).[126] An analogous 2-*O*-alkylated scaffold was also developed (Figure 1.23c), which provided significant structural and conformational insight into this class of mimetics.[127, 128]

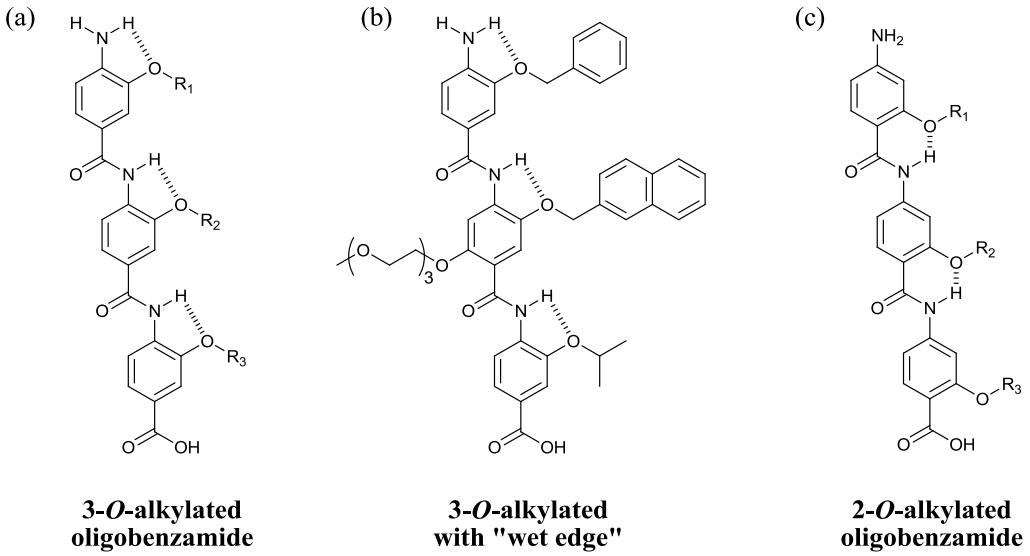


Figure 1.23 *O*-alkylated oligobenzamide scaffolds: (a) 3-*O*-alkylated; (b) 3-*O*-alkylated with "wet edge"; (c) 2-*O*-alkylated.

In parallel, nitro-acid derivatives of the 3-*O*-alkylated scaffold were reported by the Ahn[129] and Boger[130] groups. In particular, Boger and co-workers assembled a large library (>8000 members), which resulted in low affinity inhibitors of the p53/*h*DM2 interaction and inhibitors with low μM activity against gp41/HIV-1.[131]

Additionally, the *N*-alkylated oligobenzamide scaffold (Figure 1.24a) was reported in 2010 by Wilson and co-workers as the first helix mimetic family which could be accessed by solid phase methodology. It generated low μM inhibitors of the p53/*h*DM2 and Mcl-1/NOXA B interactions, which have been studied in depth in a biophysical and cellular context.[132-134] Recently, the Wilson group reported a hybrid oligobenzamide scaffold (Figure 1.24b), where the central building block is replaced by an amino acid residue. This structural change provided a significant increase in backbone flexibility, which is envisioned to aid induce-fit recognition. This template also benefits from a highly accessible modular synthesis and has provided low μM inhibitors of the p53/*h*DM2 and Mcl-1/NOXA B interactions, which have also proven to be active in cells. Importantly, selective molecular recognition determined by the stereochemistry of the helix mimetic has been achieved for first time using this scaffold. In addition, quantitative structure-activity relationship (QSAR) analyses have also been reported to aid the elucidation of the non-covalent contributions in molecular recognition.[135, 136]

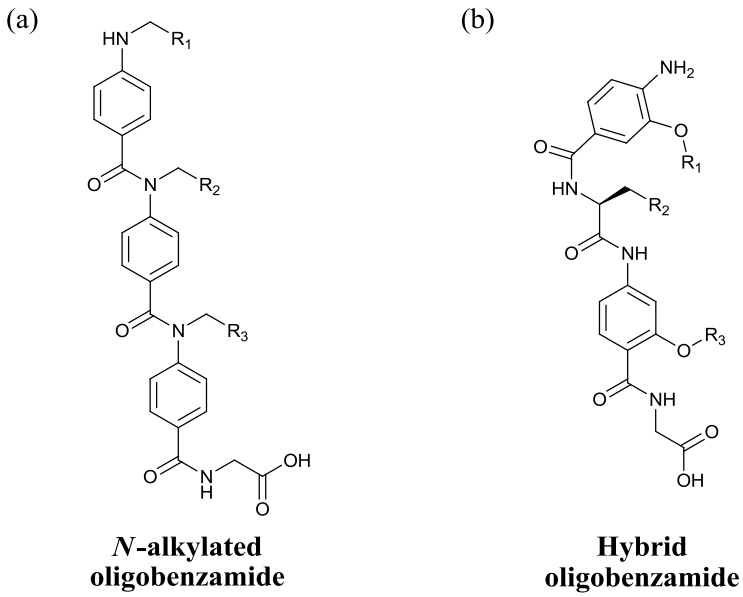


Figure 1.24 Other oligobenzamide scaffolds: (a) *N*-alkylated; (b) Hybrid.

Bifacial scaffolds

The development of scaffolds that can mimic residues located on more than one face of an α-helix is an important step towards successfully controlling the modulation of biologically relevant PPIs. In that context, the Hamilton group designed the pyridylpyridone scaffold (Figure 1.25a) to mimic key side chain residues of an α-helical LXXLL (where L is leucine and X any amino acid) motif from the co-activator peptide of the ER, which resulted in inhibitors with affinities within the low μM range.[137] Later work focused on modifying the benzoylurea scaffold (Figure 1.25b) to achieve bifacial mimicry of the residues at *i*, *i*+4, *i*+8 and *i*+1, *i*+6 positions of an α-helical strand.[138] Ahn and co-workers reported a bis-benzamide scaffold to create amphiphilic α-helix mimetics with four side chains at *i*, *i*+2, *i*+5, and *i*+7 positions of a helix (Figure 1.25c).[139]

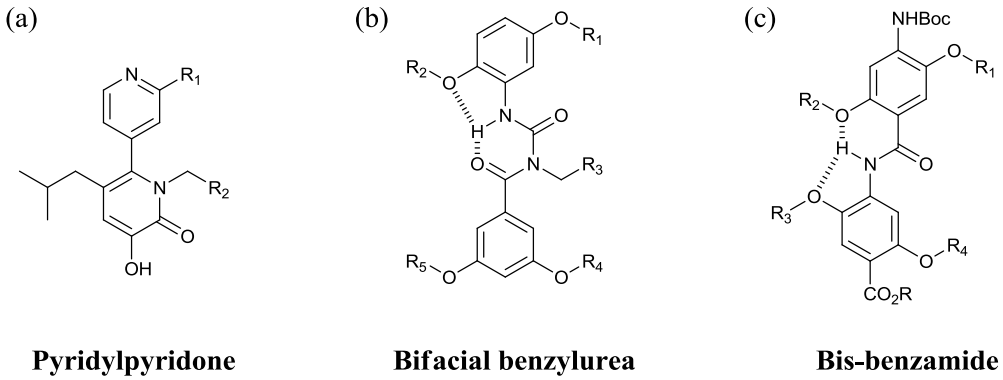


Figure 1.25 Bifacial scaffolds: (a) Pyridylpyridone; (b) Bifacial benzoylurea; (c) Bis-benzamide.

Using a combination of computational design and experimental SAR data, Arora and co-workers described the oxopiperazine scaffold (Figure 1.26a), identifying sub-μM inhibitors of the p53/*h*DM2 and p300/CREB-binding protein (CPB) interactions.[140] The Fletcher group designed a 1,2-diphenylacetylene scaffold (Figure 1.26b) that mimic the *i*, *i*+7 and *i*+2, *i*+5 side chains on opposite faces of an α-helix, facilitating amphipathic α-helix mimicry.[141] The Lim group further exploited this approach with the introduction of two-face amphiphilic α-helix mimetics based on the triazine-piperazinetriazine scaffold (Figure 1.26c). This class of compounds have the potential to generate combinatorial libraries and has already lead to nM inhibitors of the Mcl-1/NOXA B and Bcl-x_L/Bak interactions.[142]

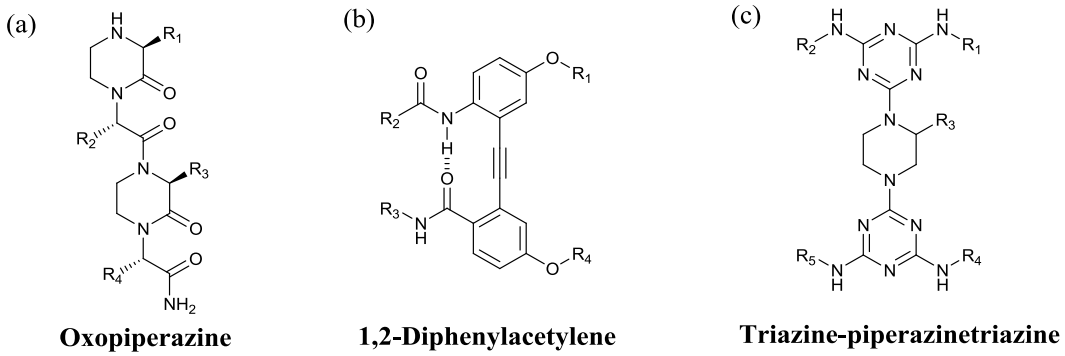


Figure 1.26 Other bifacial scaffolds: (a) Oxopiperazine; (b) 1,2-Diphenylacetylene; (c) Triazine-piperazinetriazine.

1.4. Project Aims

As illustrated in *Chapter 1*, the use of small-molecules as PPI modulators is fundamental for understanding the mechanisms that govern many diseases as well as to develop new therapeutic approaches. Among the wide range of topographies that PPIs display, the α-helix is the most common secondary structure in nature and thus represents a good generic template for inhibitor design.[1] The aim of this project was focused on developing a better understanding of the key features that play a vital role in modulating protein recognition in order to reproduce the functional role of α-helices and achieve specificity and selectivity towards different PPIs.

The Wilson group has focused on the inhibition of α-helix mediated PPIs using aromatic oligoamide proteomimetics, and most recently constrained peptides, to target different PPIs involved in cancer. Building on this previous work, a novel bifacial bis-benzamide scaffold to target the ER/co-activator PPI was to be designed and synthesised. Hence, expanding the scope of the existing scaffolds and allowing them to target multiple faces of an α-helix mediated PPI.

In addition, the insights of molecular recognition of the recently reported oligobenzmide hybrid scaffold were to be further explored using a SAR study in order to achieve more potent and selective hybrid compounds against p53/*h*DM2 and Mcl-1/NOXA B.

Finally, a novel PPI (Asf1/H3) was to be investigated using both stapled peptides and proteomimetics to determine advantages and disadvantages of each strategy when applied to the same interaction. Moreover, as each class of compound exhibit unique features, both strategies could be used to provide a more detailed understanding about the protein-protein interface.

Chapter 2. Design, Synthesis and Conformational Analyses of Bifacial Benzamide Based Foldamers

The work reported in this chapter formed the basis of the following publication:
S. Rodriguez-Marin, N. S. Murphy, H. J. Shepherd, A.J. Wilson, *RSC Adv.*, **2015,** 5, 104187-104192.[143]

2.1. Introduction

Most published studies on the development of α-helix mimetics to date focus on the design of oligobenzamides mimicking the key residues located on one face of the α-helix e.g. at the *i*, *i*+3/4, *i*+7/8 positions and so on. However, there are also examples of these scaffolds mimicking more than one face, as described in *Chapter 1*.[139, 144, 145] In the context of foldamer synthesis and structure,[146] the construction of backbones functionalised with different side-chains on multiple faces of the scaffold represents an as yet unrealised approach to achieve control over secondary conformation and higher order tertiary/quaternary organisation.

Similarly, there is an obvious need for PPI inhibiting helix mimetics that target more than one face of an interaction as shown by Arora and co-workers in their computational analysis of the PPIs on the Protein Data Bank (PDB).[5] This study revealed that helices are present at the interface of 62% of known multiprotein complexes, highlighting the importance of α-helices in PPIs. Furthermore, within this helical interface subset, 60% interact through residues on a single face of the helix, 30% contain hot spot residues on two faces and around 10% require all three faces for interaction with the target protein. These results manifest the therapeutic relevance of multifaceted helix mimetics to target biological systems e.g. the estrogen receptor (ER) is a ligand-activated transcription factor that plays a key role in the development of certain cancers and recruits a bifacial helical ligand for co-activation/repression.[147, 148]

2.2. Nuclear hormone receptor superfamily

Nuclear hormone receptors control the development, homeostasis, and metabolism of organisms. Their mode of action is based on the functional regulation of their ligands:

hormones, such as steroids and thyroids, as well as retinoids and vitamin D. In response to the binding with their natural ligands, resulting ligand/receptor dimers undergo structural changes that promote interactions with DNA and other molecules, such as cofactors. These interactions affect the transcriptional machinery, thereby up- or down-regulating the expression of specific genes.[149, 150]

2.2.1. Estrogen receptor

Estrogen receptor (ER) is a ligand-activated transcription factor that belongs to the nuclear hormone receptor superfamily. Its natural ligand is the steroid hormone 17-β-estradiol (E2),[151] and thus it is involved in controlling adolescence and reproductive processes whilst it participates in health preservation, such as maintenance of bone density and cholesterol levels.[152] In addition, it has long been implicated in human diseases such as schizophrenia, osteoporosis and cancers of the breast, colon and ovarian tissues. In particular, estrogenic signalling processes are crucial in the development of breast cancer.[148] The basis of its mechanism involves the binding of a hormone, which initiates changes in the receptor conformation and allows the ER to interact with specific DNA binding partners and other cellular transcription elements, such as co-regulators (Figure 2.1). These interactions result in the activation or repression of target genes.[152] This receptor was considered as the only ER until a second ER was reported by Kuiper *et al.* in 1996.[153] These two main groups are known as ERα and ERβ and can be detected in a broad spectrum of tissues. [151]

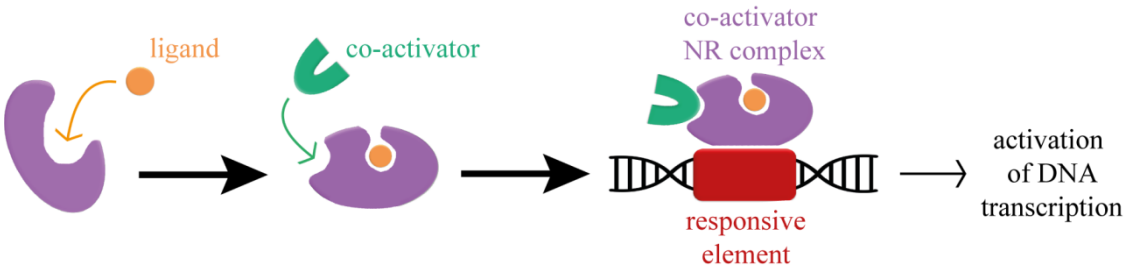


Figure 2.1 General mode of action of the Estrogen Receptor.

Overall structural features

ERα and ERβ are coded in different genes and are not splice variants. They are formed by six structural domains (termed domains A-F) (Figure 2.2). There is a high homology between domain C (96%), and domains E/F (53%), whereas domains A, B and hinge D are not that well conserved between ERα and ERβ.

Figure 2.2 Diagram representing the domain construction of nuclear receptors. The A/B domain contains AF-1 site that binds to other transcription factors. The C domain contains the two zinc fingers structure that binds to DNA. The E/F domain contains the ligand binding domain and the AF-2 site that interacts with peptidic co-activators.[154]

In addition to their structural domains, ERs enclose defined functional domains. An *N* terminal transcriptional activation function (AF-1) domain (NTD) is located in regions A and B. A DNA-binding domain (DBD) is located in region C and consists of two non-equivalent zinc fingers; one is responsible for recognizing the estrogen-responsive element (ERE) and the second stabilizes non-specific interactions with DNA segments. Region D is a hinge region and F is a variable domain. Region E functions as the ligand binding domain (LBD) and is the basis for the second AF domain (AF-2). The majority of co-activators bind the ER at AF-2. This binding interaction occurs through short amphipathic α-helix sequences contained in the co-factor structure. They contain multiple copies of a signature LXXLL (L = leucine, X = any amino acid) motif, also known as nuclear receptor box (NR box). These helical common motifs are recognized by a complementary groove formed by four α-helices on the surface of the ER whilst a "charge clamp", involving residues Lys^{362} and Glu^{542}, stabilizes them. In the co-activator helix, the leucine side chains in position *i* and *i*+4 are projected into a hydrophobic grove on the receptor surface, whereas the leucine side chain at position *i*+3 is located into an opposite hydrophobic pocket. Sequences flanking the core motif LXXLL are also found to be important in receptor selectivity (Figure 2.3).[148, 151, 152, 155]

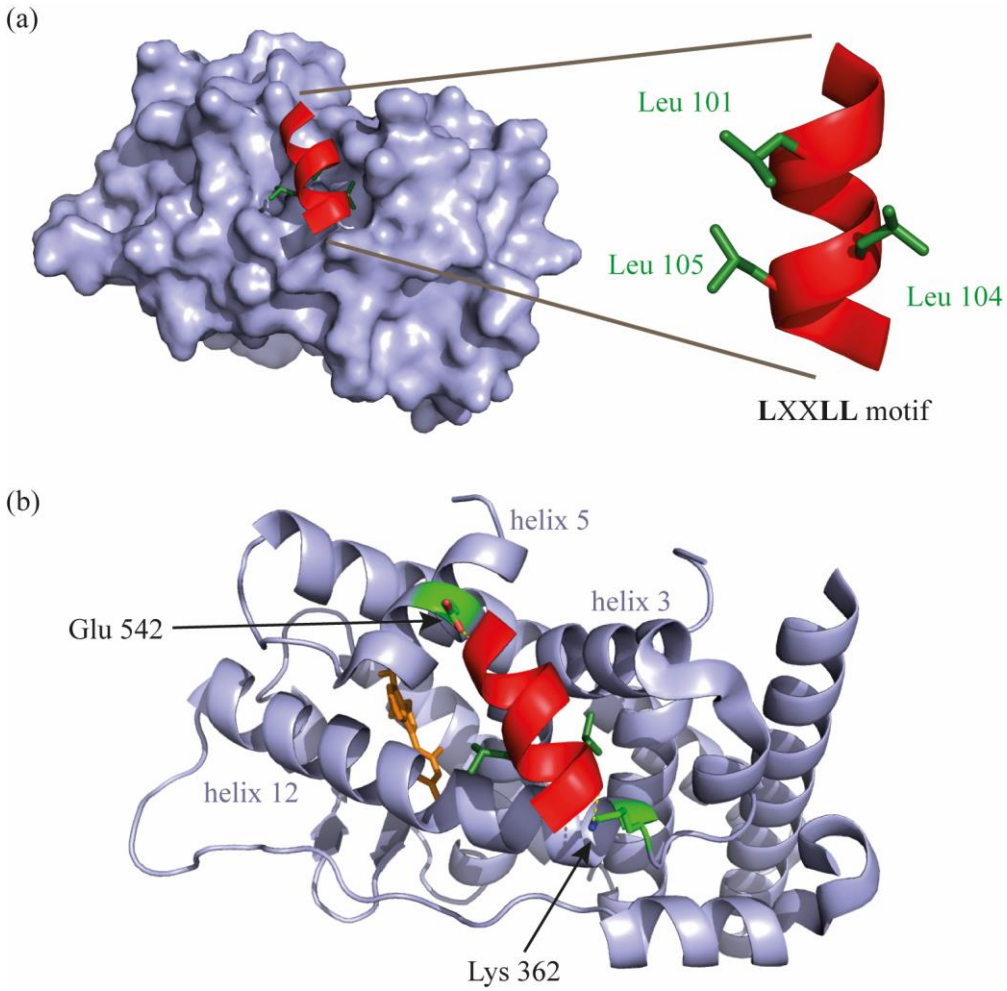


Figure 2.3 Crystal structure of the ERα (purple) bound to an LXXLL co-activator motif (red) (PDB ID: 2QZO). (a) The co-activator binding grove is shown and the key side chains on the helix are highlighted.[156] (b) The electrostatic interactions between the co-activator and the "charge clamp" residues (green) are shown (yellow) and the α-helices forming the binding groove are highlighted; the bound estrogen analogue ligand is also shown (orange).

The p160 protein family constitute one of the most relevant family of ER co-activators and consists of three members, SRC-1, SRC-2 and SRC-3, all of them containing the common LXXLL motif (NR box) in the nuclear receptor interaction domains.[154, 157] Direct inhibition of the receptor/co-activator protein-protein interaction,[137, 158-162] notably using helix mimetics[137, 161, 163] is of potential therapeutic interest as an alternative to the use of competitive inhibitors for the ligand binding site.[164]

2.2.2. Androgen receptor

The androgen receptor (AR) is a ligand-dependent transcription factor that belongs to the

nuclear hormone receptor superfamily. Its natural ligands are testosterone and dihydrotestosterone (DHT) and its main function is the regulation and maintenance of the male sexual phenotype. AR has a critical role in prostate cancer development and progression, even in the terminal stages of refractory forms of the disease. Ligand binding leads to conformational changes in the AR and its translocation from the cytoplasm into the nucleus, where it binds to androgen response elements and regulates transcription. This mechanism is analogous to the mode of action of the ER (Figure 2.1).[147, 165]

Overall structural features

AR contains the same structural and functional domains as the other members of the steroid hormone receptor family, such as ER (Figure 2.2). In AR, ligand binding induces structural modifications that reveal a groove in the AF-2 domain. This region binds short amphipathic helical peptides containing the NR box with the common recognition motifs LXXLL and FXXFF (L = leucine, F = phenylalanine, X = any amino acid). A distinctive feature of the AR is that the interaction can occur either with co-activator proteins (containing the LXXLL or FXXFF motifs) or intramolecularly with the FXXLF or WXXLF motifs located in the *N*-terminal region of the protein itself. The interaction between the hydrophobic leucine and phenylalanine residues from the common helical motif and the AR occur in the same manner as in the ER and it also involves the so-called "charge clamp" (residues Lys^{720} and Glu^{897}) which bracket the cleft. Alternatively, AR can interact with the *N*-terminal motifs *via* a glutamine-rich region rather than the LXXLL NR box motif as in the estrogen receptor.[147] Nevertheless, the interaction of ligand-bound AR with its co-activators is still not fully understood (Figure 2.4).[147, 166]

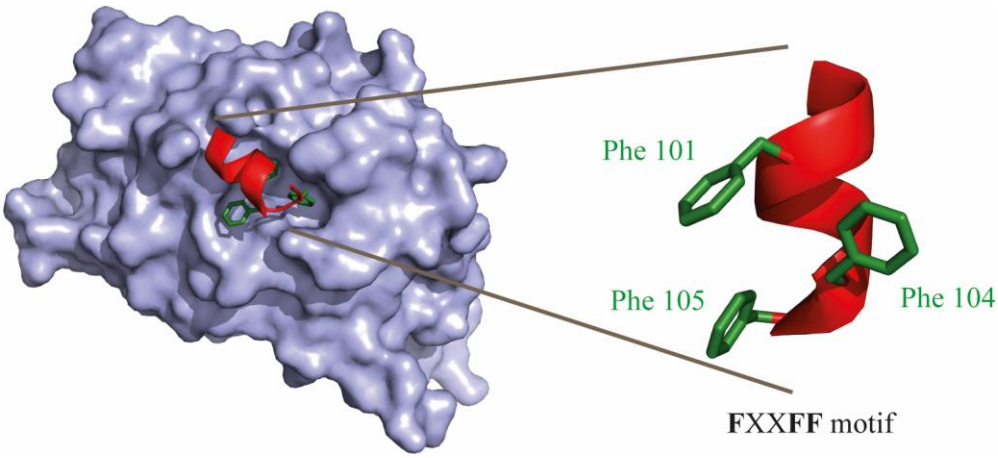
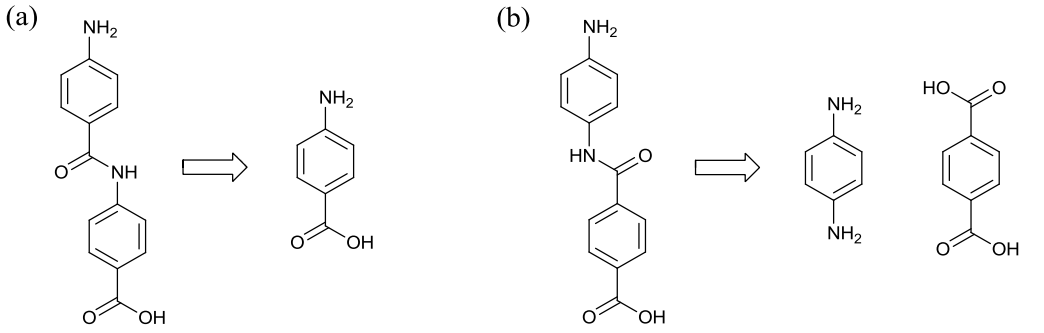


Figure 2.4 Crystal structure of the AR bound to an FXXFF co-activator motif (PDB ID: 1T73).[166] Key side chains on the binding surface of the helix are highlighted.

2.3. Design and Synthesis of bifacial proteomimetic scaffolds

In this chapter we introduce two bifacial proteomimetic scaffolds, bis-benzamide and *N*-(4-aminophenyl)terephthalamidate (Figure 2.5). The novel foldamers were anticipated as tools to (a) enhance our understanding of aromatic oligoamide foldamer conformation and (b) ligands that could mimic the key side chains at *i*, *i*+3, *i*+4 positions of α-helices that participate in PPIs mediated by such a side chain constellation.

The design of these α-helix mimetics envisioned the use of the bis-benzamide and *N*-(4-aminophenyl)terephthalamidic acid molecules as dimeric backbones which can be functionalised at different positions using *O*-alkylated monomers with the final goal of mimicking the key side chains of the co-activator recognition motif LXXLL.



Bis-benzamide backbone N-(4-aminophenyl)terephthalamidic backbone

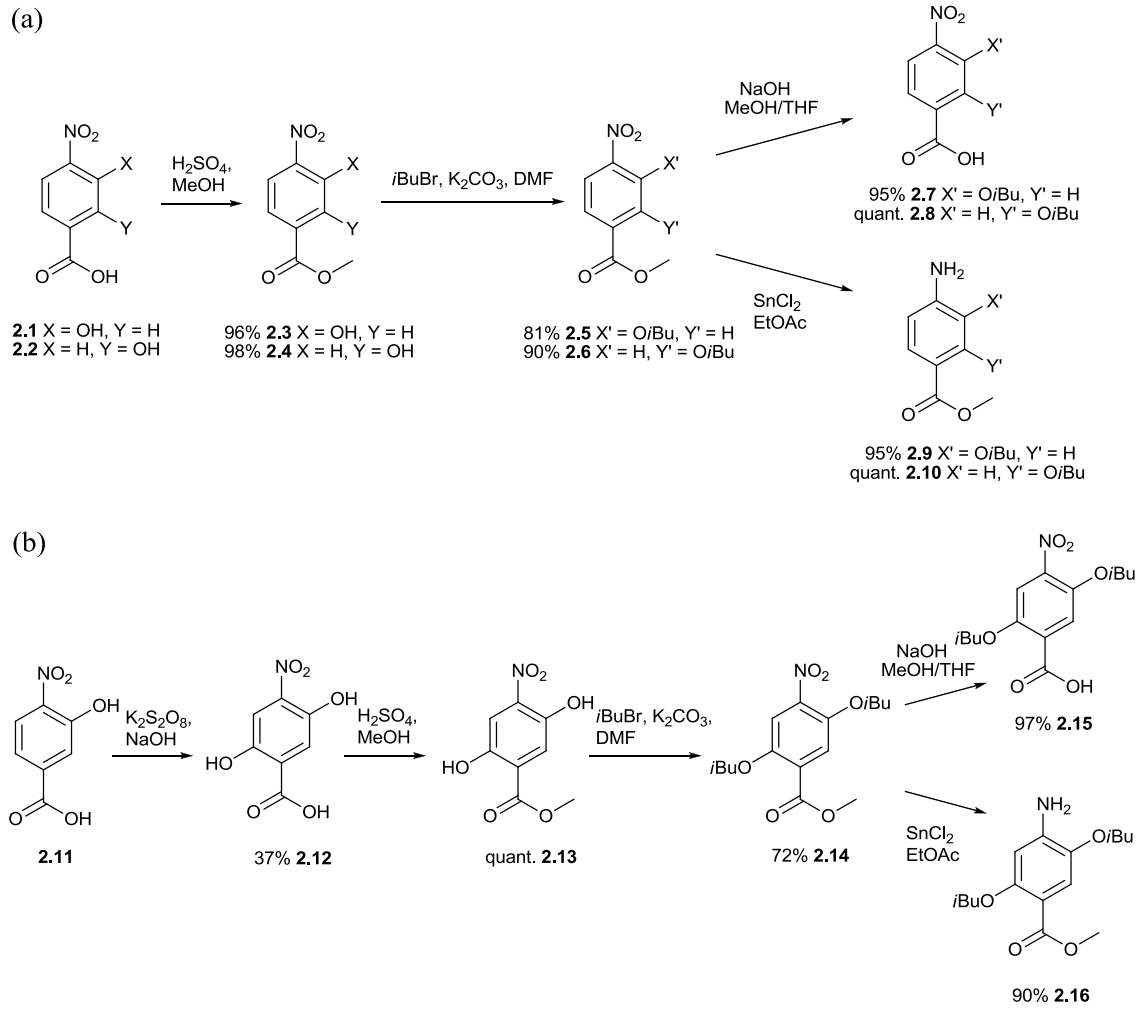
Figure 2.5 (a) Bis-benzamide scaffold with the corresponding *para*-aminobenzoic acid constituent building block; (b) *N*-(4-aminophenyl)terephthalamidic scaffold with the corresponding *para*-phenylenediamine and terephthalate constituent building blocks.

A first generation of bifacial proteomimetic inhibitors, based on the bis-benzamide backbone, was synthesised using the work previously developed in the group on the 3-*O*-alkylated, 2-*O*-alkylated and 2,5-*O*-dialkylated building blocks (Figure 2.5a).[122, 126, 127] This work continued the synthesis started by a previous member of the group Dr Natasha S. Murphy. Preliminary *in silico* studies performed on modified versions of the bis-benzamide scaffold lead to a second generation of inhibitors. The novel scaffold was designed based on a *N*-(4-aminophenyl)terephthalamidic acid backbone, where the central amide bond is inverted in comparison with the bis-benzamide scaffold. This modification made necessary the development of novel synthetic routes to the *para*-phenylenediamine and terephthalate building blocks together with new monomer coupling conditions (Figure 2.5b).

2.3.1. Synthesis of the first generation of a bifacial proteomimetic scaffold

2.3.1.1. Bis-benzamide monomer building blocks synthesis

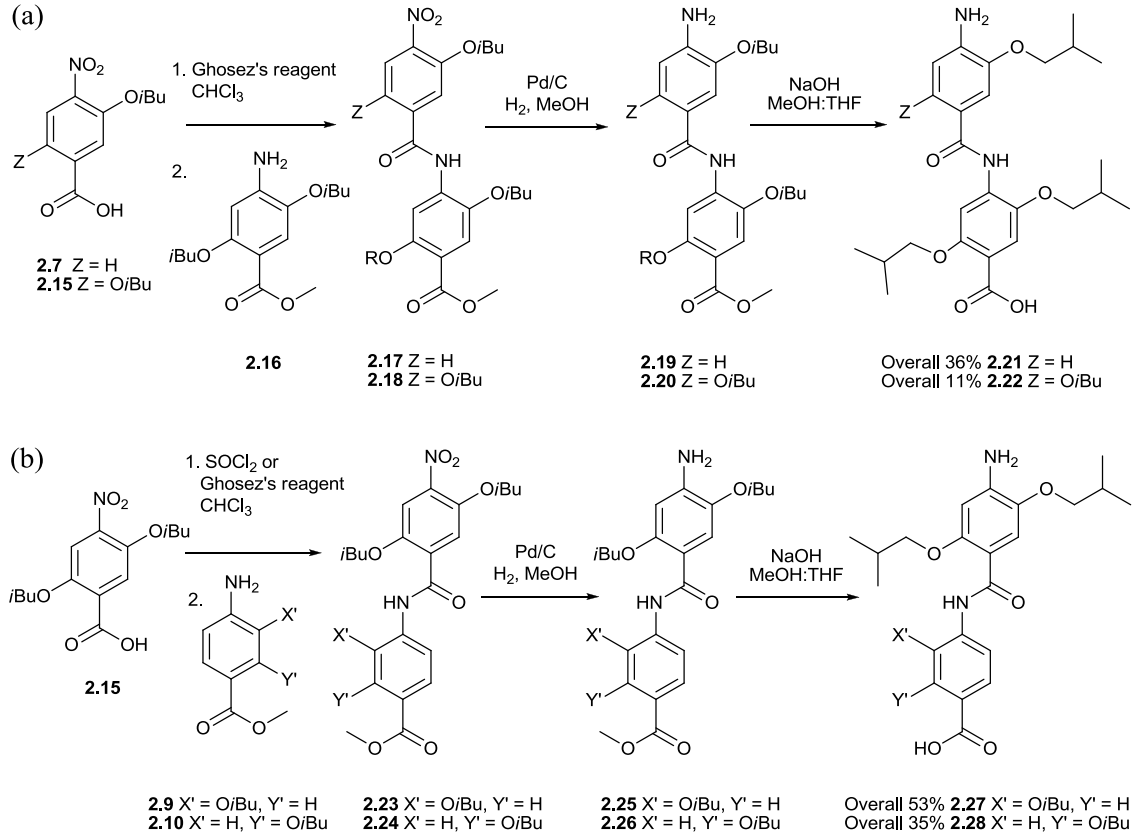
The synthesis for the bis-benzamide monomer building blocks used methods previously developed in the group (Scheme 2.1).[122] Fisher esterification of the commercially available **2.1** and **2.2** followed by alkylation using isobutyl bromide generated intermediates **2.5** and **2.6** in excellent yields. The alkylated products were then either reduced to the amino derivatives **2.9** and **2.10** using tin (II) chloride or hydrolysed to the acid product **2.7** and **2.8** in basic conditions. The synthetic route to the dialkylated building blocks **2.15** and **2.16** added a prior dihydroxylation reaction, which follows a procedure described by the Ahn group (Scheme 2.1b).[139]



Scheme 2.1 Synthesis of 3-*O*, 2-*O*, 2,5-*O* alkylated *p*-aminobenzoic acid monomers for bis-benzamide derived foldamers.

2.3.1.2. Bis-benzamide dimers synthesis

The synthesis of the bis-benzamide dimers followed the same synthetic strategy previously developed in the group (Scheme 2.2). The nitro-acid monomers **2.7** and **2.15** were reacted with thionyl chloride or Ghosez's reagent to form the corresponding acyl chloride before coupling them with an appropriate aniline partner **2.16**, **2.9** or **2.10**. Subsequent reduction of the nitro compounds **2.17**, **2.18**, **2.23** and **2.24** to the corresponding amine **2.19**, **2.20**, **2.25** or **2.26** with Pd catalysed hydrogenation followed by basic ester hydrolysis gave the final foldamers **2.21**, **2.22**, **2.27** and **2.28**.



Scheme 2.2 Synthesis of the bis-benzamide foldamers.

The regioisomer of compound **2.21** could not be obtained due to unsuccessful coupling between methyl 4-amino-2,5-diisobutoxybenzoate **2.16** and 4-amino-2-isobutoxybenzoic acid **2.8** under multiple conditions. The tetrasubstituted scaffold **2.22** was also synthesised to explore the role of a 4th side chain in helix mimicry (Figure 2.6).

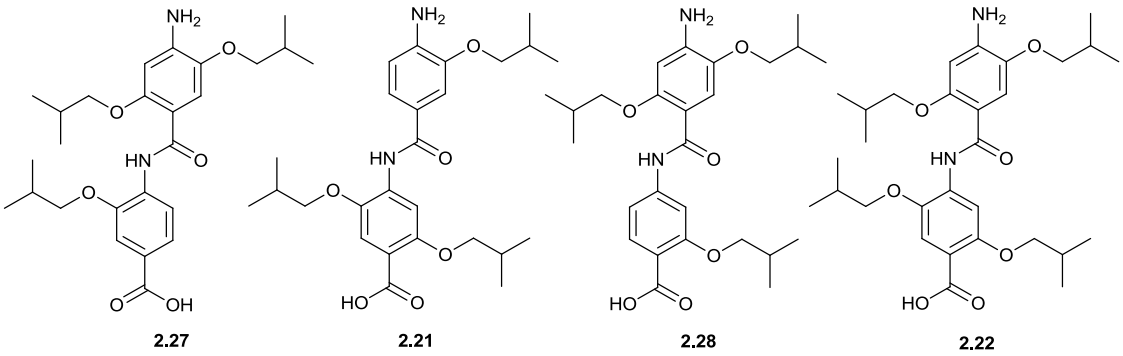


Figure 2.6 Bis-benzamide foldamers comprising 3-*O*, 2-*O*, 2,5-*O* alkylated *p*-aminobenzoic acid monomers.

2.3.2. Synthesis of the second generation of a bifacial proteomimetic scaffold

A novel second generation scaffold was designed based on a *N*-(4-aminophenyl)terephthalamidic acid backbone. The dimer is formed from a *para*-phenylenediamine monomer linked to a terephthalate monomer through an amide bond and the backbone can be functionalized at different positions using a variety of *O*-alkylated monomers (Figure 2.7).

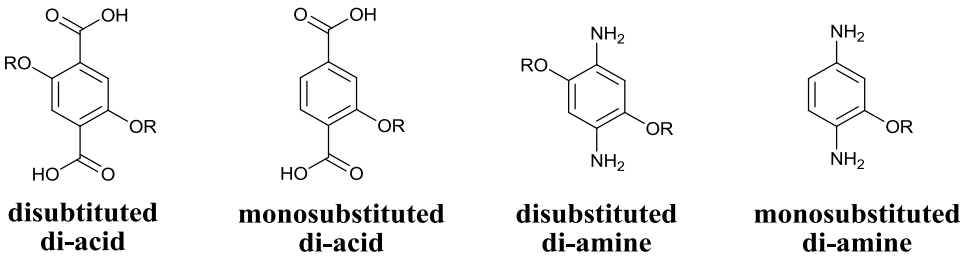
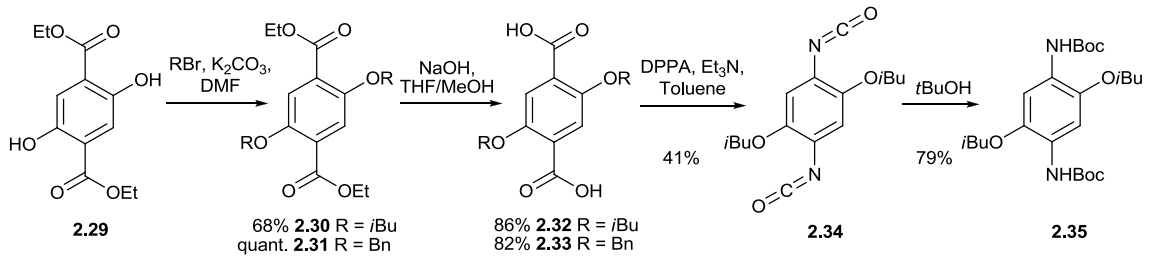


Figure 2.7 Building blocks for the *N*-(4- aminophenyl)terephthalamidic scaffold.

2.3.2.1. Disubstituted di-acid and di-amine monomer synthesis

For the disubstituted di-acid monomer, double alkylation of the commercially available diethyl 2,5-dihydroxyterephthalate **2.29** with isobutyl or benzyl bromide gave access to intermediates **2.30** and **2.31** containing isobutyl and phenyl side-chains respectively. Subsequent base-catalysed hydrolysis of the esters provided the final di-acid building blocks **2.32** and **2.33** in excellent yields (Scheme 2.3).



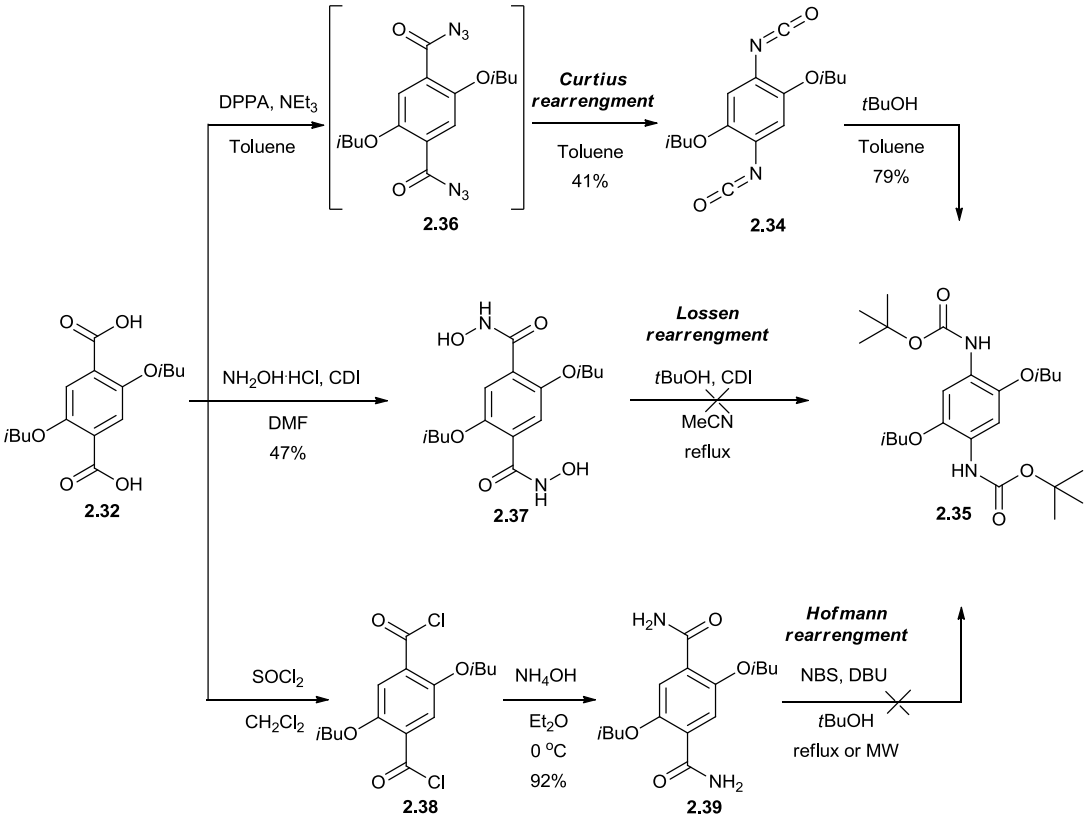
Scheme 2.3 Synthesis of disubstituted di-acid and di-amine monomers for *N*-(4-aminophenyl)terephthalamidic derived foldamers.

The initial synthetic strategy to obtain the disubstituted di-amine building block consisted of forming the isocyanate intermediate derived from **2.32**, which could then be hydrolysed to the desired amine or trapped with alcohols to afford the corresponding carbamate **2.35**.[167] Three different rearrangement reactions were investigated (Scheme 2.4):

Hofmann rearrangement:[168] The acyl chloride intermediate **2.38**, resulting from the reaction of **2.32** with thionyl chloride, was reacted without further purification with an aqueous ammonia solution to afford the desired primary carboxamide **2.39** in excellent yield. Subsequent conversion into the corresponding carbamates **2.35** was attempted using N-bromosuccinimide (NBS) as a bromine source and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) as a base, either under reflux or microwave conditions. Unfortunately, formation of the expected product **2.35** was not observed.

Lossen rearrengement:[169] Compound **2.32** was reacted with 1,1'-carbonyldiimidazole (CDI) and hydroxylammonium hydrochloride to afford the desired hydroxamic acid **2.37** in moderate yield. The rearrangement into its corresponding carbamate **2.35** was attempted using 1,1'-carbonyldiimidazole (CDI) and tert-butanol under reflux. Unfortunately, formation of the expected product **2.35** was not observed.

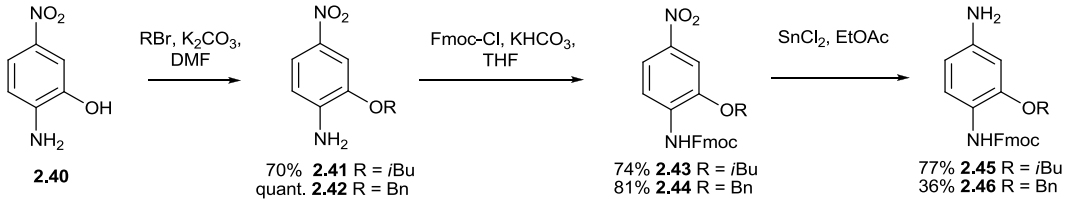
Curtius rearrengement:[170] Monomer **2.32** was reacted with diphenylphosphoryl azide (DPPA) under basic conditions to afford the desired di-isocyanate **2.34** in moderate yield. Surprisingly, the acyl azide intermediate **2.36** was not isolated as expected, which suggests that it decomposed at low temperature (ca. 30 °C). Low temperature Curtius rearrangements are unusual; however, some examples are described in the literature.[171, 172] The di-isocyanate **2.34** was treated with tert-butanol in toluene affording the corresponding carbamate **2.35** in good yield. The corresponding unprotected di-amine building block was found to be unstable upon exposure to air and/or aqueous media, possibly due to polymerization and oxidation processes.[173]



Scheme 2.4 Synthetic routes investigated from monomer **2.32** to **2.35**. Different types of rearrengment reactions are highlighted.

2.3.2.2. Monosubstituted di-amine monomer synthesis

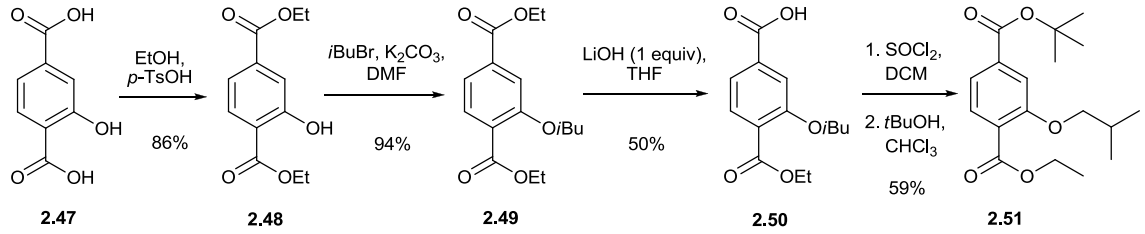
A noteworthy feature of the monosubstituted di-amine monomer synthesis is the use of the common starting material **2.40** to construct two different classes of building blocks: compounds **2.41** and **2.42**; and compounds **2.45** and **2.46**, which are suitable for the assembly of regioisomeric foldamers (Scheme 2.5). Alkylation of the commercially available compound **2.40** with isobutyl or benzyl bromide provided monomers **2.41** and **2.42** respectively. Subsequent Fmoc protection of the amino group followed by the reduction of the nitro group to the corresponding amine with tin (II) chloride gave access to the regioisomers **2.45** and **2.46**.



Scheme 2.5 Synthesis of monosubstituted di-amine monomers for *N*-(4-aminophenyl)terephthalamidic derived foldamers.

2.3.2.3. Monosubstituted di-acid monomer synthesis

For the synthesis of monosubstituted alkoxy derivatives of 2-hydroxyterephthalic acid **2.47** (Scheme 2.6), double Fisher esterification followed by alkylation gave intermediate **2.49** in high yield. It was then necessary to perform a sequence of protecting group manipulations; selective hydrolysis of the most electron-deficient ethyl ester in compound **2.49** yielded the monosubstituted di-acid building block **2.50**. Subsequent *tert*-butyl esterification of the free carboxylic acid lead to compound **2.51**.



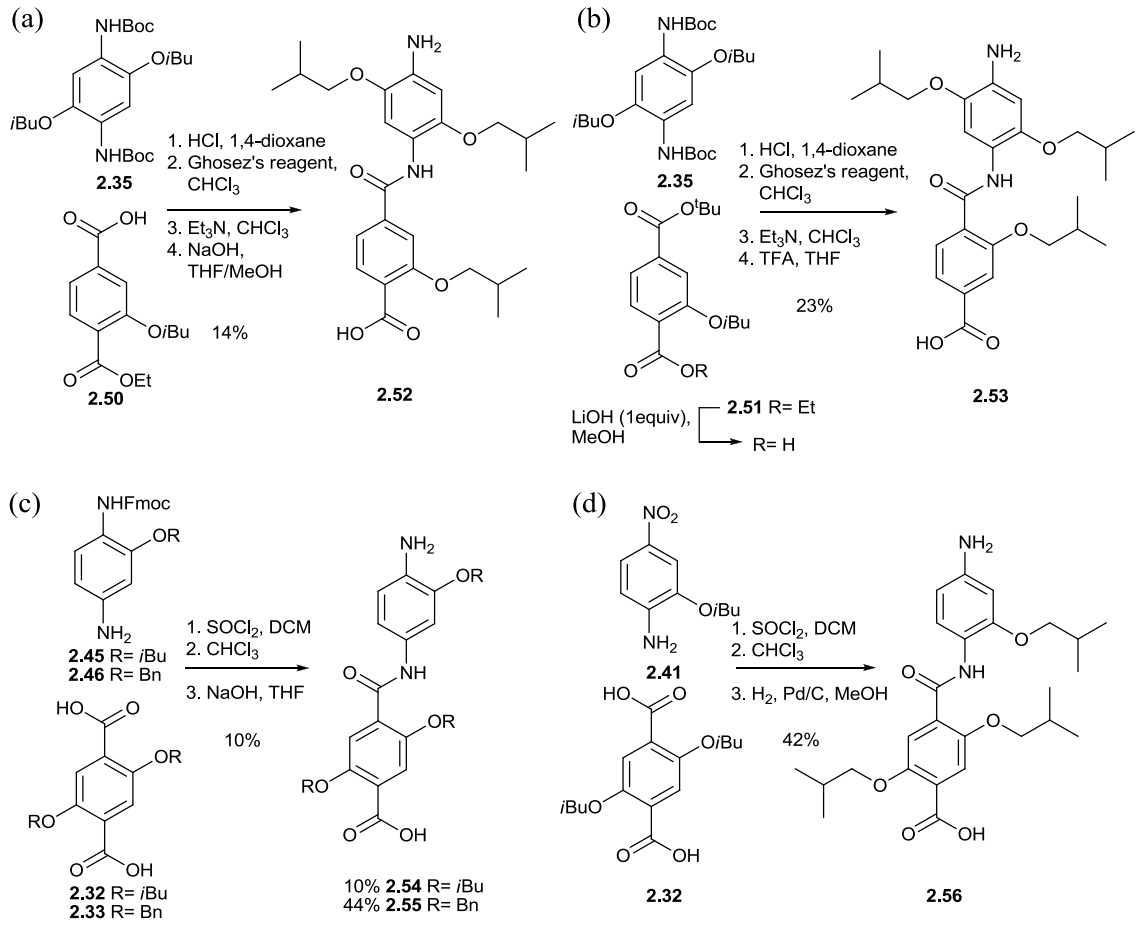
Scheme 2.6 Synthesis of monosubstituted di-acid monomers for *N*-(4-aminophenyl)terephthalamidic derived foldamers.

2.3.2.4. *N*-(4- aminophenyl)terephthalamidic dimer synthesis

To effect amide bond formation, the acyl chloride of the di-acid monomer **2.32** and **2.33** was obtained using thionyl chloride before coupling to its amino-monomer partners **2.45**, **2.46** and **2.41** (Scheme 2.7c and d). By using an excess of the di-acid **2.32** and **2.33** it was possible to bias the product distribution towards the monoamide. The final products **2.56**, **2.54** and **2.55** were obtained by hydrogentation of the nitro group or hydrolysis of the Fmoc group respectively.

Due to the oxidation upon exposure to air as mentioned above, the di-amine derivative of compound **2.35** was obtained through *in situ* Boc deprotection and direct reaction with the acid chloride derivative of **2.50**, which was obtained by *in situ* activation using Ghosez's reagent (Scheme 2.7a). Alternatively, the ethyl ester of compound **2.51** was selectively hydrolysed in basic conditions and directly transformed to the acid chloride by *in situ* activation using Ghosez's reagent before coupling to the di-amine derivative of compound **2.35** (Scheme 2.7b). Again, the monoamide product was biased by using the starting di-amine **2.35** in excess. The final compounds **2.52** and **2.53** were obtained after appropriate deprotection sequences (Scheme 2.7a and b) either basic or acid hydrolysis, respectively.

Despite numerous efforts, we were unable to obtain the dimer derived from **2.42** and **2.33**. The amide formation was performed successfully; however, the following reduction step to transform the nitro to the amino group was ineffective and cleavage of the side chains was observed under forcing conditions, such as high temperatures and long reaction times. The synthetic route for the tetrasubstituted foldamer derived from **2.32** and **2.35** was proven successful following the above described methodology. Unfortunately, the dimer was unstable upon air exposure and could not be isolated.

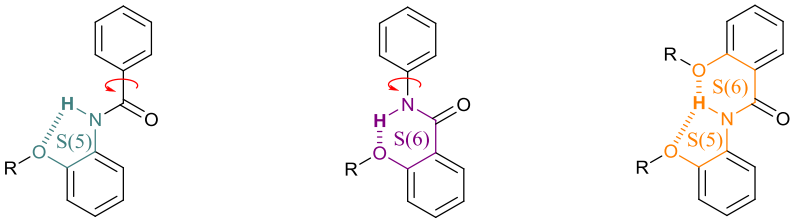


Scheme 2.7 Synthesis of *N*-(4-aminophenyl)terephthalamidic foldamers.

2.4. Conformational analyses

2.4.1. 2D NMR studies

Previous studies performed in the group on 2- and 3-*O*-alkylated trimers and model dimers on the oligobenzamide scaffold revealed intramolecular pseudo-six- or five-membered rings hydrogen bonding between the NH and adjacent *O*-alkyl group (Figure 2.8).[127, 128] This resulted in restricted rotation around one of the Ar-CO or Ar-NH bonds leaving the other free to rotate. The conformation of such scaffolds can be further restricted by introduction of a second alkoxy group leading to a "bifurcated" hydrogen bonding interaction, where the NH is located between two phenolic oxygens from adjacent monomers forming pseudo-six- and five-membered rings.[126, 174, 175]



pseudo-5-membered H bond ring | *pseudo*-6-membered H bond ring | "bifurcated" H bond ring

Figure 2.8 Illustration of the types of intramolecular hydrogen bonding interactions on the oligobenzamide scaffold.

In principle, the set of compounds discussed here can display a similar array of conformations as those described above. Therefore, structural and conformational analyses were performed on each compound using NMR spectroscopy. In particular, ^{1}H-^{1}H NOESY analyses were used to determine their preferred conformation in solution by identifying the interactions of the dimer amide NH with the adjacent aromatic protons.

Compounds **2.21** and **2.52** formed pseudo five-membered hydrogen bonded rings, whilst compounds **2.28**, **2.54** and **2.55** formed pseudo six-membered hydrogen bonded rings, as expected in each case, as the NH could only form a single type of intramolecular hydrogen bond. A representative example is shown in Figure 2.9 for compound **2.21**. The amide NH displays nOe correlations with the adjacent aromatic protons 1-H2 and 1-H6 suggesting free rotation around the Ar-CO axis, whereas absence of cross peaks with the aromatic proton 2-H3 indicated that rotation was constrained around the Ar-NH axis. Thus, confirming that the amide proton was locked in an S(5) intramolecular H-bonded ring.

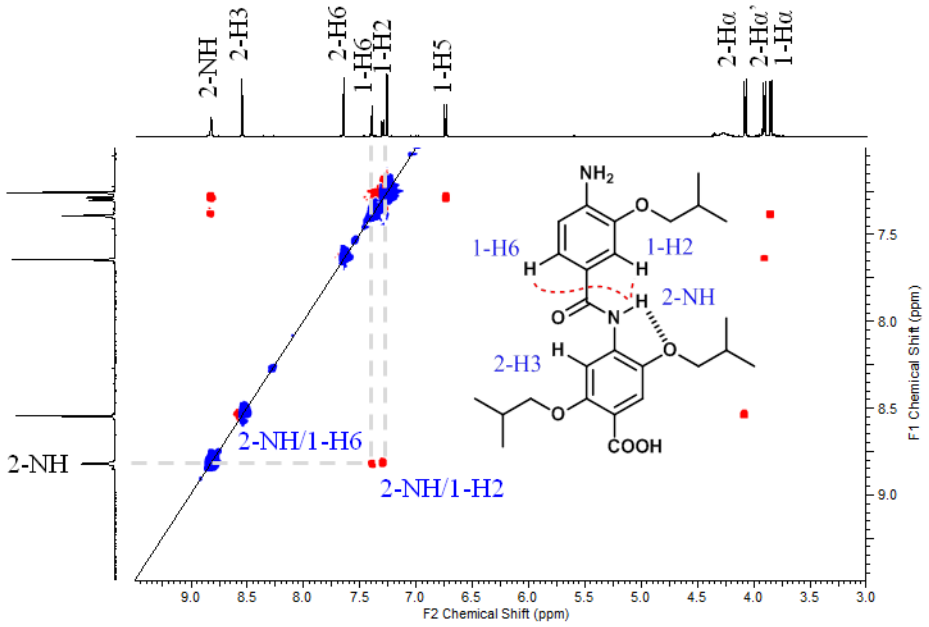


Figure 2.9 ^{1}H-^{1}H NOESY (500 MHz, $CDCl_3$) spectra of dimer **2.21** at 5 mM. Structures and ^{1}H proton assignments are shown and relevant nOe signals are highlighted.

More interestingly, compounds **2.27**, **2.22**, **2.53**, **2.56** could potentially display either isolated S(5) and S(6) or "bifurcated" S(5)/S(6) hydrogen bonding systems. Compound **2.27** was indicative of both pseudo-five- and six-membered hydrogen-bonded rings being populated in solution. The absence of nOe correlations between the amide and the adjacent aromatic protons 1-H6 and 2-H3 suggested that rotation was constrained around both the Ar-CO and Ar-NH axes (Figure 2.10a). An X-ray crystal structure of compound **2.27** previously obtained by Dr Natasha. S. Murphy support this result with NH to O distances of 2.007 and 2.223 Å for the S(6) and S(5) H-bonded rings respectively (Figure 2.10b).[143]

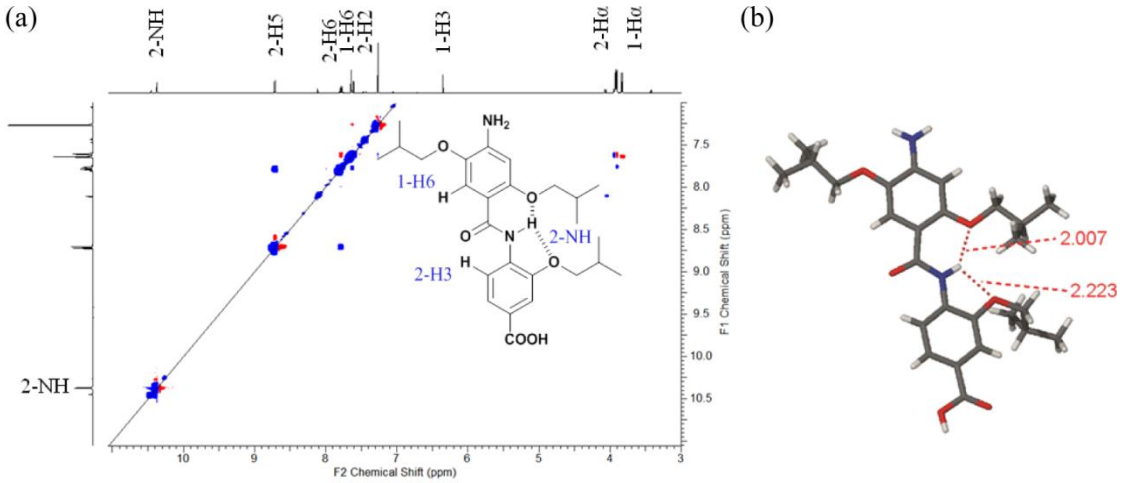


Figure 2.10 (a) ^{1}H-^{1}H NOESY (500 MHz, $CDCl_3$) spectra of dimer **2.27** at 5 mM. Structures and ^{1}H proton assignments are shown and relevant nOe signals are highlighted; (b) X-ray structure of **2.27**, H-bonding distances (Å) are shown in red.

Intriguingly, compounds **2.22**, **2.53** and **2.56** showed evidence of only pseudo-six-membered intramolecular hydrogen bonded ring formation in solution. A representative example is shown in Figure **2.11** for compound **2.56**. The amide displays nOe correlations with the adjacent aromatic proton 1-H6 suggesting free rotation around the Ar-NH axis, whereas absence of cross peaks with the aromatic proton 2-H6 and nOe signals with the side chain proton 2-Hα′ proposed that rotation was constrained around the Ar-CO axis. Thus, confirming that the amide proton was locked into an S(6) intramolecular H-bonded ring.

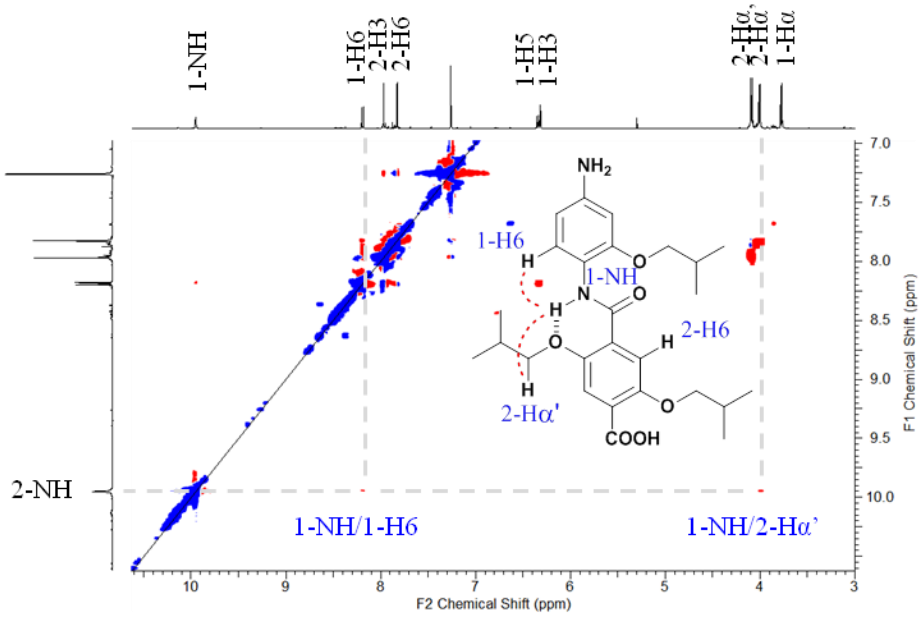


Figure 2.11 ^{1}H-^{1}H NOESY (500 MHz, $CDCl_3$) spectra of dimer **2.56** at 5 mM. Structures and ^{1}H proton assignments are shown and relevant nOe signals are highlighted.

Initially, DMSO was chosen as an appropriate solvent model for comparison with the aqueous media in biological systems. Unfortunately, the complete conformational analysis of the compounds in DMSO proved problematical due to indistinguishable peaks and weak signals. Nevertheless, the results that were obtained in DMSO were comparable to those obtained in $CDCl_3$.

2.4.2. H/D Exchange studies

H/D exchange studies were performed to further characterise the hydrogen bonding interactions involved in controlling the conformation of our set of compounds. It is worth noting that the relative rates of this exchange depend on different factors, such as the acidity of the NH proton, which will be affected by its electronic environment; the steric accessibility of the NH group and the strength of the hydrogen bonding. The H atoms are anticipated to exchange more

rapidly the more acidic they are, the less steric hindrance they present and the weaker hydrogen bonds they form.[176] Nevertheless, the correlation of the rate of exchange with the strength of the hydrogen bonding can be used to obtain additional information about these interactions.

The H/D exchange experiments were performed on compounds **2.27**, **2.21** and **2.56**, as models of the three different types of hydrogen bonding interactions. A 10% $CD_3OD/CDCl_3$ system was used to ensure pseudo first order kinetics. A distinct non-exchanging signal was used as an internal integration reference in order to minimize variability. The rate constant was determined from the slope of a non-linear least squares fit to the graph following Equation 1 and the half-life of the H/D exchange determined using Equation 2.

Equation 1 $A_t = A_o e^{-kt}$

A_t : Integral of amide proton at time t

A_o : Integral of amide proton at time zero (fixed at 1)

k : reaction rate coeficient

Equation 2 $t_{½} = \frac{\ln 2}{k}$

The resulting graph is shown in Figure **2.12**. The extracted $k_{H/D}$ and $t_{½}$ values for compounds **2.27**, **2.21** and **2.56** are listed in Table **2.1**, together with the values for two other previously reported compounds **2.57** and **2.58** from our group for comparison (Figure **2.13**).[127, 128] The order of magnitude for the amide proton exchange rate constants suggest the presence of a S(5), S(6) and bifurcated S(5)/S(6) H-bonding for compounds **2.21**, **2.56** and **2.27** respectively. These values are consistent with previous studies[128] and propose a (S)6 hydrogen bond with greater stability than its (S)5 analogue, the stabilization is even higher in the "bifurcated" S(6)/S(5) system. Importantly, these results are in agreement with the proposed conformations from the ^1H-^1H NOESY analyses (See section 2.4.1).

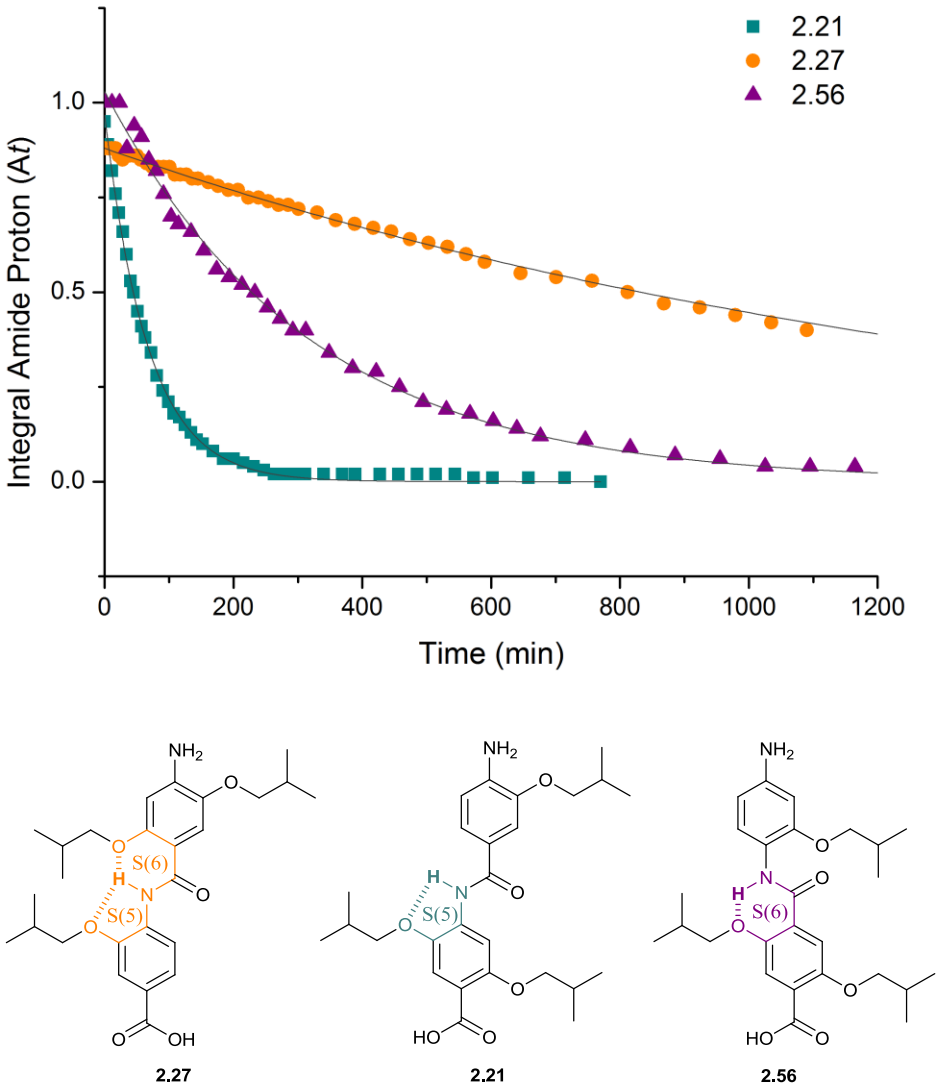
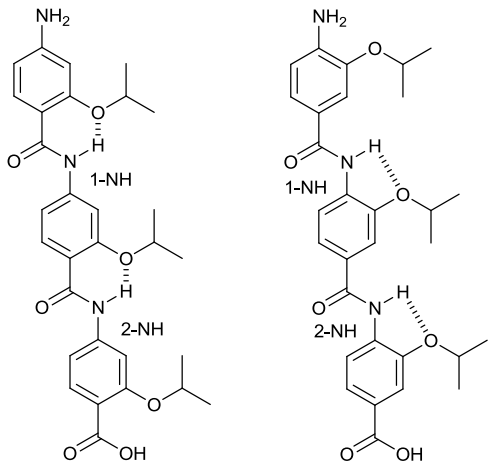


Figure 2.12 H/D exchange kinetics of compounds **2.27**, **2.21** and **2.56** at 10 mM in 10% $CD_3OD/CDCl_3$.

Table 2.1 Kinetic constants and $t_{1/2}$ based on H/D exchange in 10% $CD_3OD/CDCl_3$.

	$k_{H/D}$ (min^{-1})	$t_{½}$ (min)	H bonding
2.56	0.00305 ± 0.00005	228 ± 3	S(6)
2.27	6.7857×10^{-4} ± 0.0000093	1021.5 ± 14	S(5)/S(6)
2.21	0.01485 ± 0.00017	46.7 ± 0.5	S(5)
2.57 (1-NH)	0.00176 ± 0.00005	394 ± 12	S(6)
2.57 (2-NH)	0.00230 ± 0.00005	301 ± 6	S(6)
2.58 (1-NH)	0.0212 ± 0.0004	32.7 ± 0.6	S(5)
2.58 (2-NH)	0.0225 ± 0.0005	30.8 ± 0.7	S(5)



2.57 **2.58**

Figure 2.13 Reported reference compounds **2.57** and **2.58**, intramolecular hydrogen bonding interactions are shown.[127, 128]

2.5. Molecular modelling

A conformational search was performed on the entire set of compounds. The structure was minimised performing a full *Monte Carlo* search with the MMFFs method and using the software Macromodel®.[177] Water was chosen as implicit solvent and free rotation around the amide bonds was allowed in order to increase the accuracy of the conformational search. All the conformations within 1.5 kJ/mol of the lowest energy conformation were selected for further analysis. In the lowest energy conformation all the compounds adopt an extended structure, where the amide bond is *trans*. Importantly, the conformations for each compound are consistent with those that are accessible in solution phase according to the NOESY data and H/D exchange experiments (Figure 2.14).

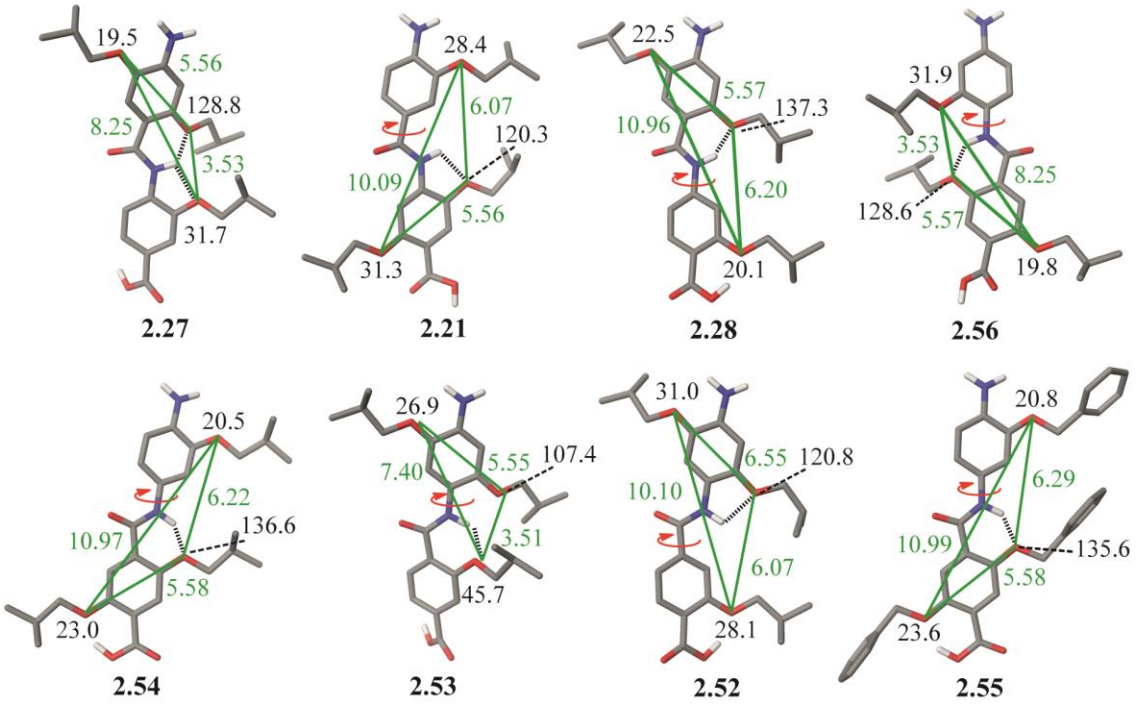


Figure 2.14 Preferred conformation and intramolecular hydrogen-bonding interactions of the compounds **2.21**, **2.27**, **2.28** and **2.52** to **2.56** supported by molecular modelling, 2D NMR studies and H/D exchange experiments. Distances and angles between side chains (green and black respectively), H-bonds (dashed black line) and free rotation axes (red arrow) are shown.

The nature of the structure permits the superimposition in both parallel and antiparallel *N*-to-*C* orientation with respect to an α-helical peptide.[125] Accordingly, both alignments were analysed using an ERα co-activator sequence (PDB ID: 2QZO). The match was assessed on the basis of the RMSD between α-carbons on the helix and oxygen atoms on the foldamer together with an evaluation on the quality of the backbone orientation with respect to the helical axis of the peptide.

The overlay of the first generation of foldamers with the native co-activator peptide is shown in Figure 2.15. Compounds **2.21**, **2.27** and **2.28** present a good overlay, where the three side chains overlap reasonably well with the leucine residues at positions *i*, *i+3* and *i+4* of the co-activator helix and the distances between the oxygens of the dimers match the distance between the α-CH of these residues.

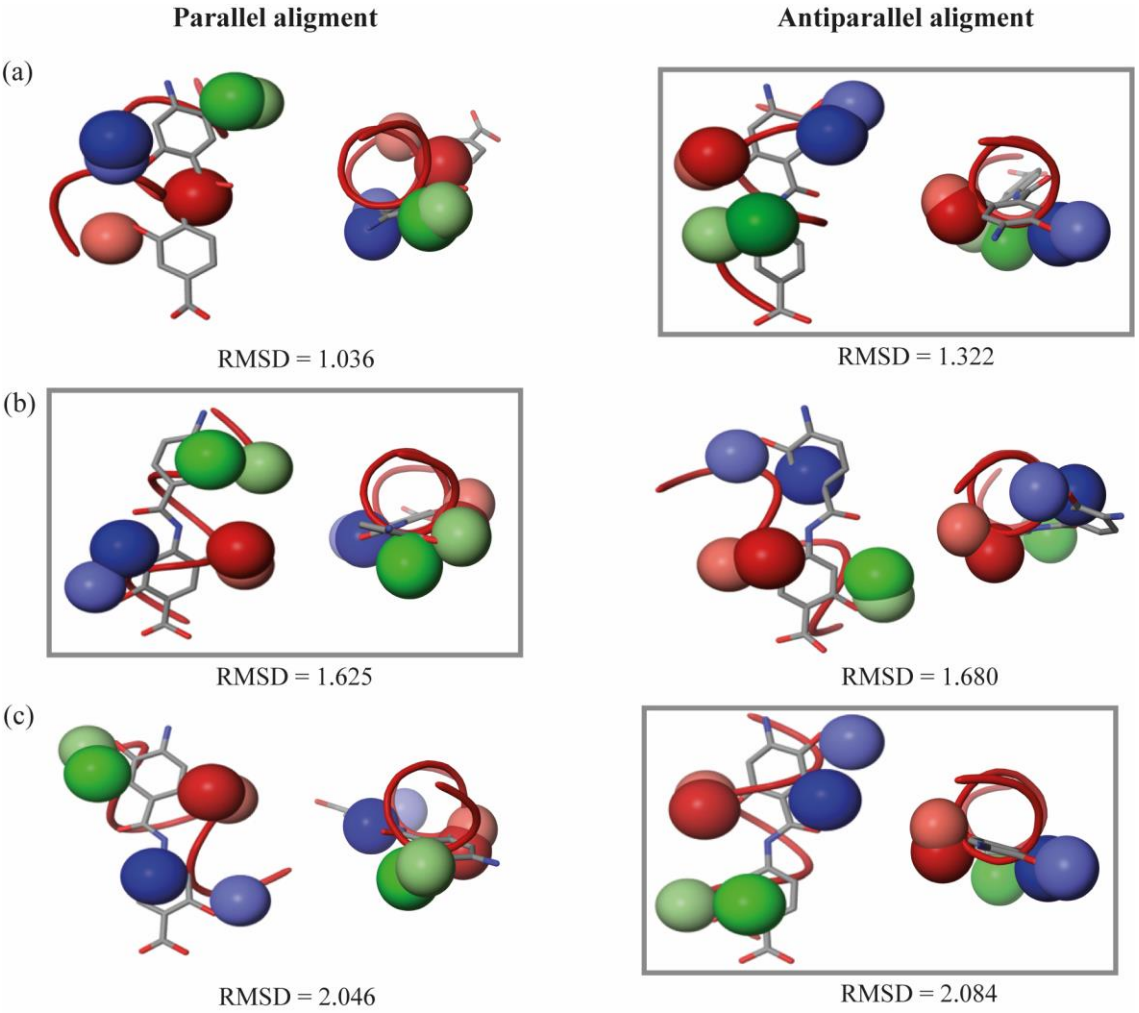


Figure 2.15 Overlay of the first generation of foldamers with the native co-activator peptide. Co-activator residues are in dark colours and helix mimetic residues are in light colours. Parallel (right) and antiparallel (left) alignment with the peptide dipole moment are shown (RMSD values are given for both alignments and the best alignment shown in a box): (a) compound **2.27**; (b) compound **2.21**; (c) compound **2.28**.

The overlay of the second generation of foldamers with the native co-activator peptide is shown in Figure 2.16. Compounds **2.52** and **2.54** present a good overlay, where the three side chains overlap reasonably well with the leucine residues at positions *i*, *i+3* and *i+4* of the co-activator helix and the distances between the oxygens of the dimers match the distance between the α-CH of these residues. Compounds **2.53** and **2.56**, matched less well in terms of alignment with the helical backbone.

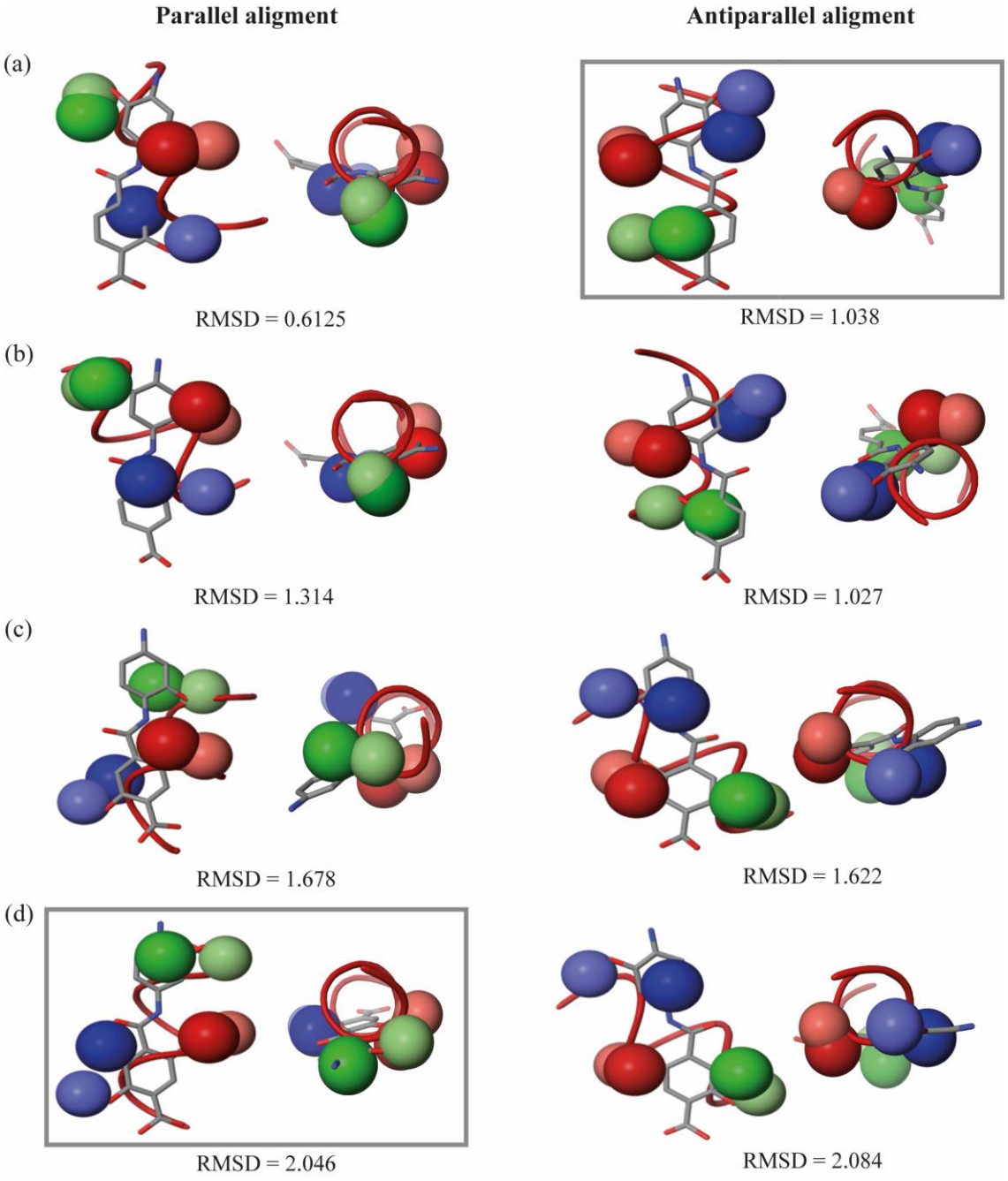


Figure 2.16 Overlay of the second generation of foldamers with the native co-activator peptide. Co-activator residues are in dark colours and helix mimetic residues are in light colours. Parallel (right) and antiparallel (left) alignment with the peptide dipole moment are shown (RMSD values are given for both alignments and the best alignment shown in a box): (a) compound **2.52**; (b) compound **2.53**; (c) compound **2.56**; (d) compound **2.54**.

The results of the molecular modelling analyses for the full set of helix mimetic analogues are summarised in Table 2.2. The best alignment with the native helical peptide and the corresponding RMSD value are shown for each compound. It is worth noting that in some cases the aligment (parallel or antiparallel) with the lowest RMSD value was not chosen as the best overlay, as the backbone of the compounds did not match the helical axis of the peptide.

Table 2.2 Summary of Molecular Modelling Analyses

	alignment[a]	RMSD
2.27	Antiparallel	1.322
2.21	Parallel	1.625
2.28	Antiparallel	2.084
2.52	Antiparallel	1.038
2.53	no good alignment	1.622
2.56	no good alignment	1.027
2.54	Parallel	2.046

[a] where *N* and *C* termini of the benzamide and helix match, they are defined as being parallel and where they oppose, they are defined as being antiparallel.

2.6. Docking studies

To ascertain the extent to which the set of foldamers might act as ERα/co-activator inhibitors, the lowest energy conformations within 1.5 kJ/mol of each compound were docked with the crystal structure of ERα (PDB ID: 2QZO) using the software Glide®. The results from the docking analyses reveal binding poses that display favourable interaction of all the foldamers **2.21**, **2.27**, **2.28** and **2.52** to **2.56** with the co-activator binding groove. Shown in Figure **2.17**a is a good pose for **2.52**; the three hydrophobic side chains of the foldamer occupy the hydrophobic space normally occupied by the co-activator peptide as shown in Figure **2.17**b.

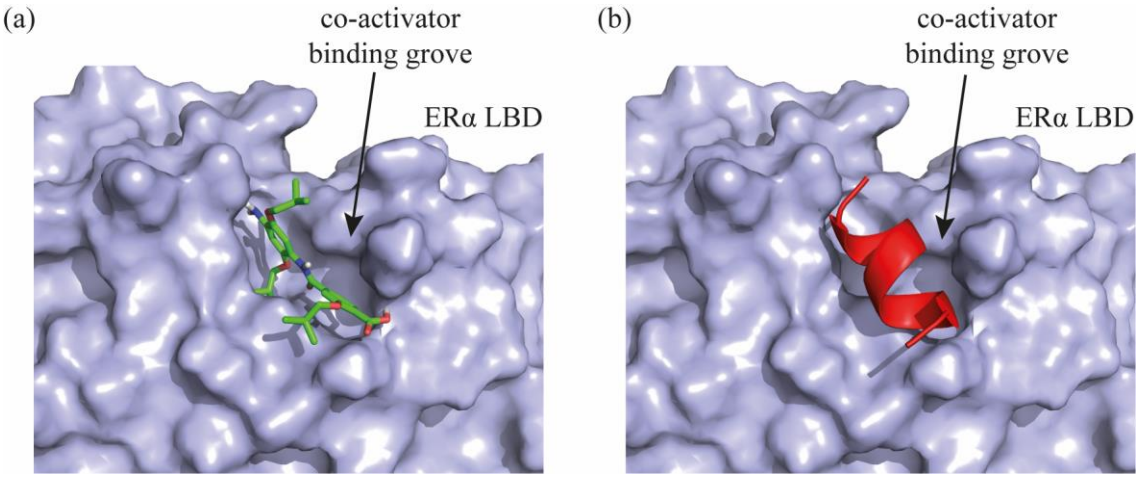


Figure 2.17 (a) Proposed binding mode of compound **2.52** in the ER co-activator binding groove (b) Native co-activator helix in the ER co-activator binding cleft.

Furthermore, the docking studies also show electrostatic interactions for both termini of the foldamer with the ER surface residues; however, in most cases only one of these involves the precise "charge clamp" residues from ER exploited by co-activator ligands (Figure 2.18a). A representative example is shown in Figure 2.18b, where the terminal carboxylic acid and aniline groups of dimer **2.52** are suitably positioned to form electrostatic interactions in the region of the "charge clamp". In particular, these interactions occur between i) the *N* terminus of the foldamer and the Glu^{542} from the ER native "charge clamp"; and ii) the *C* terminus of the foldamer and the Lys^{362} from the ER native "charge clamp", or the Gln^{372}, which is a neighbouring residue. This behaviour is reproduced for the other compounds.

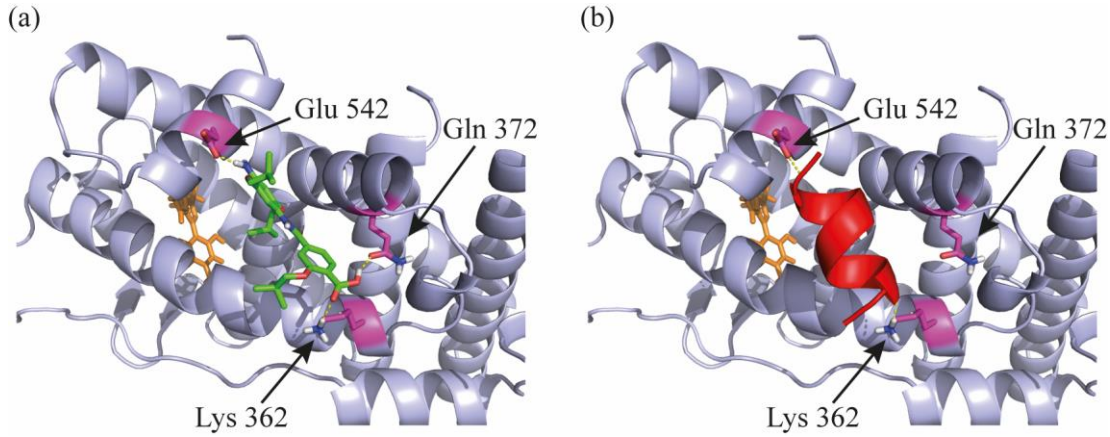


Figure 2.18 (a) Proposed hydrogen bonding interactions between compound **2.52** and ER "charge clamp" residues (b) Hydrogen bonding interactions between the native co-activator and ER "charge clamp" residues.

2.7. Biophysical assays

To perform a preliminary assessment of the ability of our set of helix mimetics to act as PPI inhibitors, we carried out fluorescence polarization competition assays against three nuclear receptor/co-activator interactions (ERα/SrcBox2, ERβ/Src1B2 and RXRα/D22) in the laboratory of Prof. Luc Brunsveld (Technische Universiteit Eindhoven). Brundsveld and co workers[161] recently reported a family of molecules with potential to change their activity as agonist and/or antagonist over small chemical modifications, thus generating opposite effects in the receptor biological functions. Accordingly, our set of proteomimetic compounds were screened in agonistic (binding at the ligand binding pocket) and antagonistic mode (binding at the coactivator binding grove) as shown in Figure 2.19. Unfortunately, our compounds were not sufficiently potent to show a significant effect in these assays.

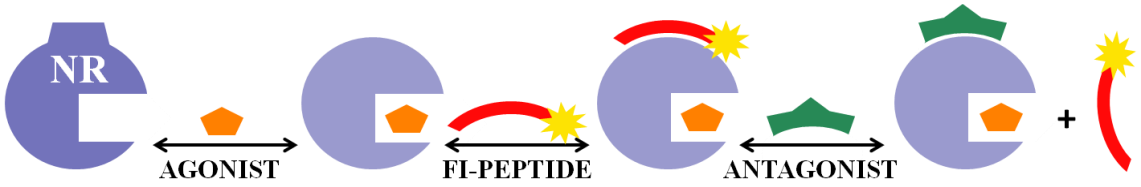


Figure 2.19 Mode of action of the Nuclear Receptor (NR) agonist dependent coactivator peptide recruitment assay.

2.8. Summary and future work

In summary, the design and syntheses of two new bifacial proteomimetic scaffolds based on bis-benzamide and *N*-(4-aminophenyl)terephthalamidic backbones have been described. A complete structural analysis in combination with *in silico* studies revealed that different combinations of monomers leads to a plethora of side chain spatial relationships, which effectively mimic the intended side-chains on an α-helix. Significant conformational knowledge has been gained and added to the already existing data on aromatic oligoamide foldamers. Unfortunately, preliminary evaluation of the new scaffolds against nuclear receptor/co-activator interactions was not able to show binding of our helix mimetics to the protein surface, highlighting the complex relationship between helix mimetic conformation and molecular recognition. Taking into account the flexible nature of the coactivator binding grove in the ER surface, we hypothesise that the rigidity of our scaffolds, introduced by the discussed intramolecular H-bonding (See section 2.4), might be unfavourable for the interaction and final binding to the protein surface.

To study this hypothesis, future work will focus on the assessment of more flexible compounds containing just two side chains and without intramolecular H-bonding constraints. In addition, to increase the binding of our proteomimetics to the protein targets, polar groups (i.e. alkyl amines or alkyl carboxylic acids) will be introduced at the *N* and/or *C* terminus of the scaffold to enhance the electrostatic interactions with the residues of the ER "charged clamp". Finally, the synthesis of libraries bearing different side-chain arrays will be required in order to exploit the potential of the scaffold to target other PPIs containing essential residues in more than one face of an α-helix.

Chapter 3. Optimization of the hybrid oligoamide proteomimetic scaffold

3.1. Introduction

The design and synthesis of proteomimetic scaffolds that target PPIs with high potencies and are amenable to library assembly has advanced significantly in the last few years.[125, 142] However, the next main breakthrough in the field seems to be pointing towards molecules with enhanced pharmacokinetic properties and that permit greater control over target selectivity. Commonly, structural constraints involving covalent or non-covalent interactions have been introduced in the proteomimetic scaffold backbone in order to reproduce the topography of the "hot spot" residues from the native α-helix and thus favour bioactive conformations. However, many studies to date have underlined the complex relationships between molecular rigidity, target plasticity and activity, which all participate in surface recognition processes.[8, 70]

In that context, the Wilson group recently reported the design and synthesis of a hybrid oligobenzamide α-helix mimetic formed by a combination of previously reported aryl building blocks and natural amino acids.[135] This scaffold is capable of reproducing the side chains at *i*, *i*+4 and *i*+7 positions of an α-helix and shows high functional group tolerance combined with a simple synthetic route. In particular, the scaffold was based on modifications of the *O*-alkylated oligobenzamide previously reported by the group (Figure 3.1). In this case the structural rigidity of the oligobenzamide backbone[127] was broken through substitution of the central aryl-unit with an α-amino acid residue.

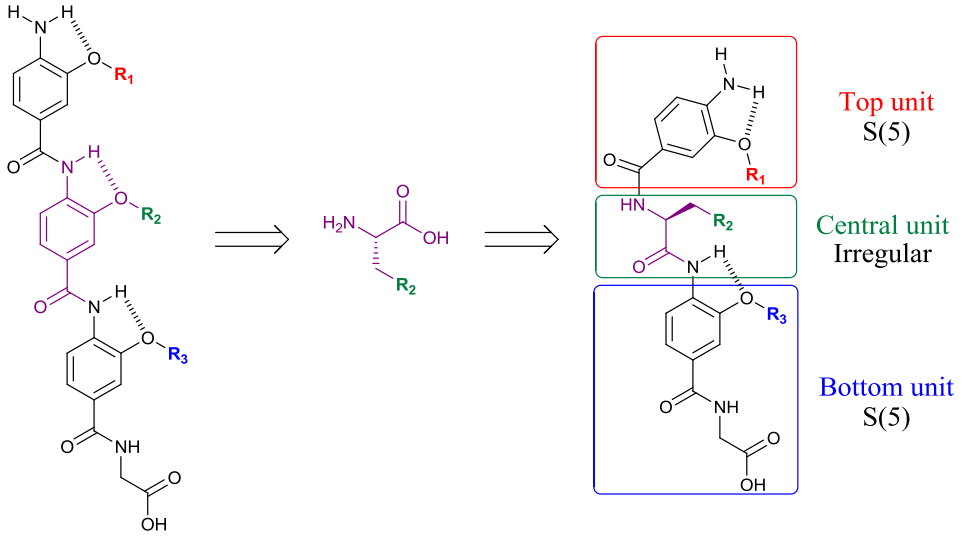


Figure 3.1 Design of the hybrid oligobenzamide α-helix mimetic by modification of the 3-*O*-alkylated oligobenzamide scaffold.

This modification generated a different intramolecular hydrogen bonding arrangement with hydrogen bonds at the top and bottom of the sequence, which maintain the potential to adopt well-defined conformations. Importantly, the irregular nature of the backbone increased the degree of flexibility and gave access to a wider conformational space, thus perhaps facilitating an induce-fit type of interaction with the target protein. This fact was illustrated by early studies where the energy of the different scaffold structures **3.1** and **3.2** was minimised in Macromodel and the conformers within 1.5 kJ/mol were superimposed without further manipulation (Figure **3.2**). These simulations also highlighted the increased conformational plasticity of the new hybrid scaffold over the original 3-*O*-alkylated oligobenzamide scaffold.[135] Importantly, previous work also discovered that switching the side chain of the bottom aryl unit from the 3-*O* position to the 2-*O* position, as in compound **3.2**, increases the binding affinity of the scaffold to the target proteins, in most cases.[135, 136] Therefore, all the work of this chapter focusses on hybrid compounds incorporating these modifications.

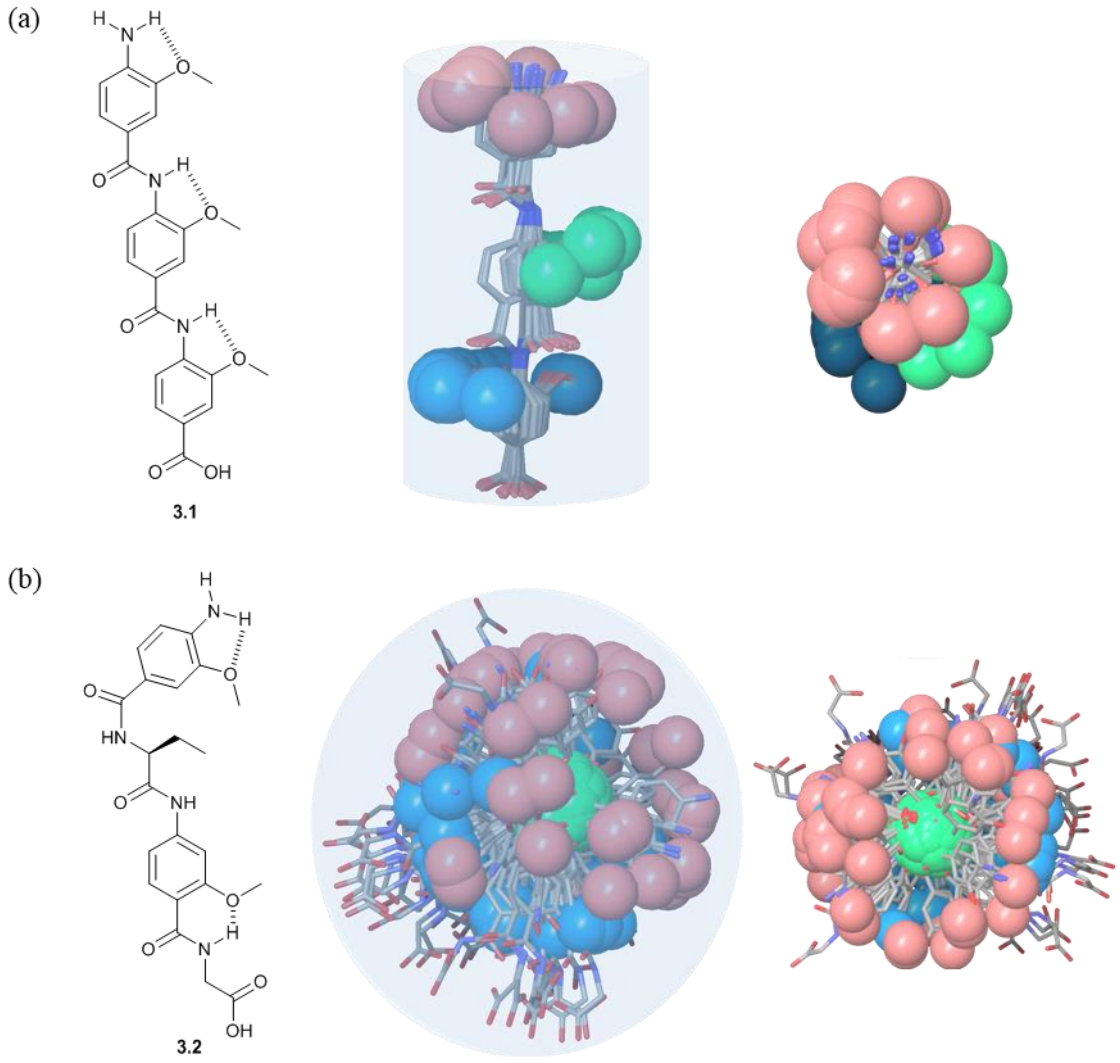


Figure 3.2 Illustrations of the accessible conformational space (shown as a shaded 3D object) highlighting the orientation of the side-chains[135] (shown in CPK format): (a) Structure of the 3-*O*-alkylated trimer model **3.1** and corresponding side (top) and top (bottom) view; (b) Structure

of the hybrid trimer model **3.2** and corresponding side (top) and top (bottom) view. Energy minimisation by Macromodel and superimposition of the conformers within 1.5 kJ/mol.

Furthermore, SAR[136] studies identified μM inhibitors of the p53/*h*DM2 and Mcl-1/NOXA B interactions, which are highly involved in cancer development. Nevertheless, the most important feature displayed by the hybrid oligobenzamide α-helix mimetic was the unprecedented stereodependent selectivity imported by the chiral nature of the scaffolds. In particular, the substitution of the central L-amino acid by its D-enantiomer switched the selectivity of the scaffold from *h*DM2 only to both *h*DM2 and Mcl-1 proteins. This attribute permitted the synthesis of chiral structures that can be tuned to achieve enantioselective recognition by the protein partner.

3.2. Interactions of interest

3.2.1. p53/*h*DM2-*h*DMX

The p53 tumour suppressor stimulates the intrinsic apoptosis pathway, in response to severe cellular stresses, such as DNA damage or hypoxia. Therefore, its activity is vital for the maintenance of the genomic integrity of the cell. Inactivation of p53 occurs due to mutations in over half of all cancer cases; whilst in most of the remaining cases it is mediated by genomic amplification of the *h*DM2 oncoprotein (Figure 3.3). In normal cells, *h*DM2 down-regulates p53, which modulates its growth-suppressing activity. Alternatively, when DNA is damaged, p53 levels increase and activate the expression of *h*DM2. The *h*DM2 protein can, in turn, bind to the transactivation domain of p53 and inhibit further activity of p53 as a transcription factor. In tumours, gene amplification and other processes can lead to *h*DM2 amplification and consequently p53 inhibition.[14, 178]

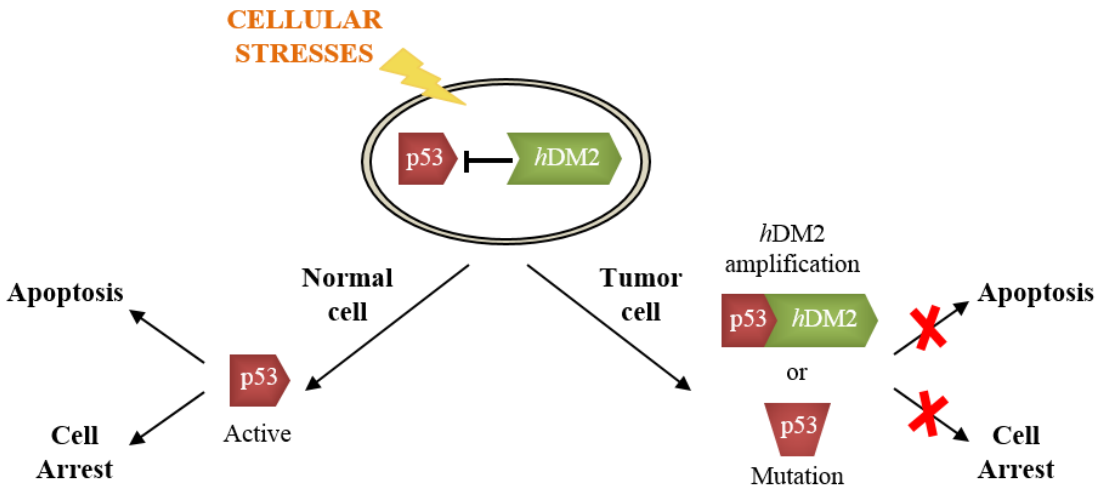


Figure 3.3 Diagram of the apoptotic pathway regulated by p53 in normal and tumor cells.

Overall structural features

The interacting region of *m*DM2 (*h*DM2 homolog protein from mouse) consists of a structural domain located in its *N*-terminal part. Alternatively, the p53 recognition motif involves a short, linear sequence of 11 amino acids (residues 17 to 27), which comprises one of the conserved regions of p53 and contains sequences responsible for transactivation (Figure 3.4).[178]

The *m*DM2 domain is divided in two structurally similar parts with low sequence similarity. The two halves form a small globular structure with a hydrophobic core. When the two repeats are joined across their hydrophobic side, they form a cleft at their interface. The cleft is about 25 Å long, 10 Å wide near the surface but narrowing toward the bottom, and up to 10 Å deep. It is asymmetric and is composed of two helices forming the sides, two short helices making the bottom and a pair of three-stranded β-sheets capping each end. The p53 peptide forms an amphipathic α-helix of about 2.5 turns, which is followed by an extended region of three residues. The α-helix has a hydrophobic face formed by three hydrophobic key amino acids (Phe^{19}, Trp^{23}, Leu^{26}), at *i*, *i*+4 and *i*+7 positions, that interact with the protein cleft through a sequence of van der Waals contacts.[178]

The p53/*m*DM2 interface has an area of 1498 $Å^2$ and is mostly hydrophobic in nature. The position of the p53 helix allows Phe^{19}, Trp^{23} and Leu^{26} residues to sit deep inside the *m*DM2 cleft in a complementary fashion. The van der Waals contacts at the interface are augmented only by two intermolecular hydrogen bonds. One occurs between the Phe^{19} backbone amide of

p53 and the Gln[72] side chain of *m*DM2 at the entrance of the cleft; the second is between the p53 Trp[23] indole group and the *m*DM2 Leu[54] carbonyl, deep inside the cleft.[178]

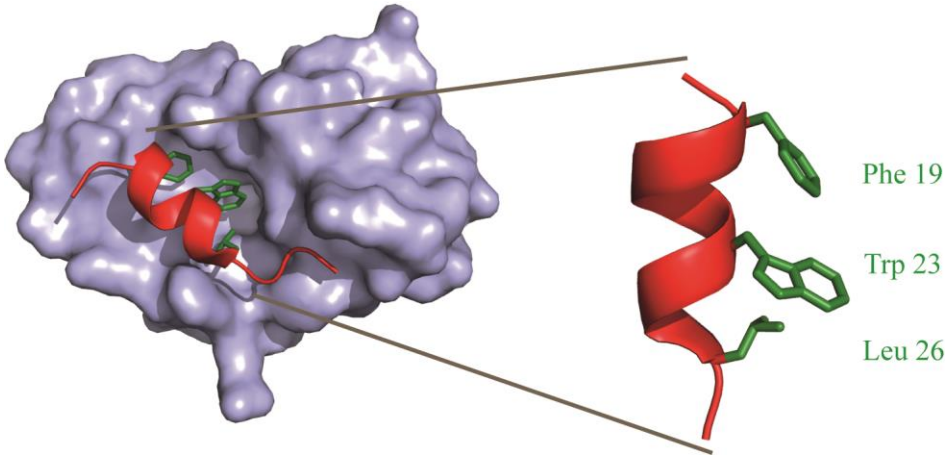


Figure 3.4 Crystal structure of p53/*m*DM2 PPI (PDB ID: 1YCR). Key side chains on the binding surface of the helix are highlighted.

3.2.2. Bcl-2 family

The members of the B-cell lymphoma 2 (Bcl-2) family play a central role as regulators of apoptotic cell death, in response to a wide variety of stimuli (Figure 3.5). These molecules can be combined with themselves or other family members to form homo-dimers and hetero-dimers, and produce several pro-apoptotic and/or anti-apoptotic entities. For example, Bcl-2, Bcl-x_L and Mcl-1 inhibit programmed cell death, and Bak and Bax (Bcl-2 homologous antagonist killers) can promote apoptosis.[179] Consistently, members that inhibit apoptosis are over expressed in many cancers and contribute to tumour initiation, progression and resistance to therapy.[62]

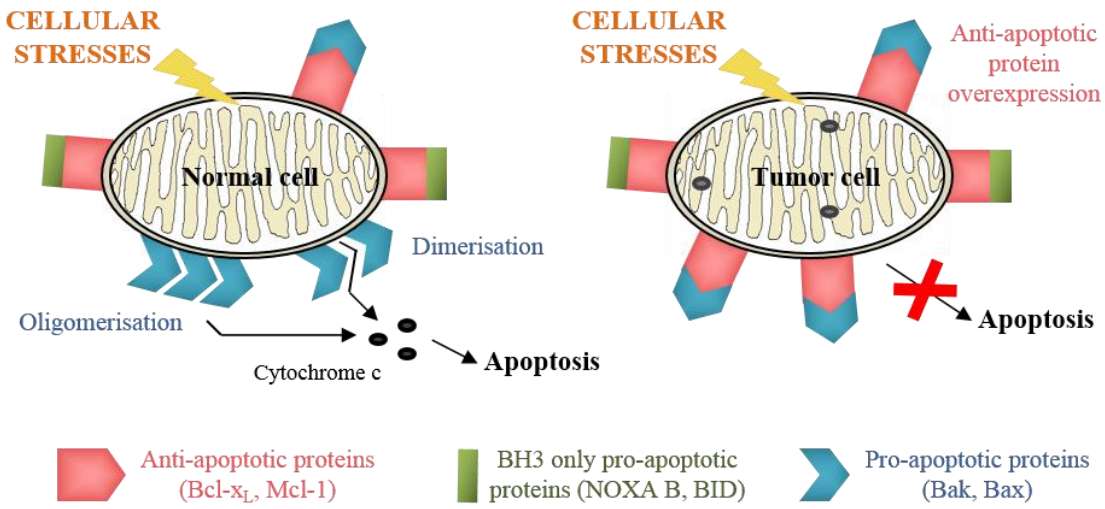


Figure 3.5 Diagram of the apoptotic pathway regulated by Bcl-2 family members in the mitochondria in normal and tumor cells.

Overall structural features

All the Bcl-2 family members are formed by up to four conserved Bcl-2 homology (BH) domains, all of which include α-helical segments (Table 3.1). Anti-apoptotic proteins exhibit conservation of the sequence in all BH domains. Conversely, pro-apoptotic proteins are divided into multi-domain members that are formed by the BH1, BH2 and BH3 domains, such as Bax and Bak; and BH3-only members, such as Bid and Bad. There is a common region in the BH3 α-helical domain, which binds to the hydrophobic groove formed by the structural connection of BH1, BH2 and BH3 domains of anti-apoptotic multidomain members. This common helical segment is necessary for the promotion of cell death.[46]

Table 3.1 Members of the Bcl-2 family[180, 181]

Multidomain Anti-apoptotic proteins	**Multidomain Pro-apoptotic proteins**	**BH3-only Pro-apoptotic proteins**
Bcl-2	Bak	BID
Bcl-x_L	Bax	BIM
Mcl-1		NOXA B
Bcl-w		PUMA
		BAD

The structural insights of the recognition process between these protein families were first reported by Fesik and co-workers with the structure of the Bcl-x_L/Bak peptide complex (Figure 3.6).[179] The structure of the Bcl-x_L protein consists of two central hydrophobic α-helices surrounded by five amphipathic helices. The Bak peptide binds in a hydrophobic cleft formed by the BH1, BH2 and BH3 domains of Bcl-x_L, where the *N*-terminal residues interact with the BH1 region, whereas the *C*-terminal end interacts mostly with the BH2 and BH3 domains. In the Bak helix, the hydrophobic face is projected into the hydrophobic cleft, stabilizing the complex, whilst the charged side chains are close to oppositely charged residues of the Bcl-x_L protein.[179]

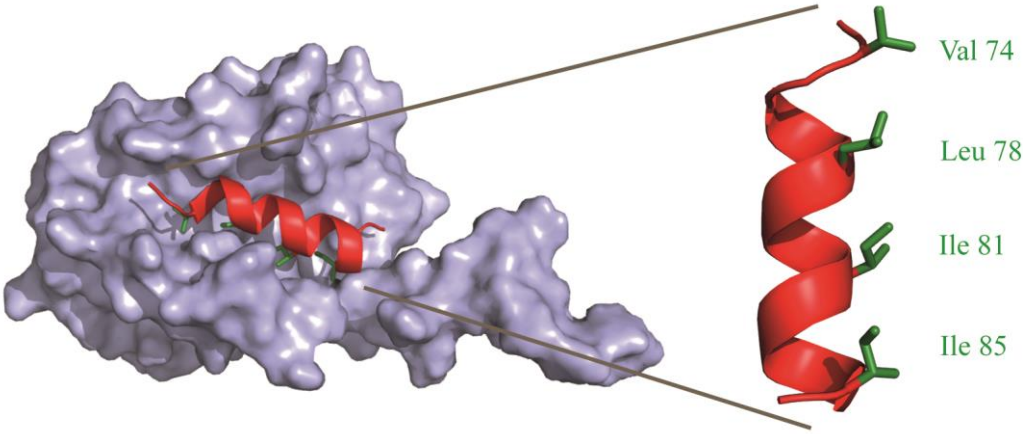


Figure 3.6 Crystal structure of Bcl-x_L/Bak PPI (PDB ID: 1BXL). Key side chains on the binding surface of the helix are highlighted.

Another member of the Bcl-2 family is Mcl-1, which has a central and non-redundant role in the maintenance of progenitor and stem cells (Figure 3.7). Its overexpression has been related to the development of a variety of resistant cancers, including multiple myeloma, acute myeloid leukemia, melanoma and poor-prognosis breast cancer.[34] Mcl-1 is neutralized by the BH3-only proteins NOXA, Puma, Bim and Bak.[34, 63] Certain residues within the Mcl-1 protein sequence are determinant for its binding activity; the conserved amino acids Leu^{213}, Arg^{214}, Gly^{217} and Asp^{218}, which are shared among many BH3 domains, and the discrete residue Val^{220}, which is responsible for the selectivity (Fig. 10).[63]

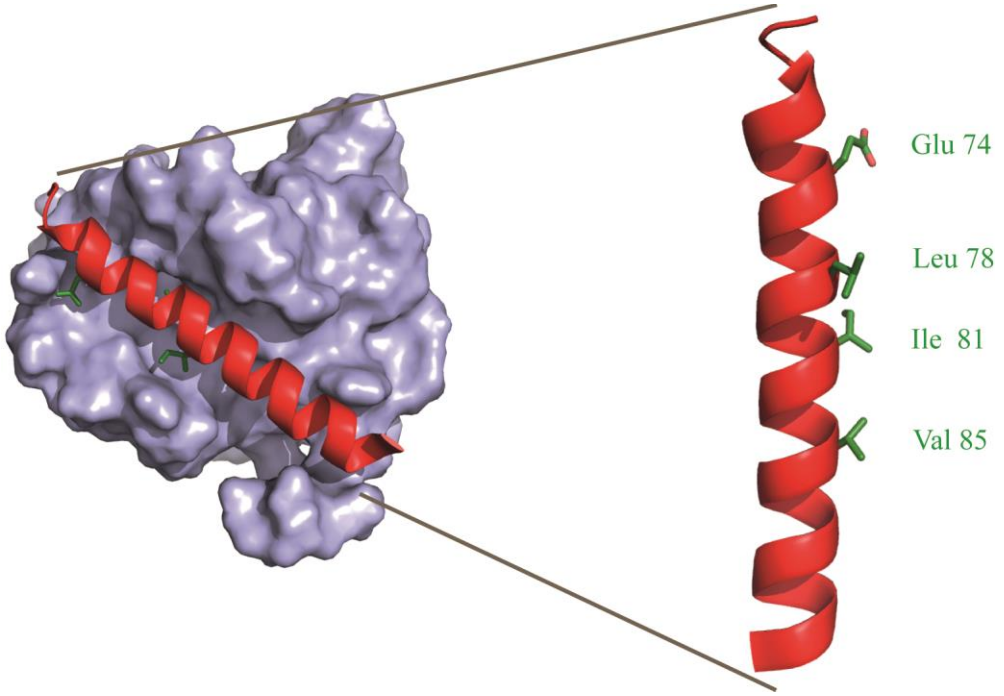


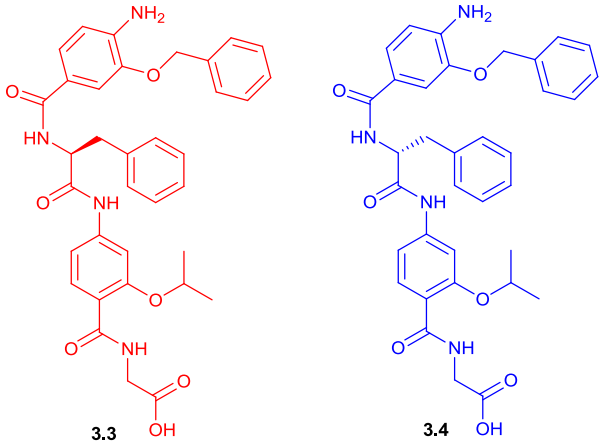
Figure 3.7 Crystal structure of Mcl-1/NOXA B PPI (PDB ID: 2JM6). Key side chains on the binding surface of the helix are highlighted.

3.3. Description of the scaffold

The results obtained using the hybrid oligobenzamide α-helix mimetic represented a good starting point to further investigate the rules that govern molecular recognition and work into establishing guidelines for the synthesis of functional proteomimetics. Therefore, in this chapter an attempt to further study and optimize the hybrid oligobenzamide scaffold towards more potent and selective inhibitors of PPIs is presented.

In particular, compound **3.3** (Figure 3.8) was chosen as initial template for the optimization process for the following reasons:

- It was one of the most potent analogues of the family against p53/*h*DM2 with an IC_{50} of 11.9 μM.
- It was selective for *h*DM2 over Mcl-1 (IC_{50} >100 μM).
- Selectivity for Mcl-1 could be tuned by changing the stereochemistry of the central aa: L- (IC_{50} >100 μM) compound **3.3**, whilst D- (IC_{50} 27.1 μM) compound **3.4**.
- It was formed using some of the most accessible building blocks.



	3.3	**3.4**
p53/*h*DM$_2$ IC_{50}	**11.9 ± 0.6 μM**	**9.2 ± 0.4 μM**
Mcl-1/NOXA B IC_{50}	**>100 μM**	**27.1 ± 1.1 μM**

Figure 3.8 Structures and inhibitory activity against p53/*h*DM2 and Mcl-1/NOXA B of hybrids **3.3** (L-Phe) and **3.4** (D-Phe).

Subsequently, five modification sites (Figure 3.9) were selected to introduce a series of chemical changes into the structure: the *N* terminus, the top bottom aryl building blocks, the central amino acid, and the *C* terminal amino acid. Each manipulation was selected to change the size, shape, flexibility and/or solubility of the compounds, which should lead to a better understanding of the SAR of these molecules.

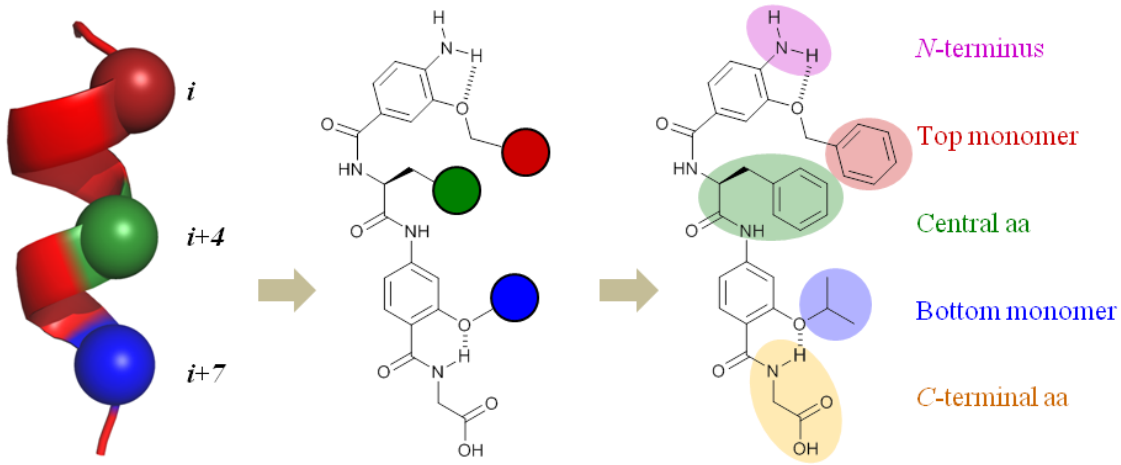


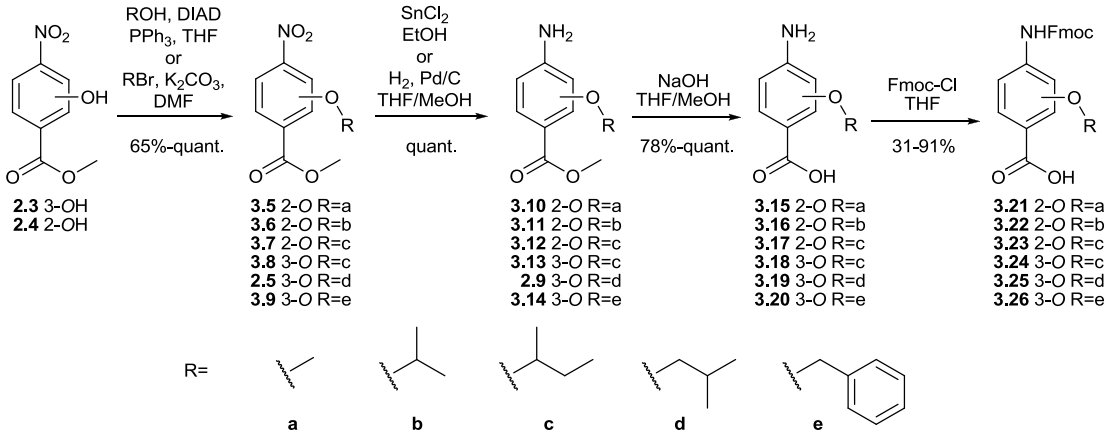
Figure 3.9 Scaffold optimization process schematics. Modification sites of compound **3.3** are highlighted.

3.4. Synthesis of the hybrid α-helix mimetic scaffold

The synthesis of the hybrid α-helix mimetic scaffold was based on an Fmoc (9-fluorenylmethyloxycarbonyl) solid phase peptide synthesis (SPPS) strategy to enable library generation and reduce the preparation time.[135]

3.4.1. Monomer synthesis

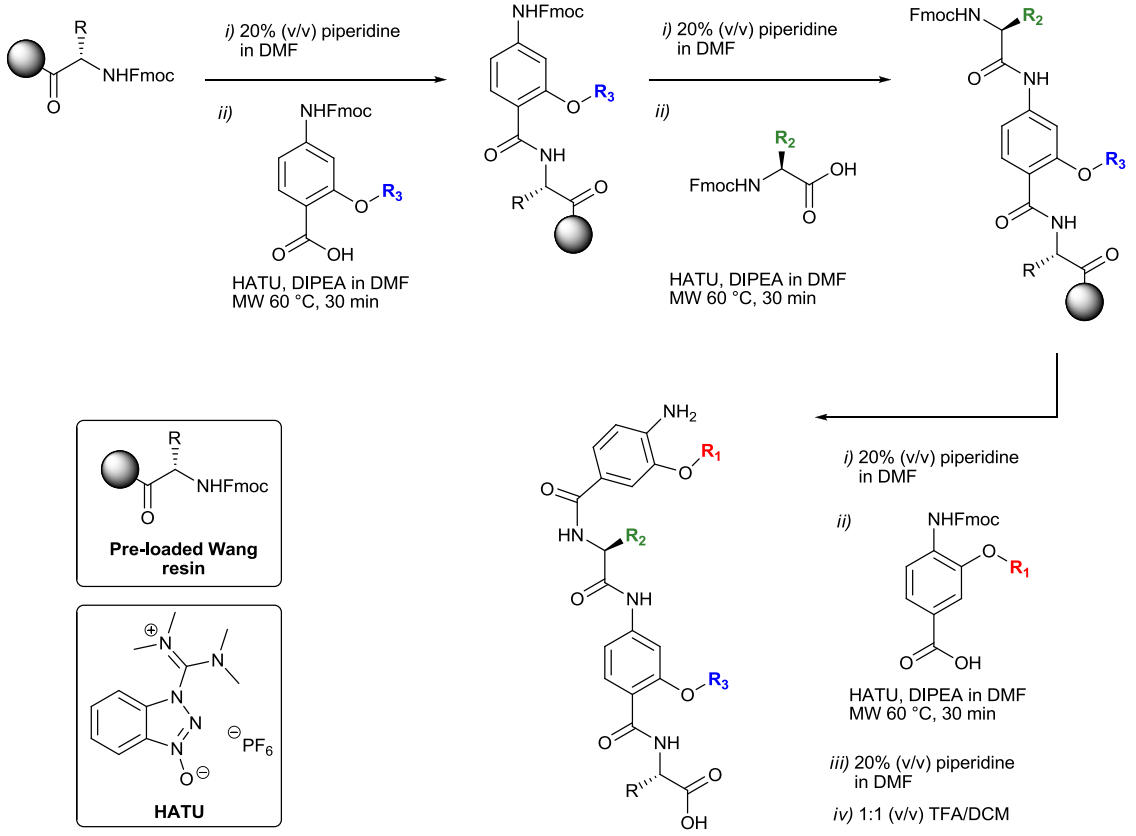
The synthesis of the 2-*O*-alkylated and the 3-*O*-alkylated Fmoc-protected monomers followed a synthetic route previously described by the group (Scheme 3.1).[125, 127] It consisted of an initial alkylation step, followed by reduction of the nitro to the amino group. Subsequent hydrolysis of the ester and final Fmoc-protection produced the monomers in good yields.



Scheme 3.1 Synthetic route to 2-*O* and 3-*O* Fmoc-protected building blocks

3.4.2. General synthetic scheme for oligomers

The hybrid α-helix mimetics were synthesised using an adapted version of the microwave-assisted automated SPPS (Scheme 3.2) originally developed by Dr V. Azzarito.[135] It consisted of a series of deprotection and coupling steps, using Fmoc-aa- (where aa can be any amino acid) preloaded Wang resins as solid support and HATU as a coupling reagent for the protected monomers.



Scheme 3.2 Solid phase synthetic route for the hybrid α-helix mimetics.

3.4.3. Side chain diversification

The methodology described above (Scheme 3.2) was used to synthesise a library of hybrid compounds (**3.27** to **3.48**). The design behind the multitude of modifications that were introduced is detailed in the following sections 3.4.3.1 to 3.4.3.5.

3.4.3.1. Modification of the top aryl unit

The benzyl substituent of compound **3.3** mimicked the original phenylalanine residue from the p53 α-helix. The modifications incorporated here (Table 3.2, Figure 3.10) mainly focused on studying the effect of:

- *para*- substituted phenyl rings to improve interactions with the Phe hydrophobic binding pocket, compounds **3.27** to **3.29**.
- Extended aromatic ring systems to improve interactions with the Phe hydrophobic binding pocket, compound **3.30**.
- Aliphatic side chains to investigate the effect of different hydrophobic groups, compounds **3.31** to **3.33**.

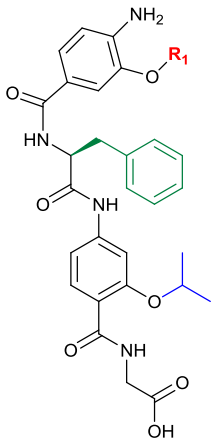


Table 3.2 Library of hybrid α-helix mimetics with modifications on the top aryl unit illustrating side-chain sequence.

Hybrid	R_1	aa	R_3	resin
3.27	3-*O*-*p*-Cl-Bn	L-Phe	2-*O*-*i*Pr	Gly
3.28	3-*O*-*p*-CF_3-Bn	L-Phe	2-*O*-*i*Pr	Gly
3.29	3-*O*-*p*-*t*Bu-Bn	L-Phe	2-*O*-*i*Pr	Gly
3.30	3-*O*-2-Nph	L-Phe	2-*O*-*i*Pr	Gly
3.31	3-*O*-*s*Bu	L-Phe	2-*O*-*i*Pr	Gly
3.32	3-*O*-(MeS-3-Pr)	L-Phe	2-*O*-*i*Pr	Gly
3.33	3-*O*-Methylcyclopropane	L-Phe	2-*O*-*i*Pr	Gly

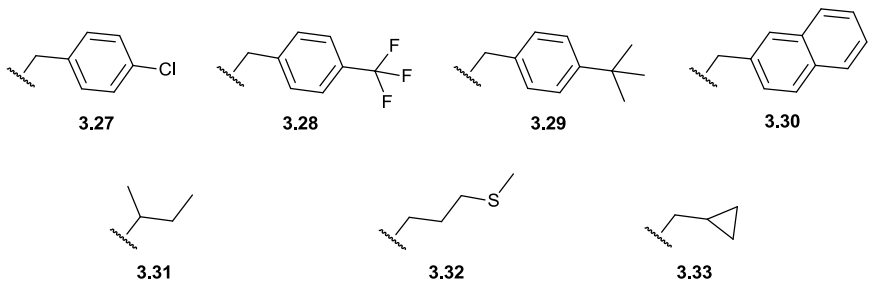


Figure 3.10 Side chains incorporated in the top aryl unit of the hybrid proteomimetics. The 3-*O*-alkylated monomers used in the synthesis of the oligomers **3.27** to **3.33** were provided by Dr N. S. Murphy.

3.4.3.2. Modification of the central amino acid

The Phe aa in compound **3.3** mimicked the original Trp residue from the p53 α-helix. The modifications incorporated here (Table 3.3, Figure 3.11) mainly focused on studying the effect of:

- Extended aromatic ring systems to improve interactions with the deep Trp hydrophobic binding pocket, compounds **3.34** and **3.35**.
- Different aa enantiomers to exploit further the stereodependent selectivity of the scaffold, compounds **3.34** and **3.35**.

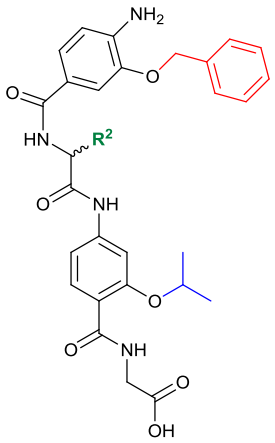
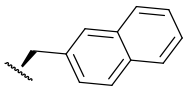
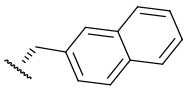


Table 3.3 Library of hybrid α-helix mimetics with modifications on the central aa illustrating side-chain sequence.

Hybrid	R_1	aa	R_3	resin
3.34	3-*O*-Bn	L-2-Nal	2-*O*-*i*Pr	Gly
3.35	3-*O*-Bn	D-2-Nal	2-*O*-*i*Pr	Gly



3.34



3.35

Figure 3.11 Aa side chains incorporated in the central position of the hybrid proteomimetics.

Structural modifications were also explored by the introduction of the following building blocks in the central position of the hybrid scaffold (Table 3.4, Figure 3.12):

- α,α'-Disubstituted amino acids[182] **3.37** and **3.38** to study the restriction of conformational freedom of side chains and the introduction of a complementary side chain.
- β-Amino acids[183] **3.39** and **3.40** to study the access to a new group of different pharmacophores with distinct flexibility on the central backbone.
- *N*-methylated Phe to study the effect of potential new bioactive conformations, compound **3.36**.

Table 3.4 Hybrid α-helix mimetic with structural modifications on the central aa illustrating side-chain sequence.

Hybrid	R_1	aa	R_3	resin
3.36	3-*O*-Bn	*N*-Me-Phe	2-*O*-*i*Pr	Gly

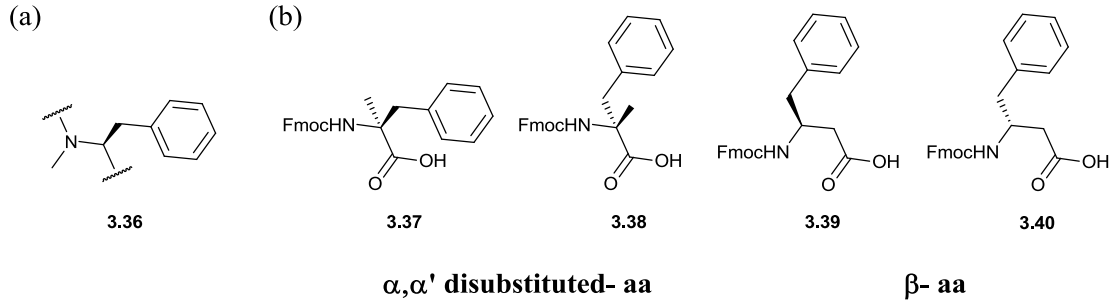
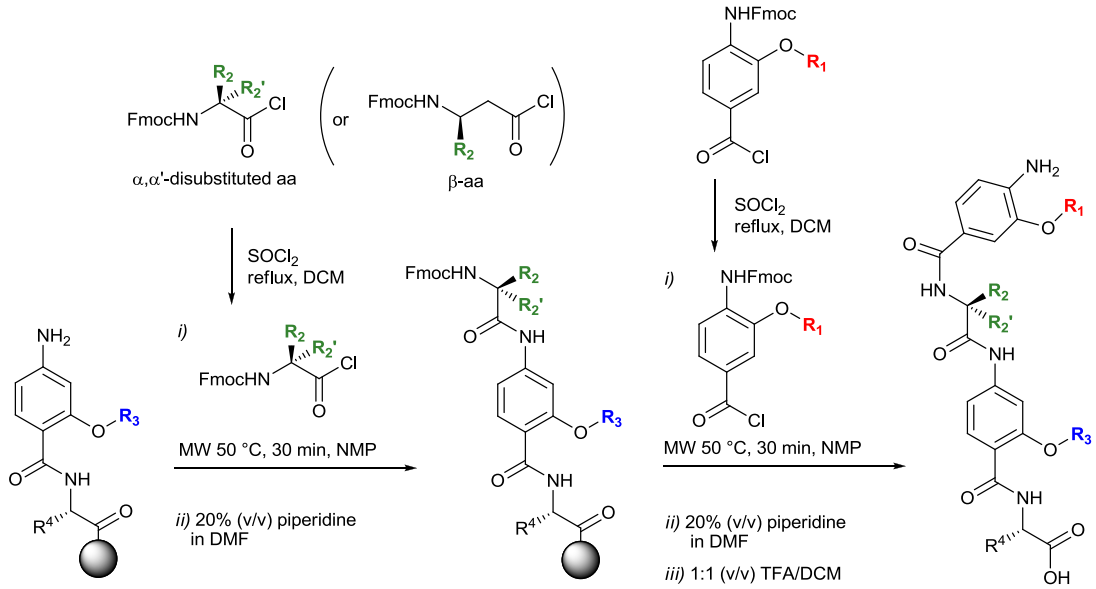


Figure 3.12 (a) *N*-Me aa incorporated in the central position of the hybrid proteomimetic; (b) Different α,α'-disubstituted and β-amino acid residues which failed to be incorporated in the central position of the hybrid oligomers.

The incorporation of the aa residues **3.37** to **3.40**, which were used as test substrates (Figure 3.12), and the aryl building block **3.26** proved unsuccessful using the standard synthetic methodology for the hybrids (Scheme 3.2). In order to increase their reactivity the acid chlorides from the corresponding monomers were pre-formed and then coupled to the growing chain (Scheme 3.3). Unfortunately, this methodology only permitted incomplete coupling of the monomers, even after repeated coupling cycles. The poor reactivity of these monomers might be due to the steric hindrance of the α,α'-disubstituted amino acids. Alternatively, the β-amino acid residues may have adopted unreactive conformations where the free amine was inaccessible for coupling, as indicated by subsequent ineffective acetylations attempts using acetic anhydride.



Scheme 3.3 Acid chloride based methodology for the incorporation of challenging building blocks.

3.4.3.3. Modification of the bottom aryl unit

The *i*Pr side chain in compound **3.3** mimicked the original Leu residue from the p53 α-helix. The modifications incorporated here (Table 3.5, Figure 3.13) mainly focused on studying the effect of:

- Different aliphatic side chains to assess the effect of size and shape of the aliphatic substituents in the Leu binding site, compounds **3.41** to **3.43**.

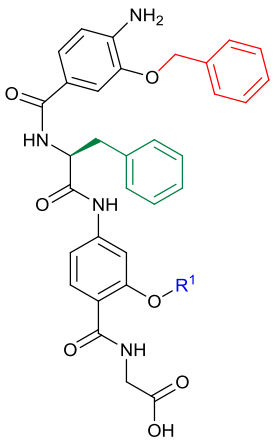


Table 3.5 Library of hybrid α-helix mimetics with modifications on the bottom aryl unit illustrating side-chain sequence.

Hybrid	R_1	aa	R_3	resin
3.41	3-*O*-Bn	L-Phe	2-*O*-*i*Bu	Gly
3.42	3-*O*-Bn	L-Phe	2-*O*-*s*Bu	Gly
3.43	3-*O*-Bn	L-Phe	2-*O*-Me	Gly

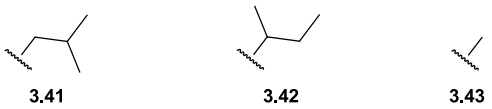


Figure 3.13 Side chains incorporated in the bottom aryl unit of the hybrid proteomimetics. The 2-*O*-alkylated monomer used in the synthesis of the oligomer **3.41** was provided by Dr V. Azzarito.

3.4.3.4. Modification of the *N*-terminus

One of the main disadvantages of this type of proteomimetics is their limited solubility in aqueous media.[126] In order to address this issue, solubilising groups could be attached onto the scaffold; however, the site of modification had to be chosen carefully as it could disrupt the activity of the compound. The *N*-terminus of the molecule was anticipated to allow further modification of the scaffold without altering the key binding side chains. Furthermore, the intra-molecular hydrogen bond between the *N*H and the alkoxy group of the top monomer would position the *N*-linked side chain towards the solvent exposed face of the compounds (Figure 3.14).

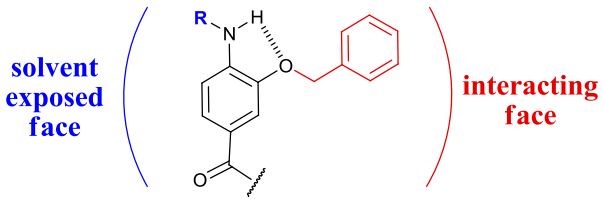
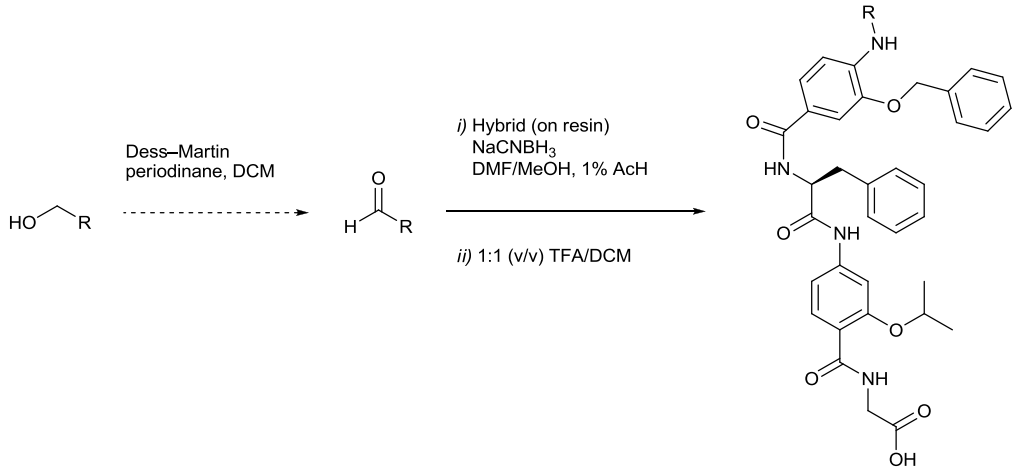


Figure 3.14 Possible projection of the *N*-linked side chain towards the solvent exposed face of the hybrid compounds.

Accordingly, a new modification strategy was developed based on a reductive amination reaction on resin (Scheme 3.4). The reaction provided the modified hybrid in good yields. An *in*

situ oxidation of the alcohol to the corresponding aldehyde was performed when necessary. The advantages of this reaction are: i) reduction of preparation time and potential for library generation resulting from the nature of a late stage derivatization; ii) solid phase chemistry permits the use of a large excess of reagents, which results in high modification yields.

The modifications incorporated at the *N*-terminus (Table 3.6, Figure 3.15) are mainly focused on the introduction of polar side chains to increase the water solubility of the analogues.



Scheme 3.4 Reductive amination on resin of the hybrids *N*-terminus.

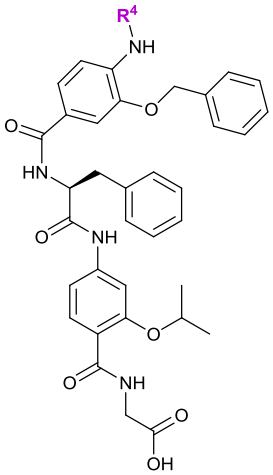
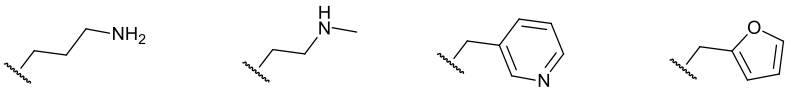


Table 3.6 Library of hybrid α-helix mimetics with modifications on the *N*-terminus illustrating side-chain sequence.

Hybrid	R_1	aa	R_3	R_4	resin
3.44	3-*O*-Bn	L-Phe	2-*O*-*i*Pr	4-*N*-3-NH_2-Pr	Gly
3.45	3-*O*-Bn	L-Phe	2-*O*-*i*Pr	4-*N*-2-MeNH-Et	Gly
3.46	3-*O*-Bn	L-Phe	2-*O*-*i*Pr	4-*N*-3-Pyridyl	Gly
3.47	3-*O*-Bn	L-Phe	2-*O*-*i*Pr	4-*N*-2-Furanyl	Gly



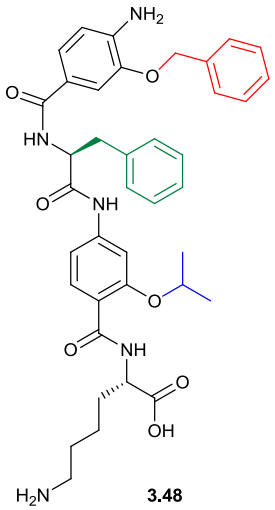
3.44 3.45 3.46 3.47

Figure 3.15 Side chains incorporated in the *N*-terminus of the hybrid protomimetics.

The *N*-terminal modification of compounds **3.44** and **3.45** was performed with Boc protected building blocks (Scheme 3.4): the Boc protected amino alcohols were oxidised to the corresponding Boc protected amino aldehydes, before being used in the reductive amination reaction. The subsequent cleavage and deprotection step in acidic conditions removed the Boc group to yield the final compounds **3.44** and **3.45**.

3.4.3.5. Modification of the *C*-terminal amino acid

The aa at the *C*-terminal position of the hybrid compounds was also modified in order to increase the solubility of the proteomimetics in aqueous media. In particular, the standard glycine residue was substituted by a lysine aa in compound **3.48**, introducing an extra polar amino group into the scaffold (Table 3.7).



3.48

Table 3.7 Hybrid α-helix mimetics with modifications on the *C*-terminal aa illustrating side-chain sequence.

Hybrid	R_1	aa	R_3	resin
3.48	3-*O*-Bn	L-Phe	2-*O*-*i*Pr	Lys

3.5. Biophysical testing: Fluorescence Anisotropy Competition Assays

The use of fluorescence anisotropy (FA) to determine binding constants is well established within the Wilson group.[123, 124, 132, 134, 135] Consequently, to test the potential of the new hybrid analogues to selectively inhibit the p53/*h*DM2 and Mcl-1/NOXA B PPIs, their activity was tested in a FA competition assay targeting those PPIs, as shown in Figure 3.16.

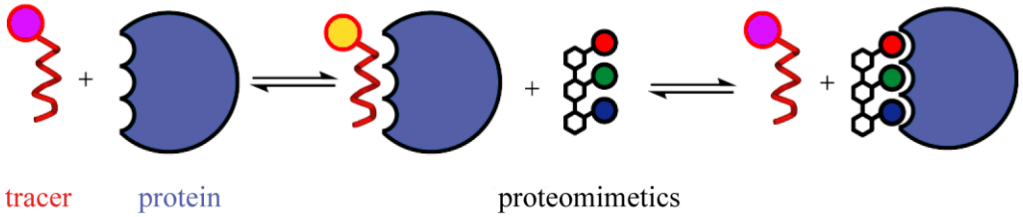


Figure 3.16 Mode of action of the fluorescence anisotropy competition assay.

In the FA competition assay the protein in complex with the corresponding peptide tracer, a binding peptide bearing a fluorophore group, is titrated with an increasing concentration of proteomimetic compound. Initially, the peptide tracer is bound to the protein and tumbles slowly in solution due to the big size of the peptide-protein complex, which produces a high anisotropy signal. This interaction can be disrupted by the binding of the proteomimetic compounds to the protein, which displace the peptide tracer from the protein surface to the solution. This process results in a decrease of anisotropy due to the fast tumble of the peptide now in solution.

It is worth noting that this assay was used to determine the binding affinities of our benchmark compound **3.3**, which displayed an IC_{50} of 11.9 μM for p53/*h*DM2 and >100 μM for Mcl-1/NOXA B (shown in Figure 3.8).

3.5.1. Modification of the top aryl unit

The FA competition assay targeting the p53/*h*DM2 and Mcl-1/NOXA B PPI for the top aryl modified series of hybrids and the corresponding IC_{50} values are shown in Figure 3.17.

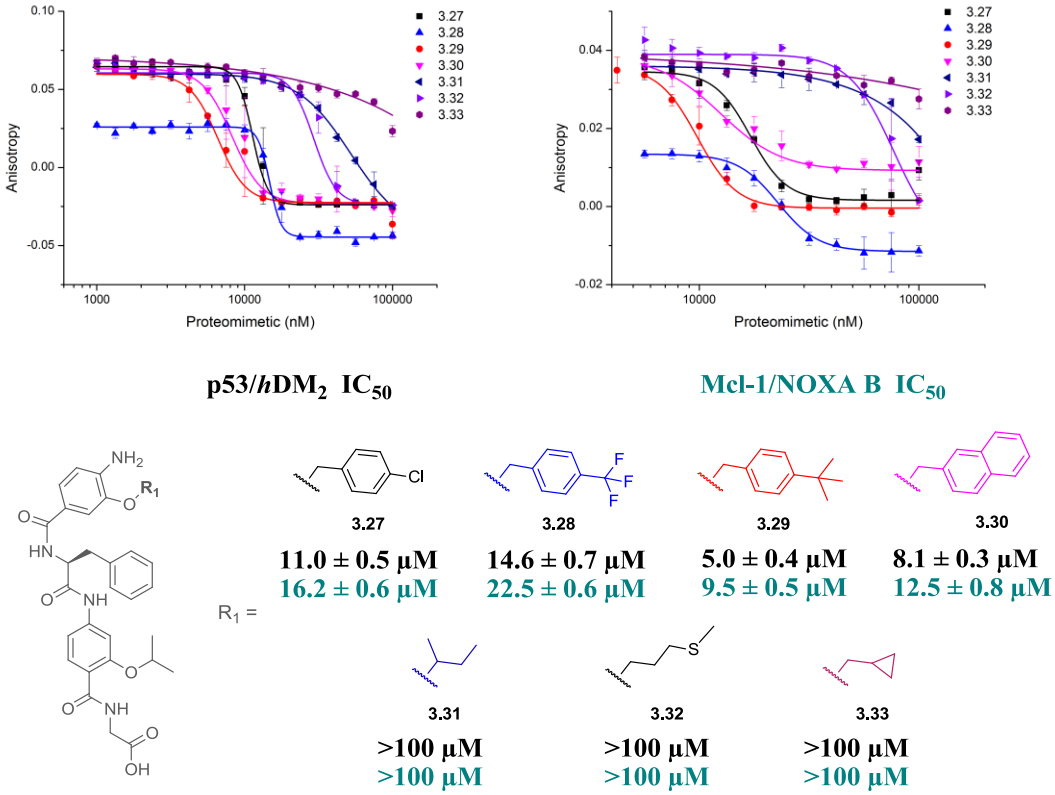


Figure 3.17 Top aryl modifications series: Dose-response curves against the p53/*h*DM2 (left) and Mcl-1/NOXA B (right) PPI (40 mM phosphate buffer pH 7.50, 200 mM sodium chloride, 0.02 mg mL^{-1} BSA).

The results illustrate that the *para* substitution of the benzene ring with -Cl and -CF_3 groups in compounds **3.27** (11.0 ± 0.5 µM) and **3.28** (14.6 ± 0.7 µM) respectively, and the replacement by a naphtyl side chain in compound **3.30** (8.0 ± 0.3 µM) did not have a significant effect on the binding affinity for *h*DM2. However, a slight improvement on potency of two fold was achieved with the introduction of a *t*Bu group in the *para*-Bn position in compound **3.29**, resulting in the most potent hybrid obtained to date against *h*DM2 with an IC_{50} of 5.0 ± 0.4 µM.

Unfortunately, all the changes resulted in compounds with binding affinities for Mcl-1 throughout the low µM range, thus eliminating the selectivity towards *h*DM2 from the original compound **3.3**. Furthermore, the substitution of the aromatic system by aliphatic side chains, compounds **3.31** to **3.33**, resulted in a decrease of binding affinity to >100 µM for both *h*DM2 and Mcl-1 proteins. This result highlights the importance of the aromatic ring in the top position.

3.5.2. Modification of the central amino acid

The FA competition assay targeting the p53/*h*DM2 and Mcl-1/NOXA B PPI for the central aa modified series of hybrids and the corresponding IC_{50} values are shown in Figure 3.18.

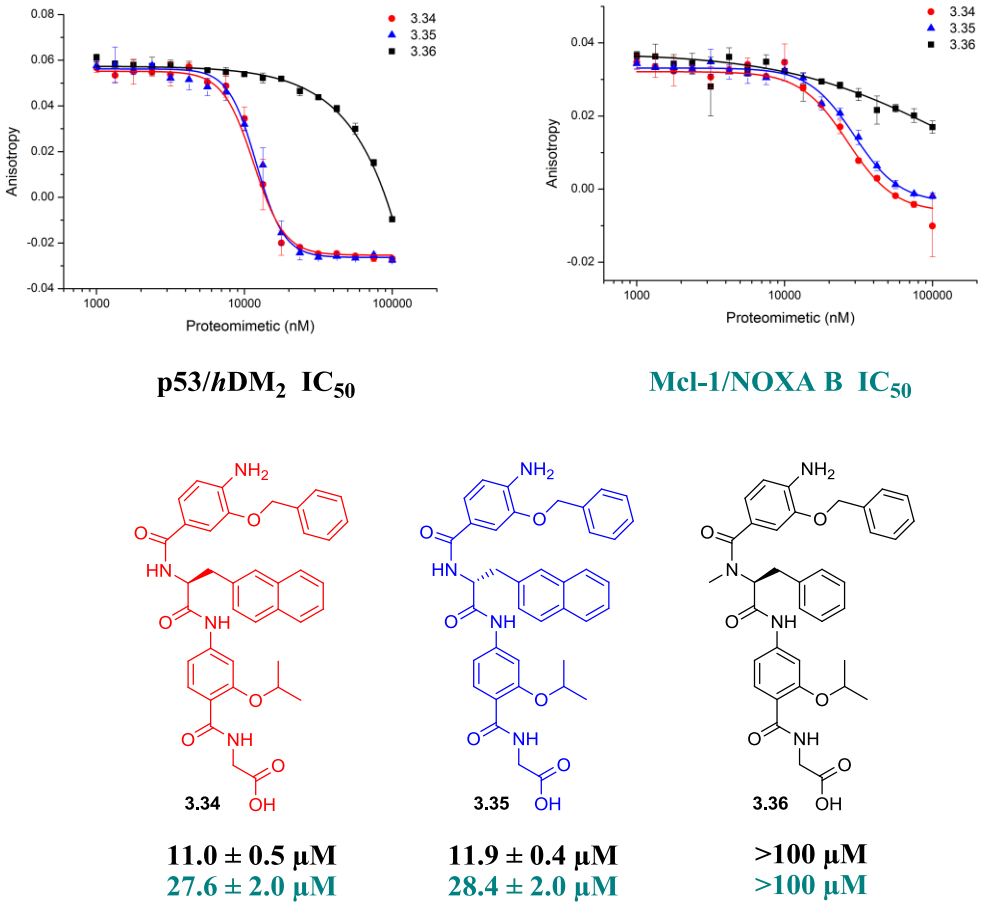


Figure 3.18 Central aa modifications series: Dose-response curves against the p53/*h*DM2 (left) and Mcl-1/NOXA B (right) PPI (40 mM phosphate buffer pH 7.50, 200 mM sodium chloride, 0.02 mg mL^{-1} BSA).

The results showed how the substitution of the central Phe ring from the original compound **3.3** by a naphtyl system, compounds **3.34** and **3.35**, did not produce more potent hybrids, and hence did not improve contacts with the deep Trp binding pocket on the *h*DM2 surface as envisioned when they were designed.

Furthermore, both L- and D- enantiomers were obtained, compounds **3.34** and **3.35** respectively. However, surprisingly the enantiomer dependent selectivity towards Mcl-1 shown by the original compounds **3.3** and **3.4** was not retained in this case.

Additionally, the introduction of an *N*-methyl group on the central Phe aa in compound **3.36** resulted in a decrease of binding affinity to >100 μM for both *h*DM2 and Mcl-1 proteins.

3.5.3. Modification of the bottom aryl unit

The FA competition assay targeting the p53/*h*DM2 and Mcl-1/NOXA B PPI for the bottom aryl modified series of hybrids and the corresponding IC_{50} values are shown in Figure 3.19.

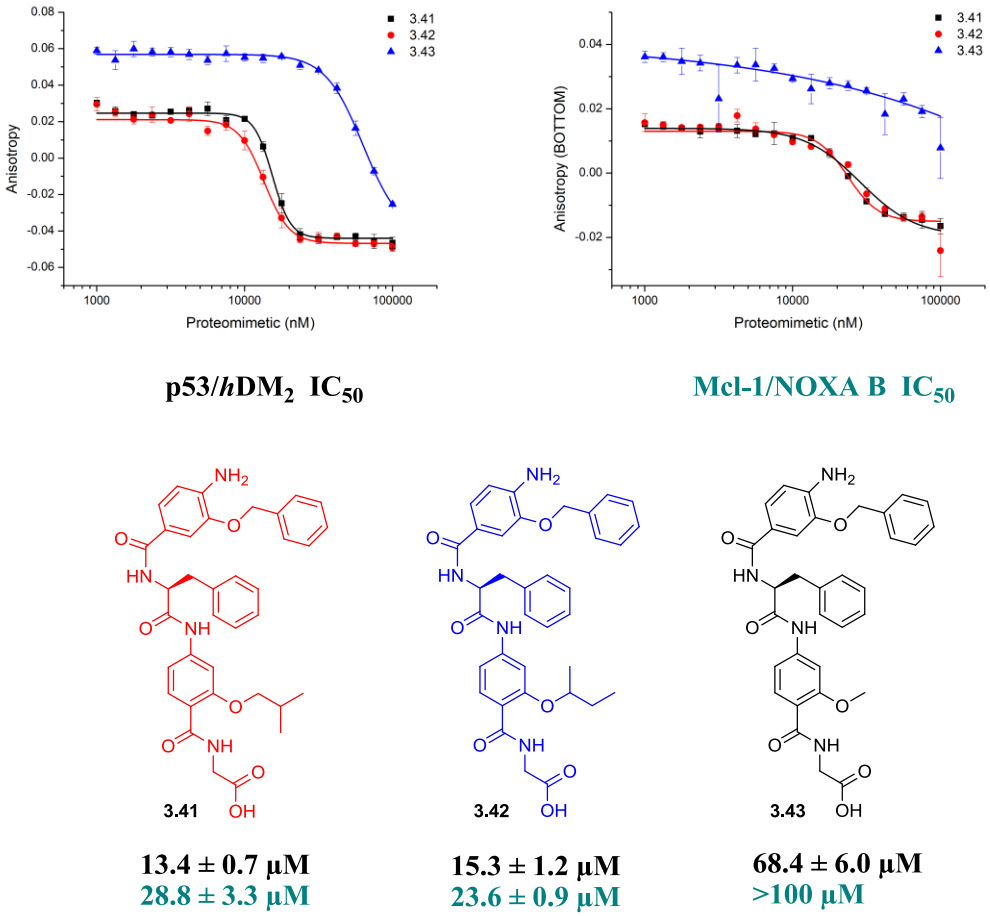


Figure 3.19 Bottom aryl modifications series: Dose-response curves against the p53/*h*DM2 (left) and Mcl-1/NOXA B (right) PPI (40 mM phosphate buffer pH 7.50, 200 mM sodium chloride, 0.02 mg mL^{-1} BSA).

Interestingly, the results showed that reducing the size of the aliphatic side chain on the bottom aryl unit in compound **3.43**, with just a methyl group on that position, allowed the selectivity towards *h*DM2 over Mcl-1 of the original compound **3.3** to be retained. However, this modification also caused a loss of potency. Conversely, the increase in size of the aliphatic chain on the same position in compounds **3.41** and **3.42**, which display an *i*Bu and sBu groups respectively, resulted in complete loss of the afore mentioned selectivity. This result could

contribute to the elucidation of the factors that drive the selectivity in this family of proteomiemtics.

3.5.4. Modifications of the *N*-terminus and *C*-terminal amino acid

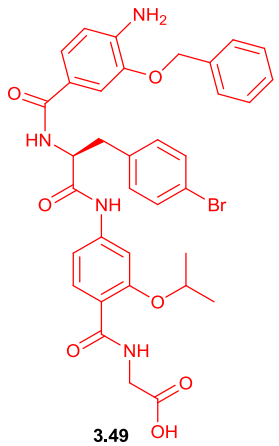
The FA competition assay targeting the p53/*h*DM2 and Mcl-1/NOXA B PPI for the *C*-terminal aa modified hybrid **3.48** showed IC_{50} values >100 μM, indicating that the substitution of the glycine by a lysine residue decreased the binding of the hybrid compound to both protein targets. Unfortunately, due to time constraints the testing of the *N*-terminal modified hybrids was not performed.

3.5.5. Investigation of additive effects

The ultimate goal of the SAR study was to identify factors that increase the potency and/or selectivity of the hybrid and combine these modifications into a new generation of PPI inhibitors. Therefore, it was necessary to investigate if a combination of different substitutions on an individual compound would result in additive effects in activity and/or selectivity.

In this context, it is important to highlight some SAR features observed in a selection of hybrid compounds:

i) Previously, the hybrids halo-substituted at the *p*-Bn position of the central aa were identified as inhibitors of the Mcl-1/NOXA B PPI. In particular, the *p*-bromo functionalised hybrid **3.49** was shown to be the best inhibitor with an IC_{50} of 13.0 ± 0.6 μM (Figure 3.20).[136]



3.49

Mcl-1/NOXA B IC_{50} 13.0 ± 0.6 μM

Figure 3.20 Structure and inhibitory activity against Mcl-1/NOXA B of *p*-bromo functionalised hybrid **3.49**.

ii) The stereochemistry of the central Phe aa in compounds **3.3** and **3.4** dictates the selectivity towards the Mcl-1 protein (See Figure 3.8).

iii) The functionalization of the top aryl unit with a *p*-*t*Bu group in compound **3.29** increase in 2-fold its binding affinity towards *h*DM2 (See Figure 3.17).

Subsequently, compounds **3.50** and **3.51**, which share a combination of substitutions were obtained (Figure 3.21). The FA competition assay targeting the p53/*h*DM2 and Mcl-1/NOXA B PPI for these hybrids and the corresponding binding affinities are shown in Figure 3.21.

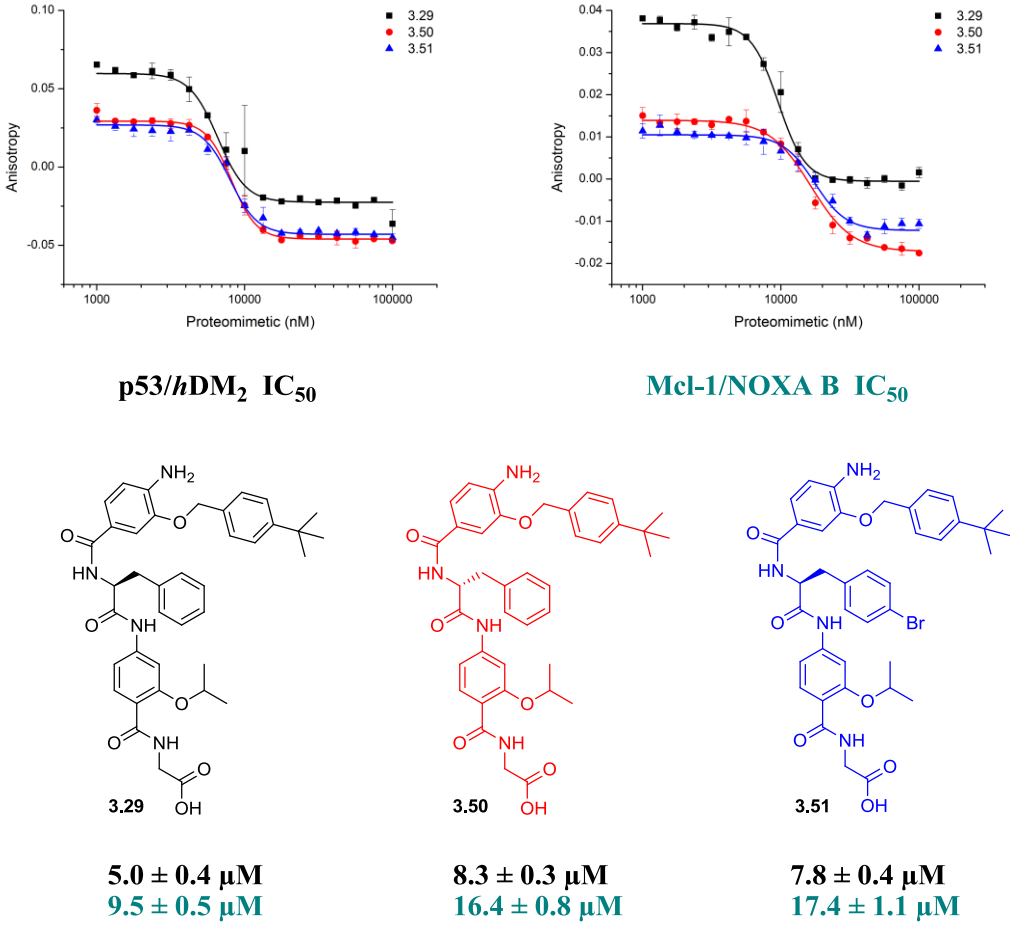


Figure 3.21 Combined modifications series: Dose-response curves against the p53/*h*DM2 (left) and Mcl-1/NOXA B (right) PPI (40 mM phosphate buffer pH 7.50, 200 mM sodium chloride, 0.02 mg mL^{-1} BSA).

Unfortunately, the resulting enantiomeric compounds **3.29** and **3.50** did not retain any selectivity between the p53/*h*DM2 and Mcl-1/NOXA B interactions, as shown by their similar binding affinities to both proteins. Furthermore, the combination of the *p*-bromo in the central aa and the *p*-*tert*-butyl group on the top aryl unit in compound **3.51** did not result in the

expected increase in potency towards any of the PPI of interest. Therefore, this set of compounds does not appear to show an additive SAR.

3.6. Conclusions and future directions

A new set of hybrid compounds derived from the benchmark compound **3.3** has been successfully synthesised allowing a SAR study of its interaction with *h*DM2 and Mcl-1. The existing synthetic strategy for this scaffold has been modified when required and also a novel methodology has been developed (i.e. reductive amination on resin) to customize the hybrid proteomimetics adequately.

Unfortunately, among the set of hybrid analogues synthesised and tested here inhibtors that display a significant increase on their binding affinity towards *h*DM2 or Mcl-1 were not identified. However, the most potent hybrid to date against *h*DM2, compound **3.29** with an IC_{50} of 5.0 ± 0.4 μM, was obtained through the introduction of a *t*Bu group in the *para*-Bn position of the top aryl unit of the original compound **3.3**. However, the enantioselectivity from the original compound **3.3** was not reproduced by compound **3.29** or any of the other chiral compounds investigated. Interestingly, the size of the aliphatic substituent at the bottom unit was found to be important for the selectivity between the p53/*h*DM2 and Mcl-1/NOXA B interactions. Finally, early studies on the SAR features within the hybrid compound library revealed that these features did not have an additive effect when combined in the same proteomimetic structure.

The next step will be to test the hybrids with modifications at the *N*-terminus and in the *C*-terminal aa against the PPIs of interest. In parallel, a quantitative analysis of the aqueous solubility of the compounds by HLPC will also be performed to determine the variations produced by the *N*-terminal and *C*-terminal aa modifications. Unfortunately, due to time constraints these experiments could not be performed. Finally, the synthetic methodology needs to be further explored to incorporate other monomers of interest, such as the already described β and α,α′-disubstituted amino acids, which could lead to a plethora of novel bioactive molecules.

Chapter 4. Design, synthesis and evaluation of inhibitors for the Asf1/H3 PPI

4.1. Introduction

Advancements in the understanding of PPIs have led to them becoming more attractive therapeutic targets.[184] There are already available drugs on the market aimed at treating a multitude of diseases linked to PPIs and an increasing number in clinical trials.[92, 185] Therefore, PPIs are one of the key targets for intervention in biological systems directly related to disease. However, the ubiquitous nature of PPIs makes challenging the complete understanding of their complex organizations. In that context, continuing investigations towards the modulation of unexplored PPIs is vital in the design of new targeted therapeutics.[184]

4.2. Interaction of interest: Asf1/H3

Chromatin - nucleosome - histones

The fundamental structural unit of chromatin is the nucleosome, which is formed by 147 base pairs of DNA wound 1.7 times around the outside of a histone octamer core.[186, 187] The histone octamer comprises two units of each of the histone proteins H2A, H2B, H3 and H4, where two H2A/H2B dimers are tethered to each side of one $(H3/H4)_2$ tetramer.[186-188] Nucleosome assembly and disassembly is a vital process that allows rapid access to specific DNA sequence during transcription, replication, repair and recombination. Hundreds of proteins regulate the folding and unfolding of chromatin allowing the DNA to be exposed and organised for each specific function.[188]

Histone chaperones are proteins that regulate the association of basic histone proteins with the DNA strand.[189, 190] Histones and DNA fail to self-assemble into nucleosomes under physiological conditions due to the strong tendency of histones to associate non-specifically with DNA and form aggregates.[191] Therefore, the assistance of chaperones permits the nucleosome to form in an ordered and controlled manner.

Asf1/H3 PPI

Histone chaperone anti-silencing function 1 (Asf1) is a highly conserved histone chaperone involved in both nucleosome assembly and disassembly.[192-195] In particular, the interaction between Asf1 and the H3 and H4 histone proteins forms Asf1-(H3/H4) complexes (Figure 4.1), which can either supply histones to the nucleosome assembly proteins chromatin assembly factor 1 (CAF-1) and histone regulator A (HIRA)[193, 196, 197] or can interact directly with the DNA.[198, 199] The primary role of Asf1 is to shield H3/H4 dimers from unfavorable interactions with the DNA whilst assisting the formation of positive histone–DNA contacts, which leads to disome [(H3/H4)-DNA] assembly.[200] Furthermore, Asf1 binds the H3/H4 dimer enveloping the *C*-terminus of histone H3 and physically blocking formation of the $(H3/H4)_2$ tetramer.[196]

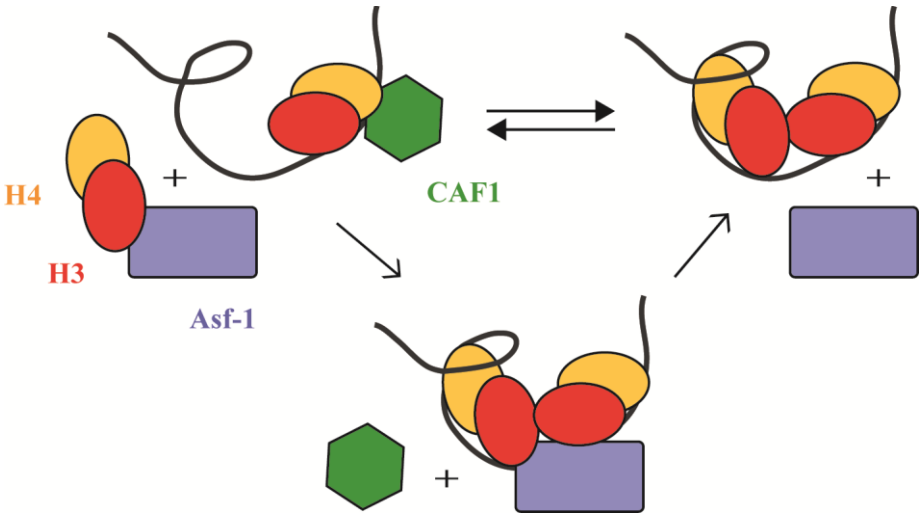


Figure 4.1 Diagram illustrating Asf1 (purple) function in nucleosome assembly by depositing an H3/H4 histone dimer (red and yellow respectively) onto DNA-complexed with the nucleosome assembly protein CAF- 1 (green).[201]

Overall structural features

The Asf1 protein comprises a conserved *N*-terminal domain of 156 amino acids, which is essential for its function *in vivo*, and a divergent unstructured *C*-terminal domain, which is believed not to be necessary for the protein function.[195, 202] The structure of the Asf1 comprises an elongated β sandwich core with three α-helices in the loops between the β strands.

The contacts between H3 and Asf1 are extensive and result in a buried surface area of 909 $Å^2$. The histone H3 binding site is located in the concave face of Asf1 (Figure 4.2) and involves β strands β3, β4, and β6-9.[195, 196, 202] Interestingly, this region of the sequence is highly conserved across species and has a distinctly negative charged nature. The main interactions occur through

the *C*-terminal helix of H3 (residues 122–134), where the key residues Leu[126] and Ile[130] form a hydrophobic clamp with the hydrophobic region of Asf1. Additionally, there is a network of electrostatic interactions within the PPI interface, such as the the salt bridge between Arg[129] from H3 and Asp[54] from Asf1.[203]

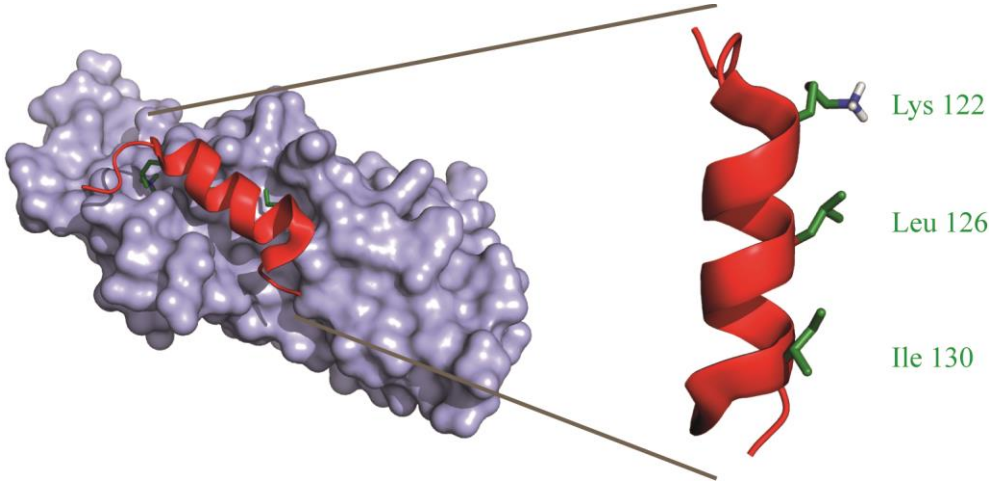


Figure 4.2 NMR structure of Asf1/H3 PPI (PDB ID: 2IIJ). Key side chains on the binding surface of the helix are highlighted.

Furthermore, the Asf1-H3/H4 structure also shows extensive contacts between Asf1 and histone H4.[196] This interface has two parts: the globular core of Asf1 interacts with the *C*-terminal tail of H4 to form a strand-swapped dimer and the *C*-terminal tail of Asf1 binds to the histone fold region of histone H4.

4.3. Inhibition of Asf1/H3 as a PPI of therapeutic interest

The histone chaperone Asf1 has emerged as a promising target for therapeutic intervention for multiple cancers.[188, 200, 204, 205] The binding of Asf1 to H3/H4 dimers promotes the acetylation of the Lys[56] residue from histone H3.[206] This acetylation process is directly related to genomic stability, DNA replication and repair, whereas decreased acetylation levels seem to sensitize the cells to DNA damaging agents. Consequently, the development of compounds able to modulate this PPI could play a key role in the development of therapeutics to treat cancer.

The NMR structure of the conserved *N*-terminal Asf1 histone-binding domain with the histone H3 *C*-terminal peptide was used as a model for the complex between Asf1 and the entire H3/H4 complex. In particular, we envisioned the *C*-terminal α-helix peptide of H3 as a template for the design of molecules able to disrupt the PPI. In order to directly compare and contrast different inhibitor designs when targeting the same PPI, two different approaches were utilised: the proteomimetic strategy and the hydrocarbon stapling peptide strategy.

4.3.1. Proteomimetic approach

The proteomimetic approach has been extensively used within the Wilson group to modulate PPIs. Among the different scaffolds available, the 3-*O*-alkylated oligobenzamide scaffold (Figure 4.3a) was chosen in this case as the proteomimetic approach for the following reasons:

- Proven ability to reproduce the side chains of residues located on one face of an α-helix.[120, 123]
- Effective inhibition of α-helix mediated PPI.[123, 126]
- Accessible synthetic route amenable to library generation.[125]

Initially, we assessed the potential of the 3-*O*-alkylated scaffold to reproduce the key binding residues of the *C*-terminal H3 peptide by minimising the energy of a model scaffold structure in Macromodel. The lowest energy conformations within 1.5 kJ/mol of the model compound adopted an extended structure where the amide bonds are *trans*. Subsequently, these structures were superimposed onto the native α-helix of histone H3, as shown in Figure 4.3b. The match was assessed on the basis of the overlapping between α-carbons on the helix and oxygen atoms on the foldamer. As anticipated, the scaffold projects the side chains in a similar orientation to the native peptide, as evidenced by their overlay (RMSD = 0.70).

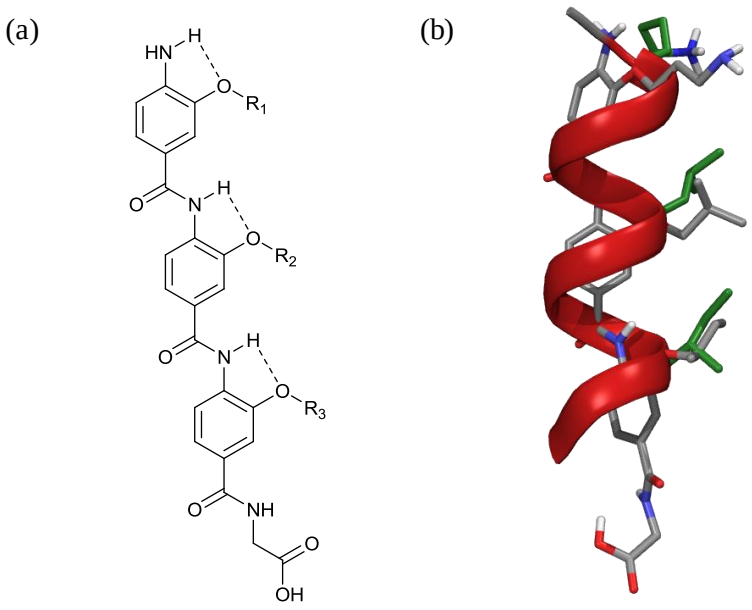


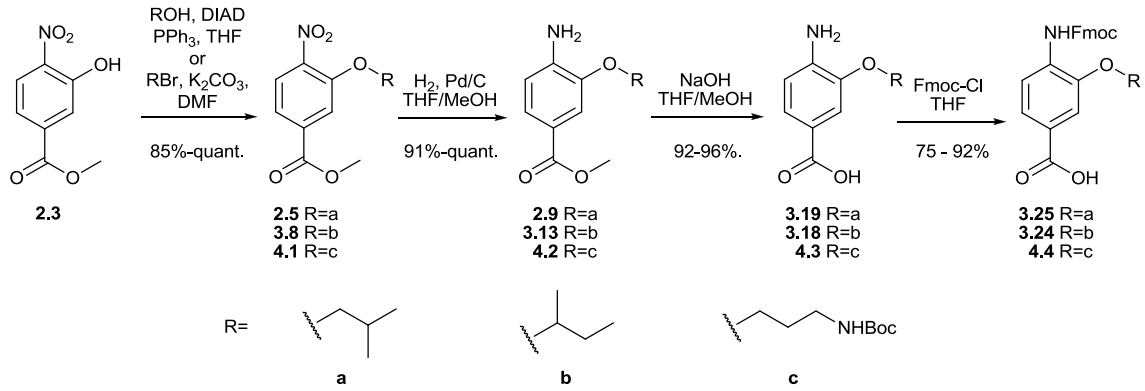
Figure 4.3 (a) structure of the 3-*O*-alkylated oligobenzamide scaffold. (b) Overlay of a model proteomimetic scaffold (grey) and the native H3 α-helix (red) with key binding residues highlighted (green).

4.3.2. Synthesis

The synthesis of the 3-*O*-alkylated oligobenzamide α-helix mimetic scaffold was based on an Fmoc (9-fluorenylmethyloxycarbonyl) solid phase peptide synthesis (SPPS) strategy previously developed in the Wilson group.[125] It comprises the synthesis of the monomeric units followed by their assembly in a sequential manner.

4.3.2.1. Monomer synthesis

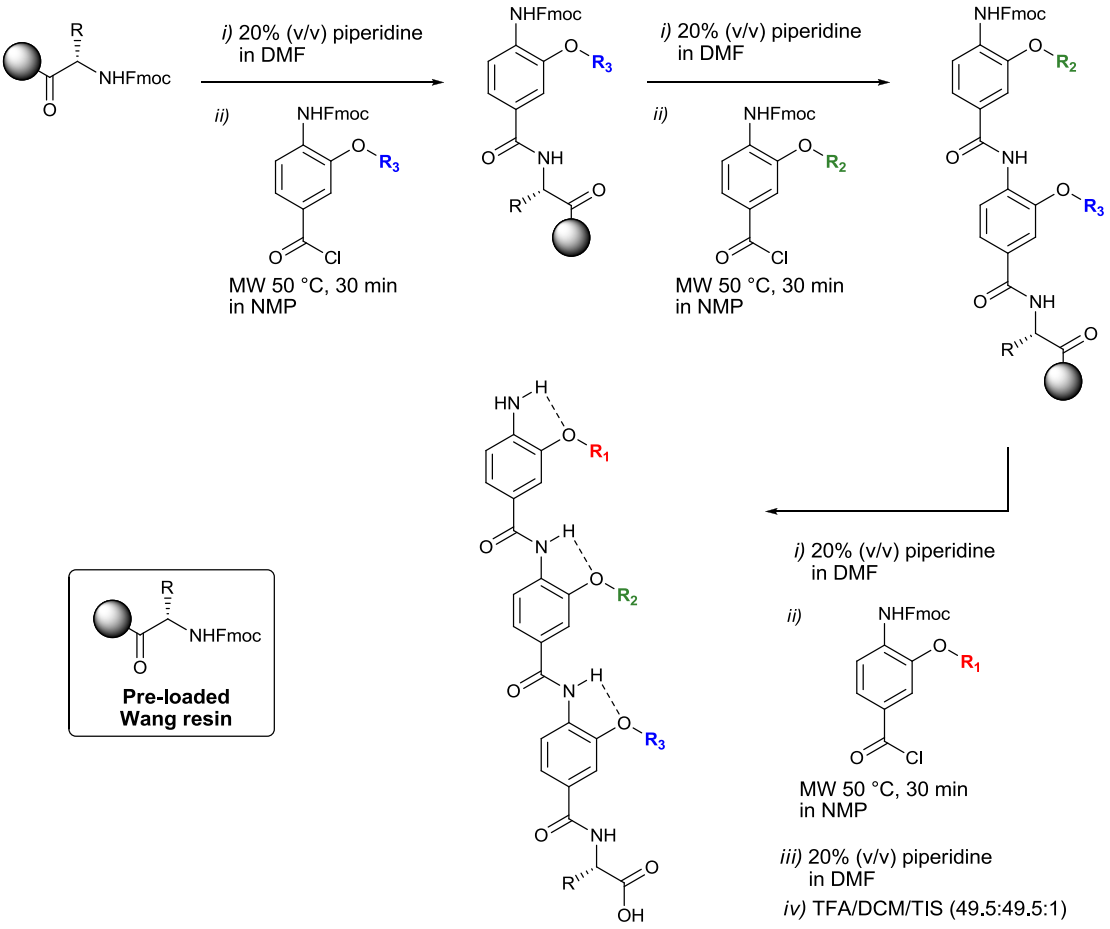
The synthesis of the 3-*O*-alkylated protected monomers followed a synthetic route previously described by the group (Scheme 3.1).[122, 125] It consisted of an initial alkylation step, followed by reduction of the nitro to the amino group. Subsequent hydrolysis of the ester and final Fmoc-protection produced the desired 3-*O* alkylated Fmoc-protected monomers in high yields.



Scheme 4.1 Synthetic route to 3-*O* alkylated Fmoc-protected building blocks. Monomer **4.4** was provided by Dr N. S. Murphy.

4.3.2.2. General synthetic scheme

The 3-*O*-alkylated α-helix mimetics were synthesised through a microwave-assisted automated SPPS (Scheme 3.2) developed by Dr N. S. Murphy.[125] It consisted of a series of deprotection and coupling steps, using Fmoc-aa- (where aa can be any amino acid) preloaded Wang resins as solid support; and the pre-formed acid chlorides from the Fmoc- monomers as building blocks.



Scheme 4.2 Solid phase synthetic route for assembly of the 3-*O*-alkylated α-helix mimetics.

It is worth noting that the coupling of monomers bearing protected polar side chains was found to be more challenging than the coupling of hydrophobic monomers. As a result of this, double coupling steps were required for these building blocks. Using the described synthetic route, a set of proteomimetics was obtained (Figure 4.4). In order to investigate the potential of the 3-*O*-alkylated scaffold to mimic the H3 α-helix, the compounds that were obtained displayed the following characteristics:

- Native sequence mimetic, compound **4.5**. The side chains from the H3 key binding residues Lys[122], Leu[126] and Ile[130] (see Figure 4.2) are mimicked by a propylamine, *iso*-butyl and a *sec*-butyl groups respectively.
- Exchange of the hydrophobic side chains, compound **4.6**. The hydrophobic Leu[126] and Ile[130] mimetics are swapped in this compound to study the subtle differences in the shape of their side chains.
- Antiparallel mimetic of the native sequence, compound **4.7**. The nature of the structure permits the superimposition in parallel and antiparallel *N*-to-*C* orientation with respect to an α-helical peptide.[125] In this case, the H3 key binding residues Lys[122], Leu[126] and

Ile[130] (See Figure 4.2) are mimicked by an ethylamine, *iso*-butyl and a *sec*-butyl groups respectively in an antiparallel orientation.

- Negative control, compound **4.8**. The three side chains displayed by this mimetic are hydrophobic *sec*-butyl groups. It is intended to be a poor mimetic of the native H3 helix.
- Control dimers, compounds **4.9** and **4.10**. These proteomimetics were generated from two building blocks bearing hydrophobic side chains and were expected to bind less tightly to the protein targets. These compounds would highlight the importance of the third side chain in the recognition process and in particular the need for an amino group to mimic the electrostatic interactions of the Lys[122] residue.

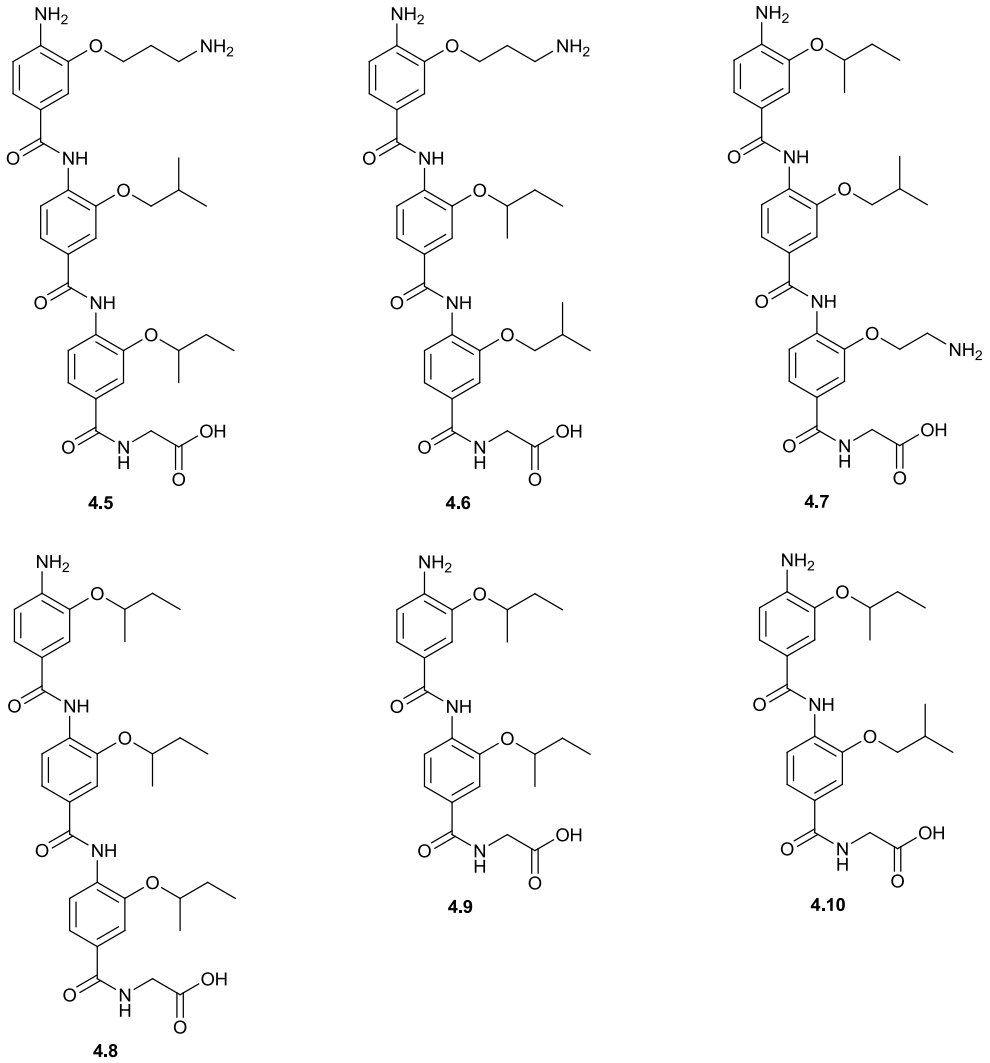
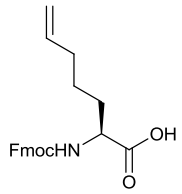


Figure 4.4 Set of 3-*O*-alkylated oligobenzamides **4.5** to **4.10** as *C*-terminal H3 α-helix proteomimetics.

4.3.3. Hydrocarbon stapled peptide approach

One of the methodologies used in the Wilson group to inhibit PPIs is with hydrocarbon stapled peptides. In recent years, Dr D. Yeo introduced a monosubstituted alkenyl amino acid **4.11** (Figure 4.5) to prepare hydrocarbon staples at *i* and *i*+4 positions of a peptide sequence. These type of stapled peptides showed an increased propensity to adopt an α-helical conformation and effectively disrupted PPIs.[69,70]



4.11

Figure 4.5 Structure of the monosubstituted alkenyl amino acid **4.11**.

In order to further expand this approach to other families of α-helix mediated PPI, the stapled variant of the *C*-terminal H3 peptide was designed. The sites to incorporate the monosubstituted amino acids within the peptide sequence were selected taking into account the following factors:

- The space between residues should be suitable to form a crosslink (usually *i* and *i*+4 positions).
- The hydrocarbon bridge should be located in a position that does not interfere in the protein-protein interface and should not modify the orientation of the key interacting residues.
- The substitutions should not replace any other key residues, such as charged residues involved in stabilizing salt bridges or peptide solubility.

Based on that criteria, the Met(120) and Ile(124) residues of the H3 peptide sequence (Figure 4.6) were substituted by the monosubstituted amino acid **4.11**.

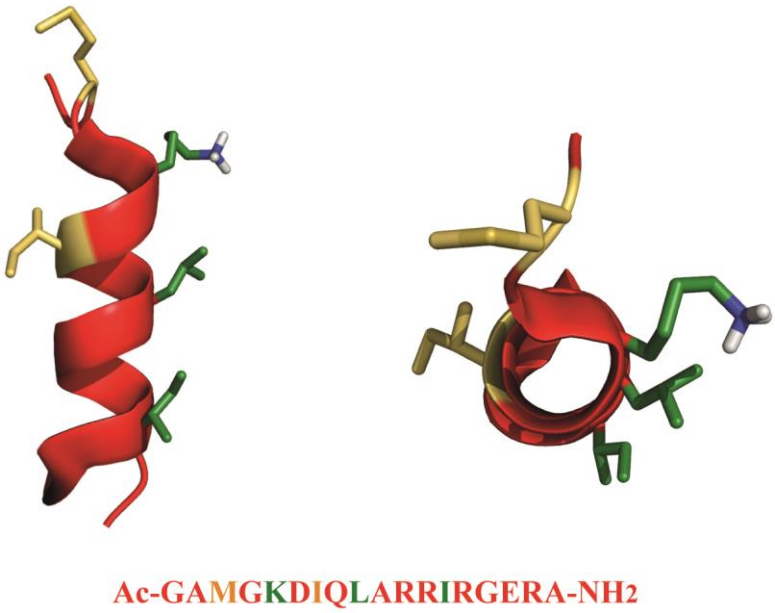
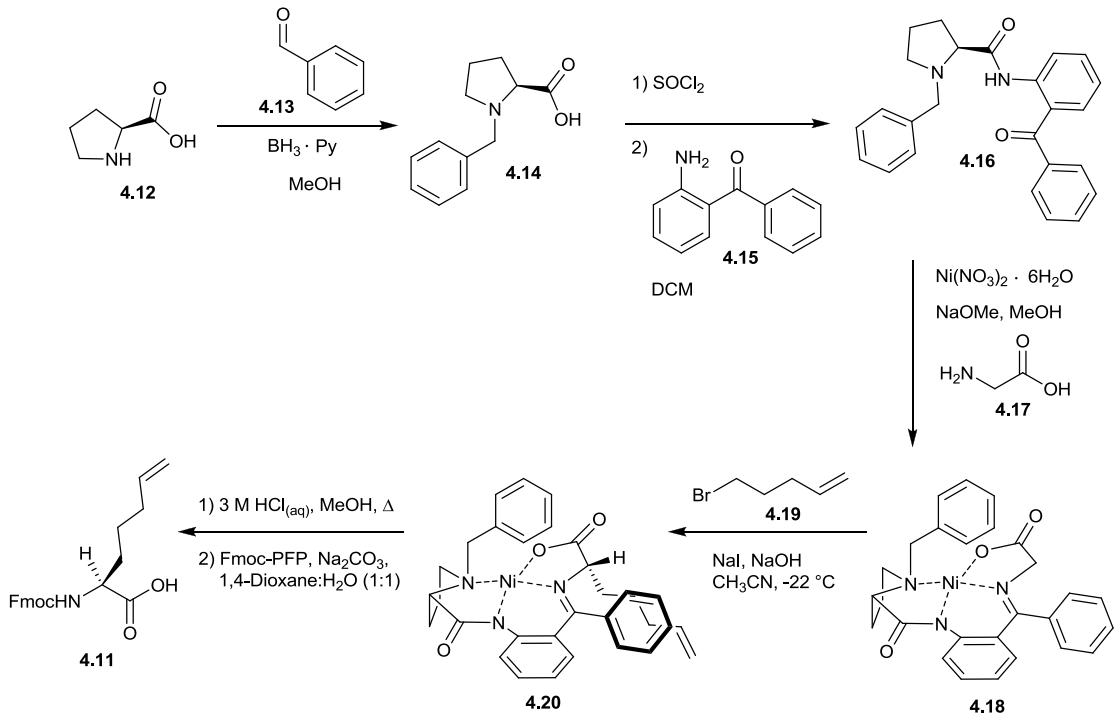


Figure 4.6 *C*-terminal H3 peptide showing the key binding residues (green) and the new stapling positions (yellow). Side (left) and top (right) views are given.

4.3.3.1. Amino acid Synthesis

The monosubstituted alkenyl amino acid incorporated in the stapled peptides was obtained using a synthetic route reported by the Wilson group (Scheme **4.3**).[69] It involves a six steps synthesis and exploits a chiral ligand to provide the desired stereochemistry in the final Fmoc-protected amino acid. The synthetic route starts with reductive amination between L-proline **4.12** and benzyl aldehyde **4.13** to give *N*-benzyl proline **4.14**, followed by in situ formation of the corresponding acyl chloride and coupling to **4.15** provides the final chiral ligand **4.16**. The complexation reaction between glycine **4.17**, the chiral ligand **4.16** and Ni(II) yields complex **4.18**, which is then enantioselectively alkylated in the α-position under basic conditions to give compound **4.20**. Final decomplexation under acid conditions followed by Fmoc-protection leads to the desired Fmoc-monosubstituted alkenyl amino acid **4.11**.



Scheme 4.3 Reaction scheme of the synthesis of Fmoc-protected monosubstituted alkenyl amino acid **4.11**.[69]

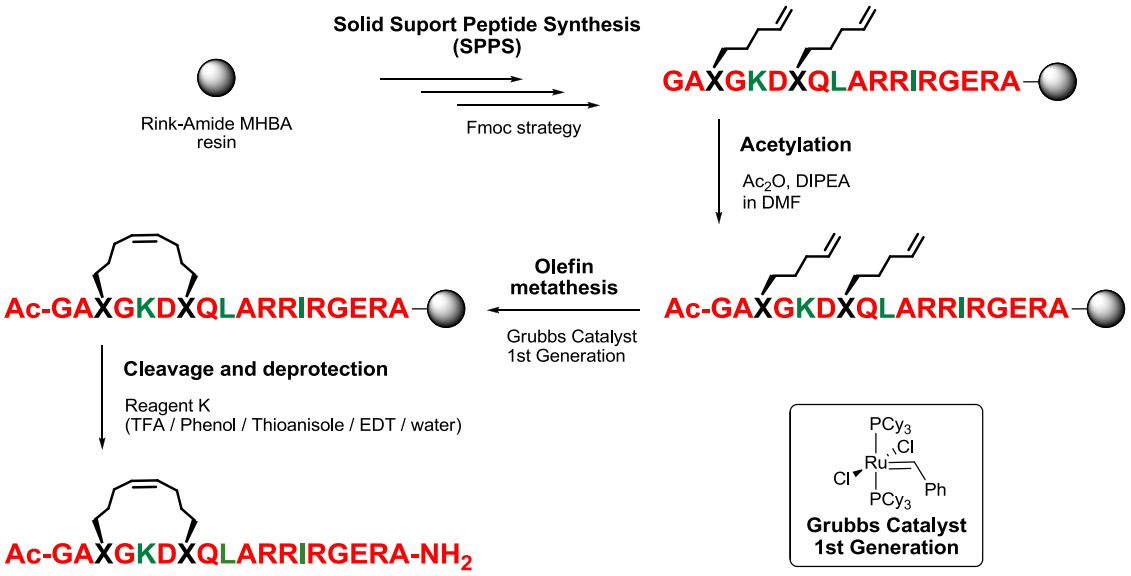
4.3.3.2. Peptide Synthesis

The peptides were synthesised on Rink Amide MBHA resin as solid support, in order to obtain *C*-terminal amidated peptides comparable to the *C*-terminal domain of H3 (Scheme 4.4). Low loading resin was used in order to reduce clustering of the growing peptide chain and hence aid the coupling and deprotection steps.

The synthesis was performed using Fmoc solid phase peptide synthesis (SPPS), with HCTU as activating reagent and DIPEA as base for peptide elongation. Double couplings were required for the coupling of the ARRIR region, after small scale test cleavages had revealed these couplings to be incomplete after 2 hours. The monosubstituted unnatural amino acid **4.11** was incorporated using HATU as a coupling reagent, with the Kaiser test being conclusive for their couplings.[207]

The alkenyl side chains of the unnatural amino acids were metathesized using the first generation Grubbs catalyst. The cleavage and side chain deprotection of the peptide from the resin was carried out with Reagent K, which contains ethanedithiol, thioanisole and phenol as scavengers to prevent the irreversible modification of nucleophilic side chains by carbocations, such as *tert*-butyl, and the oxidation of sensitive residues, such as methionine. The deprotection

times needed to be extended to 2-4 hour washes due to the multiple arginine residues bearing Pbf protecting groups, which had proven challenging to remove in this particular sequence.



Scheme 4.4 Synthetic route to the stapled peptides.

Purification of the stapled and unstapled H3 peptides was performed using reverse phase preparative HPLC. This process proved challenging, as the alternating polar and hydrophobic regions in the peptide sequence are prone to aggregate. This resulted in poor yields of 11% and 5% of pure stapled and unstapled peptide respectively.

4.3.4. Circular Dichroism (CD)

Circular dichroism (CD) was used to evaluate the effectiveness of the hydrocarbon crosslink to promote an α-helical secondary structure and hence to reduce the peptide conformational freedom in aqueous solution. CD is a technique that provides distinct spectral profiles depending on the secondary structure present on a species (Figure 4.7). In particular, α-helical conformations of peptides and peptidomimetics are identified and quantified using the amplitude of the spectrum at 222 nm.[208, 209]

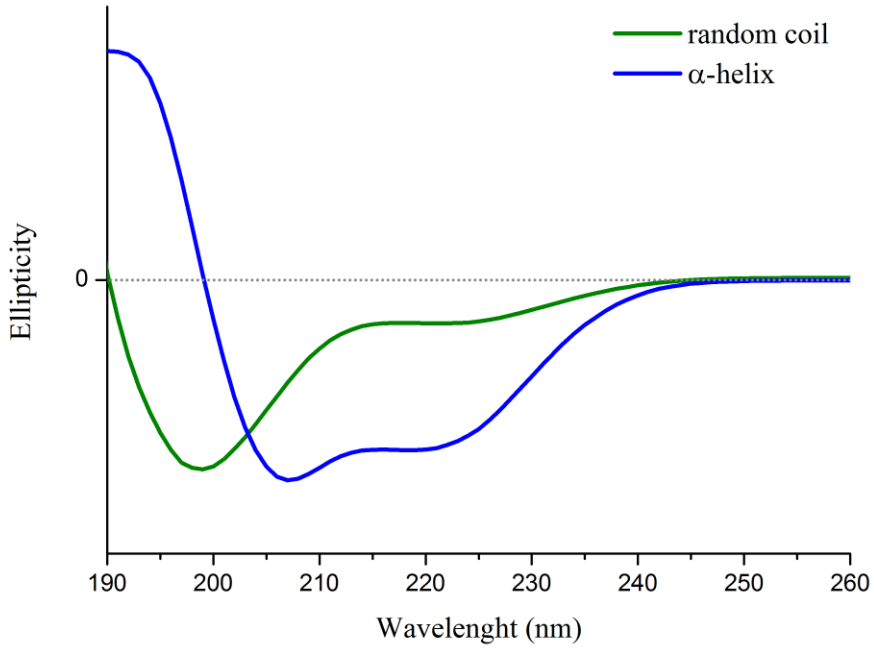


Figure 4.7 Representative example of a CD spectra of an unstructured random coil peptide (green) and an α-helical peptide (blue).

The helical character of the unstapled and stapled 1 H3 peptide variants was investigated using CD as shown in Figure 4.8. Unfortunately, the results show a predominantly unstructured random coil character for both stapled 1 and native peptide sequences, suggesting that the introduction of the hydrocarbon staple at these positions did not increased the helicity of the peptide. Additionally, trifluoroethanol (TFE) was used as a cosolvent due to its ability to promote α-helical conformations in short peptides,[210, 211] and thus providing a benchmark for the maximum helicity of the particular peptide sequence. In these conditions, both stapled 1 and native peptide showed an almost identical change in their CD profiles, matching an α-helical secondary structure. The highly similar results for both peptides also suggest that in this case the hydrocarbon staple does not interfere with the conformational flexibility of the peptide.

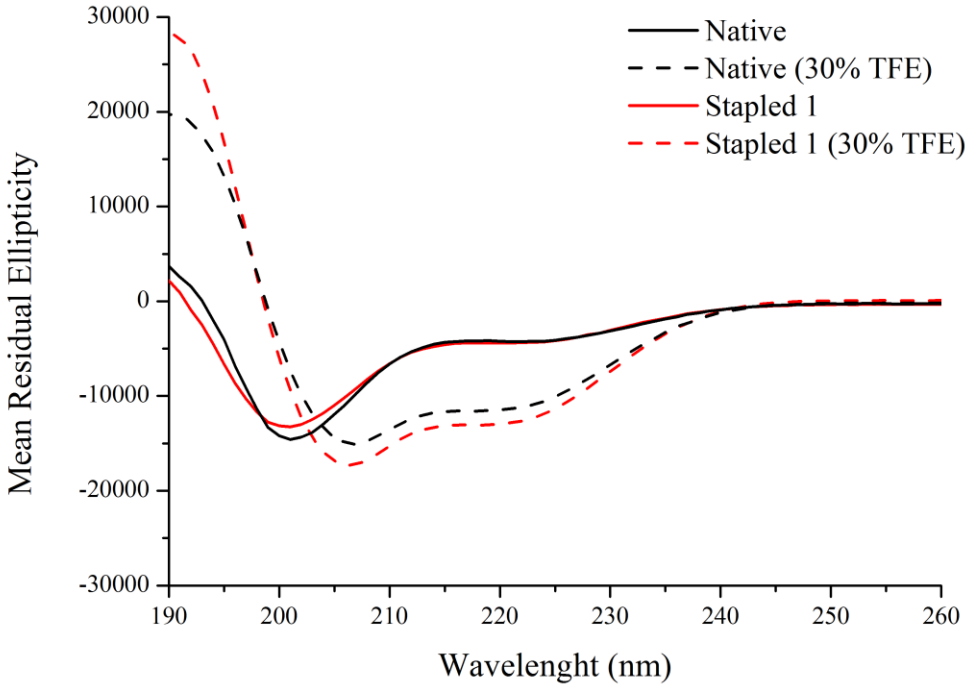


Figure 4.8 CD spectra of the native (black) and stapled 1 (red) H3 peptides in 40 mM phosphate buffer, pH 7.5 (solid line) and in 30% TFE (dash line).

4.3.5. Redesign of the stapled peptide

In order to further investigate the potential of the hydrocarbon stapling strategy to increase the helicity of the *C*-terminal H3 peptide, a different stapling position was selected (Figure 4.9). In this case, a more central position within the sequence was chosen, replacing the Asp(123) and Ala(127) residues, in the belief that it would aid the helical conformation promotion.

In addition, the extra three residue fragment GCA was added at the *C*-terminus of the peptide sequence. This additional section was introduced as a future conjugation handle via the nucleophilic thiol of the cysteine residue.

The new stapled 2 peptide was synthesised following the same procedure as described in section 4.3.3.2 and obtained in 3% yield. The CD measurements show a significant increase in the helicity of the stapled 2 peptide when compared with the unstructured unstapled variant (Figure 4.10), suggesting the hydrocarbon staple successfully promotes a more helical conformation in the peptide. Again, both peptides increased their α-helical structure when TFE was used, however on this occasion the stapled 2 variant showed a smaller increase in helicity

than the native sequence. This observation may be due to a reduction of the peptide conformational flexibility caused by the hydrocarbon cross-linkage.

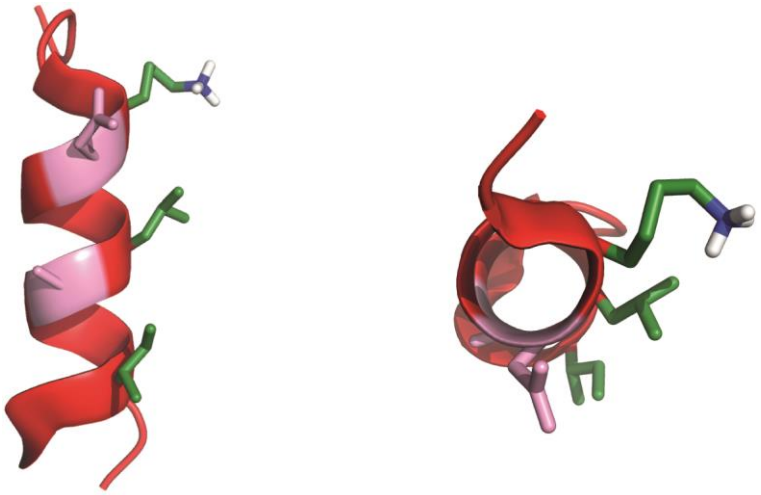


Figure 4.9 *C*-terminal H3 peptide showing the key binding residues (green) and the new stapling positions (pink). Side (left) and top (right) views are given.

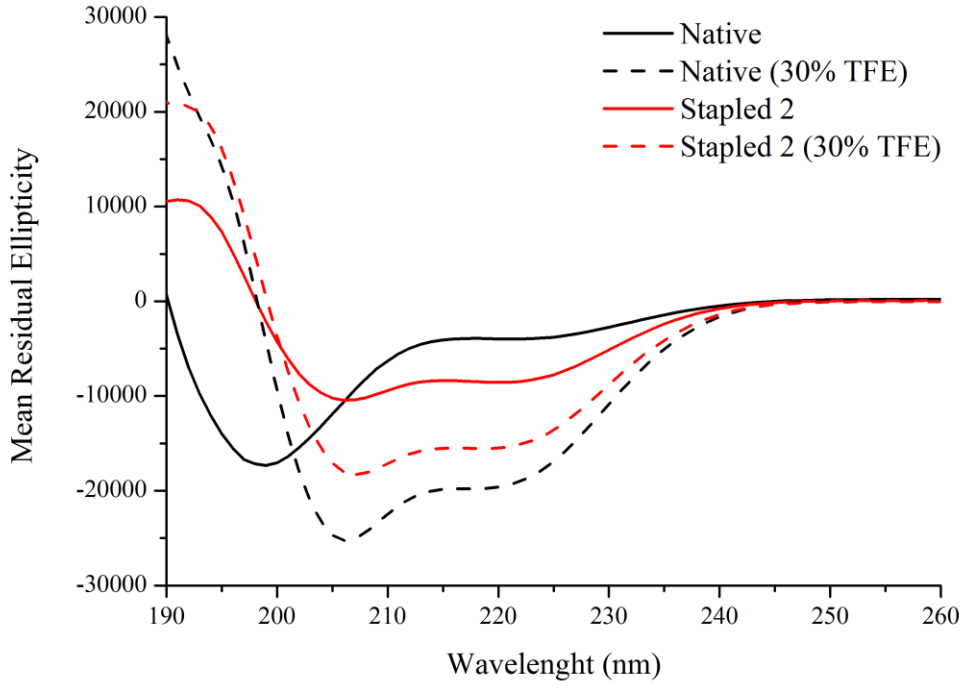


Figure 4.10 CD spectra of the native (black) and stapled 2 (red) H3 peptides in 40 mM phosphate buffer, pH 7.5 (solid line) and in 30% TFE (dash line).

4.4. Biophysical testing - Isothermal Titration Calorimetry (ITC)

The set of proteomimetic compounds and the stapled peptides were sent to our collaborators in the Ochsenbein group at the Institute of Biology and Technology of Saclay (iBiTec-S) to assess their ability to act as PPI inhibitors. The PhD student May Baikal used Isothermal Titration Calorimetry (ITC) to study the binding of the compounds to the Asf1 protein.

Isothermal titration calorimetry (ITC) is a technique used to study the thermodynamic parameters of interacting bimolecular systems.[212] In particular, this method allows determination of the thermodynamic contributions of binding events through the calculation of the K_d and ΔH values. The technique consists of measuring small heat variations in a cell as one species is titrated into another. The thermal changes reflect the enthalpy changes of the system and they decrease upon saturation of the binding sites. The main advantage of ITC compared to other biophysical techniques is that labelled substrates are not required, which avoids additional non-target interactions and equilibria.

Unfortunately, our 3-*O*-alkylated oligobenzamide proteomimetics compounds were not sufficiently potent at low concentrations to be detected using ITC; however, increasing the concentration was problematic due to the low solubility and aggregation of the compounds.

Conversely, the native and stapled 1 peptide provided clear binding curves when titrated into the Asf1 protein (Figure 4.11), displaying a 1 to 1 stoichiometry in both cases. The thermodynamic parameters extracted from the curves (Table 4.1) show marginal differences in the binding constant (K_D), and the enthalpic (ΔH) and entropic (ΔS) contributions. These results are in agreement with the previously discussed CD studies (See section 4.3.4), proposing that the hydrocarbon staple at this position did not affect the peptide conformation, and hence it did not change the binding of the peptide to the Asf1 protein. Unfortunately, we are still awaiting the ITC results for the second variant of stapled peptide.

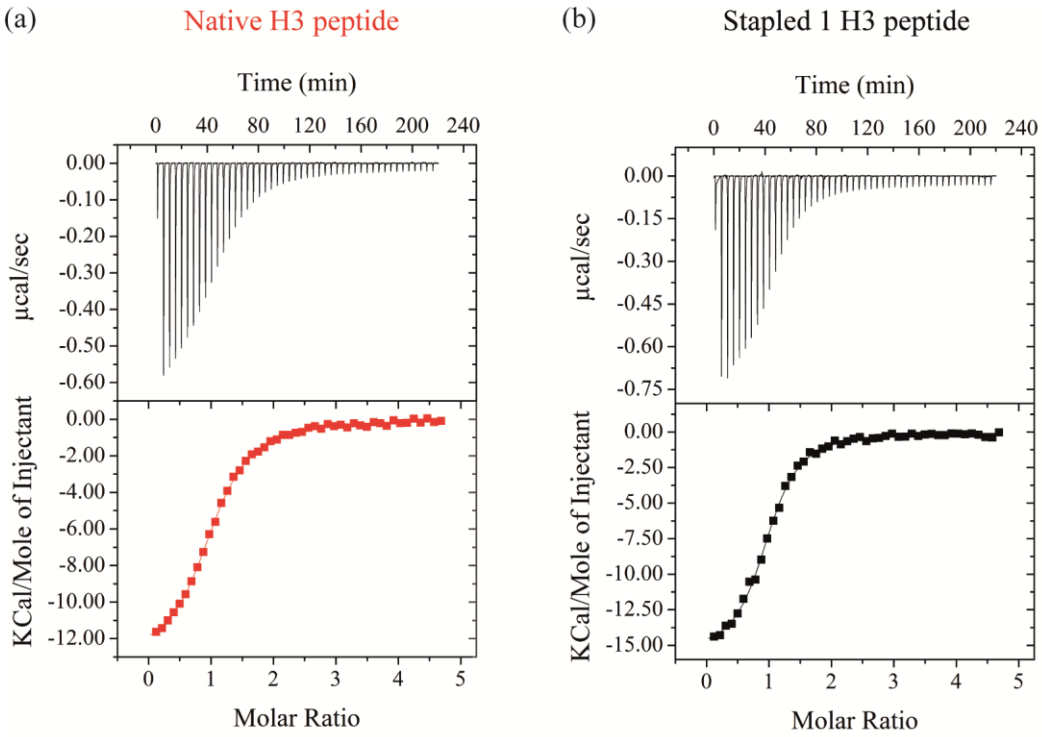


Figure 4.11 ITC thermograms for the (a) native and (b) stapled 1 H3 peptides and Asf1 protein.

Table 4.1 ITC data for the native and stapled 1 H3 peptides and Asf1 protein.

Peptide	K_D (µM)	ΔG (kcal.M^{-1})	N	ΔH (kcal.M^{-1})	-TΔS (kcal.M^{-1})
Native	1.48 (±0.46)	-7.41 (±0.17)	1.02 (±0.02)	-13.44 (±0.30)	6.03 (±0.47)
Stapled 1	0.86 (±0.11)	-7.71 (±0.07)	0.97 (±0.01)	-15 (±0.96)	7.28 (±1.03)

4.5. NMR studies – ^{1}H-^{15}N HSQC

In order to confirm the binding mode of the stapled peptide, the PhD student May Baikal from the Ochsenbein group performed ^{1}H-^{15}N HSQC perturbation shifts studies.[213] This method involves acquisition of the HSQC spectra for both the ^{15}N-labelled protein only and the ^{15}N-labelled protein in the presence of the ligand of interest. Then, the complexation-induced shifts on certain residues can be identified and mapped onto the protein structure.

Distinct complexation-induced shifts were observed upon addition of the stapled 1 H3 peptide.

Once mapped onto the structure of Asf1/H3 (PDB ID: 2IIJ), the study showed that shift changes were induced in the same area of the protein to those induced by the *C*-terminal H3 peptide, as shown in Figure 4.12. These results further confirm that the stapled 1 variant interact with Asf1 in the same region and in the same manner to the native *C*-terminal H3 peptide. Interestingly, the two main amino acids which shift differently between the stapled and the unstapled peptides are Glu^{49} and Glu^{51}, which are located in the histone binding site and face the expected position of the staple (indicated with green stars).

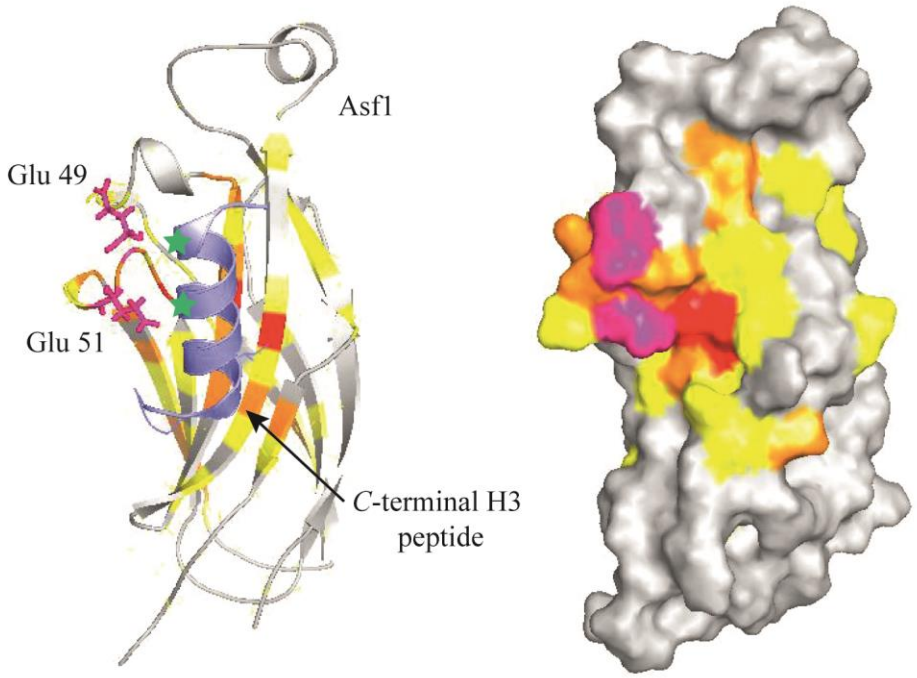


Figure 4.12 ^{1}H-^{15}N HSQC chemical shift perturbation mapping onto the structure of Asf1/H3 (PDB ID: 2IIJ) highlighting the shift changes of the residues on a red (major movement) to yellow (minor movement) gradient, grey (no movement). The position of the Glu^{49} and Glu^{51} residues are shown in magenta. The *C*-terminal H3 peptide is represented in blue (left structure) and the expected position of the staple is indicated with green stars.

4.6. Conclusions and future directions

Two different approaches for inhibition of PPIs have been used to target the Asf1/H3 interaction. The proteomimetic approach utilised the 3-*O*-alkylated oligobenzamide scaffold to synthesise a set of inhibitors, which mimic the key side chain residues of the *C*-terminal H3 helix. Unfortunately, in the biophysical assays performed by the Ochsenbein group using ITC the proteomimetic compounds did not reproducibly show binding to the Asf1 protein. These results might be due to a low binding affinity combined with low solubility and aggregation of the proteomimetics.

The hydrocarbon stapled peptide approach exploited the monosubstituted alkenyl amino acid methodology existing in the group to synthesise a first stapled variant of the *C*-terminal H3 peptide. CD analysis revealed that the staple did not increase the helicity of the peptide when compared to the native H3 sequence. Consistently, the ITC studies confirmed a 1 to 1 stoichiometry in both cases and show marginal differences in the thermodynamic parameters for the constrained peptide, such as binding constant (K_D), and the enthalpic (ΔH) and entropic (ΔS) contributions. ^{1}H-^{15}N HSQC perturbation shifts studies further suggested a similar binding mode of the stapled and unstapled peptides to the Asf1 protein. Subsequently, a second stapled peptide was synthesised bearing a novel stapling position. In this case, the promising results from the CD analysis showed a significant increase in the helical character of the stapled peptide when compared to the native H3 sequence. Unfortunately, the results from the binding studies of the second stapled peptide to the Asf1 protein have yet not been received from our colleagues in the Ochsenbein group.

The future work includes the use of other scaffolds available in the Wilson group, such as the *N*-alkylated oligobenzamide scaffold or the hybrid scaffold, with more soluble profiles. Furthermore, the introduction of polar groups, such as a guanidino functionality as an arginine mimetic, could be introduced on the top of the sequence to improve solubility and add a further binding group.

Regarding the hydrocarbon stapling strategy, ongoing work in the Ochsenbein group points towards the possibility of working with longer H3 based peptides bearing optimised sequences that bind Asf1. This opens the door to a multitude of options for stapled peptides variants, including the assessment of new stapling positions and the effect of hydrocarbon linker length. Finally, the conjugation of the H3 stapled peptides to cell penetrating peptides (CPP) using the additional GCA fragment incorporated at the *C*-terminus of the stapled peptide 2 could aid the study of our PPI inhibitors in cells.

Chapter 5. Thesis summary

Protein-protein interactions (PPIs) play a central role in the mediation of regulatory pathways and control essential cellular processes.[3, 4] Therefore, the use of small-molecules to modulate PPIs is important for the development of new therapeutic approaches, as well as to gain further insight into biological systems.[1] Since the α-helix is the most common secondary structure in nature, significant efforts have focused on targeting α-helix mediated PPIs.[5] Some of the most relevant approaches in this field are the proteomimetic approach, which recapitulate the key binding residues of an α-helix on a non-peptidic scaffold; and the constrained peptides; which pursue to reproduce the helical structure by stabilising the helical peptide. Both approaches have generated potent inhibitors of a great diversity of α-helix mediated PPIs.

Over 40 % of the α-helix mediated PPIs involve key binding residues located on more than one face of the helix. However, there are very few scaffolds designed to mimic multiple faces of the putative helix. In *Chapter 2*, we introduced the design and synthesis of two bifacial proteomimetic scaffolds, bis-benzamide and *N*-(4-aminophenyl)terephthalamidic, to target the ER/co-activator PPI by mimicking the key side chains of the co-activator recognition motif LXXLL. Subsequent structural analysis, involving 2D NMR techniques and H/D exchange experiments, in combination with *in silico* studies revealed a plethora of different side chain spatial relationships, some of which effectively mimic the intended side-chains of the co-activator α-helix. Significant conformational knowledge has been gained and added to the already existing data on aromatic oligoamide foldamers. Unfortunately, fluorescence polarization assays of the new scaffolds against nuclear receptor/co-activator interactions did not show binding of the compounds to the protein surface, highlighting the complex relationship between helix mimetic conformation and molecular recognition.

Another major challenge towards the effective modulation of PPIs are molecules with enhanced control over target selectivity. The complex relationships between molecular rigidity, target plasticity and activity all participate in surface recognition processes. Recently, the Wilson group reported a new hybrid scaffold, which reproduced the side chains at *i*, *i*+4 and *i*+7 positions of an α-helix, showed increased conformational plasticity over the original 3-*O*-alkylated oligobenzamide scaffold, and most importantly presented stereodependent target selectivity. In *Chapter 3* we attempted to further study and optimize the hybrid oligobenzamide scaffold towards more potent and selective inhibitors of the p53/*h*DM2 and the Mcl-1/NOXA B PPIs.

SAR analysis was performed on the new set of hybrid compounds obtained through modification of the *N* termini, the top or bottom aryl building blocks, the central amino acid, and the *C* terminal amino acid of the original scaffold. As a result, the most potent hybrid to date against *h*DM2, with an IC_{50} of 5.0 ± 0.4 μM, was obtained. Interestingly, the selectivity of the compounds between the p53/*h*DM2 and Mcl-1/NOXA B interactions seemed to be affected by the size of the aliphatic substituent at the bottom unit of the scaffold. However, the enantioselectivity from the original benchmark compounds was not reproduced by any of the chiral compounds investigated. In addition, the SAR features observed within the hybrid compound library did not have an additive effect when combined in the same proteomimetic structure.

Finally, it is essential to continue expanding the scope of PPI inhibition as a therapeutic tool by developing new modulators that target unexplored PPIs related with disease. In that context, the histone chaperone Asf1 has emerged as a promising target for therapeutic intervention for multiple cancers.[188, 200, 204] In *Chapter 4* we designed inhibitors of the Asf1/H3 interaction using the *C*-terminal α-helix peptide of H3 as a template. Two different approaches were used, the proteomimetic strategy and the hydrocarbon stapling peptide strategy, in order to compare the different inhibitor designs when targeting the same PPI.

The proteomimetic approach exploited the 3-*O*-alkylated oligobenzamide scaffold to synthesise a set of inhibitors. Unfortunately in the biophysical assays performed by the Ochsenbein group using ITC, the proteomimetic compounds did not show binding to the Asf1 protein. These results might be due to a low binding affinity combined with low solubility and aggregation of the proteomimetics. The hydrocarbon stapled peptide approach exploited the monosubstituted alkenyl amino acid methodology existing in the group to synthesise two stapled variants of the *C*-terminal H3 peptide. Circular Dichroism (CD) revealed that the first stapled variant did not increase the helical character of the peptide in solution when compared with the native sequence. This fact consistently resulted in identical binding affinities and binding modes to the Asf1 protein, which was determined by the Ochsenbein group using ITC and HSQC studies respectively. Promisingly, a second re-designed stapled variant did produce a peptide with an increased helical secondary structure when compared to the native helix. Unfortunately, the effect of this increased helical character on the binding affinity to Asf1 has still to be determined as we are awaiting the results from the Ochsenbein group.

During this PhD, two different approaches to inhibit PPIs have been used, the proteomimetic and the constrained peptide strategies. Three different scaffolds (bifacial, 3-*O*-alkylated and

hybrid scaffolds) and two stapled peptides have been explored to target a multitude of PPIs (ER/co-activator, *h*DM2/p53, Mcl-1/NOXAB, Asf1/H3). Both approaches have helped address key questions regarding helix mimetic conformation and molecular recognition, and selectivity between different targets. The addition of these results to the existing knowledge in the filed could aid the generation of more potent and selective PPI inhibitors in the future.

Chapter 6. Experimental Section

6.1. General experimental points

All commercial solvents were purchased and used without further purification unless stated otherwise. Commercially available starting materials and reagents were obtained from Sigma-Aldrich, Alfa Aesar or Fisher Scientific. Amino acid derivatives, coupling reagents and resins were purchased from Novabiochem. Purification by column chromatography was carried out using silica gel (40-63 μm mesh size). Analytical thin layer chromatography (TLC) was conducted using Merck 0.2 mm silica gel 60 F_{254} pre-coated aluminium sheets. ^{1}H and ^{13}C NMR spectra were measured using a Bruker DRX 500 series spectrometer. Chemical shifts are expressed as parts per million using solvent as internal standard and coupling constants (*J*) are reported to the nearest 0.1 Hz. The following abbreviations are used: s for singlet, d for doublet, t for triplet, q for quartet and m for multiplet. High resolution mass spectrometry (HRMS) was carried out using a Bruker MicroTOF mass spectrometer or a Bruker Maxis impact mass spectrometer, in both cases under electro-spray ionisation (ESI) conditions. Infra-red spectra were recorded using a Perkin-Elmer Spectrum One FT-IR spectrophotometer. Elemental combustion analyses were performed by the School of Chemistry Microanalysis facility using a Carlo Erba Elemental Analyser MOD 1106 instrument and the found composition is reported to the nearest 0.05%. LC-MS experiments were run on a Bruker Daltonics HTCUltra™ series spectrometer and were run through a C18 column on a methanol/water gradient (0-95% MeCN over 3 minutes). Analytical HPLC experiments were run on an Agilent 1290 Infinity LC series spectrometer. Mass-directed preparative HPLC experiments were run using an Agilent 1260 Infinity Preparative system and analysed by a 6120 Quadrupole detector.

6.2. Numbering system for proteomimetic scaffolds

To simplify the numbering and NMR assignment of the different proteomimetic scaffolds, a sequential nomenclature and numbering system has been created, where each of the monomer building blocks is considered separately. Assignment is as follows:

- The naming proceeds from *N* to *C* terminus. Following this order, each monomer is assigned a number corresponding to its position on the chain. This number will be added as a prefix to the individual carbon number for differentiation.

- In the 3-*O*-alkylated and hybrid scaffolds the *O*-alkylated monomers are named as [*R*-(*n*-HABA)], where *R* is the alkoxy side-chain, *n*- indicates the position of the alkoxy moiety on the aromatic ring (*e.g.* for a 2-*O*-alkylated monomer *n* = 2) and HABA is the acronym for Hydroxy Amino Benzoic Acid.
- Each alkylated monomer is numbered using the standard system, where the substituents are assigned to the lowest number, in the case of originally symmetrical monomers one of the side chains is differentiated with an apostrophe (′) after the carbon number. Side-chain assignment follows a peptide nomenclature pattern in which the carbon attached to the alkoxy oxygen is assigned as Cα and the numbering of the aliphatic part of the side chain continues with Cβ, etc. In the case of aromatic side chains, the aromatic carbons are numbered CAr1, CAr2, etc. The Fmoc carbons are differentiated by the prefix F; the CH_2 group is numbered as CFα, the neighbouring CH is CFβ, and the aromatic positions go from CF-Ar2 to CF-Ar5.
- Amino acids are numbered using the standard convention.
- The numbering of the protons is based on the carbon numbering system.

Examples of the numbering system for the proteomimetic scaffolds and monomer building blocks are given below (Figure 6.1).

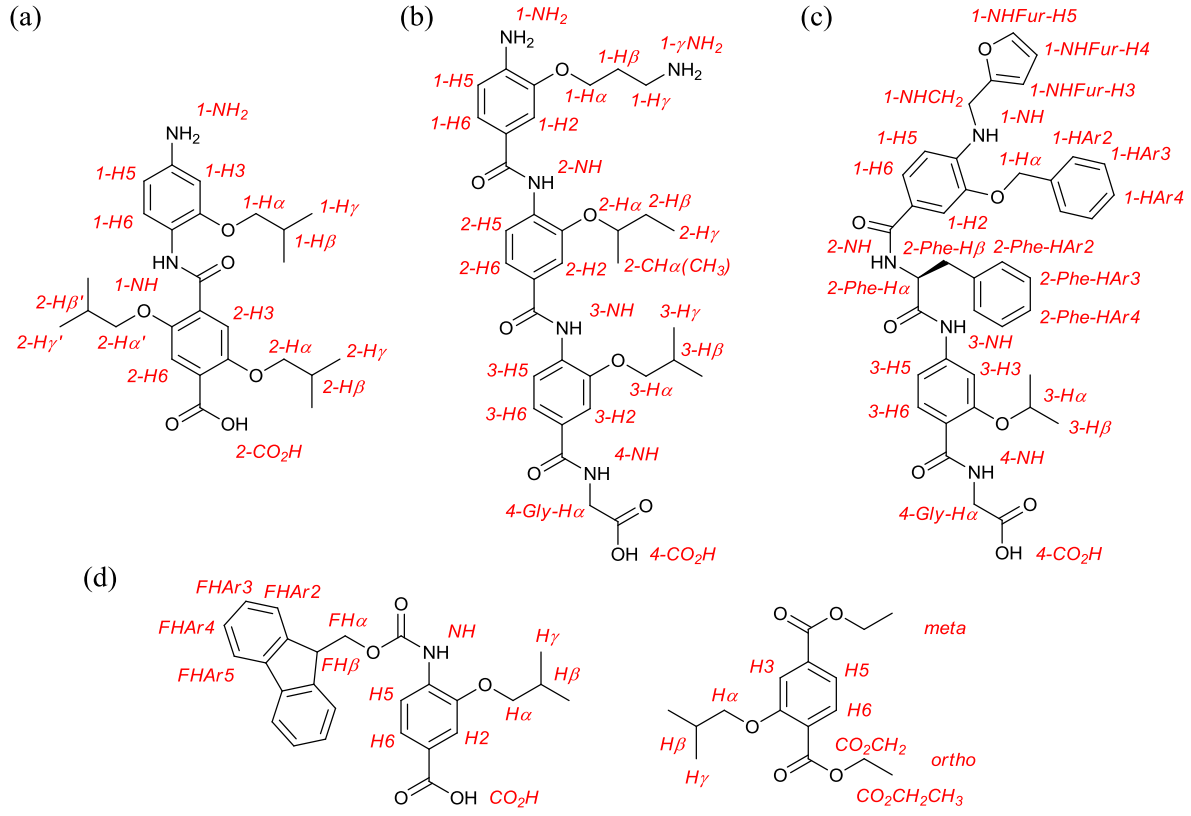
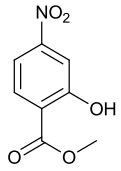


Figure 6.1 Numbering system (a) Bifacial scaffold (b) 3-*O*-alkylated scaffold (c) Hybrid scaffold (d) Monomer building blocks

6.3. Design, synthesis and conformational analyses of Bifacial Benzamide Based Foldmers *(Chapter 2)*

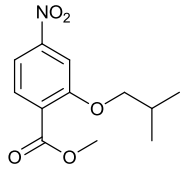
6.3.1. Monomer syntheses and characterisation

Methyl-2-hydroxy-4-nitrobenzoate (2.4)[127]



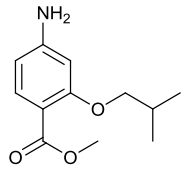
A stirred solution of 2-hydroxy-4-nitro benzoic acid **2.2** (10.0 g, 54.6 mmol) and concentrated sulfuric acid (1.0 mL, 18.8 mmol) in anhydrous methanol (200 mL) under an nitrogen atmosphere was heated at reflux. After 96 h stirring, the reaction mixture was concentrated to leave a pale yellow solid, which was poured into ethyl acetate, washed with water (2 × 100 mL) and the organic layer was dried over anhydrous Na_2SO_4. The organic solvent was removed under reduced pressure to leave the pure product **2.4** (10.6 g, 53.8 mmol, 98%) as a pale yellow powder; R_F 0.51 (30% ethyl acetate in cyclohexane); δ_H (300 MHz, MeOD) 7.97 (d, *J* = 5.4 Hz, 1H, *H*6), 7.65 (s, 1H, *H*3), 7.62 (d, *J* = 5.4 Hz, 1H, *H*5), 3.91 (s, 3H, CO_2CH_3); δ_C (75 MHz, $CDCl_3$) 170.7, 163.2, 153.8, 133.09, 119.36, 114.9, 113.8, 53.9; ν_{max}/cm^{-1} (solid state) 3627, 2965, 1733, 1661, 1665, 1558, 1440, 1387, 1237; ESI-HRMS found *m/z* 196.0253 $[M-H]^-$, $C_8H_6NO_5$ requires 197.0324; Found: C, 49.05; H, 3.65; N, 7.00; $C_8H_7NO_5$ requires: C, 48.74; H, 3.58; N, 7.10 %.

Methyl 2-isobutoxy-4-nitrobenzoate (2.6)[127]



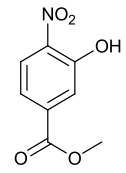
To a stirred solution of methyl-2-hydroxy-4-nitrobenzoate **2.4** (3.6 g, 18.3 mmol) and potassium carbonate (7.6 g, 54.8 mmol) in DMF (100 mL), isobutyl bromide (2.8 mL, 25.6 mmol) was added and the resulting mixture stirred at 50 °C during 20 h under a nitrogen atmosphere. The resultant suspension was allowed to cool, poured into water and extracted with ethyl acetate (3 × 150 mL). The combined organic fractions were washed with water (2 × 250 mL) and brine (4 × 300 mL) before being dried over $MgSO_4$. The organic solvent was evaporated resulting in an orange solid which was purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: ethyl acetate) to afford the product **2.6** (4.15 g, 16.4 mmol, 90%) as a bright yellow oil; R_F 0.86 (30% ethylacetate in dichloromethane); δ_H (500 MHz, MeOD) 7.77 (d, *J* = 8.5 Hz, 1 H, *H*6), 7.75 (s, 1 H, *H*3), 7.72 (d, *J* = 8.5 Hz 1 H, *H*5), 3.90 (s, 3 H, CO_2CH_3), 3.87 (d, *J* = 6 Hz, 2 H, *H*α), 2.14 – 2.09 (quin, *J* = 6 Hz, 1 H, *H*β), 1.07 (d, *J* = 6 Hz, 6H, *H*γ); δ_C (126 MHz, MeOD) 167.1 , 159.9, 152.1, 132.8, 127.5, 115.7, 108.8, 76.9, 53.0, 29.5, 19.6; ν_{max}/cm^{-1} (solid state) 3120, 2961, 1737, 1709, 1616, 1589, 1530 1489; ESI-HRMS found *m/z* 276.0853 $[M+Na]^+$, $C_{12}H_{15}NNaO_5$ requires 276.0842; Found: C, 57.15; H, 6.05; N, 5.45; $C_{12}H_{15}NO_5$ requires C, 56.91; H, 5.97; N, 5.53 %.

Methyl 4-amino-2-isobutoxybenzoate (2.10)[127]



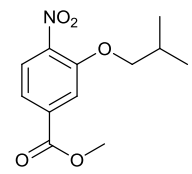
A solution containing methyl 2-isobutoxy-4-nitrobenzoate **2.6** (3.86 g, 15.2 mmol) in methanol (120 mL) and palladium on carbon (10 wt. %) was evacuated and flushed with nitrogen (3 times) and left under vacuum. Hydrogen was drawn into the flask and the reaction was left stirring at rt overnight. On completion, the reaction mixture was filtered through a celite pad and washed with methanol. The organic solvent was evaporated to dryness to yield the target product **2.10** (3.4 g, 15.2 mmol, quant.) as a grey gel; R_F 0.56 (30% ethylacetate in dichloromethane); δ_H (500 MHz, $CDCl_3$) 7.75 (d, J = 8.5 Hz, 1 H, H6), 6.21 (s, 1 H, H3), 6.18 (d, J = 8.5 Hz 1 H, H5), 4.10 (br, 2 H, NH_2), 3.84 (s, 3 H, CO_2CH_3), 3.72 (d, J = 6.5 Hz, 2 H, $H\alpha$), 2.19 – 2.11 (quin, J = 6.5 Hz, 1 H, $H\beta$), 1.04 (d, J = 6.5 Hz, 6 H, $H\gamma$); δ_C (126 MHz, $CDCl_3$) 166.6, 161.3, 152.1, 134.2, 109.2, 106.3, 98.7, 75.0, 51.3, 28.3, 19.3; ν_{max}/cm^{-1} (solid state) 3635, 3445, 2977, 1732, 1433, 1395, 1221; ESI-HRMS found m/z 246.1107 $[M+Na]^+$, $C_{12}H_{15}NNaO_5$ requires 246.1101.

Methyl 3-hydroxy-4-nitrobenzoate (2.3)[122]



Sulphuric acid (conc) (1.10 mL, 20.4 mmol) was added to a solution of 3-hydroxy-4-nitrobenzoic acid **2.1** (10.0 g, 54.6 mmol) in MeOH (100 mL) and the resulting mixture heated at reflux overnight under a nitrogen atmosphere. The reaction mixture was allowed to cool to rt and the resulting precipitate collected *via* filtration to afford the desired product **2.3** (10.37 g, 52.6 mmol, 96%) as a yellow solid; R_F: 0.61 (dichloromethane); δ_H (500 MHz, $CDCl_3$) 10.51 (s, 1 H, OH), 8.19 (d, J = 9.0 Hz, 1 H, H5), 7.84 (d, J = 1.5 Hz, 1 H, H2), 7.63 (dd, J = 9.0, 1.5 Hz, 1 H, H6), 3.98 (s, 3 H, CO_2CH_3); δ_C (125 MHz, $CDCl_3$) 164.9, 154.7, 138.0, 135.8, 125.3, 121.7, 120.6, 52.9; ν_{max}/cm^{-1} (solid state) = 3317, 2962, 1722, 1587, 1436, 1223, 743; ESI-HRMS found m/z 196.0253 $[M-H]^-$, $C_8H_6NO_5$ requires 197.0324; Found: C, 48.80; H, 3.60; N, 6.90; $C_8H_7NO_5$ requires: C, 48.74; H, 3.58; N, 7.10 %.

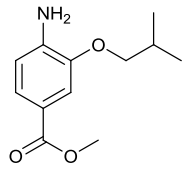
Methyl 3-isobutoxy-4-nitrobenzoate (2.5)[122]



To a stirred solution of methyl-3-hydroxy-4-nitrobenzoate **2.3** (2.00 g, 10.2 mmol) and potassium carbonate (3.52 g, 25.5 mmol) in DMF (30 mL), isobutyl bromide (1.40 mL, 12.5 mmol) was added and the resulting mixture stirred at 50 °C overnight under a nitrogen atmosphere. The reaction was incomplete, so a further aliquot of isobutyl bromide (600 µL, 5.34 mmol) was added. After 20 h stirring, the resultant suspension was allowed to cool, poured into water and extracted with ethyl acetate (3 × 50 mL); the combined organic fractions were washed with water (2 × 150 mL) and brine (4 × 200 mL) before being dried over $MgSO_4$. The organic solvents were evaporated

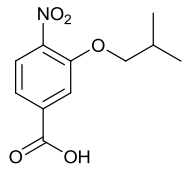
resulting in an orange solid **2.5** (2.09 g, 8.26 mmol, 81%); R_F: 0.65 (dichloromethane); δ_H (500 MHz, $CDCl_3$) 7.82 (d, *J* = 8.5 Hz, 1 H, *H*5), 7.73 (d, *J* = 1.5 Hz, 1 H, *H*2), 7.67 (dd, *J* = 8.5, 1.5 Hz, 1 H, *H*6), 3.97 (s, 3 H, CO_2CH_3), 3.93 (d, *J* = 6.5 Hz, 2 H, *H*α), 2.20 – 2.14 (m, 1 H, *H*β), 1.07 (d, *J* = 7.0 Hz, 6 H, *H*γ) ; δ_C (125 MHz, $CDCl_3$) 165.3, 152.1, 142.5, 134.7, 125.2, 121.1, 115.4, 76.0, 52.8, 28.2, 19.0; ν_{max}/cm^{-1} (solid state) = 3100, 2957, 1726, 1608, 1524, 1307, 1236, 750; ESI-HRMS found *m/z* 276.0841 $[M+Na]^+$, $C_{12}H_{15}NNaO_5$ requires 276.0842.

Methyl 4-amino-3-isobutoxybenzoate (2.9)[122]



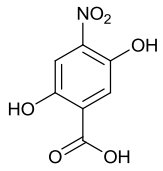
A solution containing methyl 3-isobutoxy-4-nitrobenzoate **2.5** (1.99 g, 7.86 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (40 mL) and palladium on carbon (10 wt. %) was evacuated and flushed with nitrogen (3 times) and left under vacuum. Hydrogen was drawn into the flask and the reaction was left stirring at rt overnight. On completion, the reaction mixture was filtered through a celite pad and washed with methanol and tetrahydrofuran. The organic solvents were evaporated to dryness to yield the target product **2.9** (1.66 mg, 7.44 mmol, 95%) as a beige solid; δ_H (500 MHz, $CDCl_3$) 7.54 (dd, *J* = 1.8, 8.0 Hz, 1 H, *H*6), 7.44 (d, *J* = 1.8 Hz, 1 H, *H*2), 6.67 (d, *J* = 8.0 Hz, 1 H, *H*5), 4.22 (s, broad, 2 H, NH_2), 3.87 (s, 3 H, CO_2CH_3), 3.83 (d, *J* = 6.5 Hz, 2 H, *H*α), 2.17 – 2.11 (m, 1 H, *H*β), 1.06 (d, *J* = 6.5 Hz, 6 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 167.4, 145.6, 141.3, 123.9, 119.5, 113.1, 112.1, 74.7, 51.7, 28.3, 19.3; ν_{max}/cm^{-1} (solid state) = 3461, 3343, 3201, 2951, 1687, 1622, 1270, 766; ESI-HRMS found *m/z* 224.1281 $[M+H]^+$, $C_{12}H_{18}NO_3$ requires 224.1281.

3-Isobutoxy-4-nitrobenzoic acid (2.7)[122]



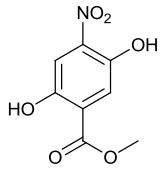
A solution of 10% aqueous sodium hydroxide (21 mL) was added to a solution of methyl 3-isobutoxy-4-nitrobenzoate **2.5** (3.50 g, 13.8 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (90 mL). On completion, the organic solvents were removed under reduced pressure and the remaining solution was poured in water (100 mL) and acidified *via* addition of hydrochloric acid (conc) to pH 1. The resulting suspension was extracted with dichloromethane (4 × 100 mL), the organic fractions were combined and washed with water (2 × 200 mL) followed by brine (200 mL) and dried over $MgSO_4$. The organic solvents were evaporated to yield the target material **2.7** (3.15 g, 13.5 mmol, 95%) as a yellow solid; δ_H (300 MHz, DMSO-d_6) 7.96 (d, *J* = 8.3 Hz, 1 H, *H*3), 7.74 (d, *J* = 1.3 Hz, 1 H, *H*6), 7.63 (dd, *J* = 1.3, 8.3 Hz, 1 H, *H*2), 4.01 (d, *J* = 6.4 Hz, 2 H, *H*α), 2.03 (m, 1 H, *H*β), 0.98 (d, *J* = 6.7 Hz, 6 H, *H*γ); δ_C (75 MHz, DMSO-d_6) 166.1, 151.4, 142.5, 136.0, 125.3, 121.5, 115.6, 75.5, 28.0, 19.0; ν_{max}/cm^{-1} (solid state) = 3088, 2531, 2342, 1817, 1693, 1310, 1015, 748; ESI-HRMS found *m/z* 238.0721 $[M-H]^-$, $C_{11}H_{12}NO_5$ requires 238.0721; Found C, 55.05; H, 5.45; N, 5.80%. $C_{11}H_{13}NO_5$ requires C, 55.23; H, 5.48; N, 5.86%.

2,5-Dihydroxy-4-nitrobenzoic acid (2.12)



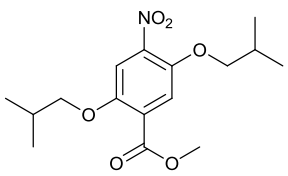
To a stirred solution of 3-hydroxy-4-nitrobenzoic acid **2.11** (40.00 g, 218.4 mmol) in 2 N aqueous sodium hydroxide solution (800 mL) was added dropwise a solution of potassium persulfate (59.00 g, 218.4 mmol) in water (1200 mL). The resulting solution was stirred at rt for 14 days. The reaction mixture was acidified *via* the addition of sulphuric acid (conc) to pH 1 and the resulting precipitate was removed by filtration. The aqueous solution was refluxed for 1 h. After cooling to rt the resulting precipitate was collected *via* filtration to yield the title compound **2.12** (16.11 g, 80.9 mmol, 37%) as gold microcrystals; δ_H (300 MHz, DMSO-d_6) 7.49 (s, 1 H, *H*6), 7.37 (s, 1 H, *H*3); δ_C (75 MHz, DMSO-d_6) 170.1, 152.3, 143.1, 141.7, 119.7, 119.1, 112.5; ν_{max}/cm^{-1} (solid state) = 3533, 3400, 2000, 1690, 1598, 1442, 1244, 760, 627; ESI-HRMS found *m/z* 198.0049 $[M-H]^-$, $C_7H_4NO_6$ requires 198.0044; Found C, 42.05; H, 2.35; N, 6.80%. $C_7H_5NO_6$ requires C, 42.22; H, 2.53; N, 7.03%.

Methyl 2,5-dihydroxy-4-nitrobenzoate (2.13)



To a stirred solution of 2,5-dihydroxy-4-nitrobenzoic acid **2.12** (5.00 g, 25.1 mmol) in methanol (200 mL) was added slowly concentrated sulphuric acid (2 mL) and the resulting solution was stirred at reflux overnight. The reaction mixture was allowed to cool to rt and sodium bicarbonate was added until carbon dioxide evolution ceased. The mixture was added to water (250 mL) and extracted with ethyl acetate (3 × 100 mL) and the combined organic fractions were washed with brine (100 mL). The organic solvent was removed by reduced pressure and the resulting orange solid was crystallised with chloroform to yield the title compound **2.13** (5.33 g, 25.0 mmol, quant.) as orange crystals; δ_H (300 MHz, $CDCl_3$) 10.19 (s, 1 H, 5-O*H*), 9.75 (s, 1H, 2-O*H*), 7.71 (s, 1 H, *H*3), 7.69 (s, 1 H, *H*6), 4.02 (s, 3 H, CO_2CH_3); δ_C (75 MHz, $CDCl_3$) 168.9, 153.5, 146.7, 137.5, 121.4, 120.8, 112.8, 53.7; ν_{max}/cm^{-1} (solid state) = 3359, 1695, 1440, 1220, 790; ESI-HRMS found *m/z* 212.0207 $[M-H]^-$, $C_8H_6NO_6$ requires 212.0273; Found C, 45.15; H, 3.25; N, 6.45%. $C_8H_7NO_6$ requires C, 45.08; H, 3.31; N, 6.57%.

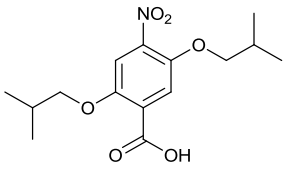
Methyl 2,5-diisobutoxy-4-nitrobenzoate (2.14)



To a stirred solution of methyl 2,5-dihydroxy-4-nitrobenzoate **2.13** (4.00 g, 18.8 mmol) and potassium carbonate (13.00 g, 93.8 mmol) in DMF (200 mL), isobutyl bromide (6.5 mL, 56.3 mmol) was added and the resulting mixture stirred at 50 °C overnight under a nitrogen atmosphere. On completion, the resulting suspension was allowed to cool to rt, poured into water and extracted with ethyl acetate (3 × 150 mL); the combined organic fractions were

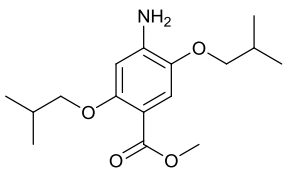
washed with water (2 × 250 mL) and brine (4 × 300 mL) before being dried over $MgSO_4$. The organic solvent was evaporated resulting in the title compound **2.14** (4.41 g, 13.6 mmol, 72%) was yielded as a yellow oil; δ_H (500 MHz, $CDCl_3$) 7.47 (s, 1 H, *H*3), 7.39 (s, 1 H, *H*6), 3.94 (s, 3 H, CO_2CH_3), 3.85 (d, *J* = 6.4 Hz, 2 H, *H*α2), 3.79 (d, *J* = 6.4 Hz, 2 H, *H*α5), 2.11 – 2.17 (m, 2 H, Hβ2, Hβ5), 1.04 – 1.07 (m, 12 H, Hγ2, Hγ5); δ_C (75 MHz, $CDCl_3$) 165.6, 151.6, 145.7, 141.7, 125.3, 118.0, 110.3, 76.6, 76.2, 52.5, 28.3, 28.3, 19.1, 19.0; ν_{max}/cm^{-1} (solid state) = 2960, 1739, 1529, 1392, 1217, 1024, 793; ESI-HRMS found *m/z* 348.1417 $[M+Na]^+$, $C_{16}H_{23}NNaO_6$ requires 348.1418.

2,5-Diisobutoxy-4-nitrobenzoic acid (2.15)



A solution of 10% aqueous sodium hydroxide (13 mL) was added to a solution of methyl 2,5-diisobutoxy-4-nitrobenzoate **2.14** (2.16 g, 8.0 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (50 mL). On completion, the organic solvents were removed under reduced pressure and the remaining solution was poured into water (100 mL) and acidified *via* addition of hydrochloric acid (conc) to pH 1. The resulting suspension was extracted with dichloromethane (4 × 100 mL), the organic fractions were combined and washed with water (2 × 200 mL) and brine (200 mL) before being dried over $MgSO_4$. The organic solvent was evaporated to yield the target material **2.15** (2.13 g, 6.84 mmol, 97%) as an amorphous yellow solid; δ_H (500 MHz, $CDCl_3$) 7.90 (s, 1 H, *H*6), 7.50 (s, 1 H, *H*3), 4.04 (d, *J* = 6.4 Hz, 2 H, *H*α2), 3.91 (d, *J* = 6.4 Hz, 2 H, *H*α5), 2.25 (m, 1 H, Hβ2), 2.14 (m, 1 H, Hβ5), 1.12 (m, 6 H, Hγ2), 1.05 (m, 6 H, Hγ5); δ_C (125 MHz, $CDCl_3$) 163.7, 150.3, 146.9, 142.4, 121.9, 119.5, 110.1, 77.6, 76.6, 28.2, 28.1, 19.1, 19.0; ν_{max}/cm^{-1} (solid state) = 3229, 2961, 1747, 1525, 1203, 1003, 803; ESI-HRMS found *m/z* 334.1267 $[M+Na]^+$, $C_{15}H_{21}NNaO_6$ requires 334.1261.

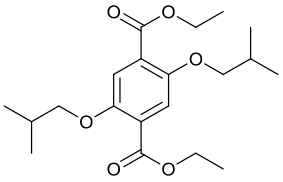
Methyl 4-amino-2,5-diisobutoxybenzoate (2.16)



To a stirred solution of methyl 2,5-diisobutoxy-4-nitrobenzoate **2.14** (2.24 g, 6.89 mmol) in ethyl acetate (50 mL), tin(II) chloride dihydrate (9.32 g, 41.33 mmol) was added and the resulting mixture stirred at 50 °C overnight under a nitrogen atmosphere. On completion, the reaction mixture was allowed to cool and poured over ice. The solution was basified to pH 8 by addition of a saturated sodium bicarbonate solution and the resulting mixture was allowed to stir for an hour. The aqueous mixture was extracted with ethyl acetate (3 × 100 mL) and the organic fractions were combined, washed with water (3 × 250 mL) and brine (2 × 250 mL), dried over $MgSO_4$ and the organic solvents were evaporated to afford the desired product **2.16** (1.83 g, 6.2 mmol, 90%) as a light brown solid; δ_H (500 MHz, $CDCl_3$) 7.33 (s, 1 H,

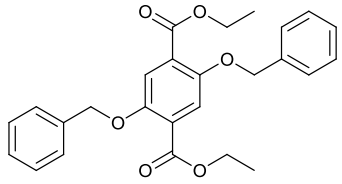
*H*6), 6.29 (s, 1 H, *H*3), 3.85 (s, 3 H, CO_2CH_3), 3.76 (d, *J* = 6.4 Hz, 2 H, *H*α2), 3.71 (d, *J* = 6.4 Hz, 2 H, *H*α5), 2.07 – 2.15 (m, 2 H, Hβ2, Hβ5), 1.03 – 1.06 (m, 12 H, Hγ2, Hγ5); δ_C (75 MHz, $CDCl_3$) 166.8, 155.7, 142.2, 139.7, 114.8, 107.7, 100.0, 76.0, 75.2, 51.5, 28.5, 28.4, 19.4, 19.3; ν_{max}/cm^{-1} (solid state) = 3492, 3368, 2957, 1704, 1621, 1523, 1445, 1252, 1210, 1035, 780; ESI-HRMS found *m/z* 318.1675 $[M+Na]^+$, $C_{16}H_{25}NNaO_4$ requires 318.1676.

Diethyl 2,5-diisobutoxyterephthalate (2.30)



To a stirred solution of diethyl 2,5-dihydroxyterephthalate **2.29** (2.02 g, 7.95 mmol) and potassium carbonate (5.00 g, 36.2 mmol) in DMF (50 mL), isobutyl bromide (2.12 mL, 18.9 mmol) was added and the resulting suspension stirred at 90 °C under a nitrogen atmosphere. After 18 h the reaction was shown to be incomplete, so a further aliquot of isobutyl bromide (1 mL, 8.91 mmol) was added. After 40 h stirring, the resulting suspension was poured into water and extracted with ethyl acetate (3 × 60 mL). The combined organic fractions were washed with water (2 × 140 mL) and brine (4 × 140 mL), then dried over $MgSO_4$. The solvent was removed under reduced pressure and the resulting oil purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: dichloromethane) to afford the desired product **2.30** (1.98 g, 5.41 mmol, 68%) as a colourless solid; R_F: 0.47 (dichloromethane); δ_H (500 MHz, $CDCl_3$) 7.34 (s, 2 H, *H*2), 4.39 (q, *J* = 7.5 Hz, 4 H, CO_2CH_2), 3.78 (d, *J* = 6.5 Hz, 4 H, *H*α), 2.14 – 2.09 (m, 2 H, *H*β), 1.40 (t, *J* = 7.5 Hz, 6 H, $CO_2CH_2CH_3$), 1.05 (d, *J* = 6.5 Hz, 12 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 166.3, 151.7, 124.6, 116.4, 76.0, 61.2, 28.4, 19.2, 14.3; ν_{max}/cm^{-1} (solid state) = 3069, 2953, 2496, 1694, 1422, 1216, 1022, 781; ESI-HRMS found *m/z* 389.1953 $[M+Na]^+$, $C_{20}H_{30}NaO_6$ requires 389.1935; Found: C, 65.55; H, 8.30; $C_{20}H_{30}O_6$ requires: C, 65.55; H, 8.25 %.

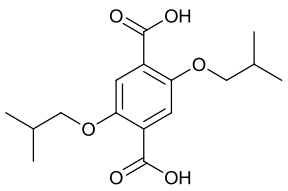
Diethyl 2,5-bis(benzyloxy)terephthalate (2.31)



To a stirred solution of diethyl 2,5-dihydroxyterephthalate **2.29** (4.00 g, 15.7 mmol) and potassium carbonate (10.00 g, 72.4 mmol) in DMF (80 mL), benzyl bromide (4.50 mL, 37.7 mmol) was added and the resulting suspension stirred at 90 °C under a nitrogen atmosphere overnight. The resulting suspension was poured into water and extracted with ethyl acetate (3 × 100 mL). The combined organic fractions were washed with water (2 × 300 mL) and brine (4 × 300 mL) before being dried over $MgSO_4$. The solvent was removed under reduced pressure to afford the desired product **2.31** (6.80 g, 15.7 mmol, quant.) as a colourless solid; δ_H (500 MHz, $CDCl_3$) 7.50 – 7.49 (m, 6 H, *H*2, *H*Ar2), 7.41 – 7.31 (m, 6 H, *H*Ar3, *H*Ar4), 5.15 (s, 4 H, *H*α), 4.39 (q, *J* = 7.0 Hz, 4 H, CO_2CH_2), 1.35 (t, *J* = 7.0 Hz, 6 H, $CO_2CH_2CH_3$); δ_C (125 MHz, $CDCl_3$) 165.8, 151.7, 136.6,

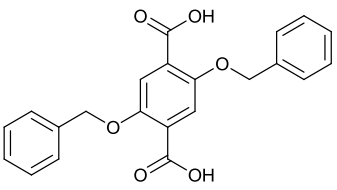
128.5, 127.9, 127.3, 125.2, 117.6, 71.8, 61.4, 14.2; (ν_{max}/cm^{-1} (solid state) = 2979, 1687, 1477, 1454, 1216, 1199, 1015, 738, 694; ESI-HRMS found *m/z* 435.1806 $[M+H]^+$, $C_{26}H_{27}O_6$ requires 435.1802.

2,5-Diisobutoxyterephthalic acid (2.32)



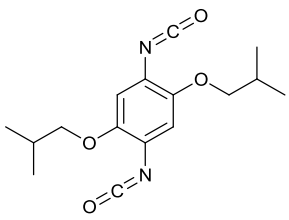
Diethyl 2,5-diisobutoxyterephthalate **2.30** (501 mg, 1.37 mmol) was dissolved in a 1:1 mixture of methanol : tetrahydrofuran (30 mL), a 10% sodium hydroxide solution (5 mL) was added and the resulting solution was stirred at rt overnight. The organic solvents were removed under reduced pressure and the remaining solution was poured into water (50 mL) and acidified *via* addition of hydrochloric acid (conc) to pH 1. The resulting suspension was extracted with dichloromethane (4 × 50 mL) and the organic fractions were combined and washed with water (2 × 150 mL), followed by brine (150 mL), and dried over $MgSO_4$. The organic solvents were evaporated to yield the target material **2.32** (365 mg, 1.18 mmol, 86%) as a colourless solid; δ_H (500 MHz, $CDCl_3$) 11.12 (s, broad, 2 H, CO_2H), 7.89 (s, 2 H, *H*2), 4.09 (d, *J* = 6.5 Hz, 4 H, *H*α), 2.27 – 2.22 (m, 2 H, *H*β), 1.11 (d, *J* = 6.5 Hz, 12 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 164.1, 151.8, 122.7, 117.4, 77.4, 28.1, 19.1; ν_{max}/cm^{-1} (solid state) = 2970, 1737, 1366, 1443, 1227, 1032, 760; ESI-HRMS found *m/z* 309.1360 $[M-H]^-$, $C_{16}H_{21}O_6$ requires 309.1344; Found: C, 62.20; H, 7.10; $C_{16}H_{22}O_6$ requires: C, 61.92; H, 7.15 %.

2,5-Bis(benzyloxy)terephthalic acid (2.33)



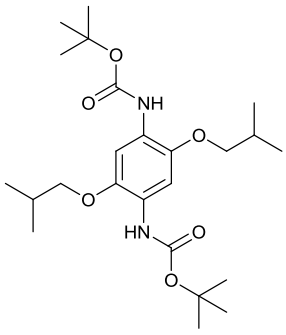
Diethyl 2,5-diisobutoxyterephthalate **2.31** (6.0 g, 13.7 mmol) was dissolved in a 1:1 mixture of methanol : tetrahydrofuran (200 mL) and a 10% sodium hydroxide solution (50 mL) was added, the resulting solution was stirred at rt overnight. The organic solvents were removed under reduced pressure and the remaining solution was poured into water (200 mL) and acidified *via* addition of hydrochloric acid (conc) to pH 1. The resulting suspension was extracted with tetrahydrofuran (4 × 150 mL); the organic fractions were combined and washed with brine (3 × 200 mL) before being dried over $MgSO_4$. The organic solvents were evaporated to yield the target material **2.33** (4.25 g, 11.2 mmol, 82%) as a colourless solid; δ_H (500 MHz, DMSO) 7.48 (d, *J* = 7.5 Hz, 4 H, *H*Ar2), 7.42 (s, 2 H, *H*2), 7.39 (app t, *J* = 7.5 Hz, 4 H, *H*Ar3), 7.32 (d, *J* = 7.5 Hz, 4 H, *H*Ar4), 5.17 (s, 4 H, *H*α); δ_C (125 MHz, DMSO) 166.7, 150.3, 137.0, 128.3, 127.7, 127.2, 125.7, 116.1, 70.5; ν_{max}/cm^{-1} (solid state) = 2859, 2644, 2528, 1673, 1444, 1378, 1220, 1021, 736; ESI-HRMS found *m/z* 377.1035 $[M-H]^-$, $C_{22}H_{17}O_6$ requires 377.1031.

1,4-Diisobutoxy-2,5-diisocyanatobenzene (2.34)



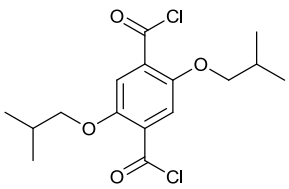
To a stirred solution of 2,5-diisobutoxyterephthalic acid **2.32** (125 mg, 0.40 mmol) in anhydrous toluene (30 mL), triethylamine (125 μL, 0.90 mmol) followed by diphenylphosphoryl azide (163 μL, 0.76 mmol) were added and the resulting solution was stirred for 2 h at rt under a nitrogen atmosphere, then the reaction mixture was left overnight without stirring. Ethyl acetate was added and the resulting solution was washed with a saturated sodium bicarbonate solution and brine before being dried over $MgSO_4$. The solvent was removed under reduced pressure and the resulting dark purple oil purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: ethyl acetate/hexane, 3:1) to afford the desired product **2.34** (50 mg, 0.41 mmol, 41%) as a dark solid; δ_H (500 MHz, $CDCl_3$) 6.55 (s, 2 H, *H*2), 3.74 (d, J = 6.5 Hz, 4 H, *H*α), 2.18 – 2.12 (m, 2 H, *H*β), 1.05 (d, J = 7.0 Hz, 12 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 147.0, 130.9, 120.8, 107.5, 75.8, 28.2, 19.2; ν_{max}/cm^{-1} (solid state) = 2957, 2254 (N=C=O stretch), 1540, 1450, 1220, 859.

tert-Butyl 2,5-diisobutoxy-1,4-phenylenedicarbamate (2.35)



A solution of 1,4-diisobutoxy-2,5-diisocyanatobenzene **2.34** (45 mg, 0.15 mmol) was dissolved in anhydrous *tert*-butanol (10 mL) and heated to reflux overnight under a nitrogen atmosphere. The solvent was evaporated under reduced pressure and the resulting oil purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: ethyl acetate/hexane, 1:3) to afford the desired product **2.35** (53 mg, 0.12 mmol, 79%) as a brown solid; δ_H (500 MHz, $CDCl_3$) 7.75 (s, 2 H, *H*2), 6.98 (s, 2 H, N*H*), 3.78 (d, J = 6.5 Hz, 4 H, *H*α), 2.13 – 2.05 (m, 2 H, *H*β), 1.53 (s, 18 H, $NHCO_2C(C\underline{H}_3)_3$), 1.04 (d, J = 6.5 Hz, 12 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 152.9, 140.9, 122.6, 103.3, 80.1, 75.8, 28.4, 28.4 19.4; ν_{max}/cm^{-1} (solid state) = 3442, 2963, 1724, 1541, 1428, 1232, 1153, 1052, 861; ESI-HRMS found *m/z* 453.2970 $[M+H]^+$, $C_{24}H_{41}N_2O_6$ requires 453.2959.

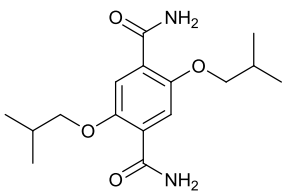
2,5-Diisobutoxyterephthaloyl dichloride (2.38)



To a stirred solution of 2,5-diisobutoxyterephthalic acid **2.32** (414 mg, 1.34 mmol) in dichloromethane (40 mL), thionyl chloride (500 μL, 6.89 mmol) was added and the resulting mixture was stirred at reflux overnight under a nitrogen atmosphere. The organic solvent and the thionyl chloride were co-evaporated under a nitrogen flow, this was repeated 3 times with the corresponding addition of further portions of dichloromethane, to yield a yellow solid **2.38**; δ_H (500 MHz, $CDCl_3$) 7.48 (s, 2 H, *H*2), 3.84 (d, J = 6.5 Hz, 4 H, *H*α), 2.19-2.13 (m, 2 H,

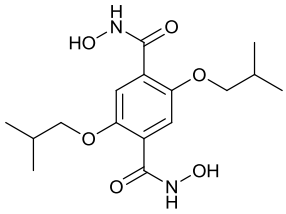
*H*β), 1.07 (d, *J* = 6.5 Hz, 12 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 163.6, 151.2, 128.3, 116.9, 76.2, 28.3, 19.1; ν_{max}/cm^{-1} (solid state) = 3127, 2963, 2875, 1770, 1492, 1386, 1226, 1154, 783. Found: C, 55.60; H, 5.75; Cl, 20.20; $C_{16}H_{20}Cl_2O_4$ requires: C, 55.34; H, 5.81; Cl, 20.42 %.

2,5-Diisobutoxyterephthalamide (2.39)



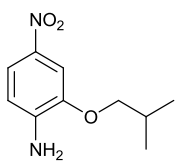
2,5-Diisobutoxyterephthaloyl dichloride **2.38** (279 mg, 0.81 mmol) was dissolved in diethyl ether (30 mL) and the resulting solution was cooled down to 0 °C in an ice bath. A 35% aqueous ammonia solution was added dropwise to the solution until the product precipitated, which was then collected *via* filtration to afford the desired product **2.39** as a colourless solid (951 mg, 2.20 mmol, 92%); δ_H (500 MHz, DMSO-d_6) 7.82 (s, broad, 2 H, N*H*), 7.66 (s, broad, 2 H, N*H*), 7.51 (s, 2 H, *H*2), 3.89 (d, *J* = 6.5 Hz, 4 H, *H*α), 2.11-2.09 (m, 2 H, *H*β), 1.00 (d, *J* = 6.5 Hz, 12 H, *H*γ); δ_C (125 MHz, DMSO-d_6) 165.1, 150.0, 125.0, 115.0, 75.4, 27.4, 18.5; ν_{max}/cm^{-1} (solid state) = 3409, 3184, 2954, 1645, 1435, 1217, 1036; ESI-HRMS found *m/z* 309.1810 $[M+H]^+$, $C_{16}H_{25}N_2O_4$ requires 309.1809.

***N*1, *N*4-Dihydroxy-2,5-diisobutoxyterephthalamide (2.37)**



To a solution of 2,5-diisobutoxyterephthalic acid **2.32** (250 mg, 0.81 mmol) in DMF (30 mL), 1,1'-carbonyldiimidazole (522 mg, 3.22 mmol) was added and the resulting solution stirred for 30 min at rt. Hydroxylammonium hydrochloride (448 mg, 6.45 mmol) was added and the reaction stirred overnight. Water (750 mL) was added to the reaction mixture and the resulting solution was stirred for 2 h. The precipitate was collected *via* filtration, washed with water and dried under vacuum to yield the target product **2.37** (130 mg, 0.38 mmol, 47%) as a beige solid; δ_H (500 MHz, DMSO-d_6) 10.53 (s, 2 H, O*H*), 9.19 (s, 2 H, N*H*), 7.20 (s, 2 H, *H*2), 3.80 (d, *J* = 6.5 Hz, 4 H, *H*α), 2.06-2.00 (m, 2 H, *H*β), 0.97 (d, *J* = 7.0 Hz, 12 H, *H*γ); δ_C (125 MHz, DMSO-d_6) 162.4, 149.6, 125.3, 114.4, 75.3, 27.6, 19.0; ν_{max}/cm^{-1} (solid state) = 3340, 2943, 1614, 1391, 1207, 1023, 805; ESI-HRMS found *m/z* 341.1711 $[M+H]^+$, $C_{16}H_{25}N_2O_6$ requires 341.1707; Found: C, 56.50; H, 7.15; N, 8.10; $C_{16}H_{24}N_2O_6$ requires: C, 56.46; H, 7.11; N, 8.23 %.

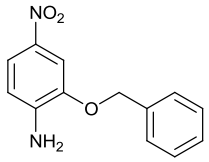
2-Isobutoxy-4-nitroaniline (2.41)



A solution of 2-amino-5-nitrophenol **2.40** (499 mg, 3.25 mmol) and potassium carbonate (1.12 g, 8.13 mmol) in DMF (15 mL) was stirred at 50 °C for 1 h. Isobutyl bromide (330 μL, 2.93 mmol) was added dropwise and the reaction was allowed to stir overnight, leading to complete conversion. The resulting suspension was poured into water and extracted with ethyl acetate (3 × 100 ml).

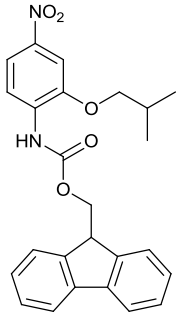
The combined organic fractions were washed with water (3 × 100 ml), brine (5 × 100 ml), and then dried over $MgSO_4$. The solvent was evaporated under reduced pressure and the resulting orange solid was dissolved in chloroform, filtered and purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: ethyl acetate) to afford the desired product **2.41** (430 mg, 2.05 mmol, 70%) as an orange solid; R_F: 0.71 (ethyl acetate); δ_H (500 MHz, $CDCl_3$) 7.81 (dd, *J* = 2.0, 8.8 Hz, 1 H, *H*5), 7.66 (d, *J* = 2.0 Hz, 1 H, *H*3), 6.65 (d, *J* = 8.8 Hz, 1 H, *H*6), 4.54 (s, broad, 2 H, N*H*$_2$), 3.86 (d, *J* = 6.5 Hz, 2 H, *H*α), 2.21 – 2.13 (m, 1 H, *H*β), 1.08 (d, *J* = 7.0 Hz, 6 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 145.0, 143.3, 138.8, 119.0, 111.8, 106.7, 75.1, 28.2, 19.3; ν_{max}/cm^{-1} (solid state) = 3505, 3392, 2915, 1615, 1312, 1235, 745; ESI-HRMS found *m/z* 211.1078 $[M+H]^+$, $C_{10}H_{15}N_2O_3$ requires 211.1077; Found: C, 57.15; H, 6.55; N, 13.45; $C_{10}H_{14}N_2O_3$ requires: C, 57.13; H, 6.71; N, 13.33 %.

2-(benzyloxy)-4-nitroaniline (2.42)



A solution of 2-amino-5-nitrophenol **2.40** (2.50 g, 16.2 mmol) and potassium carbonate (5.60 g, 40.5 mmol) in DMF (50 mL) was stirred at 50 °C for 1 h. Benzyl bromide (1.73 mL, 14.6 mmol) was added dropwise and the reaction was allowed to stir overnight. The resulting suspension was poured into water and extracted with ethyl acetate (3 × 150 ml); the combined organic fractions were washed with water (3 × 150 ml) and brine (5 × 150 ml) before being dried over $MgSO_4$. The solvent was evaporated under reduced pressure to afford the desired product **2.42** (3.60 g, 14.6 mmol, quant.) as an orange solid; R_F: 0.69 (ethyl acetate); δ_H (500 MHz, $CDCl_3$) 7.82 (dd, *J* = 2.0, 9.0 Hz, 1 H, *H*5), 7.78 (d, *J* = 2.0 Hz, 1 H, *H*3), 7.47 – 7.39 (m, 5 H, *H*Ar), 6.66 (d, *J* = 9.0 Hz, 1 H, *H*6), 5.16 (s, 4 H, *H*α), 4.58 (s, broad, 2 H, N*H*$_2$); δ_C (125 MHz, $CDCl_3$) 144.5, 143.5, 138.7, 135.8, 128.8, 128.6, 127.9, 119.4, 112.0, 107.3, 70.9; ν_{max}/cm^{-1} (solid state) = 3481, 3357, 1620, 1478, 1272, 1221, 1006, 754; ESI-HRMS found *m/z* 245.0920 $[M+H]^+$, $C_{13}H_{13}N_2O_3$ requires 245.0921.

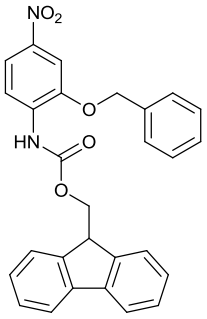
(9H-Fluoren-9-yl)methyl 2-isobutoxy-4-nitrophenylcarbamate (2.43)



A solution of 2-isobutoxy-4-nitroaniline **2.41** (505 mg, 2.40 mmol) and sodium bicarbonate (425 mg, 5.06 mmol) in tetrahydrofuran (50 mL) was stirred at reflux under a nitrogen atmosphere. A solution of 1-(9-fluorenyl)methylchloroformate (952 mg, 3.68 mmol) in tetrahydrofuran (10 mL) was added dropwise and the reaction allowed to stir at reflux overnight. The sodium bicarbonate was removed *via* hot filtration and the reaction mixture was allowed to cool to rt, concentrated under reduced pressure and the resulting precipitate collected *via* filtration to afford the desired product **2.43** (761 mg, 1.76 mmol, 74%) as a colourless solid; R_F: 0.86 (dichloromethane); δ_H (500 MHz, $CDCl_3$) 8.18 (s,

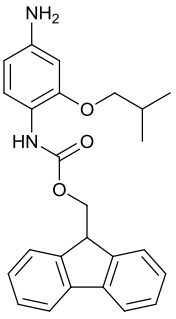
broad, 1 H, N*H*), 7.89 (d, *J* = 8.0 Hz, 1 H, *H*6), 7.80 (d, *J* = 7.5 Hz, 2 H, F*H*Ar5), 7.74 (m, 1 H, *H*5), 7.62 (d, *J* = 7.5 Hz, 2 H, F*H*Ar2), 7.53 (s, 1 H, *H*3), 7.45 – 7.42 (m, 2 H, F*H*Ar4), 7.36 – 7.33 (m, 2 H, F*H*Ar3), 4.56 (d, *J* = 6.8 Hz, 2 H, F*H*α), 4.34 (t, *J* = 6.8 Hz, 1 H, F*H*β), 3.93 (d, *J* = 6.5 Hz, 2 H, *H*α), 2.28 – 2.20 (m, 1 H, *H*β), 1.12 (d, *J* = 7.0 Hz, 6 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 152.7, 146.5, 143.5, 142.7, 141.4, 133.9, 128.0, 127.2, 125.0, 120.2, 117.6, 116.8, 106.2, 75.7, 67.7, 47.0, 28.1, 19.2; ν_{max}/cm^{-1} (solid state) = 3095, 2958, 1713, 1587, 1504, 1344, 1246, 734; ESI-HRMS found *m/z* 433.1767 $[M+H]^+$, $C_{25}H_{25}N_2O_5$ requires 433.1758; Found: C, 69.35; H, 5.50; N, 6.35; $C_{25}H_{24}N_2O_5$ requires: C, 69.43; H, 5.59; N, 6.48 %.

(9H-fluoren-9-yl)methyl 2-(benzyloxy)-4-nitrophenylcarbamate (2.44)



A solution of 2-(benzyloxy)-4-nitroaniline **2.42** (1.00 g, 4.09 mmol) and sodium bicarbonate (687 mg, 8.18 mmol) in tetrahydrofuran (90 mL) was stirred at reflux under a nitrogen atmosphere. A solution of 1-(9-fluorenyl)methylchloroformate (1.60 mg, 6.14 mmol) in tetrahydrofuran (20 mL) was added dropwise and the reaction allowed to stir at reflux overnight. The sodium bicarbonate was removed *via* hot filtration and the reaction mixture was allowed to cool to rt and concentrated under reduced pressure. The resulting precipitate was collected *via* filtration to afford the desired product **2.44** (1.54 g, 3.30 mmol, 81%) as a colourless solid; δ_H (500 MHz, DMSO) 8.25 (s, broad, 1 H, N*H*), 7.91 (d, *J* = 7.5 Hz, 2 H, F*H*Ar5), 7.88 – 7.83 (m, 2 H, *H*6, *H*5, H*3*), 7.76 (d, *J* = 7.5 Hz, 2 H, F*H*Ar2), 7.45 – 7.31 (m, 9 H, *H*Ar, F*H*Ar4, F*H*Ar3), 5.36 (s, 2 H, *H*α), 4.49 (d, *J* = 7.0 Hz, 2 H, F*H*α), 4.34 (t, *J* = 7.0 Hz, 1 H, F*H*β); δ_C (125 MHz, DMSO) 153.3, 148.1, 143.6, 142.7, 140.8, 136.1, 134.4, 128.5, 128.1, 127.7, 127.6, 127.1, 125.2, 120.2, 120.2, 116.9, 107.7, 70.4, 66.6, 46.4; ν_{max}/cm^{-1} (solid state) = 3362, 1715, 1551, 1489, 1339, 1216, 1029, 734; ESI-HRMS found *m/z* 467.1600 $[M+H]^+$, $C_{28}H_{23}N_2O_5$ requires 467.1601.

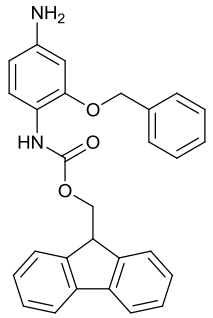
(9H-Fluoren-9-yl)methyl 4-amino-2-isobutoxyphenylcarbamate (2.45)



To a stirred solution of (9H-fluoren-9-yl)methyl 2-isobutoxy-4-nitrophenylcarbamate **2.43** (528 mg, 1.22 mmol) in ethyl acetate (15 mL) and tetrahydrofuran (3 mL), tin(II) chloride dihydrate (1.80 g, 7.04 mmol) was added and the resulting mixture stirred at 50 °C overnight under a nitrogen atmosphere. The reaction mixture was allowed to cool and poured over ice. The solution was basified to pH 8 by addition of a saturated sodium bicarbonate solution and the resulting basic mixture was allowed to stir for an hour. The aqueous mixture was extracted with ethyl acetate (3 × 100 mL) and the organic fractions were combined, washed with water (3 × 250 mL) and brine (2 × 250 mL) before being dried over $MgSO_4$. The organic solvents were evaporated and the resulting dark

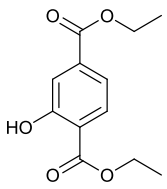
solid purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: dichloromethane to 1:1 dicholomethane / ethyl acetate) to afford the desired product **2.45** (378 mg, 0.94 mmol, 77%) as a purple solid; δ_H (500 MHz, $CDCl_3$) 7.80 – 7.78 (m, 3 H, F*H*Ar5 and *H*6), 7.64 (d, *J* = 5.5 Hz, 2 H, F*H*Ar2), 7.44 – 7.41 (m, 2 H, F*H*Ar4), 7.34 – 7.31 (m, 2 H, F*H*Ar3), 7.02 (s, broad, 1 H, N*H*), 6.32 – 6.29 (m, 2 H, *H*5 and *H*3), 4.46 (d, *J* = 6.8 Hz, 2 H, F*H*α), 4.33 (t, *J* = 6.8 Hz, 1 H, F*H*β), 3.76 (d, *J* = 6.5 Hz, 2 H, *H*α), 3.61 (s, broad, 2 H, NH_2), 2.19 – 2.14 (m, 1 H, *H*β), 1.08 (d, *J* = 6.5 Hz, 6 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 153.7, 148.9, 144.0, 142.8, 141.4, 127.7, 127.1, 125.2, 120.4, 120.0, 119.1, 107.1, 99.8, 74.9, 67.0, 47.2, 28.3, 19.3; ν_{max}/cm^{-1} (solid state) = 3312, 2960, 1705, 1534, 1448, 1224, 738; ESI-HRMS found *m/z* 403.2019 $[M+H]^+$, $C_{25}H_{27}N_2O_3$ requires 403.2016; Found: C, 74.20; H, 6.40; N, 6.80; $C_{25}H_{26}N_2O_3$ requires: C, 74.60; H, 6.51; N, 6.96 %.

(9H-fluoren-9-yl)methyl 4-amino-2-(benzyloxy)phenylcarbamate (2.46)



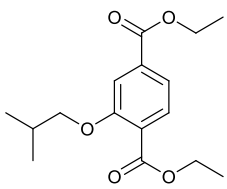
Tin(II) chloride dihydrate (3.28 g, 12.8 mmol) was added to a stirred solution of (9H-fluoren-9-yl)methyl 2-(benzyloxy)-4-nitrophenylcarbamate **2.44** (1.00 g, 2.14 mmol) in 1:1 mixture of ethyl acetate : tetrahydrofuran (100 mL) and the resulting mixture stirred at 50 °C overnight under a nitrogen atmosphere. The reaction mixture was allowed to cool to rt and poured over ice. The solution was basified to pH 8 by addition of a saturated sodium bicarbonate solution and the resulting basic mixture was allowed to stir for an hour. The aqueous suspension was extracted with ethyl acetate (3 × 100 mL) and the organic fractions were combined, washed with water (3 × 250 mL) and brine (2 × 250 mL) before being dried over $MgSO_4$. The organic solvent was evaporated and the resulting dark solid purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: dichloromethane/ethyl acetate, 9:1) to afford the desired product **2.46** (335 mg, 0.77 mmol, 36%) as a purple solid; δ_H (500 MHz, $CDCl_3$) 7.80 – 7.77 (m, 3 H, F*H*Ar5 and *H*6), 7.61 (d, *J* = 6.0 Hz, 2 H, F*H*Ar2), 7.43 – 7.28 (m, 9 H, *H*Ar, F*H*Ar4, F*H*Ar3), 7.01 (s, broad, 1 H, N*H*), 6.35 – 6.32 (m, 2 H, *H*5 and *H*3), 5.09 (s, 2 H, *H*α), 4.45 (d, *J* = 7.0 Hz, 2 H, F*H*α), 4.30 (t, *J* = 7.0 Hz, 1 H, F*H*β), 3.57 (s, broad, 2 H, NH_2); δ_C (125 MHz, $CDCl_3$) 153.7, 148.7, 144.0, 142.9, 141.3, 136.6, 128.8, 128.3, 127.7, 127.4, 127.1, 125.2, 120.9, 120.0, 119.3, 107.7, 100.3, 70.8, 67.0, 47.2; ν_{max}/cm^{-1} (solid state) = 3300, 3030, 1690, 1531, 1222, 1041, 730; ESI-HRMS found *m/z* 437.1868 $[M+H]^+$, $C_{28}H_{25}N_2O_3$ requires 437.1860.

Diethyl 2-hydroxyterephthalate (2.48)



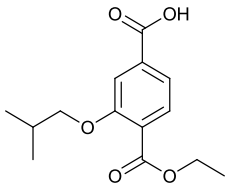
p-Toluenesulfonic acid (200 mg, 1.05 mmol) was added to a stirred solution of 2-hydroxyterephthalic acid **2.47** (2.00 g, 10.98 mmol) in anhydrous ethanol (125 mL) and the resulting solution refluxed at 90 °C under a nitrogen atmosphere. After 3 days the reaction was incomplete, so a further portion of *p*-toluenesulfonic acid (50 mg, 0.26 mmol) and anhydrous ethanol (50 mL) were added. After 3 more days, the organic solvent was evaporated under reduced pressure, the solid residue was redisolved in water and the resulting mixture basified with a saturated solution of $NaHCO_3$ to pH 8. The solution was extracted with ethyl acetate (3 × 60 mL); the combined organic fractions were washed with water (2 × 60 mL) and brine (1 × 60 mL) before being dried over $MgSO_4$. The solvent was removed under reduced pressure to afford the desired product **2.48** (2.25 g, 9.45 mmol, 86%) as light brown oil; δ_H (500 MHz, $CDCl_3$) 10.85 (s, 1 H, O*H*), 7.91 (d, *J* = 8.5 Hz, 1 H, *H*6), 7.65 (d, *J* = 1.5 Hz, 1 H, *H*3), 7.53 (dd, *J* = 1.5, 8.5 Hz, 1 H, *H*5), 4.45 (q, *J* = 7.0 Hz, 2 H, CO_2CH_2 *meta*), 4.39 (q, *J* = 7.0 Hz, 2 H, CO_2CH_2 *ortho*), 1.45 (t, *J* = 7.0 Hz, 3 H, $CO_2CH_2CH_3$ *meta*), 1.41 (t, *J* = 7.0 Hz, 3 H, $CO_2CH_2CH_3$ *ortho*); δ_C (125 MHz, $CDCl_3$) 169.6, 165.6, 161.4, 136.7, 130.0, 119.6, 118.8, 115.9, 61.9, 61.5, 14.2, 14.2; ν_{max}/cm^{-1} (solid state) = 3139, 2983, 1722, 1675, 1294, 1203, 1099, 754; ESI-HRMS found *m/z* 237.0727 $[M-H]^-$, $C_{12}H_{13}O_5$ requires 237.0768.

Diethyl 2-isobutoxyterephthalate (2.49)



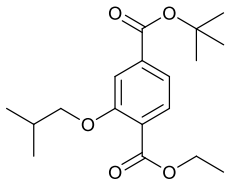
Isobutyl bromide (566 µL, 5.04 mmol) was added to a stirred solution of diethyl 2-hydroxyterephthalate **2.48** (1.00 g, 4.20 mmol) and potassium carbonate (1.28 g, 9.24 mmol) in DMF (20 mL), and the resulting suspension stirred at 50 °C overnight under a nitrogen atmosphere. The reaction was incomplete, so a further aliquot of isobutyl bromide (200 µL, 1.78 mmol) was added. After 15 h stirring, the resulting suspension was poured into water and extracted with ethyl acetate (3 × 60 mL); the combined organic fractions were washed with water (2 × 140 mL) and brine (4 × 140 mL) before being dried over $MgSO_4$. The solvent was removed under reduced pressure to afford the desired product **2.49** (1.17 g, 3.98 mmol, 94%) as a yellow oil**;** δ_H (500 MHz, $CDCl_3$) 7.78 (d, *J* = 8.0 Hz, 1 H, *H*6), 7.62 (dd, *J* = 1.5, 8.0 Hz, 1 H, *H*5), 7.59 (d, *J* = 1.5 Hz, 1 H, *H*3), 4.40 (q, *J* = 7.0 Hz, 2 H, CO_2CH_2 *meta*), 4.39 (q, *J* = 7.0 Hz, 2 H, CO_2CH_2 *ortho*), 3.87 (d, *J* = 6.5 Hz, 2 H, *H*α), 2.20 – 2.12 (m, 1 H, *H*β), 1.42 (t, *J* = 7.0 Hz, 3 H, $CO_2CH_2CH_3$ *meta*), 1.40 (t, *J* = 7.0 Hz, 3 H, $CO_2CH_2CH_3$ *ortho*), 1.08 (d, *J* = 6.5 Hz, 6 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 166.3, 165.9, 158.1, 134.5, 131.2, 124.9, 120.9, 113.6, 75.3, 61.4, 61.2, 28.4, 19.2, 14.3 (2 C); ESI-HRMS found *m/z* 295.1538 $[M+H]^+$, $C_{16}H_{23}O_5$ requires 295.1540.

4-(Ethoxycarbonyl)-3-isobutoxybenzoic acid (2.50)



Lithium hydroxide monohydrate (286 mg, 6.82 mmol) was dissolved in the minimum quantity of water and added to a solution of diethyl 2-isobutoxyterephthalate **2.49** (2.23 g, 7.58 mmol) in tetrahydrofuran (50 mL). The reaction was stirred at rt for 18 h; further portions of lithium hydroxide monohydrate were added to achieve completion. The mixture was acidified with a 10% solution of potassium bisulphate to pH 4 and was extracted with ethyl acetate (3 × 60 mL); the combined organic fractions were washed with water (2 × 60 mL) and brine (1 × 60 mL) before being dried over $MgSO_4$. The solvent was removed under reduced pressure and the solid residue subjected to high pressure liquid chromatography [(50-95% MeCN:water and 0.01% Formic acid) t = 8 min, XBridge Prep C18 column] to isolate the title compound **2.50** (1.00 g, 3.76 mmol, 50%) as a colourless solid; δ_H (500 MHz, $CDCl_3$) 7.81 (d, *J* = 8.0 Hz, 1 H, *H*6), 7.71 (dd, *J* = 1.0, 8.0 Hz , 1 H, *H*5), 7.59 (d, *J* = 1.0 Hz, 1 H, *H*3), 4.40 (q, *J* = 7.0 Hz, 2 H, CO_2CH_2), 3.87 (d, *J* = 6.5 Hz, 2 H, *H*α), 2.21 – 2.13 (m, 1 H, *H*β), 1.41 (t, *J* = 7.0 Hz, 3 H, $CO_2CH_2CH_3$), 1.08 (d, *J* = 7.0 Hz, 6 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 171.1, 166.3, 158.1, 133.1, 131.3, 125.9, 121.6, 114.0, 75.3, 61.3, 28.4, 19.2, 14.3; (ν_{max}/cm^{-1} (solid state) = 2979, 2952, 2872, 1718, 1687, 1243, 1080, 764; ESI-HRMS found *m/z* 265.1085 $[M-H]^-$, $C_{14}H_{17}O_5$ requires 265.1081.

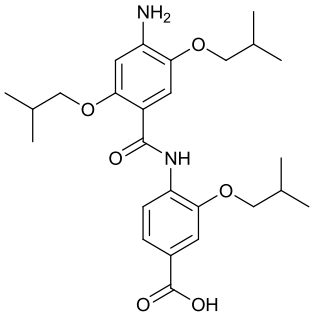
4-tert-butyl 1-ethyl 2-isobutoxyterephthalate (2.51)



Under nitrogen atmosphere, a 20% solution of Ghosez's reagent in anhydrous chloroform (1.12 mL, 1.70 mmol) was added to a solution of 4-(ethoxycarbonyl)-3-isobutoxybenzoic acid **2.50** (300 mg, 1.13 mmol) in anhydrous chloroform (30 mL). The reaction was stirred at 50 °C for 3 h before *tert*-butanol (540 µL, 5.68 mmol) was added. The resulting mixture was stirred at 50 °C overnight. The organic solvents were evaporated and the resulting oil purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: hexane / chloroform / ethyl acetate, 75:25:5) to afford the desired product **2.51** (215 mg, 0.67 mmol, 59%) as a yellow oil; δ_H (500 MHz, $CDCl_3$) 7.76 (d, *J* = 7.5 Hz, 1 H, *H*6), 7.56 (d, *J* = 1.0 Hz, 1 H, *H*3), 7.55 (dd, *J* = 7.5, 1.0 Hz , 1 H, *H*5), 4.38 (q, *J* = 7.0 Hz, 2 H, CO_2CH_2), 3.85 (d, *J* = 6.5 Hz, 2 H, *H*α), 2.11 – 2.19 (m, 1 H, *H*β), 1.61 (s, 9 H, $CO_2C(CH_3)_3$), 1.39 (t, *J* = 7.0 Hz, 3 H, $CO_2CH_2CH_3$), 1.06 (d, *J* = 6.5 Hz, 6 H, *H*γ); δ_C (125 MHz, $CDCl_3$) 166.4, 165.1, 158.2, 136.4, 131.1, 124.4, 120.7, 113.5, 81.7, 75.2, 61.1, 28.3, 28.1, 19.2, 14.3; (ν_{max}/cm^{-1} (solid state) = 2960, 2931, 2882, 1713, 1367, 1241, 1145, 962, 745; ESI-HRMS found *m/z* 323.1855 $[M+H]^+$, $C_{18}H_{27}O_5$ requires 323.1853.

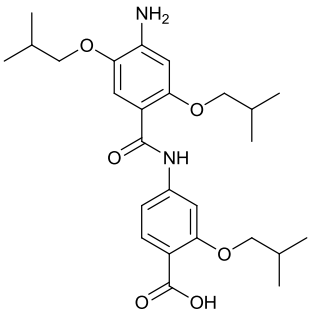
6.3.2. Dimer Syntheses and Characterisation

4-(4-amino-2,5-diisobutoxybenzamido)-3-isobutoxybenzoic acid (2.27)



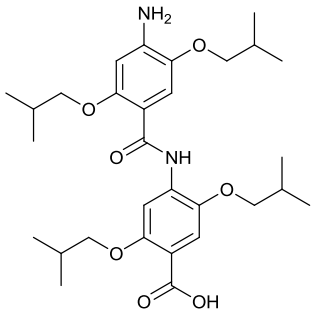
Under a nitrogen atmosphere, thionyl chloride (348 μL, 4.8 mmol, 5 equiv) was added to a stirred solution of 2,5-diisobutoxy-4-nitrobenzoic acid **2.15** (299 mg, 0.96 mmol, 1 equiv) in anhydrous dichloromethane (50 mL / g), and the resulting mixture was stirred at reflux overnight. The organic solvent and the excess thionyl chloride were co-evaporated under a nitrogen flow; this was repeated 3 times with further additions of dichloromethane to yield a yellowish solid. The resulting 2,5-diisobutoxy-4-nitrobenzoyl chloride (318 mg, 0.96 mmol, 1 equiv) was dissolved in anhydrous chloroform (50 mL / g) and methyl 4-amino-3-isobutoxybenzoate **2.9** (214 mg, 0.96 mmol, 1 equiv) was added. The resulting mixture was stirred at reflux under a nitrogen atmosphere. The organic solvent was evaporated under reduced pressure and the solid residue dissolved, without further isolation of compound **2.23**, in a 1:1 mixture of methanol : tetrahydrofuran (30 mL / g) and palladium on carbon (10 wt. %). The flask was evacuated and flushed with nitrogen (3 times) and left under vacuum, then hydrogen was drawn into the flask and the reaction was left stirring at rt overnight. On completion, the reaction mixture was filtered through a celite pad and washed with methanol and tetrahydrofuran. The resulting solid was reacted, without further purification of compound **2.25**, with a 10% sodium hydroxide solution (10 mL / g) in tetrahydrofuran (100 mL / g) at rt. On completion, the organic solvent was evaporated under reduced pressure and the solid residue subjected to HPLC [(50-95% MeCN:water and 0.1% Formic acid) t = 8 min, 20 mL min^{-1}, XBridge Prep C18 column] to yield the target dimer **2.27** (53% overall yield) as a beige amorphous solid; δ_H (500 MHz, $CDCl_3$) 10.4 (s, broad, 1 H, 2-N*H*), 8.72 (d, 1 H, *J* = 7.5 Hz, 2-*H*5), 7.80 (dd, *J* = 1.5, 7.5 Hz, 1 H, 2-*H*6), 7.66 (s, 1 H, 1-*H*6), 7.62 (d, *J* = 1.5 Hz, 1 H, 2-*H*2), 6.40 (s, 1 H, 1-*H*3), 3.93 (d, *J* = 6.5 Hz, 2 H, 2-*H*α), 3.92 (d, *J* = 7.0 Hz, 2 H, 1-*H*α'), 3.84 (d, *J* = 6.5 Hz, 2 H, 1-*H*α), 2.23 – 2.09 (m, 3 H, 1-*H*β, 1-*H*β', 2-*H*β), 1.04 (d, *J* = 6.5 Hz, 6 H, 2-*H*γ), 1.03 (d, *J* = 7.0 Hz, 6 H, 1-*H*γ), 0.98 (d, *J* = 6.5 Hz, 6 H, 1-*H*γ'); δ_C (125 MHz, $CDCl_3$) 164.3, 152.7, 147.5, 146.1, 141.1, 139.4, 134.3, 131.2, 124.2, 123.1, 120.0, 119.2, 114.1, 112.6, 75.6, 75.6, 75.2, 28.3, 28.0, 27.8, 19.3, 19.3, 19.2; ν_{max}/cm^{-1} (solid state) = 3490, 3328, 2957, 2872, 1587, 1519, 1261, 1206, 1030, 768; ESI-HRMS found *m/z* 473.2663 $[M-H]^-$, $C_{26}H_{35}N_2O_6$ requires 473.2646.

4-(4-amino-2,5-diisobutoxybenzamido)-2-isobutoxybenzoic acid (2.28)



Under a nitrogen atmosphere, Ghosez's reagent (120 μL, 0.91 mmol, 0.95 equiv) was added to a solution containing 2,5-diisobutoxy-4-nitrobenzoic acid **2.15** (299 mg, 0.96 mmol, 1 equiv) in chloroform (40 mL / g) and the resulting mixture was refluxed for 3 h. Methyl 4-amino-2-isobutoxybenzoate **2.10** (214 mg, 0.96 mmol, 1 equiv) was subsequently added and heated at reflux overnight. The solvents were removed under reduced pressure and the resulting mixture dissolved, without further isolation of compound **2.24**, in a 1:1 mixture of methanol : tetrahydrofuran (30 mL / g) and palladium on carbon (10 wt. %). The flask was evacuated and flushed with nitrogen (3 times) and left under vacuum, then hydrogen was drawn into the flask and the reaction was left stirring at rt overnight. On completion, the reaction mixture was filtered through a celite pad and washed with methanol and tetrahydrofuran. The resulting solid was reacted, without further purification of compound **2.26**, with a 10% sodium hydroxide solution (10 mL / g) in tetrahydrofuran (100 mL / g) at rt. On completion, the organic solvent was evaporated under reduced pressure and the solid residue subjected to HPLC [(50-95% MeCN:water and 0.1% Formic acid) t = 8 min, 20 mL min^{-1}, XBridge Prep C18 column] to isolate the title compound **2.28** (35% overall yield) as a light brown amorphous solid; δ_H (500 MHz, $CDCl_3$) 10.34 (s, broad, 1 H, 2-N*H*), 8.31 (d, *J* = 1.5 Hz, 1 H, 2-*H*3), 8.11 (d, *J* = 8.5, 1 H, 2-*H*6), 7.66 (s, 1 H, 1-*H*6), 6.71 (dd, *J* = 1.5, 8.5 Hz, 1 H, 2-*H*5), 6.45 (s, 1 H, 1-*H*3), 4.10 (d, *J* = 7.5 Hz, 2 H, 2-*H*α), 3.91 (d, *J* = 6.5 Hz, 2 H, 1-*H*α), 3.84 (d, *J* = 7.0 Hz, 2 H, 1-*H*α'), 2.32 – 2.21 (m, 2 H, 2-*H*β, 1-*H*β), 2.17 – 2.09 (m, 1 H, 1-*H*β'), 1.15 (d, *J* = 7.0 Hz, 6 H, 1-*H*γ), 1.11 (d, *J* = 7.0 Hz, 6 H, 2-*H*γ), 1.04 (d, *J* = 6.5 Hz, 6 H, 1-*H*γ'); δ_C (125 MHz, $CDCl_3$) 165.3, 164.4, 158.9, 153.0, 145.3, 142.1, 140.9, 134.1, 113.8, 112.1, 111.9, 109.4, 103.6, 98.5, 76.4, 76.4, 75.3, 28.6, 28.3, 28.2, 19.5, 19.3, 19.3; ν_{max}/cm^{-1} (solid state) = 3466, 3341, 2959, 2874, 1722, 1581, 1514, 1223, 1015, 760; ESI-HRMS found *m/z* 473.2659 $[M+H]^+$, $C_{26}H_{37}N_2O_6$ requires 473.2646.

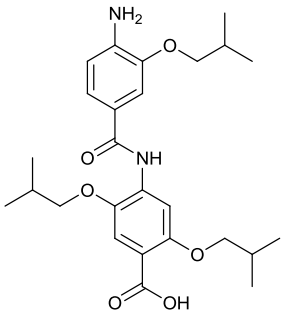
4-(4-amino-2,5-diisobutoxybenzamido)-2,5-diisobutoxybenzoic acid (2.22)



Under a nitrogen atmosphere, Ghosez's reagent (66 μL, 0.50 mmol, 1 equiv) was added to a solution containing 2,5-diisobutoxy-4-nitrobenzoic acid **2.15** (155 mg, 0.50 mmol, 1 equiv) in chloroform (40 mL / g) and the resulting mixture was refluxed for 3 h. Methyl 4-amino-2,5-diisobutoxybenzoate **2.16** (148 mg, 0.50 mmol, 1 equiv) was subsequently added and heated at reflux overnight. The solvents were removed under reduced pressure and the resulting

mixture dissolved, without further isolation of compound **2.18**, in a 1:1 mixture of methanol : tetrahydrofuran (30 mL / g) and palladium on carbon (10 wt. %). The flask was evacuated and flushed with nitrogen (3 times) and left under vacuum, then hydrogen was drawn into the flask and the reaction was left stirring at rt overnight. On completion, the reaction mixture was filtered through a celite pad and washed with methanol and tetrahydrofuran. The resulting solid was reacted, without further purification of compound **2.20**, with a 10% sodium hydroxide solution (10 mL / g) in tetrahydrofuran (100 mL / g) at rt. On completion, the organic solvent was evaporated under reduced pressure and the solid residue subjected to HPLC [(50-95% MeCN:water and 0.1% Formic acid) t = 8 min, 20 mL min^{-1}, XBridge Prep C18 column] to isolate the title compound **2.22** (11% overall yield) as a brownish amorphous solid; δ_H (500 MHz, $CDCl_3$) 10.52 (s, broad, 1 H, 2-N*H*), 8.62 (s, 1 H, 2-*H*3), 7.66 (s, 1 H, 2-*H*6), 7.61 (s, 1 H, 1-*H*6), 6.43 (s, 1 H, 1-*H*3), 4.09 (d, *J* = 7.0 Hz, 2 H, 2-*H*α), 3.97 (d, *J* = 7.0 Hz, 2 H, 1-*H*α), 3.89 (d, *J* = 7.0 Hz, 2 H, 2-*H*α'), 3.83 (d, *J* = 6.5 Hz, 2 H, 1-*H*α'), 2.27 – 2.09 (m, 4 H, 1-*H*β, 1-*H*β', 2-*H*β, 2-*H*β'), 1.09 (d, *J* = 6.5 Hz, 6 H, 2-*H*γ), 1.04 (d, *J* = 6.5 Hz, 6 H, 1-*H*γ'), 1.02 (d, *J* = 6.5 Hz, 6 H, 2-*H*γ'), 0.97 (d, *J* = 7.0 Hz, 6 H, 1-*H*γ); δ_C (125 MHz, $CDCl_3$) 166.7, 164.6, 152.8, 152.7, 142.5, 141.4, 135.1, 114.6, 113.8, 110.5, 105.0, 100.6, 77.2, 76.9, 76.0, 75.3, 28.3, 28.3, 28.0, 27.7, 19.3, 19.3, 19.2, 19.2 (two quaternary carbons were not observed); ν_{max}/cm^{-1} (solid state) = 3489, 3329, 2958, 2872, 1587, 1582, 1259, 1199, 1025, 765; ESI-HRMS found *m/z* 545.3237 $[M-H]^-$, $C_{30}H_{44}N_2O_7$ requires 545.3221.

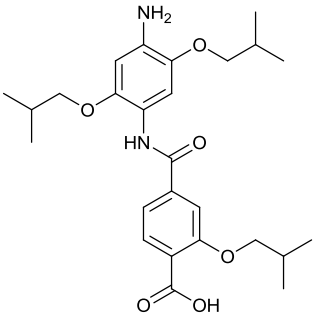
4-(4-amino-3-isobutoxybenzamido)-2,5-diisobutoxybenzoic acid (2.21)



Under a nitrogen atmosphere, Ghosez's reagent (145 µL, 1.08 mmol, 0.95 equiv) was added to a solution containing 3-isobutoxy-4-nitrobenzoic acid **2.7** (250 mg, 1.12 mmol, 1 equiv) in chloroform (40 mL / g) and the resulting mixture was refluxed for 3 h. Methyl 4-amino-2,5-diisobutoxybenzoate **2.16** (331 mg, 1.12 mmol, 1 equiv) was subsequently added and heated at reflux overnight. The solvents were removed under reduced pressure and the resulting mixture dissolved, without further isolation of compound **2.17**, in a 1:1 mixture of methanol : tetrahydrofuran (30 mL / g) and palladium on carbon (10 wt. %). The flask was evacuated and flushed with nitrogen (3 times) and left under vacuum, then hydrogen was drawn into the flask and the reaction was left stirring at rt overnight. On completion, the reaction mixture was filtered through a celite pad and washed with methanol and tetrahydrofuran. The resulting solid was reacted, without further purification of compound **2.19**, with a 10% sodium hydroxide solution (10 mL / g) in tetrahydrofuran (100 mL / g) at rt. On completion, the organic solvent was evaporated under reduced pressure and the solid residue subjected to HPLC [(50-95%

MeCN:water and 0.1% Formic acid) t = 8 min, 20 mL min^{-1}, XBridge Prep C18 column] to isolate the title compound **2.21** (36 % overall yield) as a beige amorphous solid; δ_H (500 MHz, $CDCl_3$) 8.84 (s, broad, 1 H, 2-N*H*), 8.55 (s, 1 H, 2-*H*3), 7.66 (s, 1 H, 2-*H*6), 7.43 (d, *J* = 2.0 Hz, 1 H, 1-*H*2), 7.32 (dd, *J* = 2.0, 8.0 Hz, 1 H, 1-*H*6), 6.88 (d, *J* = 8.0 Hz, 1 H, 1-*H*5), 4.09 (d, *J* = 6.5 Hz, 2 H, 2-*H*α), 3.92 (d, *J* = 6.5 Hz, 2 H, 2-*H*α'), 3.87 (d, *J* = 6.5 Hz, 2 H, 1-*H*α), 2.27 – 2.16 (m, 3 H, 1-*H*β, 2-*H*β, 2-*H*β'), 1.11 (d, *J* = 6.5 Hz, 6 H, 2-*H*γ'), 1.10 (d, *J* = 7.0 Hz, 6 H, 2-*H*γ), 1.07 (d, *J* = 7.0 Hz, 6 H, 1-*H*γ); δ_C (125 MHz, $CDCl_3$) 165.5, 165.1, 152.9, 141.9, 139.5, 134.1, 132.6, 120.1, 119.7, 114.0, 110.9, 110.8, 110.5, 103.7, 75.5, 75.1, 75.0, 28.4, 28.3, 28.2, 19.4, 19.3, 19.2; ν_{max}/cm^{-1} (solid state) = 3494, 3354, 2960, 2873, 1717, 1513, 1433, 1257, 1192, 1022, 751; ESI-HRMS found *m/z* 473.2664 $[M-H]^-$, $C_{26}H_{35}N_2O_6$ requires 473.2646.

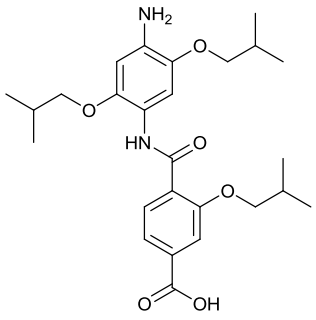
4-(4-amino-2,5-diisobutoxyphenylcarbamoyl)-2-isobutoxybenzoic acid (2.52)



tert-Butyl 2,5-diisobutoxy-1,4-phenylenedicarbamate **2.35** (300 mg, 0.66 mmol, 1 equiv) was dissolved in anhydrous hydrogen chloride (4M) solution in 1,4- dioxane (30 mL / g) and stirred for 3 h at rt. The solvent was evaporated under reduced pressure and the solid residue redissolved in anhydrous chloroform (150 mL / g) before the addition of triethylamine (115 μL, 0.83 mmol, 1.25 equiv). The resulting mixture was heated to reflux and a solution of 4-(ethoxycarbonyl)-3-isobutoxybenzoic acid **2.50** (88 mg, 0.33 mmol, 0.5 equiv) and Ghosez's reagent (218 μL, 0.33 mmol, 0.5 equiv) in anhydrous chloroform (100 mL / g), previously stirred at 50 °C for 3 h, was added dropwise *via* a cannula. The reaction was stirred at reflux overnight. Up to this point, air free conditions were required. The organic solvent was evaporated under reduced pressure and the resulting solid was reacted, without further purification, with a 10% sodium hydroxide solution (10 mL / g) in tetrahydrofuran (100 mL / g) at rt. On completion, the organic solvent was evaporated under reduced pressure and the solid residue subjected to HPLC [(50-95% MeCN:water and 0.1% Formic acid) t = 8 min, 20 mL min^{-1}, XBridge Prep C18 column] to yield the target dimer **2.52** (14% overall yield) as a brown solid; δ_H (500 MHz, $CDCl_3$) 8.56 (s, broad, 1 H, 1-N*H*), 8.30 (d, *J* = 8.0 Hz, 1 H, 2-*H*6), 8.15 (s, 1 H, 1-*H*6), 7.72 (d, *J* = 1.0 Hz, 1 H, 2-*H*3), 7.45 (dd, *J* = 1.0, 8.0 Hz, 1 H, 2-*H*5), 6.40 (s, 1 H, 1-*H*3), 4.13 (d, *J* = 6.5 Hz, 2 H, 2-*H*α), 3.82 (d, *J* = 6.5 Hz, 2 H, 1-*H*α'), 3.78 (d, *J* = 6.5 Hz, 2 H, 1-*H*α), 2.31 – 2.23 (m, 1 H, 2-*H*β), 2.18 – 2.08 (m, 2 H, 1-*H*β, 1-*H*β'), 1.12 (d, *J* = 7.0 Hz, 6 H, 2-*H*γ), 1.07 (d, *J* = 7.0 Hz, 6 H, 1-*H*γ), 1.05 (d, *J* = 7.0 Hz, 6 H, 1-*H*γ'); δ_C (125 MHz, $CDCl_3$) 164.6, 162.3, 158.1, 142.4, 141.4, 140.2, 134.3, 133.1, 120.0, 118.5, 118.4, 112.3, 105.4, 100.1, 75.6, 75.6, 75.5, 28.5, 28.4, 28.1, 19.4, 19.3, 19.1; ν_{max}/cm^{-1} (solid state) = 3364,

2958, 2872, 1660, 1530, 1431, 1217, 1196, 1026, 743; ESI-HRMS found *m/z* 473.2659 $[M-H]^-$, $C_{26}H_{35}N_2O_6$ requires 473.2646.

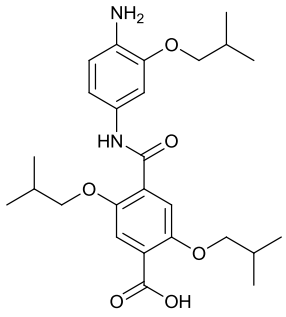
4-(4-amino-2,5-diisobutoxyphenylcarbamoyl)-3-isobutoxybenzoic acid (2.53)



Initial hydrolysis step: A lithium hydroxide (14 mg, 0.33 mmol, 1 equiv) solution in water was added to a solution of 4-tert-butyl 1-ethyl 2-isobutoxyterephthalate **2.51** (106 mg, 0.33 mmol, 1 equiv) in tetrahydrofuran (25 mL / g) and the resulting mixture was stirred at rt overnight. The organic solvent was removed under reduced pressure, an additional amount of water was added and then acidified *via* the addition of 1 M potassium bisulfate solution to pH 5. The resulting precipitate was extracted into dichloromethane, evaporated to dryness to yield a 4-(*tert*-butoxycarbonyl)-2-isobutoxybenzoic acid crude mixture which was used without further purification.

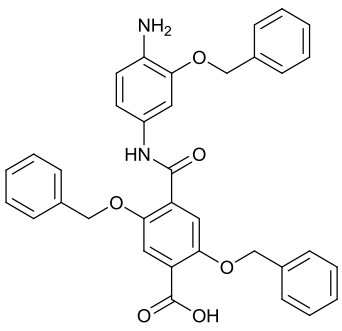
Coupling step: tert-Butyl 2,5-diisobutoxy-1,4-phenylenedicarbamate **2.35** (300 mg, 0.66 mmol, 1 equiv) was dissolved in anhydrous hydrogen chloride (4M) solution in 1,4- dioxane (30 mL / g) and stirred for 3 h at rt. The solvent was evaporated under reduced pressure and the solid residue redissolved in anhydrous chloroform (150 mL / g) before the addition of triethylamine (115 µL, 0.83 mmol, 1.25 equiv). The resulting mixture was heated to reflux and a solution of crude 4-(*tert*-butoxycarbonyl)-2-isobutoxybenzoic acid derived from **2.51** (97 mg, 0.33 mmol, 0.5 equiv) and Ghosez's reagent (218 µL, 0.33 mmol, 0.5 equiv) in anhydrous chloroform (100 mL / g), previously stirred at 50 °C for 3 h, was added dropwise *via* a cannula. The reaction was stirred at reflux overnight. Up to this point, air free conditions were required. The organic solvent was evaporated under reduced pressure and the resulting solid was reacted, without further purification with a 10% TFA solution in tetrahydrofuran (100 mL / g) at rt. On completion, the organic solvent was evaporated under reduced pressure and the solid residue subjected to HPLC [(50-95% MeCN:water and 0.1% Formic acid) t = 8 min, 20 mL min^{-1}, XBridge Prep C18 column] to yield the target dimer **2.53** (23% overall yield) as a brown solid; δ_H (500 MHz, $CDCl_3$) 9.89 (s, broad, 1 H, 1-N*H*), 8.32 (d, *J* = 8.0 Hz, 1 H, 2-*H*5), 8.15 (s, 1 H, 1-*H*6), 7.82 (dd, *J* = 1.0, 8.0 Hz, 1 H, 2-*H*6), 7.73 (d, *J* = 1.0 Hz, 1 H, 2-*H*2), 6.41 (s, 1 H, 1-*H*3), 4.04 (d, *J* = 7.0 Hz, 2 H, 2-*H*α), 3.83 (d, *J* = 6.5 Hz, 2 H, 1-*H*α'), 3.74 (d, *J* = 7.0 Hz, 2 H, 1-*H*α), 2.30 – 2.22 (m, 1 H, 2-*H*β), 2.17 – 2.03 (m, 2 H, 1-*H*β, 1-*H*β'), 1.05 (d, *J* = 6.5 Hz, 6 H, 1-*H*γ'), 1.03 (d, *J* = 6.5 Hz, 6 H, 2-*H*γ), 0.98 (d, *J* = 7.0 Hz, 6 H, 1-*H*γ); δ_C (125 MHz, $CDCl_3$) 165.1, 162.0, 156.6, 143.1, 140.5, 132.6, 132.4, 127.7, 122.6, 119.3, 114.4, 110.1, 107.1, 101.1, 76.1, 76.3, 75.4, 28.5, 28.2, 27.9, 19.4, 19.3, 19.2; ν_{max}/cm^{-1} (solid state) = 3421, 3345, 2964, 1535, 1230, 1177, 1029, 824, 514; ESI-HRMS found *m/z* 473.2647 $[M+H]^+$, $C_{26}H_{37}N_2O_6$ requires 473.2646.

4-(4-amino-3-isobutoxyphenylcarbamoyl)-2,5-diisobutoxybenzoic acid (2.54)



Under a nitrogen atmosphere, thionyl chloride (586 μL, 8.06 mmol, 5 equiv) was added to a stirred solution of 2,5-diisobutoxyterephthalic acid **2.32** (500 mg, 1.62 mmol, 1 equiv) in anhydrous dichloromethane (50 mL / g), and the resulting mixture was stirred at reflux overnight. The organic solvent and the excess thionyl chloride were co-evaporated under a nitrogen flow; this was repeated 3 times with further additions of dichloromethane to yield a yellow solid. The resulting 2,5-diisobutoxyterephthaloyl dichloride (558 mg, 1.61 mmols, 1 equiv) was dissolved in anhydrous chloroform (100 mL / g) and a solution of (9H-Fluoren-9-yl)methyl 4-amino-2-isobutoxyphenylcarbamate **2.45** (216 mg, 0.54 mmol, 0.3 equiv) in anhydrous chloroform (30 mL / g) was added dropwise. The resulting mixture was stirred at reflux under a nitrogen atmosphere. The organic solvent was evaporated under reduced pressure and the solid residue reacted, without further purification, with a 10% sodium hydroxide solution (10 mL / g) in tetrahydrofuran (100 mL / g) at rt. On completion, the organic solvent was evaporated under reduced pressure and the solid residue subjected to HPLC [(50-95% MeCN:water and 0.1% Formic acid) t = 8 min, 20 mL min^{-1}, XBridge Prep C18 column] to afford the desired product **2.54** (10% overall yield) as a yellow solid; δ_H (500 MHz, $CDCl_3$) 10.04 (s, broad, 1 H, 1-N*H*), 8.04 (s, 1 H, 2-*H*3), 7.83 (s, 1 H, 2-*H*6), 7.61 (d, *J* = 2.0 Hz, 1 H, 1-*H*5), 6.78 (dd, *J* = 2.0, 8.0 Hz, 1 H, 1-*H*6), 6.75 (d, *J* = 8.0 Hz, 1 H, 1-*H*2), 4.11 (d, *J* = 7.5 Hz, 2 H, 2-*H*α), 4.04 (d, *J* = 6.5 Hz, 2 H, 2-*H*α'), 3.83 (d, *J* = 6.5 Hz, 2 H, 1-*H*α), 2.33 – 2.12 (m, 3 H, 1-*H*β, 2-*H*β, 2-*H*β'), 1.16 (d, *J* = 7.0 Hz, 6 H, 2-*H*γ'), 1.10 (d, *J* = 7.0 Hz, 6 H, 2-*H*γ), 1.06 (d, *J* = 7.0 Hz, 6 H, 1-*H*γ); δ_C (125 MHz, $CDCl_3$) 164.6, 161.0, 151.6, 151.1, 147.0, 132.8, 130.0, 127.3, 120.5, 117.0, 116.5, 115.0, 112.4, 105.2, 77.2, 76.7, 74.7, 28.5, 28.4, 28.1, 19.4, 19.4, 19.2; ν_{max}/cm^{-1} (solid state) = 3355, 2959, 2875, 1738, 1652, 1532, 1417, 1196, 1015, 741; ESI-HRMS found *m/z* 473.2664 $[M-H]^-$, $C_{26}H_{35}N_2O_6$ requires 473.2646.

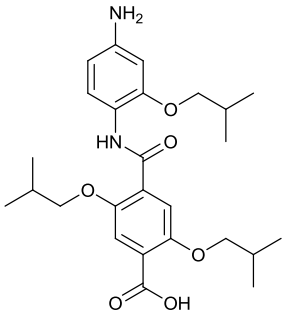
4-(4-amino-3-(benzyloxy)phenylcarbamoyl)-2,5-bis(benzyloxy)benzoic acid (2.55)



Under a nitrogen atmosphere, thionyl chloride (479 μL, 6.61 mmol, 5 equiv) was added to a stirred solution of 2,5-bis(benzyloxy)terephthalic acid **2.33** (500 mg, 1.32 mmol, 1 equiv) in anhydrous dichloromethane (50 mL / g), and the resulting mixture was stirred at reflux overnight. The organic solvent and the excess thionyl chloride were co-evaporated under a nitrogen flow; this was repeated 3 times with further additions of dichloromethane to yield a yellow solid. The resulting 2,5-

bis(benzyloxy)terephthaloyl dichloride (548 mg, 1.32 mmol, 1 equiv) was dissolved in anhydrous chloroform (100 mL / g) and a solution of (9H-fluoren-9-yl)methyl 4-amino-2-(benzyloxy)phenylcarbamate **2.46** (192 mg, 0.44 mmol, 0.3 equiv) in anhydrous chloroform (30 mL / g) was added dropwise. The resulting mixture was stirred at reflux under a nitrogen atmosphere. The organic solvent was evaporated under reduced pressure and the solid residue reacted, without further purification, with a 10% solution of piperidine in tetrahydrofuran (100 mL / g) at rt. On completion, the organic solvent was evaporated under reduced pressure and the solid residue subjected to HPLC [(50-95% MeCN:water and 0.1% Formic acid) t = 8 min, 20 mL min^{-1}, XBridge Prep C18 column] to afford the desired product **2.55** (44% overall yield) as a yellow solid; δ_H (500 MHz, $CDCl_3$) 9.99 (s, broad, 1 H, 1-N*H*), 8.23 (s, 1 H, 2-*H*3), 7.99 (s, 1 H, 2-*H*6), 7.56 – 7.55 (m, 2 H, 2-*H*Ar'), 7.50 – 7.35 (m, 13 H, 2-*H*Ar, 2-*H*Ar', 1-*H*Ar), 7.24 (d, *J* = 1.5 Hz, 1 H, 1-*H*2), 6.60 – 6.55 (m, 2 H, 1-*H*5, 1-*H*6), 5.37 (s, 2 H, 2-*H*α), 5.27 (s, 2 H, 2-*H*α'), 4.94 (s, 2 H, 1-*H*α); δ_C (125 MHz, $CDCl_3$) 164.3, 160.5, 151.6, 151.1, 146.3, 136.8, 134.8, 134.1, 133.4, 129.6, 129.4, 129.3, 129.2, 129.2, 128.9, 128.6, 128.4, 128.1, 127.8, 127.6, 121.1, 117.4, 117.1, 114.8, 112.9, 105.0, 73.1, 72.7, 70.5; ν_{max}/cm^{-1} (solid state) = 3469, 3359, 1730, 1515, 1413, 1219, 1000, 735; ESI-HRMS found *m/z* 575.2191 $[M+H]^+$, $C_{35}H_{30}N_2O_6$ requires 575.2177.

4-(4-amino-2-isobutoxyphenylcarbamoyl)-2,5-diisobutoxybenzoic acid (2.56)



Under a nitrogen atmosphere, thionyl chloride (586 μL, 8.06 mmol, 5 equiv) was added to a stirred solution of 2,5-diisobutoxyterephthalic acid **2.32** (500 mg, 1.62 mmol, 1 equiv) in anhydrous dichloromethane (50 mL / g), and the resulting mixture was stirred at reflux overnight. The organic solvent and the excess thionyl chloride were co-evaporated under a nitrogen flow; this was repeated 3 times with further additions of dichloromethane to yield a yellow solid. The resulting 2,5-diisobutoxyterephthaloyl dichloride (558 mg, 1.61 mmols, 1 equiv) was dissolved in anhydrous chloroform (100 mL / g) and a solution of 2-isobutoxy-4-nitroaniline **2.41** (169 mg, 0.81 mmol, 0.5 equiv) in anhydrous chloroform (30 mL / g) was added dropwise. The resulting mixture was stirred at reflux under a nitrogen atmosphere. The organic solvent was evaporated under reduced pressure and the solid redisolved in a 1:1 mixture of methanol : tetrahydrofuran (30 mL / g), palladium on carbon (10 wt. %) was added and the flask evacuated and flushed with nitrogen (3 times) and left under vacuum. Hydrogen was drawn into the flask and the reaction was left stirring at rt overnight. On completion, the reaction mixture was filtered through a celite pad and washed with methanol and tetrahydrofuran. The organic solvents were evaporated to dryness to yield the target dimer **2.56**

(42% overall yield) as a orange solid; δ_H (500 MHz, $CDCl_3$) 9.97 (s, broad, 1 H, 1-N*H*), 8.20 (d, *J* = 8.5 Hz, 1 H, 1-*H*6), 7.98 (s, 1 H, 2-*H*3), 7.83 (s, 1 H, 2-*H*6), 6.33 (dd, *J* = 2.0, 8.5 Hz, 1 H, 1-*H*5), 6.31 (d, *J* = 2.0 Hz, 1 H, 1-*H*3), 4.09 (d, *J* = 6.5 Hz, 2 H, 2-*H*α), 4.01 (d, *J* = 7.0 Hz, 2 H, 2-*H*α'), 3.78 (d, *J* = 7.0 Hz, 2 H, 1-*H*α), 2.26 – 2.09 (m, 3 H, 1-*H*β, 2-*H*β, 2-*H*β'), 1.09 (d, *J* = 7.0 Hz, 6 H, 2-*H*γ), 1.02 (d, *J* = 6.5 Hz, 6 H, 1-*H*γ), 0.99 (d, *J* = 6.5 Hz, 6 H, 2-*H*γ'); δ_C (125 MHz, $CDCl_3$) 164.8, 161.2, 151.6, 151.1, 150.0, 144.0, 128.7, 123.1, 120.2, 119.3, 117.8, 116.4, 107.0, 99.8, 77.1, 77.0, 75.2, 28.1 (2 C), 27.9, 19.2, 19.2, 19.1; ν_{max}/cm^{-1} (solid state) = 3355, 2959, 2875, 1738, 1652, 1532, 1417, 1196, 1015, 741; ESI-HRMS found *m/z* 471.2493 $[M-H]^-$, $C_{26}H_{35}N_2O_6$ requires 471.2493; Found: C, 66.15; H, 7.40; N, 5.90; $C_{26}H_{36}N_2O_6$ requires: C, 66.08; H, 7.68; N, 5.93%.

6.3.3. Molecullar modeling

A conformational search was performed on the complete set of dimers. The structures were minimised by employing a full Monte Carlo search in the software Macromodel® using the Merk Molecular Force Fields (MMFFs) method and sampling a total of 50,000 structures. Water was chosen as implicit solvent and free rotation around the amide bonds was allowed in order to increase the accuracy of the conformational search.

All the conformations within 1.5 kJ/mol from the lowest energy conformation were superimposed with the ER co-activator helix (PDB: 2QZO) and AR co-activator helices (PDB: 1T7F and 1T73 for compound **2.55**). A mean value of the Root Mean Square Deviation (RMSD) was calculated from the superimposition of the oxygen of the alkoxy group and the alpha carbon of the key aminoacids of the co-activator helix. The alignment was also investigated in the reverse dipole sequence.

6.3.4. Docking studies

The LBDs of ERα (PDB ID: 2QZO) and AR (PDB: 1T7F) were prepared for docking using the Protein Preparation Wizard (Schrödinger) function in Maestro. Once the protein was refined, Glide (Schrödinger) was then used to create a grid box for docking. The structures of the ER and AR are dimeric; therefore, only one monomer was used within the docking grid. The dimensions and position of the grid box were adjusted according to the coactivator binding pose, which must be centered and lie fully within the grid. The LigPrep (Schrödinger) function

was then used to prepare the set of compounds for docking. Once arranged, the resulting compounds were docked into the LBDs of the prepared protein using Glide XP (extra precision) mode.

6.3.5. Fluorescence Polarization assays

All proteins (Sumo-Erα, His-RXRα, Erβ), SRC-Box2 peptide and its fluorescein-labelled analogue FITC-SRC-Box2 were obtained in the laboratory of Prof. Luc Brunsveld (Technische Universiteit Eindhoven). Fluorescein-labelled FITC-D22 peptide was purchased from ThermoFisher Scientific.

The Fluorescence Polarization Assays were performed in triplicate and were measured with a Tecan Safire monochromator microplate reader. The assays were performed in freshly prepared HTRF buffer (10 mM HEPES buffer at pH 7.50, containing 150 mM NaCl, 0.01 mg mL^{-1} *bovine serum albumin* (BSA) and 5 mM DTT). Serial dilutions of the compounds in DMSO (starting point: 10 mM; 12-points, 1/3.16 dilution) were performed. Then, a Mastermix solution was prepared containing the following components:

- *Assay in Agonistic mode:* Mastermix (×10) Protein + tracer + buffer (+ DMSO).
- *Assay in Antagonistic mode:* Mastermix (×10) Protein + tracer + ligand + buffer (+ DMSO)

The preparation of the wells goes as follows:

1) 99 µL of Mastermix were added to each well (96 well plates, white).
2) 1 µL of compound stock solution was added to each well (96 well plates, white).
3) The plates were incubated for 1 h at 4 °C and protected from light.
4) The assay mixture was transferred (3 × 30 µL) to the final wells (384 well plates, low volume black) for measurements.

All the assays contain a final concentration of 5% DMSO.

Processing of the Fluorescence Polarization data

The data obtained for both the P (perpendicular intensity) and S (parallel (same) intensity) channels was used to calculate the polarization for each well using the following equation:

$$p = \frac{S - P}{S + P}$$

Where: p = polarization, P = perpendicular intensity, S = parallel intensity.

6.4. Optimization of the hybrid oligoamide proteomimetic scaffold (*Chapter 3*)

6.4.1. General procedures for Solid Phase Synthesis of the hybrid scaffold

A generic procedure was followed by adapting the previously reported solid phase synthesis of 3-*O*-alkylated oligobenzamides[135] (See section 6.5.2).

Resin preparation and Fmoc deprotection

Fmoc protected pre-loaded Wang resin (0.1 mmol, 1 equiv) was loaded onto a Liberty CEM™ microwave peptide synthesiser after being swelled for a total of 30 min in dichloromethane. Standard washing cycles were carried out and two deprotection cycles using 6 mL of a 20% piperidine in DMF solution, under microwave at 75 °C, performed before coupling of each monomer and after the last coupling reaction.

Coupling of 2-*O* and 3-*O*-alkylated monomers

Each protected monomer (2 equiv per coupling) was dissolved in anhydrous DMF (2.5 mL per coupling) and pre-activated with HATU (3 equiv per coupling) and DIPEA (5 equiv per coupling) at rt. A double coupling method of 30 min was carried out under microwave at 60 °C.

Coupling of central amino acids

Each protected amino acid (2.5 equiv per coupling) was dissolved in anhydrous DMF (2.5 mL per coupling) and pre-activated with HATU (3 equiv per coupling) and DIPEA (5 equiv per coupling) at rt. For natural amino acid derivatives, a double coupling of 30 minutes was carried out under microwave at 60 °C. For non-natural amino acids, a triple coupling of 30 minutes was carried out under microwave at 60 °C.

Coupling of challenging monomers and amino acids

Monomers and amino acids that could not be coupled using the standard methodology were activated using procedure I: acyl chloride formation (See section 6.5.2). NMP was used as a solvent for these couplings.

Reductive amination on resin

The resin bound hybrid was transferred to a reservoir and washed with DMF (2 × 3 mL × 2 min). A solution of aldehyde (5 equiv) in 2.5 mL of 1:1 methanol:DMF (1% acetic acid) followed by a suspension of $NaBH_3CN$ (5 equiv) in 1 mL of methanol was added to the resin beads and the mixture stirred at rt. Reaction times varied from 2 to 15 h and further portions of aldehyde and reducing agent $NaBH_3CN$ were added when the reaction was shown incomplete by LC-MS.

Cleavage

After complete synthesis, the resin was transferred to a reservoir, washed with dichloromethane (2 × 3 mL × 2 min) and cleaved with a cleavage cocktail (3 mL, 2 × 30 min) consisting of a 1:1 mixture of TFA:dichloromethane. After the reaction had reached completion, the cleavage cocktail was evaporated under N_2 (g).

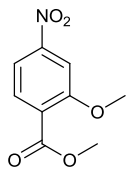
Mass – Directed HPLC Purification

The resulting crude product was dissolved in either DMSO or methanol at an approximate concentration of 20 mg mL^{-1} and purified using reversed phase mass directed HPLC [Agilent XBridge C18 preparative column; variable gradient of MeCN to water (plus 0.1% formic acid v/v in both solvents) and flow rate of 20 mL min^{-1} during 8 min]. The resulting fractions are concentrated by centrifugal evaporation (Genevac).

6.4.2. *O*-Alkylated monomer syntheses and characterization

The *O*-alkylated monomers employed in the hybrid scaffold syntheses were obtained using the same general procedures as for the 3-*O*-alkylated monomers unless stated otherwise (See section 6.5.1). Only the characterization of the novel monomers has been included in the experimental section; the monomers which have been routinely synthesised and reported previously in the group have not been included.[125, 127, 135]

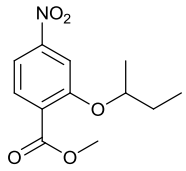
Methyl 2-methoxy-4-nitrobenzoate (3.5)



Procedure A; methyl 2-hydroxy-4-nitrobenzoate **2.4** (2.0 g, 10.1 mmol), potassium carbonate (3.5 g, 25.3 mmol), in dimethylformamide (120 mL), methyl iodide (7.7 mL, 50.5 mmol). Work up afforded the title compound **3.5** (2.08 g, 9.9 mmol, 98 %) as a pale yellow solid; δ_H (500 MHz, $CDCl_3$) 7.90 (d, *J* = 8.2 Hz, 1 H, *H*6), 7.85

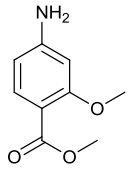
(dd, *J* = 8.2, 1.8 Hz, 1 H, *H*5), 7.83 (d, *J* = 1.8 Hz, 1 H, *H*3), 4.01 (s, 3 H, *H*α), 3.95 (s, 3 H, CO_2CH_3); δ_C (125 MHz, $CDCl_3$) 165.3, 159.3, 150.8, 132.1, 126.0, 115.0, 107.0, 56.6, 52.6; ν_{max}/cm^{-1} (solid state) = 3121, 2952, 2848, 1733, 1517, 1245, 1079, 735; ESI-HRMS found *m/z* 212.0552 $[M+H]^+$, $C_9H_{10}NO_5$ requires 212.0553.

Methyl 2-sec-butoxy-4-nitrobenzoate (3.7)



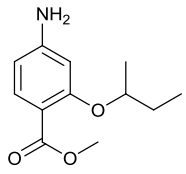
Procedure B; methyl 2-hydroxy-4-nitrobenzoate **2.4** (1.0 g, 5.1 mmol), *sec*-butanol (415 mg, 5.6 mmol), triphenylphosphine (2.0 g, 7.6 mmol) in anhydrous tetrahydrofuran (40 mL) and diisopropyl azodicarboxylate (1.5 mL, 7.6 mmol). An extra portion of *sec*-butanol (1 equiv), triphenylphosphine (1 equiv) and diisopropyl azodicarboxylate (1 equiv) was added to bring the reaction to completion. The reaction mixture was purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: ethyl acetate/hexane, 1:1) to afford the desired product **3.7** (850 mg, 3.4 mmol, 66%) as a yellow oil; δ_H (500 MHz, $CDCl_3$) 7.84 (d, *J* = 9.0 Hz, 1 H, *H*6), 7.75 – 7.81 (m, 2 H, *H*3, *H*5), 4.47 – 4.53 (m, 1 H, *H*α), 3.92 (s, 3 H, CO_2CH_3), 1.86 – 1.78 (m, 1 H, *H*β), 1.69 – 1.78 (m, 1 H, *H*β′), 1.38 (d, *J* = 6.2 Hz, 3 H, CHα($C\underline{H}_3$)), 1.03 (t, *J* = 7.4 Hz, 3 H, Hγ); δ_C (125 MHz, $CDCl_3$) 165.7, 157.9, 150.6, 131.9, 127.4, 114.5, 109.1, 52.4, 33.6, 29.1, 18.8, 9.4; ν_{max}/cm^{-1} (solid state) = 3108, 2973, 2880, 1736, 1526, 1346, 1245, 1077, 736; ESI-HRMS found *m/z* 254.1025 $[M+H]^+$, $C_{12}H_{16}NO_5$ requires 254.1023.

Methyl 4-amino-2-methoxybenzoate (3.10)



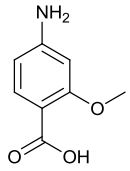
Procedure D; methyl 2-methoxy-4-nitrobenzoate **3.5** (2.08 g, 9.9 mmol) in a 1:1 mixture of tetrahydrofuran/methanol (20 mL). After work up, the reaction mixture was purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: ethyl acetate/hexane, 3:1) to yield the title compound **3.10** (1.78 g, 9.9 mmol, quant.) as a pale orange solid; δ_H (500 MHz, $CDCl_3$) 7.75 (d, *J* = 8.5 Hz, 1 H, *H*6), 6.24 (dd, *J* = 8.5, 1.9 Hz, 1 H, *H*5), 6.21 (d, *J* = 1.9 Hz, 1 H, *H*3), 4.05 (s, broad, 2 H, NH_2), 3.87 (s, 3 H, *H*α), 3.83 (s, 3 H, CO_2CH_3); δ_C (125 MHz, $CDCl_3$) 166.2, 161.8, 152.0, 134.2, 109.1, 106.4, 97.7, 55.8, 51.4; ν_{max}/cm^{-1} (solid state) = 3461, 3364, 3236, 2948, 2837, 1698, 1606, 1434, 1254, 1217, 1086, 830, 774; ESI-HRMS found *m/z* 182.0814 $[M+H]^+$, $C_9H_{12}NO_3$ requires 182.0812.

Methyl 4-amino-2-sec-butoxybenzoate (3.12)



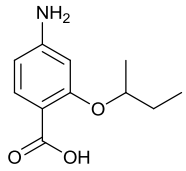
Procedure D; methyl 2-sec-butoxy-4-nitrobenzoate **3.7** (850 mg, 3.4 mmol) in a 1:1 mixture of tetrahydrofuran/methanol (40 mL). Work up yielded the title compound **3.12** (747 mg, 3.4 mmol, quant.) as a purple brown oil; δ_H (500 MHz, $CDCl_3$) 7.72 (d, *J* = 8.5 Hz, 1 H, *H*6), 6.23 (dd, *J* = 8.5, 2.1 Hz, 1 H, *H*5), 6.20 (d, *J* = 2.1 Hz, 1 H, *H*3), 4.27 – 4.33 (m, 1 H, *H*α), 3.97 (s, broad, 2 H, NH_2), 3.82 (s, 3 H, CO_2CH_3), 1.74 – 1.86 (m, 1 H, *H*β), 1.62 – 1.74 (m, 1 H, *H*β′), 1.33 (d, *J* = 6.2 Hz, 3 H, CHα($C\underline{H}_3$)), 1.01 (t, *J* = 7.4 Hz, 3 H, Hγ); δ_C (125 MHz, $CDCl_3$) 166.5, 160.5, 151.6, 134.2, 110.8, 106.7, 100.9, 76.5, 51.3, 29.2, 19.1, 9.6; ν_{max}/cm^{-1} (solid state) = 3463, 3365, 3234, 2971, 2946, 2878, 1699, 1598, 1448, 1432, 1246, 1140, 1080, 773; ESI-HRMS found *m/z* 246.1103 $[M+Na]^+$, $C_{12}H_{17}NNaO_3$ requires 246.1101.

4-Amino-2-methoxybenzoic acid (3.15)



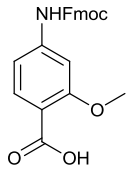
Procedure E; methyl 4-amino-2-methoxybenzoate **3.10** (1.60 g, 8.8 mmol) in a 1:1 mixture of tetrahydrofuran/methanol (40 mL), 10% aqueous sodium hydroxide (30 mL). Additional 1:1 mixture of tetrahydrofuran/methanol (150 mL) was added to aid solubility over the course of the 4 day reaction. Work up yielded the title compound **3.15** (1.15 g, 6.9 mmol, 78%) as a colourless solid; δ_H (500 MHz, MeOD) 7.69 (d, *J* = 8.5 Hz, 1 H, *H*6), 6.34 (d, *J* = 1.9 Hz, 1 H, *H*3), 6.28 (dd, *J* = 8.5, 1.9 Hz, 1 H, *H*5), 3.87 (s, 3 H, *H*α); δ_C (125 MHz, MeOD) 169.4, 162.8, 156.8, 135.6, 107.9, 106.2, 97.3, 56.3; ν_{max}/cm^{-1} (solid state) = 3420, 3337, 3226, 1698, 1603, 1333, 1267, 1022, 823; ESI-HRMS found *m/z* 190.0476 $[M+Na]^+$, $C_8H_9NNaO_3$ requires 190.0475.

4-Amino-2-sec-butoxybenzoic acid (3.17)



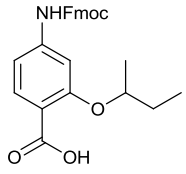
Procedure E; methyl 4-amino-2-sec-butoxybenzoate **3.12** (750 mg, 3.4 mmol) in a 1:1 mixture of tetrahydrofuran/methanol (100 mL), 10% aqueous sodium hydroxide (30 mL). Further aliquots of 10% aqueous sodium hydroxide were added to bring the reaction to completion. Work up yielded the title compound **3.17** (700 g, 3.3 mmol, quant.) as an orange oil; δ_H (500 MHz, $CDCl_3$) 10.94 (s, broad, 1 H, CO_2H), 7.96 (d, *J* = 8.5 Hz, 1 H, *H*6), 6.36 (dd, *J* = 8.5, 2.1 Hz, 1 H, *H*5), 6.23 (d, *J* = 2.1 Hz, 1 H, *H*3), 4.53 – 4.59 (m, 1 H, *H*α), 4.23 (s, broad, 2 H, NH_2), 1.81 – 1.92 (m, 1 H, *H*β), 1.71 – 1.81 (m, 1 H, *H*β′), 1.41 (d, *J* = 6.2 Hz, 3 H, CHα($C\underline{H}_3$)), 1.03 (t, *J* = 7.6 Hz, 3 H, Hγ); δ_C (125 MHz, $CDCl_3$) 165.9, 158.5, 152.7, 135.5, 108.6, 108.3, 98.7, 78.5, 29.1, 19.2, 9.6; ν_{max}/cm^{-1} (solid state) = 3468, 3358, 3233, 2973, 2936, 2880, 1707, 1601, 1458, 1394, 1268, 988; ESI-HRMS found *m/z* 210.1128 $[M+H]^+$, $C_{11}H_{16}NO_3$ requires 210.1125.

4-(((9H-Fluoren-9-yl)methoxy)carbonylamino)-2-methoxybenzoic acid (3.21)



Procedure G; 4-amino-2-methoxybenzoic acid **3.15** (1.0 g, 6.0 mmol) in tetrahydrofuran (40 mL) and fluorenylmethyloxycarbonyl chloride (2.6 g, 9.0 mmol) in tetrahydrofuran (20 mL). After work up, the reaction mixture was purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: A gradient of dichloromethane to dichloromethane/ethylacetate 1:1 to ethylacetate/methanol 9.5:0.5) to yield the title compound **3.21** (1.13 g, 2.9 mmol, 48%) as a beige solid; δ_H (500 MHz, $CDCl_3$) 10.60 (s, broad, 1 H, CO_2H), 8.09 (d, J = 8.5 Hz, 1 H, *H*6), 7.79 (d, J = 8.5 Hz, 2 H, F*H*Ar5), 7.65 (s, 1 H, *H*3), 7.64 (d, J = 7.5 Hz, 2 H, F*H*Ar2), 7.43 (t, J = 7.5 Hz, 2 H, F*H*Ar4), 7.34 (t, J = 7.5 Hz, 2 H, F*H*Ar3), 7.16 (s, broad, 1 H, N*H*), 6.81 (dd, J = 8.5, 1.5 Hz, 1 H, *H*5), 4.59 (d, J = 6.5 Hz, 2 H, F*H*α), 4.28 (t, J = 6.5 Hz, 1 H, F*H*β), 4.06 (s, 3 H, *H*α); δ_C (125 MHz, $CDCl_3$) 165.1, 159.2, 152.9, 144.3, 143.4, 141.4, 134.6, 128.0, 127.4, 124.8, 120.2, 112.1, 111.4, 101.1, 67.2, 56.8, 47.0; ν_{max}/cm^{-1} (solid state) = 3276, 3018, 2948, 1714, 1594, 1530, 1215, 755, 740; ESI-HRMS found *m/z* 390.1340 $[M+H]^+$, $C_{23}H_{20}NO_5$ requires 390.1336.

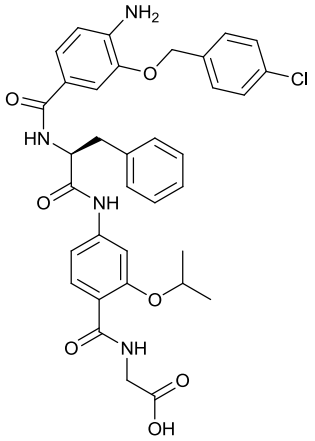
4-(((9H-Fluoren-9-yl)methoxy)carbonylamino)-2-sec-butoxybenzoic acid (3.23)



Procedure G; 4-amino-2-sec-butoxybenzoic acid **3.17** (700 mg, 3.3 mmol) in tetrahydrofuran (50 mL) and fluorenylmethyloxycarbonyl chloride (1.3 g, 5.0 mmol) in tetrahydrofuran (10 mL). Work up yielded the title compound **3.23** (435 mg, 1.0 mmol, 31%) as an off-white solid; δ_H (500 MHz, $CDCl_3$) 11.06 (s, broad, 1 H, CO_2H), 8.10 (d, J = 8.5 Hz, 1 H, *H*6), 7.80 (d, J = 7.5 Hz, 2 H, F*H*Ar5), 7.68 (s, 1 H, *H*3), 7.63 (d, J = 7.5 Hz, 2 H, F*H*Ar2), 7.44 (t, J = 7.5 Hz, 2 H, F*H*Ar4), 7.34 (t, J = 7.5 Hz, 2 H, F*H*Ar3), 7.02 (s, broad, 1 H, N*H*), 6.76 (dd, J = 8.5, 1.5 Hz, 1 H, *H*5), 4.65 – 4.72 (m, 1 H, *H*α), 4.58 (d, J = 6.5 Hz, 2 H, F*H*α), 4.29 (t, J = 6.5 Hz, 1 H, F*H*β), 1.85 – 1.93 (m, 1 H, *H*β), 1.74 – 1.84 (m, 1 H, *H*β′), 1.45 (d, J = 6.2 Hz, 3 H, CHα(C$\underline{H}_3$)), 1.04 (t, J = 7.4 Hz, 3 H, Hγ); δ_C (125 MHz, $CDCl_3$) 165.3, 157.8, 152.9, 149.6, 143.9, 143.4, 141.4, 134.6, 128.0, 127.2, 124.8, 120.2, 111.4, 103.2, 79.0, 67.2, 47.0, 29.0, 19.2, 9.6; ν_{max}/cm^{-1} (solid state) = 3273, 2972, 2938, 1713, 1592, 1531, 1449, 1216, 736; ESI-HRMS found *m/z* 432.1810 $[M+H]^+$, $C_{26}H_{26}NO_5$ requires 432.1805.

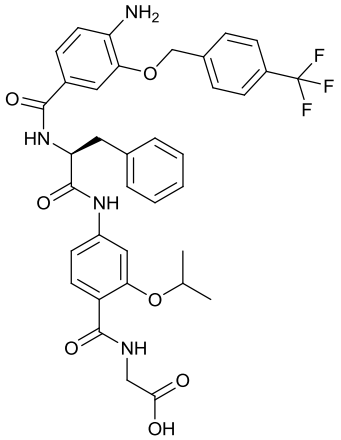
6.4.3. Hybrids characterization

H_2N-[*O*-4-Cl-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.27)



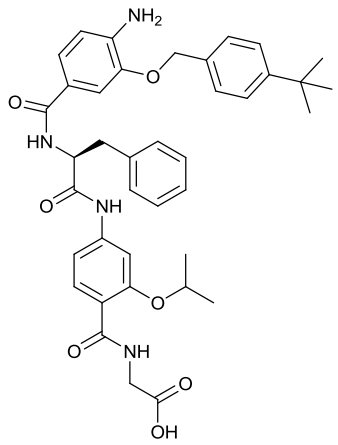
Pale brown solid; isolated yield: 23 mg, 35%; δ_H (500 MHz, MeOD) 7.95 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.60 (d, *J* = 1.7 Hz, 1 H, 3-*H*3), 7.44 (d, *J* = 8.3 Hz, 2 H, 1-*H*Ar2), 7.37 – 7.36 (m, 3 H, 1-*H*Ar3, 1-*H*2), 7.33 – 7.28 (m, 3 H, 2-Phe-*H*Ar2, 1-*H*6), 7.25 – 7.23 (m, 2 H, 2-Phe-*H*Ar3), 7.21 – 7.16 (m, 1 H, 2-Phe-*H*Ar4), 7.07 (dd, *J* = 8.5, 1.7 Hz, 1 H, 3-*H*5), 6.72 (d, *J* = 8.3 Hz, 1 H, 1-*H*5), 5.10 (s, 2 H, 1-*H*α), 4.90 (dd, *J* = 8.3, 6.6 Hz, 1 H, 2-Phe-*H*α), 4.77 (spt, *J* = 6.1, 1 H, 3-*H*α), 4.16 (s, 2 H, 4-Gly-*H*α), 3.27 (dd, *J* = 13.7, 6.6 Hz, 1 H, 2-Phe-*H*β), 3.17 (dd, *J* = 13.7, 8.3 Hz, 1 H, 2-Phe-*H*β′), 1.46 (d, *J* = 6.1 Hz, 3 H, 3-*H*β), 1.44 (d, *J* = 6.1 Hz, 3 H, 3-*H*β′); δ_C (125 MHz, MeOD) 172.9, 170.1, 167.2, 158.3, 146.4, 144.2, 143.1, 138.3, 137.3, 134.7, 133.1, 127.8, 122.9, 117.9, 114.5, 113.0, 112.4, 106.5, 73.6, 70.6, 57.6, 42.6, 38.9, 22.1; ν_{max}/cm^{-1} (solid state) = 3357, 2979, 2930, 1683, 1522, 1407, 1211, 698; ESI-HRMS found *m/z* 659.2278 $[M+H]^+$, $C_{35}H_{36}ClN_4O_7$ requires 659.2267.

H_2N-[*O*-4-CF_3-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.28)



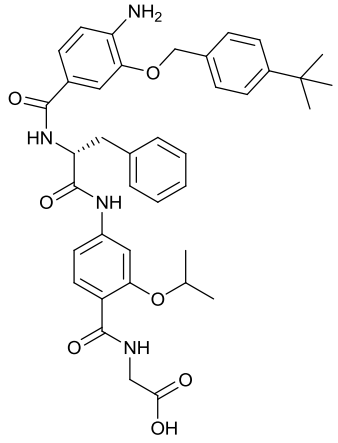
Beige solid; isolated yield: 29 mg, 42%; δ_H (500 MHz, DMSO-d_6) 10.42 (s. broad, 1 H, 4-Gly-CO_2H), 8.46 (t, *J* = 4.9 Hz, 1 H, 4-Gly-N*H*), 8.32 (d, *J* = 8.0 Hz, 1 H, 2-Phe-N*H*), 7.89 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.88 – 7.81 (m, 2 H, 1-*H*Ar2), 7.70 – 7.64 (m, 1 H, 1-*H*Ar3), 7.61 (d, *J* = 1.9 Hz, 1 H, 3-*H*3), 7.42 (d, *J* = 1.8 Hz, 1 H, 1-*H*2), 7.38 (d, *J* = 7.3 Hz, 2 H, 2-Phe-*H*Ar2), 7.33 (dd, *J* = 8.5, 1.9 Hz, 1 H, 3-*H*5), 7.28 – 7.22 (m, 3 H, 1-*H*6, 2-Phe-*H*Ar3), 7.17 (t, *J* = 8.5 Hz, 1 H, 2-Phe-*H*Ar4), 6.65 (d, *J* = 8.2 Hz, 1 H, 1-*H*5), 5.40 (s. broad, 2 H, 1-NH_2), 5.24 (s, 2 H, 1-*H*α), 4.84 – 4.76 (m, 2 H, 2-Phe-*H*α), 4.71 (quin, *J* = 5.9 Hz, 1 H, 3-*H*α), 4.04 (d, *J* = 5.0 Hz, 2 H, 4-Gly-*H*α), 3.16 – 3.04 (m, 2 H, 2-Phe-*H*β), 1.41 (dd, *J* = 8.5, 6.5 Hz, 6 H, 3-*H*β); δ_C (125 MHz, DMSO-d_6) 171.4, 171.2, 166.3, 163.9, 156.2, 143.9, 143.0, 141.7, 138.8, 138.1, 131.9, 131.4, 129.4, 129.2, 128.0, 126.3, 123.9, 122.0, 120.8, 116.6, 112.4, 111.6, 111.3, 104.7, 71.9, 68.7, 55.8, 41.6, 37.1, 21.6, 21.6; ν_{max}/cm^{-1} (solid state) = 3340, 2979, 2923, 1615, 1595, 1497, 1213, 1122, 699; ESI-HRMS found *m/z* 693.2549 $[M+H]^+$, $C_{36}H_{36}F_3N_4O_7$ requires 693.2531.

H_2N-[*O*-4-*t*Bu-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.29)



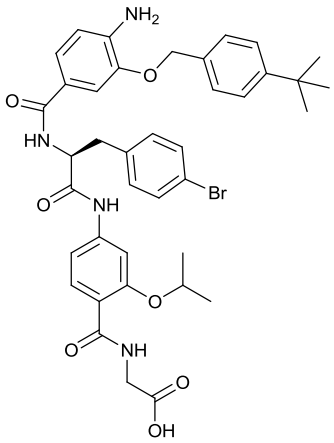
Beige solid; isolated yield: 24 mg, 35%; δ_H (500 MHz, MeOD) 7.95 (d, *J* = 8.7 Hz, 1 H, 3-*H*6), 7.59 (d, *J* = 1.8 Hz, 1 H, 3-*H*3), 7.43– 7.38 (m, 3 H, 1-*H*2, 1-*H*Ar2), 7.38 – 7.34 (m, 2 H, 1-*H*Ar3), 7.33 – 7.28 (m, 3 H, 2-Phe-*H*Ar2, 1-*H*6), 7.25 – 7.23 (m, 2 H, 2-Phe-*H*Ar3), 7.21 – 7.16 (m, 1 H, 2-Phe-*H*Ar4), 7.08 (dd, *J* = 8.7, 1.8 Hz, 1 H, 3-*H*5), 6.71 (d, *J* = 8.1 Hz, 1 H, 1-*H*5), 5.05 (s, 2 H, 1-*H*α), 4.92 (dd, *J* = 8.2, 6.9 Hz, 1 H, 2-Phe-*H*α), 4.74 (spt, *J* = 6.1, 1 H, 3-*H*α), 4.16 (s, 2 H, 4-Gly-*H*α), 3.27 (dd, *J* = 13.8, 6.9 Hz, 1 H, 2-Phe-*H*β), 3.17 (dd, *J* = 13.8, 8.3 Hz, 1 H, 2-Phe-*H*β′), 1.42 (d, *J* = 6.1 Hz, 3 H, 3-*H*β), 1.44 (d, *J* = 6.1 Hz, 3 H, 3-*H*β), 1.31 (s, 9 H, 1-HAr4-C(C$\underline{H}_3$)$_3$); δ_C (125 MHz, MeOD) 173.0, 172.6, 170.2, 167.2, 158.3, 152.1, 146.7, 144.2, 143.0, 138.3, 135.4, 133.1, 130.3, 129.4, 128.6, 127.8, 126.4, 123.0, 122.7, 117.9, 114.4, 113.0, 112.2, 106.5, 73.6, 71.3, 57.6, 42.6, 39.0, 35.3, 31.7, 22.1; ν_{max}/cm^{-1} (solid state) = 3356, 2962, 2870, 1596, 1493, 1408, 1213, 1004, 699; ESI-HRMS found *m/z* 703.3104 $[M+Na]^+$, $C_{39}H_{44}N_4NaO_7$ requires 703.3102.

H_2N-[*O*-4-*t*Bu-Bn-(3-HABA)]-D-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.50)



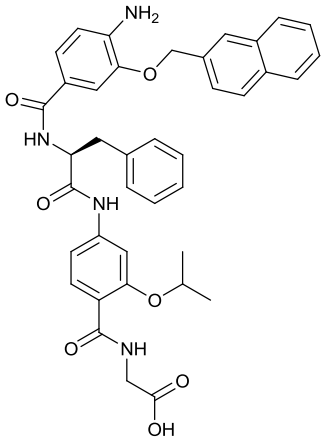
Beige solid; isolated yield: 34 mg, 50%; δ_H (500 MHz, MeOD) 7.94 (d, *J* = 8.6 Hz, 1 H, 3-*H*6), 7.58 (d, *J* = 1.9 Hz, 1 H, 3-*H*3), 7.42 – 7.38 (m, 3 H, 1-*H*2, 1-*H*Ar2), 7.37 – 7.33 (m, 2 H, 1-*H*Ar3), 7.32 – 7.29 (m, 3 H, 2-Phe-*H*Ar2, 1-*H*6), 7.25 – 7.23 (m, 2 H, 2-Phe-*H*Ar3), 7.19 – 7.17 (m, 1 H, 2-Phe-*H*Ar4), 7.08 (dd, *J* = 8.6, 1.9 Hz, 1 H, 3-*H*5), 6.70 (d, *J* = 8.2 Hz, 1 H, 1-*H*5), 5.03 (s, 2 H, 1-*H*α), 4.94 (dd, *J* = 8.2, 7.0 Hz, 1 H, 2-Phe-*H*α), 4.74 (spt, *J* = 6.3 Hz, 1 H, 3-*H*α), 4.15 (s, 2 H, 4-Gly-*H*α), 3.27 (dd, *J* = 13.9, 7.0 Hz, 1 H, 2-Phe-*H*β), 3.17 (dd, *J* = 13.9, 8.2 Hz, 1 H, 2-Phe-*H*β′), 1.41 (d, *J* = 6.3 Hz, 3 H, 3-*H*β), 1.42 (d, *J* = 6.3 Hz, 3 H, 3-*H*β′), 1.30 (s, 9 H, 1-HAr4-C(C$\underline{H}_3$)$_3$); δ_C (125 MHz, MeOD) 173.0, 171.4, 170.2, 167.2, 158.3, 152.1, 146.7, 144.2, 143.0, 138.3, 135.4, 133.1, 130.3, 129.4, 128.6, 127.8, 126.4, 122.9, 122.8, 117.9, 114.4, 113.0, 112.2, 106.5, 73.6, 71.3, 57.6, 42.6, 39.0, 35.3, 31.7, 22.1; ν_{max}/cm^{-1} (solid state) = 3362, 2960, 2929, 2868, 1651, 1496, 1445, 1404, 1218, 818; ESI-HRMS found *m/z* 681.3297 $[M+H]^+$, $C_{39}H_{45}N_4O_7$ requires 681.3283.

H_2N-[*O*-4-*t*Bu-Bn-(3-HABA)]-*p*-Br-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.51)



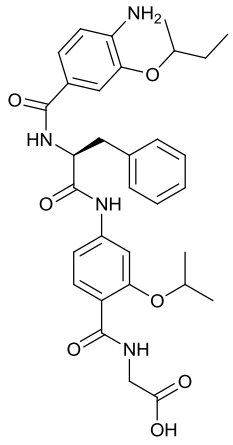
Beige solid; isolated yield: 23 mg, 30%; δ_H (500 MHz, DMSO-d_6) 10.43 (s. broad, 1 H, 4-Gly-CO_2H), 8.46 (t, J = 4.9 Hz, 1 H, 4-Gly-N*H*), 8.34 (d, J = 8.1 Hz, 1 H, 2-Phe-N*H*), 7.90 (d, J = 8.8 Hz, 1 H, 3-*H*6), 7.60 (d, J = 1.5 Hz, 1 H, 3-*H*3), 7.46 (d, J = 8.3 Hz, 2 H, 1-*H*Ar2), 7.44 – 7.38 (m, 5 H, 1-*H*2, 1-*H*Ar3, 2-Phe-*H*Ar3), 7.36 (d, J = 7.3 Hz, 2 H, 2-Phe-*H*Ar2), 7.30 (dd, J = 8.3, 1.7 Hz, 1 H, 3-*H*5), 7.25 (dd, J = 8.5, 1.7 Hz, 1 H, 1-*H*6), 6.63 (d, J = 8.3 Hz, 1 H, 1-*H*5), 5.30 (s. broad, 2 H, 1-NH_2), 5.11 (d, J = 2.1 Hz, 2 H, 1-*H*α), 4.84 – 4.76 (m, 2 H, 2-Phe-*H*α), 4.71 (quin, J = 5.9 Hz, 1 H, 3-*H*α), 4.04 (d, J = 5.1 Hz, 2 H, 4-Gly-*H*α), 3.16 – 3.04 (m, 2 H, 2-Phe-*H*β), 1.41 (dd, J = 8.5, 6.5 Hz, 6 H, 3-*H*β), 1.28 (s, 9 H, 1-HAr4-C(C$\underline{H}_3$)$_3$); δ_C (125 MHz, DMSO-d_6) 171.2, 171.2, 166.4, 163.9, 156.2, 150.1, 144.1, 143.0, 141.6, 137.7, 134.2, 132.0, 131.5, 130.9, 127.4, 125.1, 121.7, 120.7, 119.5, 116.6, 112.3, 111.3, 111.1, 104.7, 71.9, 69.1, 55.6, 41.6, 36.4, 34.3, 31.1, 21.6; ν_{max}/cm^{-1} (solid state) = 2960, 2925, 2864, 1594, 1508, 1489, 1260, 1215, 1011; ESI-HRMS found *m/z* 759.2398 $[M+H]^+$, $C_{39}H_{44}BrN_4O_7$ requires 759.2388.

H_2N-[*O*-2-Nph-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.30)



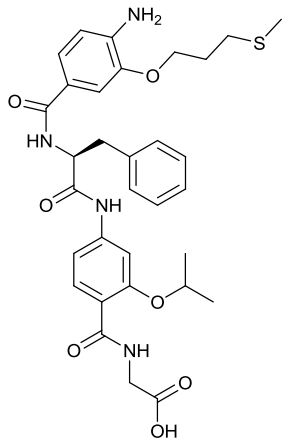
Pale brown; isolated yield: 11 mg, 16%; δ_H (500 MHz, MeOD) 7.95 (d, J = 8.6 Hz, 1 H, 3-*H*6), 7.92 – 7.81 (m, 4 H, 1-*H*Ar4, 1-*H*Ar5, 1-*H*Ar8, 1-*H*Ar7), 7.59 – 7.54 (m, 2 H, 1-*H*Ar6, 3-*H*3), 7.49 - 7.42 (m, 3 H, 1-*H*Ar1, 1-*H*Ar3, 1-*H*2), 7.33 – 7.26 (m, 3 H, 2-Phe-*H*Ar2, 1-*H*6), 7.24 – 7.21 (m, 2 H, 2-Phe-*H*Ar3), 7.19 – 7.14 (m, 1 H, 2-Phe-*H*Ar4), 7.07 (dd, J = 8.6, 1.8 Hz, 1 H, 3-*H*5), 6.73 (d, J = 8.3 Hz, 1 H, 1-*H*5), 5.27 (s, 2 H, 1-*H*α), 4.91 (dd, J = 8.3, 6.6 Hz, 1 H, 2-Phe-*H*α), 4.74 (spt, J = 6.1, 1 H, 3-*H*α), 4.15 (s, 2 H, 4-Gly-*H*α), 3.26 (dd, J = 13.8, 6.6 *H*z, 1 H, 2-Phe-*H*β), 3.17 (dd, J = 13.8, 8.3 Hz, 1 H, 2-Phe-*H*β′), 1.42 (d, J = 6.1 Hz, 3 H, 3-*H*β), 1.43 (d, J = 6.1 Hz, 3 H, 3-*H*β′); δ_C (125 MHz, MeOD) 172.9, 170.1, 167.2, 158.3, 146.7, 144.2, 143.6, 143.1, 138.3, 135.9, 134.7, 134.5, 133.1, 130.3, 129.4, 129.2, 128.9, 128.6, 127.8, 127.4, 127.2, 127.0, 126.4, 123.0, 122.9, 117.9, 114.5, 113.0, 112.5, 106.5, 73.6, 71.5, 57.6, 42.7, 39.0, 22.1; ν_{max}/cm^{-1} (solid state) = 3357, 3276, 2979, 2929, 2872, 1594, 1507, 1496, 1213; ESI-HRMS found *m/z* 675.2827 $[M+H]^+$, $C_{39}H_{39}N_4O_7$ requires 675.2813.

H_2N-[*O*-*s*Bu-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.31)



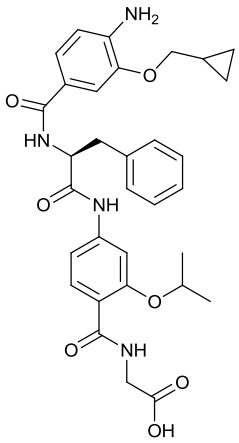
Beige solid; isolated yield: 22 mg, 37%; δ_H (500 MHz, MeOD) 7.95 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.60 (d, *J* = 1.8 Hz, 1 H, 3-*H*3), 7.34 – 7.28 (m, 3 H, 2-Phe-*H*Ar2, 1-*H*2), 7.28 – 7.23 (m, 3 H, 2-Phe-*H*Ar3, 1-*H*6), 7.21 – 7.16 (m, 1 H, 2-Phe-*H*Ar4), 7.07 (dd, *J* = 8.5, 1.8 Hz, 1 H, 3-*H*5), 6.71 (d, *J* = 8.3 Hz, 1 H, 1-*H*5), 4.90 (dd, *J* = 8.2, 6.8 Hz, 1 H, 2-Phe-*H*α), 4.77 (spt, *J* = 6.1 Hz, 1 H, 3-*H*α), 4.41 – 4.37 (m, 1 H, 1-*H*α), 4.16 (s, 2 H, 4-Gly-*H*α), 3.27 (dd, *J* = 13.5, 6.8 Hz, 1 H, 2-Phe-*H*β), 3.17 (dd, *J* = 13.5, 8.2 Hz, 1 H, 2-Phe-*H*β′), 1.80 – 1.71 (m, 1 H, 1-*H*β), 1.69 – 1.62 (m, 1 H, 1-*H*β′), 1.45 (d, *J* = 6.1 Hz, 3 H, 3-*H*β), 1.44 (d, *J* = 6.1 Hz, 3 H, 3-*H*β′), 1.28 (dd, *J* = 6.0, 4.3 Hz, 3 H, 1-CHα(C*H*$_3$)), 1.00 – 0.91 (m, 3 H, 1-*H*γ); δ_C (125 MHz, MeOD) 172.9, 172.6, 170.3, 167.2, 158.3, 145.8, 144.2, 138.3, 133.1, 130.3, 129.4, 127.8, 123.1, 122.3, 117.9, 114.6, 113.6, 113.6, 113.0, 106.5, 77.0, 73.6, 57.6, 42.5, 38.9, 30.2, 22.1, 19.6, 10.0; ν_{max}/cm^{-1} (solid state) = 3350, 3034, 2973, 2929, 1670, 1497, 1215, 699; ESI-HRMS found *m/z* 591.2824 $[M+H]^+$, $C_{32}H_{39}N_4O_7$ requires 591.2813.

H_2N-[*O*-CH2-CH2-CH2-S-CH2-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.32)



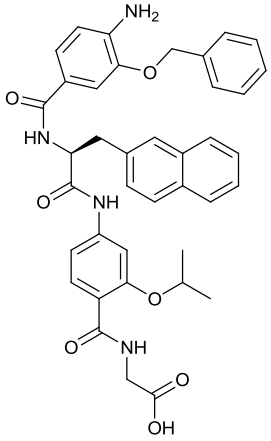
Dark yellow solid; isolated yield: 60 mg, 96%; δ_H (500 MHz, MeOD) 7.95 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.60 (d, *J* = 1.8 Hz, 1 H, 3-*H*3), 7.33 – 7.28 (m, 3 H, 2-Phe-*H*Ar2, 1-*H*2), 7.28 – 7.23 (m, 3 H, 2-Phe-*H*Ar3, 1-*H*6), 7.21 – 7.16 (m, 1 H, 2-Phe-*H*Ar4), 7.06 (dd, *J* = 8.5, 1.8 Hz, 1 H, 3-*H*5), 6.71 (d, *J* = 7.9 Hz, 1 H, 1-*H*5), 4.90 (dd, *J* = 8.3, 6.8 Hz, 1 H, 2-Phe-*H*α), 4.77 (spt, *J* = 6.1, 1 H, 3-*H*α), 4.16 (s, 2 H, 4-Gly-*H*α), 4.12 (t, *J* = 6.0 Hz, 2 H, 1-*H*α), 3.27 (dd, *J* = 13.5, 6.8 Hz, 1 H, 2-Phe-*H*β), 3.17 (dd, *J* = 13.5, 8.3 Hz, 1 H, 2-Phe-*H*β′), 2.69 (t, *J* = 7.2 Hz, 2 H, 1-*H*γ), 2.09 (m, 5 H, 1-Hβ, 1-SC*H*$_3$), 1.45 (d, *J* = 6.1 Hz, 3 H, 3-*H*β), 1.44 (d, *J* = 6.1 Hz, 3 H, 3-*H*β′); δ_C (125 MHz, MeOD) 172.9, 170.2, 167.2, 158.3, 152.3, 146.9, 144.2, 138.3, 133.1, 130.3, 129.4, 127.8, 123.3, 122.6, 117.9, 116.2, 114.5, 113.0, 111.8, 106.5, 73.6, 67.9, 57.6, 42.5, 38.9, 31.6, 29.8, 22.1, 15.3; ν_{max}/cm^{-1} (solid state) = 3357, 2974, 2918, 1595, 1508, 1497, 1212; ESI-HRMS found *m/z* 623.2540 $[M+H]^+$, $C_{32}H_{39}N_4O_7S$ requires 623.2534.

H_2N-[*O*-4-Cl-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.33)



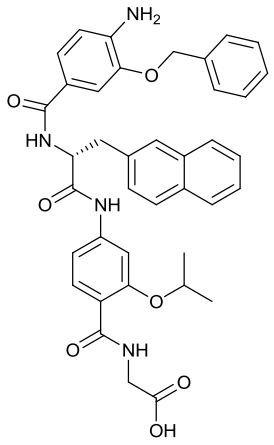
Off-white solid, > 95% pure by NMR; isolated yield: 55 mg, 93%; δ_H (500 MHz, DMSO-d_6) 10.47 (s. broad, 1 H, 4-Gly-CO_2H), 8.63 (s. broad, 1 H, 4-Gly-N*H*), 8.41 (d, *J* = 8.1 Hz, 1 H, 2-Phe-N*H*), 7.90 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.57 (s, 1 H, 3-*H*3), 7.39 (m, 3 H, 1-*H*2, 2-Phe-*H*Ar2), 7.32 – 7.22 (m, 4 H, 3-*H*5, 1-*H*6, 2-Phe-*H*Ar3), 7.17 (t, *J* = 8.5 Hz, 1 H, 2-Phe-*H*Ar4), 6.61 (d, *J* = 8.5 Hz, 1 H, 1-*H*5), 5.20 (s. broad, 2 H, 1-NH_2), 4.84 – 4.76 (m, 2 H, 2-Phe-*H*α), 4.68 (quin, *J* = 5.9 Hz, 1 H, 3-*H*α), 3.83 (d, *J* = 6.8 Hz, 2 H, 4-Gly-*H*α), 3.16 – 3.09 (m, 2 H, 2-Phe-*H*β), 1.40 (dd, *J* = 8.5, 6.5 Hz, 6 H, 3-*H*β), 0.91 – 0.82 (m, 1 H, 1-*H*α), 0.60 – 0.52 (m, 2 H, 1-*H*β), 0.37 – 0.32 (m, 2 H, 1-*H*β′); δ_C (125 MHz, DMSO-d_6) 171.4, 166.4, 163.0, 156.2, 144.4, 142.7, 141.4, 141.0, 138.3, 131.8, 129.2, 128.0, 126.3, 121.6, 121.0, 117.2, 112.2, 111.3, 111.1, 104.6, 72.5, 71.6, 56.0, 37.1, 29.0, 21.6, 10.3, 3.1; ν_{max}/cm^{-1} (solid state) = 3324, 3090, 2926, 2852, 1595, 1509, 1496, 1406, 1105, 1007, 919; ESI-HRMS found *m/z* 589.2669 $[M+H]^+$, $C_{32}H_{37}N_4O_7$ requires 589.2657.

H_2N-[*O*-Bn-(3-HABA)]-2-Nal-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.34)



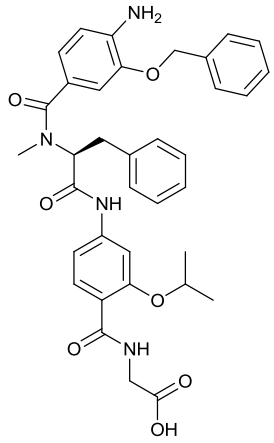
Brown solid; isolated yield: 18 mg, 27%; δ_H (500 MHz, MeOD) 7.93 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.81 – 7.68 (m, 4 H, 2-*H*Ar4, 2-*H*Ar5, 2-*H*Ar8, 2-*H*Ar7), 7.48 (d, *J* = 1.8 Hz, 1 H, 3-*H*3), 7.45 (dd, *J* = 8.5, 1.5 Hz, 1 H, 1-*H*2), 7.40 – 7.38 (m, 4 H, 1-Phe-*H*Ar2, 2-*H*Ar1, 2-*H*Ar6), 7.37 – 7.26 (m, 5 H, 2-*H*Ar3, 1-Phe-*H*Ar3, 1-Phe-*H*Ar4, 1-*H*6), 7.07 (dd, *J* = 8.7, 1.7 Hz, 1 H, 3-*H*5), 6.70 (d, *J* = 7.9 Hz, 1 H, 1-*H*5), 5.06 – 4.95 (m, 3 H, 1-*H*α, 2-Nal-*H*α), 4.65 (spt, *J* = 6.1, 1 H, 3-*H*α), 4.15 (s, 2 H, 4-Gly-*H*α), 3.44 (dd, *J* = 13.5, 8.0 Hz, 1 H, 2-Nal-*H*β), 3.34 (dd, *J* = 13.5, 7.0 Hz, 1 H, 2-Nal-*H*β′), 1.38 (d, *J* = 5.5 Hz, 3 H, 3-*H*β), 1.39 (d, *J* = 5.5 Hz, 3 H, 3-*H*β′); δ_C (125 MHz, MeOD) 172.9, 172.3, 170.2, 167.2, 158.3, 146.6, 144.1, 143.0, 138.4, 135.8, 134.9, 133.9, 133.1, 129.4, 129.0, 129.0, 128.9, 128.7, 128.6, 128.5, 128.4, 127.0, 126.6, 123.0, 122.8, 117.9, 114.4, 113.1, 112.3, 106.5, 73.6, 71.4, 57.5, 39.1, 22.1, 22.1; ν_{max}/cm^{-1} (solid state) = 3357, 2921, 2851, 1602, 1507, 1217, 697; ESI-HRMS found *m/z* 675.2827 $[M+H]^+$, $C_{39}H_{39}N_4O_7$ requires 675.2813.

H_2N-[*O*-Bn-(3-HABA)]-2-Nal-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.35)



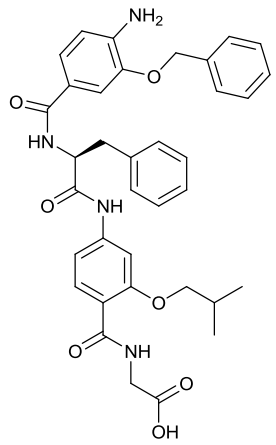
Brown solid; isolated yield: 25 mg, 37%; δ_H (500 MHz, MeOD) 7.91 (d, *J* = 8.8 Hz, 1 H, 3-*H*6), 7.77 – 7.66 (m, 4 H, 2-*H*Ar4, 2-*H*Ar5, 2-*H*Ar8, 2-*H*Ar7), 7.44 – 7.42 (m, 2 H, 3-*H*3, 1-*H*2), 7.38 – 7.34 (m, 2 H, 2-*H*Ar1, 2-*H*Ar6), 7.34 – 7.22 (m, 7 H, 1-Phe-*H*Ar2, 2-*H*Ar3, 1-Phe-*H*Ar3, 1-Phe-*H*Ar4, 1-*H*6), 7.08 (dd, *J* = 8.7, 1.8 Hz, 1 H, 3-*H*5), 6.67 (d, *J* = 7.9 Hz, 1 H, 1-*H*5), 5.10 (dd, *J* = 8.3, 6.8 Hz, 1 H, 2-Nal-*H*α), 4.88 (d, *J* = 7.9 Hz, 2 H, 1-*H*α), 4.53 (spt, *J* = 6.1, 1 H, 3-*H*α), 4.14 (s, 2 H, 4-Gly-*H*α), 3.44 (dd, *J* = 13.5, 8.0 Hz, 1 H, 2-Nal-*H*β), 3.34 (dd, *J* = 13.5, 7.0 Hz, 1 H, 2-Nal-*H*β′), 1.31 (d, *J* = 6.0 Hz, 6 H, 3-*H*β); δ_C (125 MHz, MeOD) 173.1, 172.6, 170.1, 167.1, 158.2, 146.6, 144.1, 143.0, 138.3, 135.8, 134.9, 133.8, 133.1, 129.4, 129.0, 129.0, 128.9, 128.7, 128.5, 128.5, 128.4, 127.0, 126.6, 122.9, 122.9, 117.8, 114.4, 113.0, 112.2, 106.4, 73.5, 71.3, 57.6, 39.1, 22.1, 22.0; ν_{max}/cm^{-1} (solid state) = 3376, 3054, 2979, 2929, 1603, 1443, 757, 697; ESI-HRMS found *m/z* 675.2820 $[M+H]^+$, $C_{39}H_{39}N_4O_7$ requires 675.2813.

H_2N-[*O*-Bn-(3-HABA)]-*N*-Me-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.36)



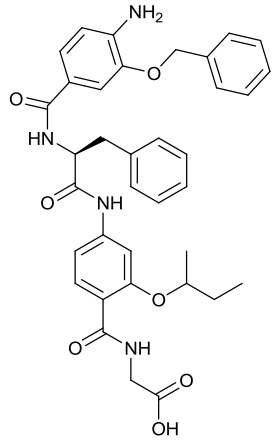
Beige solid; isolated yield: 15 mg, 23%; δ_H (500 MHz, DMSO-d_6) 10.28 (s. broad, 1 H, 4-Gly-CO_2H), 8.48 (t, *J* = 4.8 Hz, 1 H, 4-Gly-N*H*), 7.91 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.64 (s, 1 H, 3-*H*3), 7.30 – 7.20 (m, 11 H, 1-*H*Ar2, 1-*H*Ar3, 2-Phe-*H*Ar3, 2-Phe-*H*Ar2, 3-*H*5, 2-Phe-*H*Ar4, 1-*H*Ar4), 6.61 – 6.50 (m, 3 H, 1-*H*5, 1-*H*6, 1-*H*2), 5.15 (s. broad, 2 H, 1-NH_2), 4.99 – 4.83 (m, 2 H, 1-*H*α), 4.71 (spt, *J* = 5.9 Hz, 1 H, 3-*H*α), 4.00 (d, *J* = 4.7 Hz, 2 H, 4-Gly-*H*α), 3.29 (d, *J* = 12.2 Hz, 2 H, 2-Phe-*H*α), 3.12 (dd, *J* = 14.5, 10.9 Hz, 2 H, 2-Phe-*H*β), 2.88 (s, 3 H, 2-NCH_3), 1.39 (d, *J* = 6.0 Hz, 6 H, 3-*H*β); δ_C (125 MHz, DMSO-d_6) 171.9, 171.2, 170.1, 163.8, 156.2, 143.9, 142.8, 140.0, 137.6, 137.0, 131.9, 128.7, 128.3, 128.3, 127.6, 127.1, 126.4, 122.7, 121.4, 116.9, 112.7, 111.6, 105.0, 71.9, 69.3, 42.0, 34.9, 34.6, 34.5, 21.6; ν_{max}/cm^{-1} (solid state) = 3366, 3033, 2979, 2930, 1604, 1454, 1389, 1235, 699; ESI-HRMS found *m/z* 639.2820 $[M+H]^+$, $C_{36}H_{39}N_4O_7$ requires 639.2813.

H_2N-[*O*-Bn-(3-HABA)]-Phe-[*O*-*i*Bu-(2-HABA)]-Gly-CO_2H (3.41)



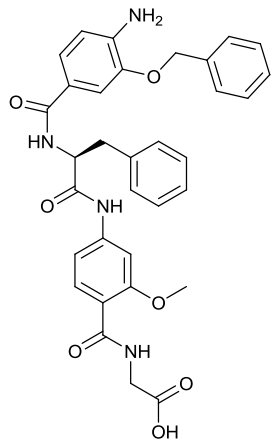
Brown solid; isolated yield: 41 mg, 63%; δ_H (500 MHz, MeOD) 7.90 (d, *J* = 8.6 Hz, 1 H, 3-*H*6), 7.47 (d, *J* = 1.7 Hz, 1 H, 3-*H*3), 7.38 (d, *J* = 1.9 Hz, 1 H, 1-*H*2), 7.37 - 7.33 (m, 2 H, 1-*H*Ar2), 7.33 - 7.24 (m, 6 H, 1-*H*Ar3, 1-*H*Ar4, 2-Phe-*H*Ar2, 1-*H*6), 7.18 - 7.24 (m, 2 H, 2-Phe-*H*Ar3), 7.14 - 7.18 (m, 1 H, 2-Phe-*H*Ar4), 7.09 (dd, *J* = 8.6, 1.7 Hz, 1 H, 3-*H*5), 6.67 (d, *J* = 8.3 Hz, 1 H, 1-*H*5), 5.00 (dd, *J* = 8.2, 6.8 Hz, 1 H, 2-Phe-*H*α), 4.96 (d, *J* = 3.2 Hz, 2 H, 1-*H*α), 4.15 (s, 2 H, 4-Gly-*H*α), 3.68 - 3.80 (m, 2 H, 3-*H*α), 3.13 - 3.29 (dd, *J* = 8.2, 6.8 Hz, 2 H, 2-Phe-*H*β), 2.17 (m, 1 H, 3-*H*β), 0.97 (dd, *J* = 4.5, 6.8 Hz, 6 H, 3-*H*γ); δ_C (125 MHz, MeOD) 173.3, 172.7, 170.0, 167.2, 159.4, 146.6, 144.3, 143.0, 138.4, 133.1, 130.3, 130.3, 129.5, 129.4, 128.9, 128.6, 127.8, 122.9, 122.8, 117.0, 114.4, 112.8, 112.3, 104.9, 76.8, 71.3, 57.8, 42.6, 38.9, 29.0, 19.6; ν_{max}/cm^{-1} (solid state) = 3370, 2956, 2924, 2873, 1600, 1498, 1394, 1215, 1016, 697; ESI-HRMS found *m/z* 639.2830 $[M+H]^+$, $C_{36}H_{39}N_4O_7$ requires 639.2813.

H_2N-[*O*-Bn-(3-HABA)]-Phe-[*O*-*s*Bu-(2-HABA)]-Gly-CO_2H (3.42)



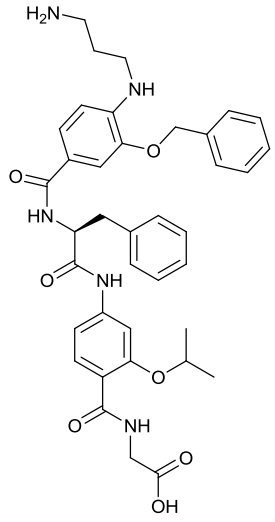
Brown solid; isolated yield: 15 mg, 23%; δ_H (500 MHz, MeOD) 7.94 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.53 (d, *J* = 1.6 Hz, 1 H, 3-*H*3), 7.40 – 7.39 (m, 3 H, 1-*H*2, 1-*H*Ar2), 7.36 – 7.25 (m, 6 H, 1-*H*Ar3, 1-*H*Ar4, 2-Phe-*H*Ar2, 1-*H*6), 7.23 (t, *J* = 7.4 Hz, 2 H, 2-Phe-*H*Ar3), 7.20 – 7.14 (m, 1 H, 2-Phe-*H*Ar4), 7.10 (dd, *J* = 8.5, 1.6 Hz, 1 H, 3-*H*5), 6.73 (d, *J* = 8.3 Hz, 1 H, 1-*H*5), 5.02 (s, 2 H, 1-*H*α), 4.99 – 4.96 (m, 1 H, 2-Phe-*H*α), 4.50 – 4.41 (m, 1 H, 3-*H*α), 4.15 (s, 2 H, 4-Gly-*H*α), 3.28 – 3.16 (m, 2 H, 2-Phe-*H*β), 1.91 – 1.83 (m, 1 H, 3-*H*β), 1.73 – 1.65 (m, 1 H, 3-*H*β′), 1.34 (dd, *J* = 11.2, 6.1 Hz, 3 H, 3-CHα(C<u>*H*</u>$_3$)), 1.00 – 0.91 (m, 3 H, 3-*H*γ); δ_C (125 MHz, MeOD) 173.1, 172.5, 170.0, 167.2, 158.5, 147.0, 144.2, 141.8, 138.3, 138.3, 133.2, 130.3, 129.5, 128.9, 128.7, 127.8, 123.7, 122.8, 117.7, 115.0, 112.9, 112.3, 106.3, 78.6, 71.2, 57.7, 42.5, 38.9, 29.8, 19.5, 10.1; ν_{max}/cm^{-1} (solid state) = 3334, 2966, 2924, 2853, 1696, 1495, 1219, 1201, 650; ESI-HRMS found *m/z* 639.2826 $[M+H]^+$, $C_{36}H_{39}N_4O_7$ requires 639.2813.

H_2N-[*O*-Bn-(3-HABA)]-Phe-[*O*-Me-(2-HABA)]-Gly-CO_2H (3.43)



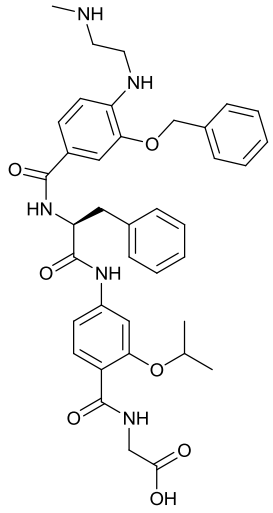
Brown solid; isolated yield: 30 mg, 50%; δ_H (500 MHz, MeOD) 7.93 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.58 (d, *J* = 1.8 Hz, 1 H, 3-*H*3), 7.47 – 7.45 (m, 2 H, 1-*H*Ar2), 7.41 – 7.34 (m, 3 H, 1-*H*2, 1-*H*Ar3), 7.34 – 7.28 (m, 4 H, 1-*H*Ar4, 2-Phe-*H*Ar2, 1-*H*6), 7.27 – 7.24 (m, 2 H, 2-Phe-*H*Ar3), 7.22 – 7.17 (m, 1 H, 2-Phe-*H*Ar4), 7.09 (dd, *J* = 8.6, 1.7 Hz, 1 H, 3-*H*5), 6.72 (d, *J* = 7.9 Hz, 1 H, 1-*H*5), 5.12 (s, 2 H, 1-*H*α), 4.91 (dd, *J* = 8.1, 6.9 Hz, 1 H, 2-Phe-*H*α), 4.11 (s, 2 H, 4-Gly-*H*α), 3.97 (s, 3 H, 3-*H*α), 3.27 (dd, *J* = 13.5, 6.8 Hz, 1 H, 2-Phe-*H*β), 3.17 (dd, *J* = 13.5, 8.3 Hz, 1 H, 2-Phe-*H*β′); δ_C (125 MHz, MeOD) 172.9, 170.2, 169.6, 167.2, 160.0, 146.7, 144.3, 143.0, 138.4, 138.3, 133.0, 130.3, 129.5, 129.4, 128.9, 128.7, 127.8, 123.0, 122.8, 117.3, 114.4, 112.9, 112.3, 104.2, 79.4, 71.4, 57.6, 56.5, 38.9; ν_{max}/cm^{-1} (solid state) = 3370, 3029, 2924, 2852, 1601, 1498, 1452, 1405, 1217, 1026, 698; ESI-HRMS found *m/z* 597.2355 $[M+H]^+$, $C_{33}H_{33}N_4O_7$ requires 597.2344.

3-NH_2-Pr-HN-[*O*-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.44)



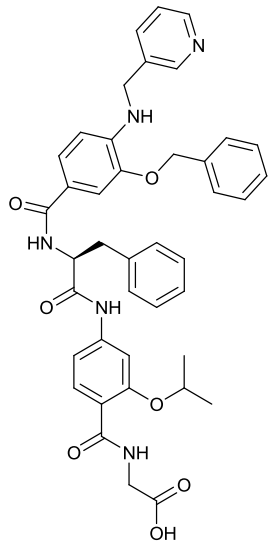
Off-white solid; isolated yield: 26 mg, 38%; δ_H (500 MHz, MeOD) 7.92 (d, *J* = 8.6 Hz, 1 H, 3-*H*6), 7.48 (d, *J* = 1.7 Hz, 1 H, 3-*H*3), 7.40 (dd, *J* = 8.4, 1.7 Hz, 1 H, 1-*H*6), 7.36 (d, *J* = 1.7 Hz, 1 H, 1-*H*2), 7.35 – 7.25 (m, 7 H, 2-Phe-*H*Ar2, 1-*H*Ar2, 1-*H*Ar3, 1-*H*Ar4), 7.24 – 7.21 (m, 2 H, 2-Phe-*H*Ar3), 7.18 – 7.15 (m, 1 H, 2-Phe-*H*Ar4), 7.10 (dd, *J* = 8.6, 1.7 Hz, 1 H, 3-*H*5), 6.54 (d, *J* = 8.4 Hz, 1 H, 1-*H*5), 5.06 (dd, *J* = 8.5, 7.0 Hz, 1 H, 2-Phe-*H*α), 4.94 (d, *J* = 4.5 Hz, 2 H, 1-*H*α), 4.58 (sept, *J* = 6.1 Hz, 1 H, 3-*H*α), 3.95 (s, 2 H, 4-Gly-*H*α), 3.29 – 3.25 (m, 3 H, 2-Phe-*H*β, 1-NHC$\underline{H_2}$), 3.23 (dd, *J* = 7.5, 13.5 Hz, 1 H, 2-Phe-*H*β′), 2.96 (t, *J* = 7.5, 2 H, 1-$NHCH_2CH_2C\underline{H_2}$), 1.95 – 1.86 (m, 2 H, 1-$NHCH_2C\underline{H_2}$), 1.35 (d, *J* = 10.9 Hz, 3 H, 3-*H*β), 1.34 (d, *J* = 10.6 Hz, 3 H, 3-*H*β′); ν_{max}/cm^{-1} (solid state) = 3355, 3033, 2979, 2928, 1594, 1509, 1491, 1256, 1211, 1105, 697; ESI-HRMS found *m/z* 341.6654 $[M+2H]^{2+}$, $C_{38}H_{45}N_5O_7$ requires 341.6654.

2-MeNH-Et-HN-[*O*-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.45)



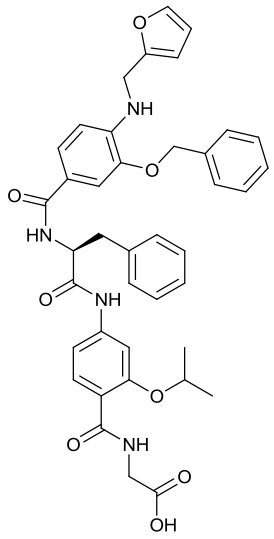
Pale orange solid; isolated yield: 11 mg, 16%; δ_H (500 MHz, DMSO-d_6) 10.47 (s. broad, 1 H, 4-Gly-CO_2H), 8.57 – 8.35 (m, 2 H, 4-Gly-N*H*, 2-Phe-N*H*), 7.90 (d, *J* = 8.7 Hz, 1 H, 3-*H*6), 7.64 (s, 1 H, 3-*H*3), 7.50 (d, *J* = 8.5 Hz, 2 H, 2-Phe-*H*Ar2), 7.45 (s, 1 H, 1-*H*2), 7.43 – 7.37 (m, 5 H, 1-*H*Ar2, 1-*H*Ar3, 1-*H*Ar4), 7.34 (t, *J* = 8.5 Hz, 2 H, 2-Phe-*H*Ar3), 7.26 (m, 3 H, 1-*H*6, 3-*H*5, 3-N*H*), 7.17 (t, *J* = 8.5 Hz, 1 H, 2-Phe-*H*Ar4), 6.68 (d, *J* = 8.7 Hz, 1 H, 1-*H*5), 5.71 (s. broad, 1 H, 1-N*H*), 5.19 (s, 2 H, 1-*H*α), 4.88 – 4.77 (m, 2 H, 2-Phe-*H*α), 4.71 (quin, *J* = 5.9 Hz, 1 H, 3-*H*α), 4.05 (d, *J* = 4.8 Hz, 2 H, 4-Gly-*H*α), 3.47 (m, 2 H, 1-NHC$\underline{H_2}$), 3.19 – 3.03 (m, 4 H, 2-Phe-*H*β, 1-$NHCH_2C\underline{H_2}$), 2.61 (s, 3 H, 1-$NHCH_2CH_2NHC\underline{H_3}$), 1.41 (dd, *J* = 8.5, 6.5 Hz, 6 H, 3-*H*β); δ_C (125 MHz, DMSO-d_6) 171.4, 171.2, 166.2, 163.9, 156.2, 144.6, 143.1, 140.5, 138.1, 137.0, 131.9, 129.2, 128.4, 128.1, 127.8, 127.5, 126.3, 121.9, 121.2, 116.6, 111.3, 110.5, 108.0, 104.7, 71.9, 69.6, 55.9, 47.1, 41.6, 38.8, 37.1, 32.8, 21.6; ESI-HRMS found *m/z* 682.3253 $[M+H]^+$, $C_{38}H_{44}N_5O_7$ requires 682.3235.

3-Pyridyl-CH_2-HN-[*O*-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.46)



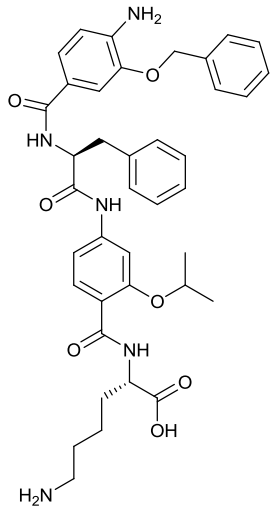
Pale brown solid; isolated yield: 12mg, 17%; δ_H (500 MHz, MeOD) 8.71 (s, 1 H, 1-NHPyr-*H*2), 8.49 (d, *J* = 8.0 Hz, 1 H, 1-NHPyr-*H*6), 7.98 (dd, *J* = 8.0, 5.8 Hz, 1 H, 1-NHPyr-*H*4), 7.93 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.53 (d, *J* = 1.7 Hz, 1 H, 3-*H*3), 7.48 – 7.40 (m, 3 H, 1-*H*6, 1-*H*2, 1-NHPyr-*H*5), 7.40 - 7.27 (m, 7 H, 2-Phe-*H*Ar2, 1-*H*Ar2, 1-*H*Ar3, 1-*H*Ar4), 7.25 - 7.22 (m, 2 H, 2-Phe-*H*Ar3), 7.21 - 7.15 (m, 1 H, 2-Phe-*H*Ar4), 7.10 (dd, *J* = 8.6, 1.7 Hz, 1 H, 3-*H*5), 6.41 (d, *J* = 8.5 Hz, 1 H, 1-*H*5), 5.12 (d, *J* = 4.6 Hz, 2 H, 1-*H*α), 4.94 (dd, *J* = 8.2, 6.9 Hz, 1 H, 2-Phe-*H*α), 4.76 – 4.63 (m, 3 H, 2-*H*α, 1-NHC$\underline{H_2}$), 4.16 (s, 2 H, 4-Gly-*H*α), 3.27 (dd, *J* = 13.9, 6.8 Hz, 1 H, 2-Phe-*H*β), 3.18 (dd, *J* = 7.5, 13.5 Hz, 1 H, 2-Phe-*H*β′), 1.43 (d, *J* = 6.2 Hz, 3 H, 3-*H*β), 1.40 (d, *J* = 6.2 Hz, 3 H, 3-*H*β′); δ_C (125 MHz, MeOD) 173.0, 172.6, 169.9, 167.2, 158.3, 146.7, 144.2, 143.7, 143.5, 143.4, 142.1, 138.3, 138.1, 133.1, 130.3, 129.5, 129.4, 129.1, 129.0, 128.7, 127.8, 127.4, 123.0, 122.8, 117.9, 113.0, 111.7, 109.9, 106.5, 73.6, 71.7, 57.6, 44.6, 42.5, 38.9, 22.1; ν_{max}/cm^{-1} (solid state) = 3383, 2980, 2930, 1603, 1441, 1406, 1239, 1202, 1119, 700; ESI-HRMS found *m/z* 716.3085 $[M+H]^+$, $C_{41}H_{41}N_5O_7$ requires 716.3079.

2-Furyl-CH_2-HN-[*O*-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Gly-CO_2H (3.47)



Brown solid; isolated yield: 14 mg, 20%; δ_H (500 MHz, MeOD) 7.93 (d, *J* = 8.7 Hz, 1 H, 3-*H*6), 7.54 (d, *J* = 1.6 Hz, 1 H, 3-*H*3), 7.41 – 7.26 (m, 10 H, 1-*H*6, 1-*H*2, 2-Phe-*H*Ar2, 1-*H*Ar2, 1-*H*Ar3, 1-*H*Ar4, 1-NHFur-*H*5), 7.24 - 7.21 (m, 2 H, 2-Phe-*H*Ar3), 7.20 - 7.15 (m, 1 H, 2-Phe-*H*Ar4), 7.08 (dd, *J* = 8.6, 1.7 Hz, 1 H, 3-*H*5), 6.62 (d, *J* = 8.2 Hz, 1 H, 1-*H*5), 6.30 (dd, *J* = 3.1, 1.8 Hz, 1 H, 1-NHFur-*H*4), 6.19 (d, *J* = 3.1 Hz, 1 H, 1-NHFur-*H*3), 5.03 (d, *J* = 2.7 Hz, 2 H, 1-*H*α), 4.95 (dd, *J* = 8.3, 6.9 Hz, 1 H, 2-Phe-*H*α), 4.67 (sept, *J* = 6.1 Hz, 1 H, 3-*H*α), 4.36 (s, 2 H, 1-NHC*H*$_2$), 4.15 (s, 2 H, 4-Gly-*H*α), 3.27 (dd, *J* = 6.9, 13.6 Hz, 1 H, 2-Phe-*H*β), 3.17 (dd, *J* = 8.3, 13.6 Hz, 1 H, 2-Phe-*H*β′), 1.40 (d, *J* = 6.1 Hz, 3 H, 3-*H*β), 1.37 (d, *J* = 6.1 Hz, 3 H, 3-*H*β′); δ_C (125 MHz, MeOD) 173.1, 172.6, 170.1, 167.2, 158.3, 154.1, 146.6, 144.2, 143.0, 142.9, 138.3, 138.2, 133.1, 130.3, 129.5, 129.4, 129.0, 128.8, 127.8, 123.0, 122.0, 117.9, 113.0, 111.5, 111.2, 110.0, 107.8, 106.5, 73.6, 71.6, 57.6, 42.5, 41.0, 38.9, 22.1; ν_{max}/cm^{-1} (solid state) = 3423, 3032, 2980, 2930, 1603, 1441, 1406, 1228, 699; ESI-HRMS found *m/z* 705.2920 $[M+H]^+$, $C_{40}H_{41}N_4O_8$ requires 705.2919.

H_2N-[*O*-Bn-(3-HABA)]-Phe-[*O*-*i*Pr-(2-HABA)]-Lys-CO_2H (3.48)



Brown solid; isolated yield: 41 mg, 59%; δ_H (500 MHz, MeOD) 7.94 (d, *J* = 8.5 Hz, 1 H, 3-*H*6), 7.55 (d, *J* = 1.8 Hz, 1 H, 3-*H*3), 7.45 – 7.44 (m, 2 H, 2-Phe-*H*Ar2), 7.41 (d, *J* = 1.7 Hz, 1 H, 1-*H*2), 7.37 – 7.33 (m, 2 H, 1-*H*Ar2), 7.33 – 7.24 (m, 4 H, 1-*H*Ar3, 1-*H*Ar4, 1-*H*6), 7.27 – 7.24 (m, 2 H, 2-Phe-*H*Ar3), 7.21 – 7.18 (m, 1 H, 2-Phe-*H*Ar4), 7.12 (dd, *J* = 8.5, 1.8 Hz, 1 H, 3-*H*5), 6.76 (d, *J* = 8.3 Hz, 1 H, 1-*H*5), 5.11 (d, *J* = 1.5 Hz, 2 H, 1-*H*α), 4.90 (dd, *J* = 8.1, 7.1 Hz, 1 H, 2-Phe-*H*α), 4.80 – 4.74 (m, 1 H, 3-*H*α), 4.69 (dd, *J*=6.8, 5.6 Hz, 1 H, 4-Lys-*H*α), 3.29 – 3.23 (dd, 2 H, 2-Phe-*H*β), 2.92 (t, *J* = 7.0 Hz, 2 H, 4-Lys-*H*ε), 2.11 – 1.81 (m, 2 H, 4-Lys-*H*β), 1.79 – 1.64 (m, 2 H, 4-Lys-*H*δ), 1.58 – 1.48 (m, 2 H, 4-Lys-*H*γ), 1.47 (d, *J* = 6.2 Hz, 3 H, 3-*H*β) 1.40 (d, *J* = 6.0 Hz, 3 H, 3-*H*β′); δ_C (125 MHz, MeOD) 174.7, 173.0, 170.1, 166.8, 158.2, 147.1, 144.3, 141.8, 138.3, 133.1, 130.3, 129.6, 129.5, 129.4, 129.0, 128.7, 127.8, 123.9, 122.7, 117.8, 115.1, 113.0, 112.4, 106.3, 73.5, 71.5, 57.7, 53.5, 40.4, 38.9, 33.2, 28.1, 23.3, 22.3, 22.0; ν_{max}/cm^{-1} (solid state) = 3032, 2932, 1670, 1626, 1600, 1495, 1198, 1182, 1132, 698, 681; ESI-HRMS found *m/z* 696.3410 $[M+H]^+$, $C_{39}H_{46}N_5O_7$ requires 696.3392.

6.4.4. Fluorescence anisotropy assays

WT-p53(15-31) transactivation domain peptide and its fluorescein-labelled analogue p53 (15-31) Flu were purchased from Peptide Protein Research Ltd and used without further purification. WT-NOXA B (68-87) and its fluorescein-labelled analogue FITC-NOXA B (68-87) were synthesised and purified by Dr P. Prabhakaran and Dr. D. J. Yeo. Molecular cloning and expression of *h*DM2(17-126) L33E (no tag) and Mcl-1 (172-327) proteins were performed by Dr A. Bartlett and Dr J. Miles.

The Fluorescence Anisotropy Assays were run in 384 well plates (Greiner Bio-one) in triplicate and were scanned using a Perkin Elmer EnVisionTM 2103 MultiLabel plate reader. Fluorescein labeled peptides used an excitation and emission wavelength of 480 nm (30 nm bandwidth) and 535 nm (40 nm bandwidth) respectively. All the assays were performed in phosphate buffer (40 mM phosphate buffer at pH 7.50, containing 200 mM NaCl and 0.02 mg mL^{-1} *bovine serum albumin* (BSA).

Fluorescence Anisotropy Competition Assay

For the fluorescence anisotropy competition assays, compound stock solutions (400 μM in 90:10 (v/v) assay buffer:DMSO) were used to carry out the serial dilutions across the plate (starting point: 100 μM; 24-points, 3/4 serial dilution). The fluorescently labeled peptide and protein were added to each well to give a final concentration of 50 nM and 150 nM, respectively. The assays consist of four rows (three containing the compound of interest and one control row). For control wells, the tracer peptide was replaced with an identical volume of assay buffer. The additions to the wells go as follows:

1) 20 μL of assay buffer
2) 80 μL of compound stock solution were added to the first well of each row, mixed and 80 μL transferred to the next well to complete the serial dilution.
3) 20 μL of fluorescently labeled peptide solution to the wells with compound and 20 μL of assay buffer to the control wells.
4) 20 μL of protein solution

The plates were incubated for 2 h at rt and protected from light.

Processing of the fluorescent anisotropy data

The data obtained for both the *P* (perpendicular intensity) and *S* (parallel (same) intensity) channels was corrected by subtracting the corresponding control wells. The resulting values were used to calculate the intensity and anisotropy for each well using the following equations:

$$I = 2PG + S$$

$$r = \frac{S - PG}{I}$$

Where: r = anisotropy, I = total intensity, P = perpendicular intensity, S = parallel intensity, G = 1 (instrumental factor).

The average anisotropy (across three replicates) and the standard deviation for these values were calculated and fit to a logisitic model (for calculation of IC_{50} values) using Origin Pro 9.0:

$$y = r_{min} + \frac{r_{max} - r_{min}}{1 + 10^{(x - logx_0)}}$$

Where for the logisitic model; $y = r$ = anisotropy, x_0 = mid-point of the curve between the r_{max} and r_{min} plateaux.

6.5. Design, synthesis and evaluation of inhibitors for the Asf1/H3 PPI (*Chapter 4*)

6.5.1. General procedures for 3-*O*-alkylated monomer synthesis[125]

<u>Procedure A:</u> RBr Alkylation

To a stirred solution of methyl-3-hydroxy-4-nitrobenzoate **2.3** (1 equiv) and potassium carbonate (3 equiv) in DMF (20 mL / g), was added RBr (1.2 equiv) and the resulting mixture stirred at 50 ºC overnight under a nitrogen atmosphere. Further portions of the RBr were added when the reaction was shown incomplete by TLC. The resultant mixture was allowed to cool to rt, poured into water (40 mL / g) and extracted with ethyl acetate. The combined organic fractions were thoroughly washed with water and further washed with brine, dried with $MgSO_4$ and evaporated to dryness.

Procedure B: Mitsunobu

A stirred solution containing methyl-3-hydroxy-4-nitrobenzoate **2.3** (1 equiv), ROH (1.1 equiv) and triphenylphosphine (1.5 equiv) in anhydrous tetrahydrofuran (30 mL / g) was cooled to 0 °C. Diisopropyl azodicarboxylate (1.5 equiv) was added and the resulting solution allowed to warm to rt and left stirring overnight under a nitrogen atmosphere. Organic solvents were removed under reduced pressure and the product was purified *via* column chromatography.

Procedure C: Tin Reduction

To a stirred solution containing either i) nitro/ester or ii) nitro/acid (1 equiv) in ethyl acetate (20 mL / g), tin(II) chloride dihydrate (6 equiv) was added and the resulting mixture stirred at 50 ºC overnight, under a nitrogen atmosphere. On completion, the reaction mixture was allowed to cool to rt and poured over ice. The pH was made slightly basic (~pH 8) by addition of a 1 M NaOH solution and the resulting basic mixture was filtered under vacuum to remove $SnO{\cdot}H_2O(s)$ in the suspension. The aqueous mixture was extracted with ethyl acetate and the combined organic fractions washed thoroughly with brine, dried with $MgSO_4$ and evaporated to dryness.

Procedure D: Hydrogenation

A solution containing either i) nitro/ester or ii) nitro/acid (1 equiv) in methanol (20 mL / g) and palladium on carbon (10 wt %) was evacuated and flushed with nitrogen (3 times) and left under vacuum. Hydrogen was drawn into the flask and the reaction left stirring at rt overnight. On completion, the reaction mixture was filtered through a celite pad and evaporated to dryness.

Procedure E: NaOH Saponification

To a solution containing either i) amine/ester or ii) nitro/ester (1 equiv) in a 1:1 mixture of methanol: tetrahydrofuran (25 mL / g), a 10 % sodium hydroxide solution (5 mL / g) was added and the resulting mixture was allowed to stir at rt overnight. Further portions of the hydroxide solution were added when the reaction was shown incomplete by TLC. The organic solvents were removed under reduced pressure, the remaining solution was poured into water and extracted with dichloromethane (unreacted starting material). The aqueous layer was acidified *via* the addition of hydrochloric acid (conc) to pH 4 and the resulting precipitate was extracted into dichloromethane. The combined organic extracts were washed with water and brine, dried with $MgSO_4$ and evaporated to dryness.

Procedure F: **LiOH Saponification**

To a solution containing either i) amine/ester or ii) nitro/ester (1 equiv) in a 1:1 mixture of tetrahydrofuran / water (25 mL / g), a lithium hydroxide (1 equiv) solution in water was added and the resulting mixture was stirred at rt overnight. The organic solvent was removed under reduced pressure and an additional amount of water was added. The resulting solution was extracted with dichloromethane (unreacted starting material) and the aqueous layer acidified *via* the addition of 1 M potassium bisulfate solution to pH 4. The resulting precipitate was extracted into dichloromethane and the combined organic extracts were washed with water and brine, dried with $MgSO_4$ and evaporated to dryness.

Procedure G: **Fmoc protection**

A solution of amine/acid (1 equiv) in anhydrous tetrahydrofuran (20 mL / g) was held at a reflux under a nitrogen atmosphere. A solution of fluorenylmethyloxycarbonyl chloride (1.5 equiv) in anhydrous tetrahydrofuran (10 mL / g) was then added dropwise and the resulting solution was stirred at reflux overnight. The reaction mixture was evaporated to dryness, the resulting solid was crystallized from a chloroform/hexane mixture and the precipitate collected *via* filtration.

Procedure H: **Fmoc protection**

A solution of amine/acid (1 equiv) and sodium bicarbonate (3 equiv) in anhydrous tetrahydrofuran (20 mL / g) was held at a reflux under a nitrogen atmosphere. A solution of fluorenylmethyloxycarbonyl chloride (1.5 equiv) in anhydrous tetrahydrofuran (10 mL / g) was then added dropwise and the resulting solution was stirred at reflux overnight. Sodium bicarbonate was removed *via* hot filtration and the reaction mixture was evaporated to dryness. The resulting solid was crystallized from a chloroform/hexane mixture and the precipitate collected *via* filtration.

6.5.2. General procedures for Solid Phases Synthesis of the 3-*O*-alkylated scaffold[125]

General Points for Solid Phase Synthesis: Fmoc-Gly-Wang resin (0.79 mmol/g, 100-200 mesh; carrier: polystyrene, crosslinked with 1% DVB), was purchased from Merck. All solvents used were HPLC grade. Anhydrous *N*-methyl-2-pyrrolidone was purchased from Acros Organics. 1-Chloro-*N*, *N*, 2-trimethyl-1-propenylamine (Ghosez's reagent) was purchased from Sigma-Aldrich and stored in a schlenk tube under a nitrogen atmosphere in the freezer as a 20% solution in chloroform. Oligomer formation was carried out on a CEM™ Liberty automated

microwave peptide synthesiser. The volume of the reaction mixture in the reaction vessel was 2.5 mL.

Procedure I: Acyl Chloride Formation

To a stirred solution of an Fmoc protected building block in anhydrous dichloromethane (20 mL / g), thionyl chloride (10 equiv) was added and the resulting mixture refluxed overnight. The organic solvent and thionyl chloride were removed under N_2 (g) to yield the corresponding acyl chloride monomer, which was stored under an inert atmosphere for a maximum of 12 h before being re-dissolved in anhydrous NMP (2.5 mL) and added to the resin for microwave treatment.

Procedure J: Acyl Chloride Preactivation

To a solution containing Fmoc protected monomers (1 equiv) functionalised with acid sensitive protecting groups in anhydrous NMP (2.5 mL), 0.9 equiv of Ghosez's reagent was added. The resulting mixture was stirred under an inert atmosphere for 3 h at 50 ºC before the addition to the resin and microwave treatment.

Procedure K: Oligomer Formation

Fmoc-glycine pre-loaded Wang resin (127 mg, 0.1 mmol, 1 equiv) was loaded onto a Liberty CEM™ microwave peptide synthesiser after being swelled for a total of 30 min in dichloromethane. A series of washes (3 × NMP), deprotection (2 × 20 % Piperidine/NMP, total of 3.5 min at 75 °C) and further washes (5 × NMP) prepared the resin for coupling.

Single coupling: Fmoc protected acyl chloride (0.4 mmol, 4 equiv) obtained by pre-activation or prepared separately was dissolved in NMP (2.5 mL), delivered to the reaction vessel and submitted to microwave irradiation at 50 °C for 30 min.

Double coupling: Fmoc protected acyl chloride (0.2 mmol, 2 equiv) obtained by pre-activation or prepared separately was dissolved in NMP (2.5 mL), delivered to the reaction vessel and submitted to microwave irradiation at 50 °C for 30 min. A second solution containing Fmoc protected acyl chloride (0.2 mmol, 2 equiv) (preactivated or isolated) in NMP (2.5 mL) was delivered to the reaction vessel and submitted to microwave power at 50 °C for 30 min.

A coupling cycle finished with a final series of washes of the reaction vessel (3 × NMP). After the required number of cycles, a final Fmoc deprotection was carried out and then the resin was removed from the synthesiser and transferred to a reservoir for manual cleavage.

Procedure L: Cleavage

The resin was washed with dichloromethane (2 × 3 mL × 2 min) and cleaved with a 3 mL cleavage cocktail (2 × 30 min) consisting of TFA:dichloromethane:triisopropylsilane (49.5:49.5:1). If no protecting groups were present, a simple 1:1 TFA:dichloromethane mixture was sufficient without the need for a scavenger. Once cleavage was complete, the cleavage solution was isolated and evaporated under N_2 (g).

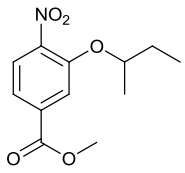
Procedure M: Mass – Directed HPLC Purification

The resulting crude mixture was dissolved in either DMSO or methanol at an approximate concentration of 20 mg mL^{-1} and purified using reversed phase mass directed HPLC [Agilent XBridge C18 preparative column; 50-95% gradient of MeCN to water (plus 0.1% formic acid v/v in both solvents) and flow rate of 20 mL min^{-1} during 8 min]. The resulting fractions were concentrated by centrifugal evaporation (Genevac).

6.5.3. 3-*O*-Alkylated monomer syntheses and characterization

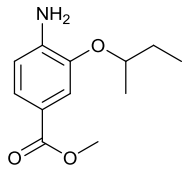
Only the characterization of the novel monomers has been included in the experimental section; the monomers which have been routinely synthesised and reported previously in the group have not been included.[125]

Methyl 3-sec-butoxy-4-nitrobenzoate (3.8)



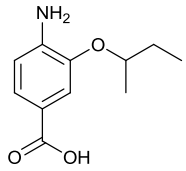
Procedure B; methyl 3-hydroxy-4-nitrobenzoate **2.3** (2.42 g, 12.3 mmol), *sec*-butanol (1.0 g, 13.5 mmol), triphenylphosphine (4.8 g, 18.4 mmol) in anhydrous tetrahydrofuran (80 mL) and diisopropyl azodicarboxylate (3.61 mL, 18.4 mmol). The reaction mixture was purified by column chromatography (*Stationary Phase*: Silica; *Mobile Phase*: ethyl acetate/hexane, 1:1) to afford the desired product **3.8** (3.10 g, 12.2 mmol, quant.) as a yellow oil; δ_H (500 MHz, $CDCl_3$) 7.76 (d, *J* = 8.2 Hz, 1 H, *H*5), 7.73 (d, *J* = 1.0 Hz, 1 H, *H*2), 7.66 (dd, *J* = 8.2, 1.0 Hz, 1 H, *H*6), 4.52 – 4.60 (m, 1 H, *H*α), 3.97 (s, 3 H, CO_2CH_3), 1.85 – 1.76 (m, 1 H, *H*β), 1.67 – 1.76 (m, 1 H, *H*β′), 1.37 (d, *J* = 6.0 Hz, 3 H, CHα($C\underline{H}_3$)), 1.01 (t, *J* = 7.4 Hz, 3 H, Hγ); δ_C (125 MHz, $CDCl_3$) 165.4, 151.1, 143.7, 134.4, 125.1, 120.9, 116.7, 77.8, 52.8, 29.0, 19.0, 9.5; ν_{max}/cm^{-1} (solid state) = 2974, 2937, 2880, 1726, 1605, 1528, 1290, 1235, 744; ESI-HRMS found *m/z* 276.0843 $[M+Na]^+$, $C_{12}H_{15}NNaO_5$ requires 276.0842.

Methyl 4-amino-3-sec-butoxybenzoate (3.13)



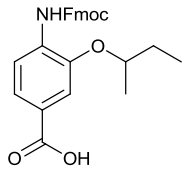
Procedure D; methyl 3-sec-butoxy-4-nitrobenzoate **3.8** (3.10 g, 12.2 mmol) in a 1:1 mixture of tetrahydrofuran/methanol (40 mL). Work up yielded the title compound **3.13** (2.72 g, 12.2 mmol, quant.) as a yellow oil; δ_H (500 MHz, $CDCl_3$) 7.52 (dd, *J* = 8.2, 1.5 Hz, 1 H, *H*6), 7.46 (d, *J* = 1.5 Hz, 1 H, *H*2), 6.67 (d, *J* = 8.2 Hz, 1 H, *H*5), 4.39 – 4.45 (m, 1 H, *H*α), 3.86 (s, 3 H, CO_2CH_3), 1.73 – 1.83 (m, 1 H, *H*β), 1.62 – 1.72 (m, 1 H, *H*β′), 1.32 (d, *J* = 6.2 Hz, 3 H, CHα($C\underline{H}_3$)), 1.00 (t, *J* = 7.4 Hz, 3 H, Hγ); δ_C (125 MHz, $CDCl_3$) 167.4, 144.4, 142.1, 123.8, 119.5, 113.9, 113.4, 75.8, 51.6, 29.2, 19.3, 9.8; ν_{max}/cm^{-1} (solid state) = 3486, 3369, 2971, 2878, 1698, 1614, 1518, 1440, 1290, 1258, 1210, 728; ESI-HRMS found *m/z* 224.1285 $[M+H]^+$, $C_{12}H_{18}NO_3$ requires 224.1281.

4-Amino-3-sec-butoxybenzoic acid (3.18)



Procedure E; methyl 4-amino-3-sec-butoxybenzoate **3.13** (2.72 g, 12.2 mmol) in a 1:1 mixture of tetrahydrofuran/methanol (40 mL), 10% aqueous sodium hydroxide (20 mL). Work up yielded the title compound **3.18** (2.34 g, 11.2 mmol, 92%) as a pale purple solid; δ_H (500 MHz, $CDCl_3$) 7.61 (dd, *J* = 8.2, 1.5 Hz, 1 H, *H*6), 7.51 (d, *J* = 1.5 Hz, 1 H, *H*2), 6.70 (d, *J* = 8.2 Hz, 1 H, *H*5), 4.43 (app sxt, *J* = 6.0 Hz, 1 H, *H*α), 1.75 – 1.85 (m, 1 H, *H*β), 1.62 – 1.74 (m, 1 H, *H*β′), 1.34 (d, *J* = 6.0 Hz, 3 H, CHα($C\underline{H}_3$)), 1.01 (t, *J* = 7.4 Hz, 3 H, Hγ); δ_C (125 MHz, $CDCl_3$) 172.0, 144.3, 143.0, 124.8, 118.3, 114.2, 113.3, 75.8, 29.2, 19.3, 9.8; ν_{max}/cm^{-1} (solid state) = 3494, 3387, 2971, 2936, 2879, 1672, 1612, 1445, 1293, 1265, 1223, 769; ESI-HRMS found *m/z* 210.1122 $[M+H]^+$, $C_{11}H_{16}NO_3$ requires 210.1125.

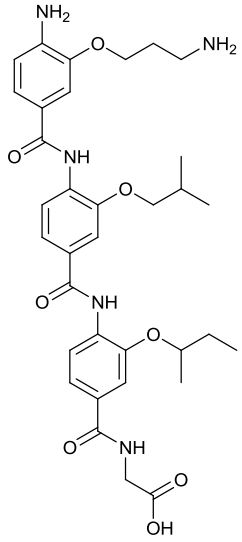
4-(((9H-Fluoren-9-yl)methoxy)carbonylamino)-3-sec-butoxybenzoic acid (3.24)



Procedure G; 4-amino-3-sec-butoxybenzoic acid **3.18** (1.0 g, 4.78 mmol) in tetrahydrofuran (40 mL) and fluorenylmethyloxycarbonyl chloride (1.86 g, 7.17 mmol) in tetrahydrofuran (20 mL). Work up yielded the title compound **3.24** (1.90 g, 4.40 mmol, 92%) as off-white solid; δ_H (500 MHz, $CDCl_3$) 8.17 (s, broad, 1 H, N*H*), 7.80 (d, *J* = 7.5 Hz, 2 H, F*H*Ar5), 7.73 (d, *J* = 8.0 Hz, 1 H, *H*5), 7.64 (d, *J* = 7.5 Hz, 2 H, F*H*Ar2), 7.59 (s, 1 H, *H*2), 7.56 (s, broad, 1 H, *H*6), 7.44 (t, *J* = 7.5 Hz, 2 H, F*H*Ar4), 7.34 (t, *J* = 7.5 Hz, 2 H, F*H*Ar3), 4.63 – 4.45 (m, 3 H, *H*α, F*H*α), 4.35 (t, *J* = 6.5 Hz, 1 H, F*H*β), 1.91 – 1.80 (m, 1 H, *H*β), 1.80 – 1.70 (m, 1 H, *H*β′), 1.39 (d, *J* = 6.2 Hz, 3 H, CHα($C\underline{H}_3$)), 1.05 (t, *J* = 7.4 Hz, 3 H, Hγ); δ_C (125 MHz, $CDCl_3$) 169.9, 156.7, 145.5, 143.7, 141.4, 133.6, 127.9, 127.2, 125.0, 124.1, 123.1, 120.1, 117.4, 113.7, 76.9, 67.5, 47.1, 29.1, 19.3, 9.9; ν_{max}/cm^{-1} (solid state) = 3431, 2958, 2928, 2874, 1743, 1687, 1594, 1534, 1188, 1026, 738; ESI-HRMS found *m/z* 432.1807 $[M+H]^+$, $C_{26}H_{26}NO_5$ requires 432.1805.

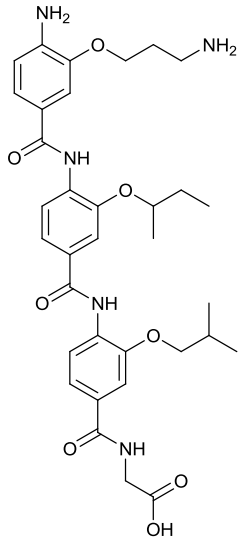
6.5.4. 3-*O*-Alkylated oligomers characterization

NH_2-[*O*-CH_2-CH_2-CH_2-NH_2-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-[*O*-*s*Bu-(3-HABA)]-Gly-CO_2H (4.5)



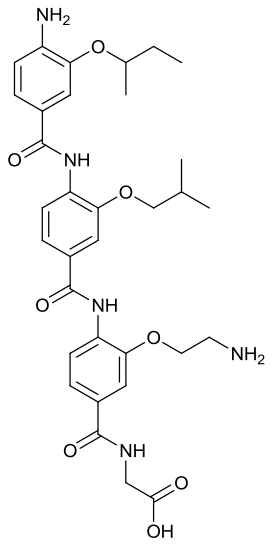
Brown solid; isolated yield: 63 mg, 97%; δ_H (500 MHz, MeOD) 8.34 (d, *J* = 8.1 Hz, 1 H, 2-*H*5), 8.30 (d, *J* = 8.3 Hz, 1 H, 3-*H*5), 7.55 (d, *J* = 1.7 Hz, 1 H, 2-*H*2), 7.49 – 7.46 (m, 3 H, 3-*H*2, 3-*H*6, 2-*H*6), 7.42 (d, *J* = 1.9 Hz, 1 H, 1-*H*2), 7.37 (dd, *J* = 8.1, 1.9 Hz, 2 H, 1-*H*6), 6.91 (d, *J* = 8.1 Hz, 1 H, 1-*H*5), 4.61 – 4.55 (m, 1 H, 2-*H*α), 4.21 (t, *J* = 5.8 Hz, 2 H, 1-*H*α), 4.10 (s, 2 H, 4-*H*α), 3.95 (d, *J* = 6.6 Hz, 2 H, 3-*H*α), 3.23 (t, *J* = 7.4 Hz, 2 H, 1-*H*γ), 2.26 – 2.16 (m, 4 H, 1-*H*β, 3-*H*β), 1.88 – 1.84 (m, 1 H, 2-*H*β), 1.84 – 1.74 (m, 1 H, 2-*H*β′), 1.41 (d, *J* = 6.0 Hz, 3 H, 2-CHα(C*H*$_3$)), 1.12 (d, *J* = 6.6 Hz, 6 H, 3-*H*γ), 1.06 (t, *J* = 7.5 Hz, 3 H, 2-*H*γ); δ_C (125 MHz, MeOD) 173.2, 169.6, 167.0, 166.5, 161.0, 160.7, 150.1, 148.9, 147.8, 133.4, 131.8, 131.1, 130.8, 125.7, 122.0, 121.3, 121.0, 120.6, 116.3, 113.1, 112.0, 111.3, 77.9, 76.2, 66.6, 42.3, 38.4, 30.2, 29.5, 28.3, 19.6, 19.5, 9.9; ν_{max}/cm^{-1} (solid state) = 3356, 2964, 2936, 1670, 1596, 1506, 1193, 1123, 1024, 721; ESI-HRMS found *m/z* 650.3198 $[M+H]^+$, $C_{34}H_{44}N_5O_8$ requires 650.3184.

NH_2-[*O*-CH_2-CH_2-CH_2-NH_2-(3-HABA)]-[*O*-*s*Bu-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-Gly-CO_2H (4.6)



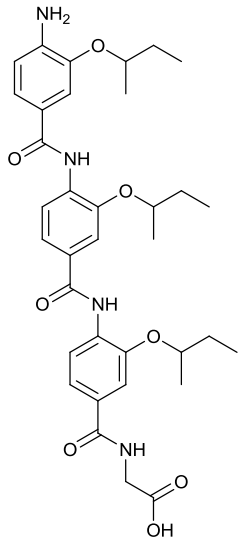
Brown solid; isolated yield: 61 mg, 94%; δ_H (500 MHz, MeOD) 8.37 (d, *J* = 8.3 Hz, 1 H, 2-*H*5), 8.26 (d, *J* = 8.1 Hz, 1 H, 3-*H*5), 7.57 (d, *J* = 1.9 Hz, 1 H, 2-*H*2), 7.53 (s, 1 H, 3-*H*2), 7.52 – 7.47 (m, 2 H, 3-*H*6, 2-*H*6), 7.44 (d, *J* = 1.9 Hz, 1 H, 1-*H*2), 7.35 (dd, *J* = 8.1, 1.9 Hz, 2 H, 1-*H*6), 6.93 (d, *J* = 8.1 Hz, 1 H, 1-*H*5), 4.64 – 4.56 (m, 1 H, 2-*H*α), 4.22 (t, *J* = 5.8 Hz, 2 H, 1-*H*α), 4.10 (s, 2 H, 4-*H*α), 3.94 (d, *J* = 6.6 Hz, 2 H, 3-*H*α), 3.23 (t, *J* = 7.3 Hz, 2 H, 1-*H*γ), 2.25 – 2.17 (m, 4 H, 1-*H*β, 3-*H*β), 1.88 – 1.84 (m, 1 H, 2-*H*β), 1.84 – 1.74 (m, 1 H, 2-*H*β′), 1.41 (d, *J* = 6.0 Hz, 3 H, 2-CHα(C*H*$_3$)), 1.11 (d, *J* = 6.8 Hz, 6 H, 3-*H*γ), 1.06 (t, *J* = 7.5 Hz, 3 H, 2-*H*γ); δ_C (125 MHz, MeOD) 173.2, 169.6, 167.0, 166.8, 161.0, 160.7, 150.3, 148.9, 147.9, 133.4, 131.8, 131.1, 131.0, 125.8, 121.8, 121.2, 121.0, 120.6, 116.4, 113.0, 112.0, 111.6, 78.1, 76.2, 66.6, 42.3, 38.4, 30.2, 29.5, 28.2, 19.6, 19.5, 9.9; ν_{max}/cm^{-1} (solid state) = 3421, 2964, 1595, 1507, 1127, 1026, 747; ESI-HRMS found *m/z* 650.3197 $[M+H]^+$, $C_{34}H_{44}N_5O_8$ requires 650.3184.

NH_2-[*O*-*s*Bu-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-[*O*-CH_2-CH_2-NH_2-(3-HABA)]-Gly-CO_2H (4.7)



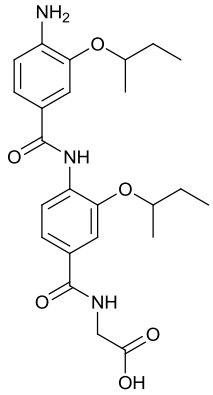
Beige solid; isolated yield: 8 mg, 13%; δ_H (500 MHz, DMSO-d_6) 8.99 (s, 1 H, 2-N*H*), 8.44 (s, broad, 1 H, 4-N*H*), 8.25 (d, *J* = 8.5 Hz, 1 H, 2-*H*5), 8.11 (d, *J* = 8.3 Hz, 1 H, 3-*H*5), 7.66 – 7.61 (m, 2 H, 2-*H*2, 2-*H*6), 7.59 (s, 1 H, 3-*H*2), 7.53 (d, *J* = 8.1 Hz, 1 H, 3-*H*6), 7.39 – 7.29 (m, 2 H, 1-*H*2, 1-*H*6), 6.72 (d, *J* = 7.9 Hz, 1 H, 1-*H*5), 5.42 (s, broad, 1 H, 1-N*H*), 4.42 – 4.37 (m, 1 H, 1-*H*α), 4.30 (s, 2 H, 4-*H*α), 4.21 (t, *J* = 5.8 Hz, 2 H, 3-*H*α), 3.97 (d, *J* = 6.0 Hz, 2 H, 2-*H*α), 3.74 (s, 2 H, 3-N*H*), 3.21 (s, 2 H, 3-*H*β), 2.21 – 2.12 (m, 1 H, 2-*H*β), 1.77 – 1.70 (m, 1 H, 1-*H*β), 1.68 – 1.60 (m, 1 H, 1-*H*β′), 1.27 (d, *J* = 6.0 Hz, 3 H, 1-CHα(C*H*$_3$)), 1.06 (d, *J* = 6.6 Hz, 6 H, 2-*H*γ), 0.96 (t, *J* = 7.5 Hz, 3 H, 1-*H*γ); δ_C (125 MHz, DMSO-d_6) 172.0, 165.0, 164.8, 164.5, 148.5, 148.4, 143.4, 143.3, 131.2, 130.6, 130.0, 129.6, 121.9, 121.3, 120.9, 120.4, 120.3, 120.1, 112.7, 112.2, 111.1, 110.9, 75.3, 74.6, 66.9, 42.8, 35.6, 28.6, 27.8, 19.1, 19.1, 9.6; ν_{max}/cm^{-1} (solid state) = 3351, 2963, 2931, 1575, 1508, 1382, 1252, 1025, 753; ESI-HRMS found *m/z* 636.3043 $[M+H]^+$, $C_{33}H_{42}N_5O_8$ requires 636.3028.

NH_2-[*O*-*s*Bu-(3-HABA)]-[*O*-*s*Bu-(3-HABA)]-[*O*-*s*Bu-(3-HABA)]-Gly-CO_2H (4.8)



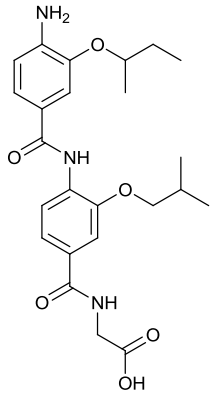
Beige solid; isolated yield: 48 mg, 74%; δ_H (500 MHz, $CDCl_3$) 8.85 (s, 1 H, 3-N*H*), 8.75 (s, 1 H, 2-N*H*), 8.68 (d, *J* = 8.3 Hz, 1 H, 2-*H*5), 8.59 (d, *J* = 8.1 Hz, 1 H, 3-*H*5), 7.58 (d, *J* = 1.0 Hz, 1 H, 2-*H*2), 7.53 (s, 1 H, 3-*H*2), 7.44 (d, *J* = 1.8 Hz, 1 H, 1-*H*2), 7. 42 – 7.37 (m, 2 H, 3-*H*6, 2-*H*6), 7.27 (dd, *J* = 8.1, 1.8 Hz, 2 H, 1-*H*6), 7.04 (s, broad, 1 H, 4-N*H*), 6.75 (d, *J* = 8.1 Hz, 1 H, 1-*H*5), 4.61 – 4.54 (m, 2 H, 3-*H*α, 2-*H*α), 4.49 – 4.43 (m, 1 H, 1-*H*α), 4.25 (s, 2 H, 4-*H*α), 1.89 – 1.65 (m, 6 H, 1-*H*β, 2-*H*β, 3-*H*β), 1.43 – 1.33 (m, 9 H, 1-CHα(C*H*$_3$), 2-CHα(C*H*$_3$), 3-CHα(C*H*$_3$)), 1.08 – 0.97 (m, 9 H, 1-*H*γ, 2-*H*γ, 3-*H*γ); δ_C (125 MHz, $CDCl_3$) 167.7, 165.3, 164.8, 146.6, 146.6, 145.1, 141.5, 132.8, 132.3, 129.1, 128.2, 126.0, 124.6, 124.1, 119.9, 119.5, 119.0, 118.7, 113.7, 112.3, 111.9, 111.8, 76.7, 76.7, 76.0, 42.1, 29.3, 29.2, 29.2, 19.4, 19.3, 19.3, 9.8, 9.7, 9.6; ν_{max}/cm^{-1} (solid state) = 3441, 3310, 3181, 2967, 2930, 1750, 1595, 1505, 1323, 1255, 1033, 746; ESI-HRMS found *m/z* 649.3245 $[M+H]^+$, $C_{35}H_{45}N_4O_8$ requires 649.3232.

NH_2-[*O*-*s*Bu-(3-HABA)]-[*O*-*s*Bu-(3-HABA)]-Gly-CO_2H (4.9)



Dark yellow solid; isolated yield: 20 mg (isolated as a side product in the synthesis of the corresponding trimer); δ_H (500 MHz, MeOD) 8.37 (d, *J* = 8.4 Hz, 1 H, 2-*H*5), 7.57 (d, *J* = 1.9 Hz, 1 H, 2-*H*2), 7.51 (dd, *J* = 8.4, 1.9 Hz, 1 H, 2-*H*6), 7.38 (d, *J* = 1.9 Hz, 1 H, 1-*H*2), 7.31 (dd, *J* = 8.2, 1.9 Hz, 1 H, 1-*H*6), 6.79 (d, *J* = 8.2 Hz, 1 H, 1-*H*5), 4.63 – 4.57 (m, 1 H, 2-*H*α), 4.48 – 4.41 (m, 1 H, 1-*H*α), 4.09 (s, 2 H, 3-*H*α), 1.90 – 1.67 (m, 4 H, 1-*H*β, 2-*H*β), 1.39 (dd, *J* = 6.1, 0.9 Hz, 3 H, 2-CHα(C*H*$_3$)), 1.35 (dd, *J* = 6.0, 1.8 Hz, 3 H, 1-CHα(C*H*$_3$)), 1.07 – 1.02 (m, 6 H, 1-*H*γ, 2-*H*γ); δ_C (125 MHz, MeOD) 173.3, 169.7, 167.6, 148.5, 146.0, 144.5, 133.3, 130.3, 123.3, 121.8, 121.0, 120.8, 114.6, 113.2, 113.2, 78.0, 77.2, 42.4, 30.2, 30.2, 19.6, 19.5, 10.1, 9.9; ν_{max}/cm^{-1} (solid state) = 3347, 2970, 2930, 2878, 612, 1503, 1476, 1250, 1202, 756; ESI-HRMS found *m/z* 458.2298 $[M+H]^+$, $C_{24}H_{32}N_3O_6$ requires 458.2286.

NH_2-[*O*-*s*Bu-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-Gly-CO_2H (4.10)



Dark yellow solid; isolated yield: 28 mg, 41%; δ_H (500 MHz, MeOD) 8.32 (d, *J* = 8.3 Hz, 1 H, 2-*H*5), 7.55 (d, *J* = 1.9 Hz, 1 H, 2-*H*2), 7.51 (dd, *J* = 8.3, 1.9 Hz, 1 H, 2-*H*6), 7.39 (d, *J* = 1.9 Hz, 1 H, 1-*H*2), 7.34 (dd, *J* = 8.2, 1.9 Hz, 1 H, 1-*H*6), 6.78 (d, *J* = 8.2 Hz, 1 H, 1-*H*5), 4.48 – 4.41 (m, 1 H, 1-*H*α), 4.09 (s, 2 H, 3-*H*α), 3.96 (d, *J* = 6.4 Hz, 2 H, 2-*H*α), 1.24 – 1.16 (m, 1 H, 2-*H*β), 1.85 – 1.77 (m, 1 H, 1-*H*β), 1.74 – 1.68 (m, 1 H, 1-*H*β′), 1.34 (d, *J* = 6.1 Hz, 3 H, 2-CHα(C*H*$_3$)), 1.11 (d, *J* = 6.7 Hz, 6 H, 2-*H*γ), 1.03 (t, *J* = 7.5 Hz, 3 H, 1-*H*γ,); δ_C (125 MHz, MeOD) 173.2, 169.8, 167.6, 149.8, 146.0, 144.5, 132.3, 130.4, 123.3, 121.9, 121.0, 121.0, 114.6, 113.2, 111.5, 77.2, 76.2, 42.3, 30.2, 29.5, 19.6, 19.5, 10.0; ν_{max}/cm^{-1} (solid state) = 3354, 2960, 2928, 2873, 1611, 1503, 1482, 1249, 1205, 756; ESI-HRMS found *m/z* 458.2294 $[M+H]^+$, $C_{24}H_{32}N_3O_6$ requires 458.2286.

6.5.5. General methods for manual Solid Phase Peptide Synthesis

Method A: Resin Swelling

Rink amide MBHA resin (0.1 mmol) was used for all syntheses unless stated otherwise. The resin was placed in a vaculate reservoir, dichloromethane (3 mL) was added and the resin was agitated on a blood-spinner for 30 min to allow its swelling.

Method B: Fmoc Deprotection

N-terminal Fmoc protecting groups were removed by the addition of 20% piperidine in DMF (5 × 2 mL × 2 min), followed by washing of the resin with DMF (5 × 2 mL × 2 min).

Method C: Kaiser Test[207]

The Kaiser Test was employed to determine the successful coupling of most of the amino acid residues. A small amount of resin beads was rinsed with dichloromethane and placed in a vial, followed by the addition of two drops of each of the following three solutions in the respective order:

1) Ninhydrin (5% w/v) in ethanol;

2) Phenol (80% w/v) in ethanol;

3) 1 mM KCN(aq.) in pyridine (2% v/v).

The solution was then heated to ca. 100 °C for 1 min. Yellow solution and no change in the colour of the beads indicate successful couplings, whereas blue solution and dark colour of the beads indicate presence of primary amines as a result of incomplete couplings. In the second case, a double coupling was necessary. This colour test can be used to identify free primary amines, however is inconclusive for Asp, Ser, Pro and Asn residues.

Method D: Chloranil Test[214]

The chloranil test was also employed to determine successful couplings of some amino acid residues. A small amount of beads was rinsed in dichloromethane and placed in a vial, followed by the addition of two drops of each of the following solutions in the respective order:

1) Acetaldehyde (2% v/v) in DMF;

2) *p*-Chloranil (2% w/v) in DMF.

The solution was left at rt for 5 min. No change in colour of the beads showed successful couplings, whereas the change of bead colour to pale green/bright blue indicate presence of primary amines as a result of incomplete couplings. In the second case, a double coupling was necessary. This test is a reliable method to detect secondary amines; therefore, it was particularly useful for Pro residues.

Method E: Chain Elongation

The coupling of the desired amino acids (5 equiv) was performed with HCTU or HATU (5 equiv) and DIPEA (5 equiv) dissolved in DMF (2 mL) and added to the resin, followed by agitation for 1 h (2 h for problematic couplings). For double couplings, this step was repeated. After draining the reagents, the resin was washed with DMF (3 × 2 mL × 2 min) and the success of coupling determined by a negative colour test (*Method C* or *D*). Once a coupling was confirmed successful, the *N*-terminal Fmoc was deprotected (*Method B*) and a subsequent coupling or *N*-terminal acetylation (*Method F*) was performed.

Method F: *N*-terminal Acetylation

After complete peptide chain elongation, *N*-terminal acetylation was performed with acetic anhydride (10 equiv) and DIPEA (10 equiv) dissolved in DMF (2 mL), which was transferred to the resin for 2 h. Later the resin was drained, washed with DMF (3 × 2 mL × 2 min) and successful capping determined by a negative colour test (*Method C* or *D*).

Method G: On-Resin Olefin Metathesis

After *N*-terminal acetylation, on-resin olefin metathesis was performed using a 10 mM solution of Grubbs 1st Generation Catalyst in degassed dichloroethanol (2 mL), which was transferred to the vessel (2 × 2 h).

Method H: Cleavage and Deprotection

After *N*-terminal acetylation (and on-resin olefin metathesis when required), the resin was washed with DMF (5 × 2 mL × 2 min.), dichloromethane (5 × 2 mL × 2 min.), and then Et_2O (3 × 2mL × 2min.). Peptides were then simultaneously cleaved and side chain deprotected with cleavage "Reagent K" TFA : EDT : Thioanisole : Phenol : H_2O, 87:3:5:5:5 (2 × 2 mL × l h). Peptides with large number of Arg(Pbf) residues required incubation times up to 3 h for complete deprotection of the side chains. The resin was washed with fresh TFA (2 mL × 2 min) and the TFA evaporated under N_2 (g).

The resulting oil was precipitated with ice-cold ether and placed in a centrifuge (3000 rpm × 2 min). The supernatants were removed, the precipitate rinsed with ice-cold ether (3 × 3000 rpm × 2 min) and dried *in vacuo*.

Method I: Mass – Directed HPLC Purification

Crude peptides were dissolved in either DMSO, 1:1 dioxane:water or 9:1 water:hexafluoroisopropanol at an approximate concentration of 20 mg mL^{-1}.

Peptides were purified using reversed phase mass directed HPLC with software Masshunter by ChemStation (Agilent). The columns used were a Jupiter Proteo or an Agilent XBridge 5 μm 19×100 mm C18 preparative column. An increasing gradient of MeCN to water (plus 0.1% formic acid v/v in both solvents) at a flow rate of 20 mL min^{-1} was used as mobile phase; in each case the gradient was optimized to obtain the best separation of the desired peptide from the rest of impurities from the crude mixture. Mass directed chromatography allows the collection of the desired peptide by mass, with the eluent split into an Agilent 6120 Quadrupole LCMS which triggers collection of eluent at a programmed *m/z*. This technique is particularly useful for peptides with weak UV traces. When the separation of the desired peptide from the impurities was problematic, the collection was performed using time slices at the required time interval.

The resulting fractions were checked on the analytical HPLC, concentrated by centrifugal evaporation (Genevac), re-suspended in water and lyophilized.

6.5.6. General methods for automated Solid Phase Peptide Synthesis

The synthesis of peptides using the microwave assisted Liberty CEM™ Peptide Synthesiser follows the cycle below:

Resin Loading

Clean reaction vessel; wash with DMF; wash with dichloromethane; transfer resin to reaction vessel; wash with DMF; wash with dichloromethane; transfer resin to reaction vessel; wash with DMF; wash with dichloromethane; vessel draining.

Deprotection and coupling

Clean resin dip tube, wash with DMF (15 mL), add 20% piperidine in DMF (6 mL), agitation at rt (10 min), wash with DMF (15 mL), add 20% piperidine in DMF (6 mL), agitation at rt (10 min), wash with DMF (15 mL), clean resin dip tube, wash with DMF (15 mL), add amino acid (2.5 mL), add coupling reagent (1 mL), add activator base (0.5 mL), agitation at rt (90 min), wash with DMF (15 mL), drain. After the final residue, the resin is ejected from the reaction vessel, *N*-terminal acetylation and cleavage/deprotection were performed manually (*Methods F* and *H*).

6.5.7. Peptides characterization data

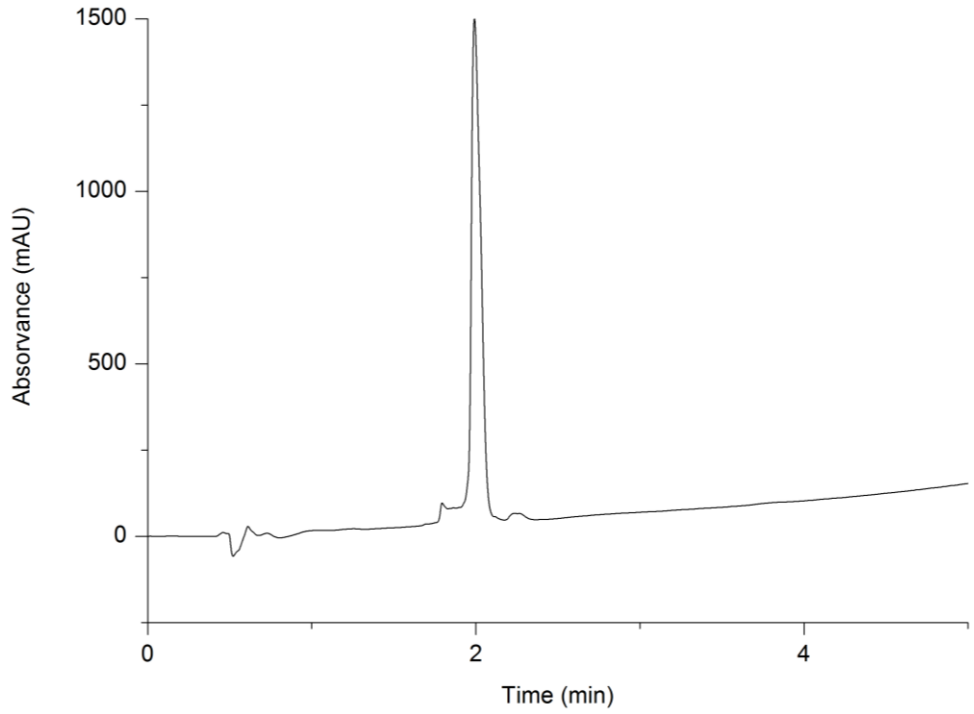
Below are the analytical HPLC and HRMS data of the peptides that have been synthesised. The chromatograms correspond to the UV signal obtained at 220 nm in the following conditions: (5-95% MeCN:water and 0.1% TFA v/v in both solvents), t = 4.91 min, 0.5 mL min^{-1}, Ascentis Express C_{18} column. Peptide identity was confirmed by the assessment of multiple charge states, which are tabulated as the *monoisotopic* peak for the Expected and Observed masses. The mass spectra shown below report the *most abundant* isotope peaks. The peptides were dissolved in pure water or DMSO for its characterization.

Native H3 peptide

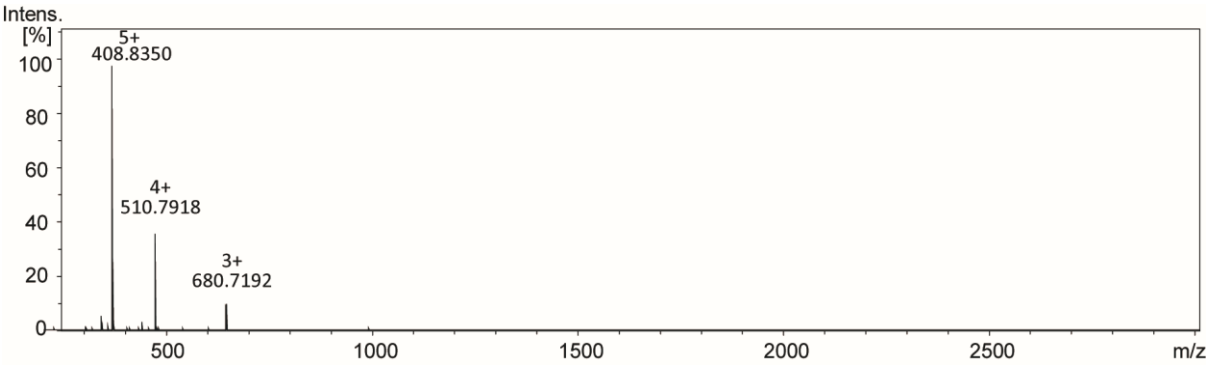
Ac - GAMGKDIQLARRIRGERA - $CONH_2$

The native H3 peptide was obtained using the general methods for SPPS (See sections 6.5.5 and 6.5.6), 11 mg (11% yield), > 97 % purity.

Analytical HPLC data for native H3 peptide



HRMS data for native H3 peptide



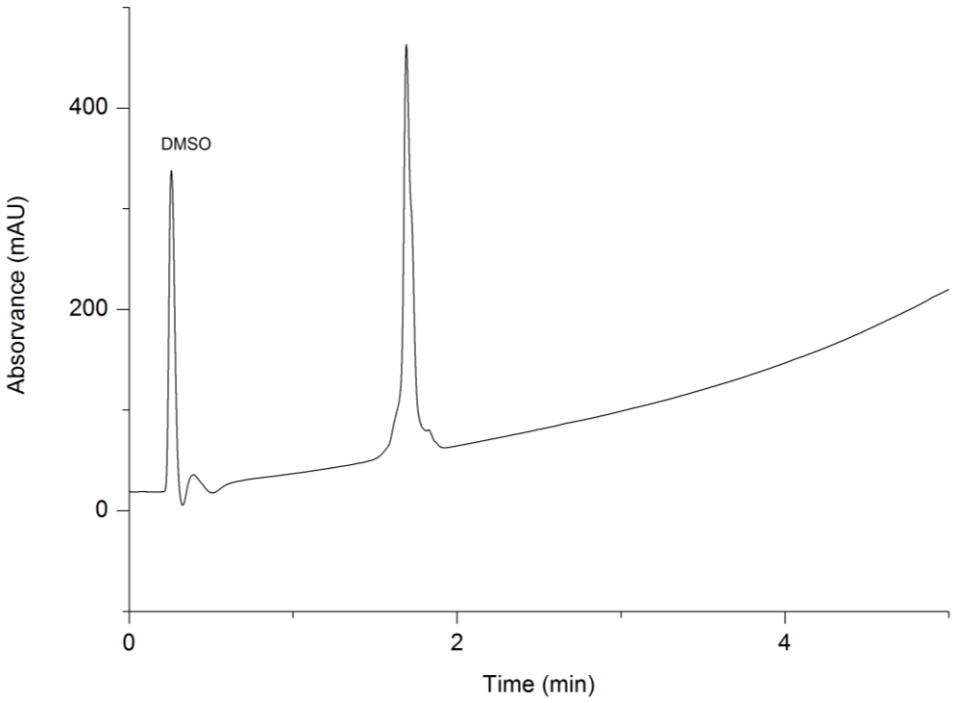
Native peptide	Observed	Expected
$[M+3H]^{3+}$	680.3848	680.3850
$[M+4H]^{4+}$	510.5413	510.5405
$[M+5H]^{5+}$	408.6344	408.6339

Stapled 1 H3 peptide

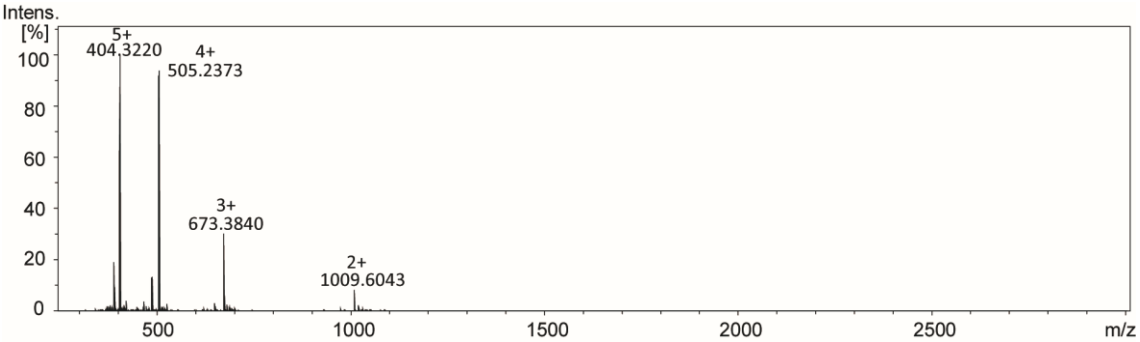
Ac - GAXGKDXQLARRIRGERA - $CONH_2$

The stapled 1 H3 peptide was obtained using the general methods for SPPS (See sections 6.5.5 and 6.5.6), 10 mg (5% yield), > 90 % purity.

Analytical HPLC data for stapled H3 position 1 peptide



HRMS data for stapled 1 H3 peptide



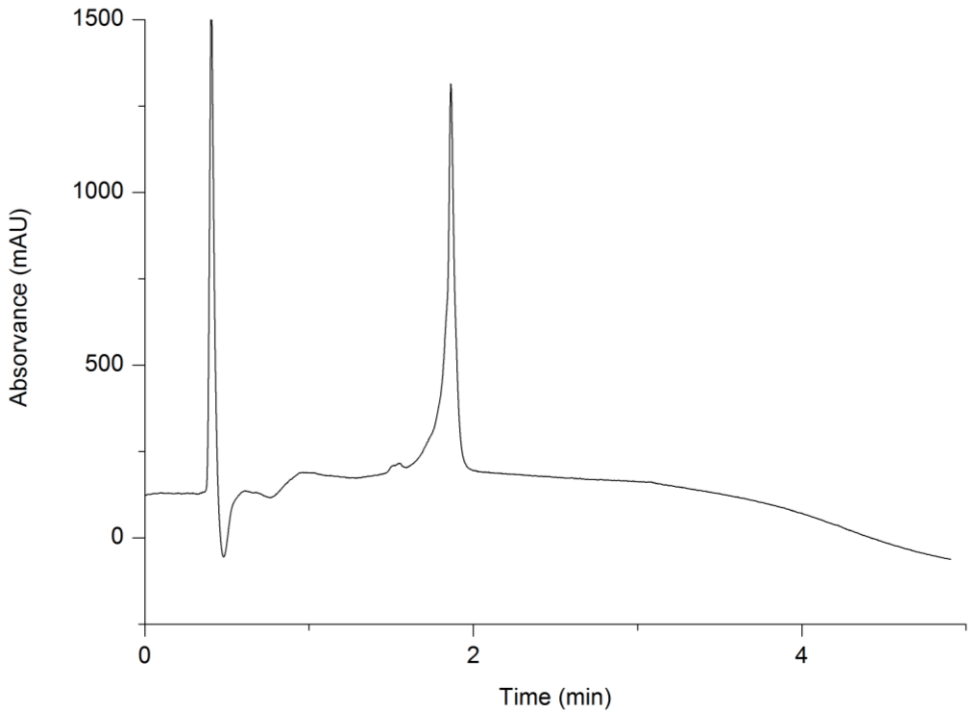
Stapled 1 peptide	Observed	Expected
$[M+2H]^{2+}$	1009.1023	1009.0799
$[M+3H]^{3+}$	673.0500	673.0557
$[M+4H]^{4+}$	504.9862	505.0436
$[M+5H]^{5+}$	404.1218	404.2363

Stapled 2 H3 peptide

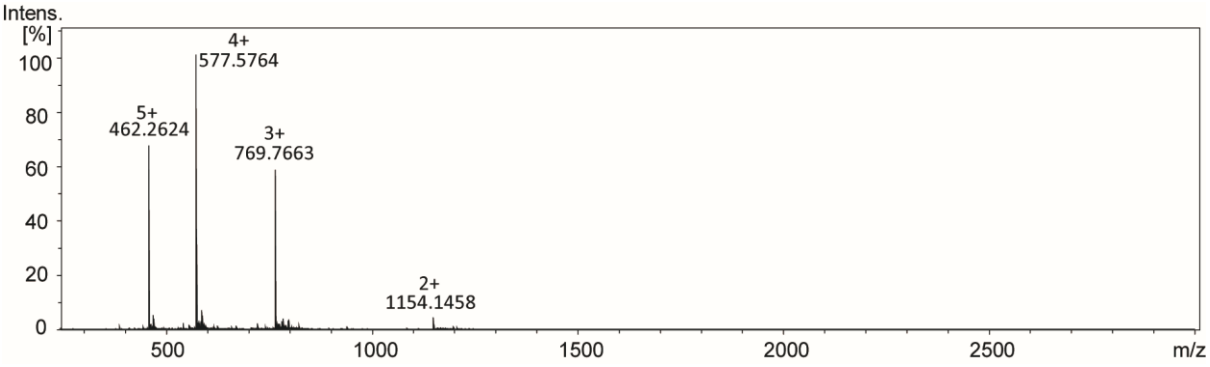
Ac - GAMGKXIQLXRRIRGERAGCA - $CONH_2$

The stapled 2 H3 peptide was obtained using the general methods for SPPS (See sections 6.5.5 and 6.5.6), 6 mg (3% yield), > 90% purity

Analytical HPLC data for stapled 2 H3 peptide



HRMS data for stapled H3 position 2 peptide



Stapled 2 peptide	Observed	Expected
$[M+2H]^{2+}$	1153.6446	1153.6441
$[M+3H]^{3+}$	769.4320	769.4318
$[M+4H]^{4+}$	577.3255	577.3257
$[M+5H]^{5+}$	462.0613	462.0620

6.5.8. Circular Dichroism

Circular Dichroism was performed on an Applied Photophysics ChiraScan Apparatus and Software. For each scan, the following parameters were used: 180-260 nm range; point time 1 s; 1 nm per point; step = 1; bandwidth 4.3 nm; path length 1 mm; temperature 20 °C. Scans were done in duplicate. Samples were dissolved in 40 mM sodium phosphate buffer pH 7.50 at concentrations between 10 - 100 μM. DMSO was not used in the samples for CD measurements due to its high absorbance below 230 nm. The solvent signal was subtracted to the raw circular dichroism data obtained for the peptides before conversion to the mean residue ellipticity:

$$[\theta] = \frac{\theta}{10 \times c \times l}$$

$$[\theta]_{MRE} = \frac{[\theta]}{(R-1)}$$

Where θ = circular dichroism at a given wavelength, *c* = molar concentration, l = path length in cm, R = number of residues in the peptide sequence.

Chapter 7. References

1. V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nature Chem.*, 2013, ***5***, 161-173.
2. J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001-1009.
3. A. J. Wilson, *Chem. Soc. Rev.*, 2009, **38**, 3289-3300.
4. S. Surade and T. L. Blundell, *Chem. Biol.*, 2012, **19**, 42-50.
5. B. N. Bullock, A. L. Jochim and P. S. Arora, *J. Am. Chem. Soc.*, 2011, **133**, 14220-14223.
6. S. Eyrisch and V. Helms, *J. Med. Chem.*, 2007, **50**, 3457-3464.
7. W. E. Stites, *Chem. Rev.*, 1997, **97**, 1233-1250.
8. O. Keskin, A. Gursoy, B. Ma and R. Nussinov, *Chem. Rev.*, 2008, **108**, 1225-1244.
9. W. L. DeLano, *Curr. Opin. Struct. Biol.*, 2002, **12**, 14-20.
10. T. Clackson and J. A. Wells, *Science*, 1995, **267**, 383-386.
11. O. Keskin, B. Y. Ma and R. Nussinov, *J. Mol. Biol.*, 2005, **345**, 1281-1294.
12. A. A. Bogan and K. S. Thorn, *J. Mol. Biol.*, 1998, **280**, 1-9.
13. A. de Vos, M. Ultsch and A. Kossiakoff, *Science*, 1992, **255**, 306-312.
14. T. A. Edwards and A. J. Wilson, *Amino Acids*, 2011, **41**, 743-754.
15. A. S. Ripka and D. H. Rich, *Curr. Opin. Chem. Biol.*, 1998, **2**, 441-452.
16. C. N. Pace, B. A. Shirley, M. McNutt and K. Gajiwala, *FASEB J.*, 1996, **10**, 75-83.
17. S. A. Marshall, G. A. Lazar, A. J. Chirino and J. R. Desjarlais, *Drug Discov. Today*, 2003, **8**, 212-221.
18. N. E. Shepherd, H. N. Hoang, G. Abbenante and D. P. Fairlie, *J. Am. Chem. Soc.*, 2005, **127**, 2974-2983.
19. S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173-180.
20. E. A. Harker, D. S. Daniels, D. A. Guarracino and A. Schepartz, *Biorg. Med. Chem.*, 2009, **17**, 2038-2046.
21. R. P. Cheng, S. H. Gellman and W. F. DeGrado, *Chem. Rev.*, 2001, **101**, 3219-3232.
22. D. Seebach, M. Overhand, F. N. M. Kühnle, B. Martinoni, L. Oberer, U. Hommel and H. Widmer, *Helv. Chim. Acta*, 1996, **79**, 913-941.
23. M. Werder, H. Hauser, S. Abele and D. Seebach, *Helv. Chim. Acta*, 1999, **82**, 1774-1783.
24. 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.
25. A. D. Bautista, O. M. Stephens, L. Wang, R. A. Domaoal, K. S. Anderson and A. Schepartz, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 3736-3738.

26. E. V. Denton, C. J. Craig, R. L. Pongratz, J. S. Appelbaum, A. E. Doerner, A. Narayanan, G. I. Shulman, G. W. Cline and A. Schepartz, *Org. Lett.*, 2013, **15**, 5318-5321.
27. E. A. Harker and A. Schepartz, *ChemBioChem*, 2009, **10**, 990-993.
28. A. D. Bautista, J. S. Appelbaum, C. J. Craig, J. Michel and A. Schepartz, *J. Am. Chem. Soc.*, 2010, **132**, 2904-2906.
29. G. V. M. Sharma, P. Nagendar, P. Jayaprakash, P. Radha Krishna, K. V. S. Ramakrishna and A. C. Kunwar, *Angew. Chem. Int. Ed.*, 2005, **44**, 5878-5882.
30. S. De Pol, C. Zorn, C. D. Klein, O. Zerbe and O. Reiser, *Angew. Chem. Int. Ed.*, 2004, **43**, 511-514.
31. A. Hayen, M. A. Schmitt, F. N. Ngassa, K. A. Thomasson and S. H. Gellman, *Angew. Chem. Int. Ed.*, 2004, **43**, 505-510.
32. W. S. Horne and S. H. Gellman, *Acc. Chem. Res.*, 2008, **41**, 1399-1408.
33. J. D. Sadowsky, M. A. Schmitt, H.-S. Lee, N. Umezawa, S. Wang, Y. Tomita and S. H. Gellman, *J. Am. Chem. Soc.*, 2005, **127**, 11966-11968.
34. P. E. Czabotar, E. F. Lee, M. F. van Delft, C. L. Day, B. J. Smith, D. C. S. Huang, W. D. Fairlie, M. G. Hinds and P. M. Colman, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 6217-6222.
35. J. D. Sadowsky, W. D. Fairlie, E. B. Hadley, H.-S. Lee, N. Umezawa, Z. Nikolovska-Coleska, S. Wang, D. C. S. Huang, Y. Tomita and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 139-154.
36. E. F. Lee, J. D. Sadowsky, B. J. Smith, P. E. Czabotar, K. J. Peterson-Kaufman, P. M. Colman, S. H. Gellman and W. D. Fairlie, *Angew. Chem. Int. Ed.*, 2009, **48**, 4318-4322.
37. W. S. Horne, M. D. Boersma, M. A. Windsor and S. H. Gellman, *Angew. Chem. Int. Ed.*, 2008, **47**, 2853-2856.
38. E. F. Lee, B. J. Smith, W. S. Horne, K. N. Mayer, M. Evangelista, P. M. Colman, S. H. Gellman and W. D. Fairlie, *ChemBioChem*, 2011, **12**, 2025-2032.
39. K. J. Peterson-Kaufman, H. S. Haase, M. D. Boersma, E. F. Lee, W. D. Fairlie and S. H. Gellman, *ACS Chem. Biol.*, 2015, **10**, 1667-1675.
40. B. J. Smith, E. F. Lee, J. W. Checco, M. Evangelista, S. H. Gellman and W. D. Fairlie, *ChemBioChem*, 2013, **14**, 1564-1572.
41. W. S. Horne, L. M. Johnson, T. J. Ketas, P. J. Klasse, M. Lu, J. P. Moore and S. H. Gellman, *Proc. Natl. Acad. Sci.*, 2009, **106**, 14751-14756.
42. L. M. Johnson, S. Barrick, M. V. Hager, A. McFedries, E. A. Homan, M. E. Rabaglia, M. P. Keller, A. D. Attie, A. Saghatelian, A. Bisello and S. H. Gellman, *J. Am. Chem. Soc.*, 2014, **136**, 12848-12851.

43. 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.
44. A. A. Kaspar and J. M. Reichert, *Drug Discov. Today*, 2013, **18**, 807-817.
45. J. E. Bock, J. Gavenonis and J. A. Kritzer, *ACS Chem. Biol.*, 2013, **8**, 488-499.
46. L. D. Walensky, A. L. Kung, I. Escher, T. J. Malia, S. Barbuto, R. D. Wright, G. Wagner, G. L. Verdine and S. J. Korsmeyer, *Science*, 2004, **305**, 1466-1470.
47. A. K. Galande, K. S. Bramlett, J. O. Trent, T. P. Burris, J. L. Wittliff and A. F. Spatola, *ChemBioChem*, 2005, **6**, 1991-1998.
48. A. Muppidi, K. Doi, S. Edwardraja, E. J. Drake, A. M. Gulick, H.-G. Wang and Q. Lin, *J. Am. Chem. Soc.*, 2012, **134**, 14734-14737.
49. A. Muppidi, Z. Wang, X. Li, J. Chen and Q. Lin, *Chem. Commun.*, 2011, **47**, 9396-9398.
50. A. M. Spokoyny, Y. Zou, J. J. Ling, H. Yu, Y.-S. Lin and B. L. Pentelute, *J. Am. Chem. Soc.*, 2013, **135**, 5946-5949.
51. Y. Wang and D. H.-C. Chou, *Angew. Chem. Int. Ed.*, 2015, **54**, 10931-10934.
52. J. Bredenbeck, J. Helbing, J. R. Kumita, G. A. Woolley and P. Hamm, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 2379-2384.
53. S. Kneissl, E. J. Loveridge, C. Williams, M. P. Crump and R. K. Allemann, *ChemBioChem*, 2008, **9**, 3046-3054.
54. R. J. Mart, D. Meah and R. K. Allemann, *ChemBioChem*, 2015, n/a-n/a.
55. M. Chorev, M. P. Caulfield, E. Roubini, R. L. McKee, S. W. Gibbons, C.-T. Leu, J. J. Levy and M. Rosenblatt, *Int. J. Pept. Protein Res.*, 1992, **40**, 445-455.
56. S. K. Sia, P. A. Carr, A. G. Cochran, V. N. Malashkevich and P. S. Kim, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 14664-14669.
57. R. S. Harrison, N. E. Shepherd, H. N. Hoang, G. Ruiz-Gómez, T. A. Hill, R. W. Driver, V. S. Desai, P. R. Young, G. Abbenante and D. P. Fairlie, *Proc. Natl. Acad. Sci.*, 2010, **107**, 11686-11691.
58. R. S. Harrison, G. Ruiz-Gómez, T. A. Hill, S. Y. Chow, N. E. Shepherd, R.-J. Lohman, G. Abbenante, H. N. Hoang and D. P. Fairlie, *J. Med. Chem.*, 2010, **53**, 8400-8408.
59. H. N. Hoang, K. Song, T. A. Hill, D. R. Derksen, D. J. Edmonds, W. M. Kok, C. Limberakis, S. Liras, P. M. Loria, V. Mascitti, A. M. Mathiowetz, J. M. Mitchell, D. W. Piotrowski, D. A. Price, R. V. Stanton, J. Y. Suen, J. M. Withka, D. A. Griffith and D. P. Fairlie, *J. Med. Chem.*, 2015, **58**, 4080-4085.
60. H. E. Blackwell and R. H. Grubbs, *Angew. Chem. Int. Ed.*, 1998, **37**, 3281-3284.
61. C. E. Schafmeister, J. Po and G. L. Verdine, *J. Am. Chem. Soc.*, 2000, **122**, 5891-5892.

62. T. Oltersdorf, S. W. Elmore, A. R. Shoemaker, R. C. Armstrong, D. J. Augeri, B. A. Belli, M. Bruncko, T. L. Deckwerth, J. Dinges, P. J. Hajduk, M. K. Joseph, S. Kitada, S. J. Korsmeyer, A. R. Kunzer, A. Letai, C. Li, M. J. Mitten, D. G. Nettesheim, S. Ng, P. M. Nimmer, J. M. O'Connor, A. Oleksijew, A. M. Petros, J. C. Reed, W. Shen, S. K. Tahir, C. B. Thompson, K. J. Tomaselli, B. L. Wang, M. D. Wendt, H. C. Zhang, S. W. Fesik and S. H. Rosenberg, *Nature*, 2005, **435**, 677-681.

63. M. L. Stewart, E. Fire, A. E. Keating and L. D. Walensky, *Nat. Chem. Biol.*, 2010, **6**, 595-601.

64. G. H. Bird, E. Gavathiotis, J. L. LaBelle, S. G. Katz and L. D. Walensky, *ACS Chem. Biol.*, 2014, **9**, 831-837.

65. J. L. LaBelle, S. G. Katz, G. H. Bird, E. Gavathiotis, M. L. Stewart, C. Lawrence, J. K. Fisher, M. Godes, K. Pitter, A. L. Kung and L. D. Walensky, *J. Clin. Invest.*, 2012, **122**, 2018-2031.

66. G. H. Bird, N. Madani, A. F. Perry, A. M. Princiotto, J. G. Supko, X. He, E. Gavathiotis, J. G. Sodroski and L. D. Walensky, *Proc. Natl. Acad. Sci.*, 2010, **107**, 14093-14098.

67. S. Baek, P. S. Kutchukian, G. L. Verdine, R. Huber, T. A. Holak, K. W. Lee and G. M. Popowicz, *J. Am. Chem. Soc.*, 2012, **134**, 103-106.

68. C. Phillips, L. R. Roberts, M. Schade, R. Bazin, A. Bent, N. L. Davies, R. Moore, A. D. Pannifer, A. R. Pickford, S. H. Prior, C. M. Read, A. Scott, D. G. Brown, B. Xu and S. L. Irving, *J. Am. Chem. Soc.*, 2011, **133**, 9696-9699.

69. D. J. Yeo, S. L. Warriner and A. J. Wilson, *Chem. Commun.*, 2013, **49**, 9131-9133.

70. J. A. Miles, D. J. Yeo, P. Rowell, S. Rodriguez-Marin, C. M. Pask, S. L. Warriner, T. A. Edwards and A. J. Wilson, *Chem. Sci.*, 2016, **7**, 3694-3702.

71. T. Okamoto, K. Zobel, A. Fedorova, C. Quan, H. Yang, W. J. Fairbrother, D. C. S. Huang, B. J. Smith, K. Deshayes and P. E. Czabotar, *ACS Chem. Biol.*, 2013, **8**, 297-302.

72. T. Okamoto, D. Segal, K. Zobel, A. Fedorova, H. Yang, W. J. Fairbrother, D. C. S. Huang, B. J. Smith, K. Deshayes and P. E. Czabotar, *ACS Chem. Biol.*, 2014, **9**, 838-839.

73. T. K. Sawyer, *Chem. Biol. Drug Des.*, 2009, **73**, 3-6.

74. E. Cabezas and A. C. Satterthwait, *J. Am. Chem. Soc.*, 1999, **121**, 3862-3875.

75. E. Cabezas, M. Wang, P. W. H. I. Parren, R. L. Stanfield and A. C. Satterthwait, *Biochemistry (Mosc).* 2000, **39**, 14377-14391.

76. R. N. Chapman, G. Dimartino and P. S. Arora, *J. Am. Chem. Soc.*, 2004, **126**, 12252-12253.

77. S. E. Miller, P. F. Thomson and P. S. Arora, *Curr. Protoc. Chem. Biol.*, 2009, **6**, 101–116.
78. D. Wang, W. Liao and P. S. Arora, *Angew. Chem. Int. Ed.*, 2005, **44**, 6525-6529.
79. L. K. Henchey, J. R. Porter, I. Ghosh and P. S. Arora, *ChemBioChem*, 2010, **11**, 2104-2107.
80. L. K. Henchey, S. Kushal, R. Dubey, R. N. Chapman, B. Z. Olenyuk and P. S. Arora, *J. Am. Chem. Soc.*, 2010, **132**, 941-943.
81. A. Patgiri, K. K. Yadav, P. S. Arora and D. Bar-Sagi, *Nat. Chem. Biol.*, 2011, **7**, 585-587.
82. A. B. Mahon and P. S. Arora, *Chem. Commun.*, 2012, **48**, 1416-1418.
83. S. E. Miller, N. R. Kallenbach and P. S. Arora, *Tetrahedron*, 2012, **68**, 4434-4437.
84. A. J. Vernall, P. Cassidy and P. F. Alewood, *Angew. Chem. Int. Ed.*, 2009, **48**, 5675-5678.
85. C. Sheng, G. Dong, Z. Miao, W. Zhang and W. Wang, *Chem. Soc. Rev.*, 2015, **44**, 8238-8259.
86. 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.
87. C. Tse, A. R. Shoemaker, J. Adickes, M. G. Anderson, J. Chen, S. Jin, E. F. Johnson, K. C. Marsh, M. J. Mitten, P. Nimmer, L. Roberts, S. K. Tahir, Y. Mao, X. F. Yang, H. C. Zhang, S. Fesik, S. H. Rosenberg and S. W. Elmore, *Cancer Res.*, 2008, **68**, 3421-3428.
88. B. Vu, P. Wovkulich, G. Pizzolato, A. Lovey, Q. Ding, N. Jiang, J.-J. Liu, C. Zhao, K. Glenn, Y. Wen, C. Tovar, K. Packman, L. Vassilev and B. Graves, *ACS Med. Chem. Lett.*, 2013, **4**, 466-469.
89. C. Tovar, B. Graves, K. Packman, Z. Filipovic, B. H. M. Xia, C. Tardell, R. Garrido, E. Lee, K. Kolinsky, K.-H. To, M. Linn, F. Podlaski, P. Wovkulich, B. Vu and L. T. Vassilev, *Cancer Res.*, 2013, **73**, 2587-2597.
90. Q. Ding, Z. Zhang, J.-J. Liu, N. Jiang, J. Zhang, T. M. Ross, X.-J. Chu, D. Bartkovitz, F. Podlaski, C. Janson, C. Tovar, Z. M. Filipovic, B. Higgins, K. Glenn, K. Packman, L. T. Vassilev and B. Graves, *J. Med. Chem.*, 2013, **56**, 5979-5983.
91. C.-M. Park, M. Bruncko, J. Adickes, J. Bauch, H. Ding, A. Kunzer, K. C. Marsh, P. Nimmer, A. R. Shoemaker, X. Song, S. K. Tahir, C. Tse, X. Wang, M. D. Wendt, X. Yang, H. Zhang, S. W. Fesik, S. H. Rosenberg and S. W. Elmore, *J. Med. Chem.*, 2008, **51**, 6902-6915.
92. A. J. Souers, J. D. Leverson, E. R. Boghaert, S. L. Ackler, N. D. Catron, J. Chen, B. D. Dayton, H. Ding, S. H. Enschede, W. J. Fairbrother, D. C. S. Huang, S. G. Hymowitz, S. Jin, S. L. Khaw, P. J. Kovar, L. T. Lam, J. Lee, H. L. Maecker, K. C. Marsh, K. D. Mason, M. J. Mitten, P. M. Nimmer, A. Oleksijew, C. H. Park, C.-M. Park, D. C.

Phillips, A. W. Roberts, D. Sampath, J. F. Seymour, M. L. Smith, G. M. Sullivan, S. K. Tahir, C. Tse, M. D. Wendt, Y. Xiao, J. C. Xue, H. Zhang, R. A. Humerickhouse, S. H. Rosenberg and S. W. Elmore, *Nat. Med.*, 2013, **19**, 202-208.

93. D. A. Erlanson, S. W. Fesik, R. E. Hubbard, W. Jahnke and H. Jhoti, *Nat. Rev. Drug Discov.*, 2016, **15**, 605-619.
94. E. F. Lee, P. E. Czabotar, B. J. Smith, K. Deshayes, K. Zobel, P. M. Colman and W. D. Fairlie, *Cell Death Differ.*, 2007, **14**, 1711-1713.
95. D. C. Horwell, W. Howson, G. S. Ratcliffe and H. M. G. Willems, *Biorg. Med. Chem.*, 1996, **4**, 33-42.
96. D. Horwell, M. Pritchard, J. Raphy and G. Ratcliffe, *Immunopharmacology*, 1996, **33**, 68-72.
97. B. P. Orner, J. T. Ernst and A. D. Hamilton, *J. Am. Chem. Soc.*, 2001, **123**, 5382-5383.
98. J. T. Ernst, O. Kutzki, A. K. Debnath, S. Jiang, H. Lu and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2002, **41**, 278-281.
99. H. Yin, G.-i. Lee, H. S. Park, G. A. Payne, J. M. Rodriguez, S. M. Sebti and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2005, **44**, 2704-2707.
100. 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.
101. A. Kazi, J. Sun, K. Doi, S.-S. Sung, Y. Takahashi, H. Yin, J. M. Rodriguez, J. Becerril, N. Berndt, A. D. Hamilton, H.-G. Wang and S. M. Sebti, *J. Biol. Chem.*, 2011, **286**, 9382-9392.
102. I. Saraogi and Andrew D. Hamilton, *Biochem. Soc. Trans.*, 2008, **36**, 1414-1417.
103. H. Yin, G. I. Lee, K. A. Sedey, J. M. Rodriguez, H. G. Wang, S. M. Sebti and A. D. Hamilton, *J. Am. Chem. Soc.*, 2005, **127**, 5463-5468.
104. J. M. Rodriguez, L. Nevola, N. T. Ross, G.-i. Lee and A. D. Hamilton, *ChemBioChem*, 2009, **10**, 829-833.
105. J. M. Rodriguez and A. D. Hamilton, *Tetrahedron Lett.*, 2006, **47**, 7443-7446.
106. J. M. Rodriguez and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2007, **46**, 8614-8617.
107. M. J. Adler, R. T. W. Scott and A. D. Hamilton, *Chem. Eur. J.*, 2012, **18**, 12974-12977.
108. J. M. Rodriguez, N. T. Ross, W. P. Katt, D. Dhar, G.-i. Lee and A. D. Hamilton, *ChemMedChem*, 2009, **4**, 649-656.
109. A. Volonterio, L. Moisan and J. Rebek, *Org. Lett.*, 2007, **9**, 3733-3736.
110. L. Moisan, S. Odermatt, N. Gombosuren, A. Carella and J. Rebek, *Eur. J. Org. Chem.*, 2008, **2008**, 1673-1676.
111. P. Restorp and J. Rebek Jr, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 5909-5911.
112. C. G. Cummings, N. T. Ross, W. P. Katt and A. D. Hamilton, *Org. Lett.*, 2009, **11**, 25-28.

113. C. G. Cummings and A. D. Hamilton, *Tetrahedron*, 2013, **69**, 1663-1668.
114. J. H. Lee, Q. Zhang, S. Jo, S. C. Chai, M. Oh, W. Im, H. Lu and H.-S. Lim, *J. Am. Chem. Soc.*, 2011, **133**, 676-679.
115. M. E. Lanning, P. T. Wilder, H. Bailey, B. Drennen, M. Cavalier, L. Chen, J. L. Yap, M. Raje and S. Fletcher, *Org. Biomol. Chem.*, 2015, **13**, 8642-8646.
116. I. Saraogi, C. D. Incarvito and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2008, **47**, 9691-9694.
117. J. T. Ernst, J. Becerril, H. S. Park, H. Yin and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2003, **42**, 535-539.
118. I. Saraogi, J. A. Hebda, J. Becerril, L. A. Estroff, A. D. Miranker and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2010, **49**, 736-739.
119. S. Kumar, Diana E. Schlamadinger, Mark A. Brown, Joanna M. Dunn, B. Mercado, James A. Hebda, I. Saraogi, E. Rhoades, Andrew D. Hamilton and Andrew D. Miranker, *Chem. Biol.*, 2015, **22**, 369-378.
120. J. L. Yap, X. Cao, K. Vanommeslaeghe, K.-Y. Jung, C. Peddaboina, P. T. Wilder, A. Nan, A. D. MacKerell, W. R. Smythe and S. Fletcher, *Org. Biomol. Chem.*, 2012, **10**, 2928-2933.
121. X. Cao, J. L. Yap, M. K. Newell-Rogers, C. Peddaboina, W. Jiang, H. T. Papaconstantinou, D. Jupitor, A. Rai, K.-Y. Jung, R. P. Tubin, W. Yu, K. Vanommeslaeghe, P. T. Wilder, A. D. MacKerell, S. Fletcher and R. W. Smythe, *Mol. Cancer*, 2013, **12**, 1-16.
122. J. Plante, F. Campbell, B. Malkova, C. Kilner, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2008, **6**, 138-146.
123. 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.
124. 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.
125. 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.
126. P. Prabhakaran, A. Barnard, N. S. Murphy, C. A. Kilner, T. A. Edwards and A. J. Wilson, *Eur. J. Org. Chem.*, 2013, 3504-3512.
127. 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.
128. 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.
129. J.-M. Ahn and S.-Y. Han, *Tetrahedron Lett.*, 2007, **48**, 3543-3547.

130. A. Shaginian, L. R. Whitby, S. Hong, I. Hwang, B. Farooqi, M. Searcey, J. C. Chen, P. K. Vogt and D. L. Boger, *J. Am. Chem. Soc.*, 2009, **131**, 5564-5572.
131. L. R. Whitby, K. E. Boyle, L. Cai, X. Yu, M. Gochin and D. L. Boger, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 2861-2865.
132. F. Campbell, J. P. Plante, T. A. Edwards, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2010, **8**, 2344-2351.
133. K. Long, T. A. Edwards and A. J. Wilson, *Biorg. Med. Chem.*, 2013, **21**, 4034-4040.
134. 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.
135. V. Azzarito, J. A. Miles, J. Fisher, T. A. Edwards, S. L. Warriner and A. J. Wilson, *Chem. Sci.*, 2015, **6**, 2434-2443.
136. V. Azzarito, P. Rowell, A. Barnard, T. A. Edwards, A. Macdonald, S. L. Warriner and A. J. Wilson, *ChemBioChem*, 2016, **17**, 768-773.
137. J. Becerril and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2007, **46**, 4471-4473.
138. S. Thompson and A. D. Hamilton, *Org. Biomol. Chem.*, 2012, **10**, 5780-5782.
139. S. Marimganti, M. N. Cheemala and J. M. Ahn, *Org. Lett.*, 2009, **11**, 4418-4421.
140. P. Tošovská and P. S. Arora, *Org. Lett.*, 2010, **12**, 1588-1591.
141. K.-Y. Jung, K. Vanommeslaeghe, M. E. Lanning, J. L. Yap, C. Gordon, P. T. Wilder, A. D. MacKerell and S. Fletcher, *Org. Lett.*, 2013, **15**, 3234-3237.
142. J. H. Lee, M. Oh, H. S. Kim, H. Lee, W. Im and H.-S. Lim, *ACS Comb. Sci.*, 2016, **18**, 36-42.
143. S. Rodriguez-Marin, N. S. Murphy, H. J. Shepherd and A. J. Wilson, *RSC Adv.*, 2015, **5**, 104187-104192.
144. S. Thompson, R. Vallinayagam, M. J. Adler, R. T. W. Scott and A. D. Hamilton, *Tetrahedron*, 2012, **68**, 4501-4505.
145. I. C. Kim and A. D. Hamilton, *Org. Lett.*, 2006, **8**, 1751-1754.
146. C. M. Goodman, S. Choi, S. Shandler and W. F. DeGrado, *Nat. Chem. Biol.*, 2007, **3**, 252-262.
147. E. P. Gelmann, *J. Clin. Oncol.*, 2002, **20**, 3001-3015.
148. M. H. Herynk and S. A. W. Fuqua, *Endocr. Rev.*, 2004, **25**, 869-898.
149. R. Evans, *Science*, 1988, **240**, 889-895.
150. J. M. Olefsky, *J. Biol. Chem.*, 2001, **276**, 36863-36864.
151. K. Dahlman-Wright, V. Cavailles, S. A. Fuqua, V. C. Jordan, J. A. Katzenellenbogen, K. S. Korach, A. Maggi, M. Muramatsu, M. G. Parker and J. A. Gustafsson, *Pharmacol. Rev.*, 2006, **58**, 773-781.
152. E. H. Kong, A. C. W. Pike and R. E. Hubbard, *Biochem. Soc. Trans.*, 2003, **31**, 56-59.

153. G. Kuiper, E. Enmark, M. PeltoHuikko, S. Nilsson and J. A. Gustafsson, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 5925-5930.
154. S. Nilsson, S. Mäkelä, E. Treuter, M. Tujague, J. Thomsen, G. Andersson, E. Enmark, K. Pettersson, M. Warner and J.-Å. Gustafsson, *Physiol. Rev.*, 2001, **81**, 1535-1565.
155. K. A. Green and J. S. Carroll, *Nat. Rev. Cancer*, 2007, **7**, 713-722.
156. J. B. Bruning, A. A. Parent, G. Gil, M. Zhao, J. Nowak, M. C. Pace, C. L. Smith, P. V. Afonine, P. D. Adams, J. A. Katzenellenbogen and K. W. Nettles, *Nat. Chem. Biol.*, 2010, **6**, 837-843.
157. X. Gao, B. W. Loggie and Z. Nawaz, *Mol. Cancer*, 2002, **1**, 1-7.
158. J. R. Gunther, T. W. Moore, M. L. Collins and J. A. Katzenellenbogen, *ACS Chem. Biol.*, 2008, **3**, 282-286.
159. A. A. Parent, J. R. Gunther and J. A. Katzenellenbogen, *J. Med. Chem.*, 2008, **51**, 6512-6530.
160. J. R. Gunther, A. A. Parent and J. A. Katzenellenbogen, *ACS Chem. Biol.*, 2009, **4**, 435-440.
161. M. Scheepstra, L. Nieto, A. K. H. Hirsch, S. Fuchs, S. Leysen, C. V. Lam, L. in het Panhuis, C. A. A. van Boeckel, H. Wienk, R. Boelens, C. Ottmann, L.-G. Milroy and L. Brunsveld, *Angew. Chem. Int. Ed.*, 2014, **53**, 6443-6448.
162. P. Ravindranathan, T. K. Lee, L. Yang, M. M. Centenera, L. Butler, W. D. Tilley, J. T. Hsieh, J. M. Ahn and G. V. Raj, *Nat. Commun.*, 2013, **4**.
163. A. B. Williams, P. T. Weiser, R. N. Hanson, J. R. Gunther and J. A. Katzenellenbogen, *Org. Lett.*, 2009, **11**, 5370-5373.
164. M.-E. Taplin, *Nat Clin Prac Oncol*, 2007, **4**, 236-244.
165. M. J. Linja, K. P. Porkka, Z. Kang, K. J. Savinainen, O. A. Jänne, T. L. J. Tammela, R. L. Vessella, J. J. Palvimo and T. Visakorpi, *Clin. Cancer Res.*, 2004, **10**, 1032-1040.
166. E. Hur, S. J. Pfaff, E. S. Payne, H. Gron, B. M. Buehrer and R. J. Fletterick, *PLoS Biol.*, 2004, **2**, 1303-1312.
167. D. Chaturvedi, *Tetrahedron*, 2012, **68**, 15-45.
168. R. H. Boutin and G. M. Loudon, *J. Org. Chem.*, 1984, **49**, 4277-4284.
169. D. G. Hoare, A. Olson and D. E. Koshland, *J. Am. Chem. Soc.*, 1968, **90**, 1638-1643.
170. G. L'Abbe, *Chem. Rev.*, 1969, **69**, 345-363.
171. M. L. Leathen and E. A. Peterson, *Tetrahedron Lett.*, 2010, **51**, 2888-2891.
172. V. Tarwade, O. Dmitrenko, R. D. Bach and J. M. Fox, *J. Org. Chem.*, 2008, **73**, 8189-8197.
173. N. Nishi, M. Tsunemi, K. Nakamura and S. Tokura, *Macromol. Chem. Phys.*, 1991, **192**, 1811-1820.

174. K. Yamato, L. Yuan, W. Feng, A. J. Helsel, A. R. Sanford, J. Zhu, J. Deng, X. C. Zeng and B. Gong, *Org. Biomol. Chem.*, 2009, **7**, 3643-3647.
175. J. Zhu, R. D. Parra, H. Zeng, E. Skrzypczak-Jankun, X. C. Zeng and B. Gong, *J. Am. Chem. Soc.*, 2000, **122**, 4219-4220.
176. L. R. Steffel, T. J. Cashman, M. H. Reutershan and B. R. Linton, *J. Am. Chem. Soc.*, 2007, **129**, 12956-12957.
177. F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440-467.
178. P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine and N. P. Pavletich, *Science*, 1996, **274**, 948-953.
179. M. Sattler, H. Liang, D. Nettesheim, R. P. Meadows, J. E. Harlan, M. Eberstadt, H. S. Yoon, S. B. Shuker, B. S. Chang, A. J. Minn, C. B. Thompson and S. W. Fesik, *Science*, 1997, **275**, 983-986.
180. A. Letai, M. C. Bassik, L. D. Walensky, M. D. Sorcinelli, S. Weiler and S. J. Korsmeyer, *Cancer Cell*, **2**, 183-192.
181. D. T. C. and and S. J. Korsmeyer, *Annu. Rev. Immunol.*, 1998, **16**, 395-419.
182. F. Urech, *Justus Liebigs Ann. Chem.*, 1872, **164**, 255-279.
183. D. H. Appella, L. A. Christianson, I. L. Karle, D. R. Powell and S. H. Gellman, *J. Am. Chem. Soc.*, 1996, **118**, 13071-13072.
184. L. Laraia, G. McKenzie, David R. Spring, Ashok R. Venkitaraman and David J. Huggins, *Chem. Biol.*, **22**, 689-703.
185. A. A. Ivanov, F. R. Khuri and H. Fu, *Trends Pharmacol. Sci.*, 2013, **34**, 393-400.
186. K. Luger, A. W. Mader, R. K. Richmond, D. F. Sargent and T. J. Richmond, *Nature*, 1997, **389**, 251-260.
187. R. D. Kornberg, *Annu. Rev. Biochem.*, 1977, **46**, 931-954.
188. C. Das, J. K. Tyler and M. E. A. Churchill, *Trends Biochem. Sci.*, 2010, **35**, 476-489.
189. C. W. Akey and K. Luger, *Curr. Opin. Struct. Biol.*, 2003, **13**, 6-14.
190. A. Loyola and G. Almouzni, *Biochim. Biophys. Acta, Gene Struct. Expression*, 2004, **1677**, 3-11.
191. J. K. Tyler, *Eur. J. Biochem.*, 2002, **269**, 2268-2274.
192. J. K. Tyler, C. R. Adams, S.-R. Chen, R. Kobayashi, R. T. Kamakaka and J. T. Kadonaga, *Nature*, 1999, **402**, 555-560.
193. H. Tagami, D. Ray-Gallet, G. Almouzni and Y. Nakatani, *Cell*, 2004, **116**, 51-61.
194. M. W. Adkins and J. K. Tyler, *J. Biol. Chem.*, 2004, **279**, 52069-52074.
195. F. Mousson, A. Lautrette, J.-Y. Thuret, M. Agez, R. Courbeyrette, B. Amigues, E. Becker, J.-M. Neumann, R. Guerois, C. Mann and F. Ochsenbein, *Proc. Nat. Acad. Sci. U.S.A.*, 2005, **102**, 5975-5980.

196. C. M. English, M. W. Adkins, J. J. Carson, M. E. A. Churchill and J. K. Tyler, *Cell*, 2006, **127**, 495-508.
197. J. A. Sharp, E. T. Fouts, D. C. Krawitz and P. D. Kaufman, *Curr. Biol.*, 2001, **11**, 463-473.
198. M. Alilat, A. Sivolob, B. Révet and A. Prunell, *J. Mol. Biol.*, 1999, **291**, 815-841.
199. V. Jackson, *Biochemistry (Mosc).* 1995, **34**, 10607-10619.
200. D. C. Donham, J. K. Scorgie and M. E. A. Churchill, *Nucleic Acids Res.*, 2011, **39**, 5449-5458.
201. S. Henikoff, *Nat. Rev. Genet.*, 2008, **9**, 15-26.
202. S. M. Daganzo, J. P. Erzberger, W. M. Lam, E. Skordalakes, R. Zhang, A. A. Franco, S. J. Brill, P. D. Adams, J. M. Berger and P. D. Kaufman, *Curr. Biol.*, 2003, **13**, 2148-2158.
203. M. Agez, J. Chen, R. Guerois, C. van Heijenoort, J.-Y. Thuret, C. Mann and F. Ochsenbein, *Structure*, 2007, **15**, 191-199.
204. Y. Bao and X. Shen, *Cell*, **127**, 458-460.
205. A. Corpet, L. De Koning, J. Toedling, A. Savignoni, F. Berger, C. Lemaître, R. J. O'Sullivan, J. Karlseder, E. Barillot, B. Asselain, X. Sastre-Garau and G. Almouzni, *EMBO J.*, 2011, **30**, 480-493.
206. C. Das, M. S. Lucia, K. C. Hansen and J. K. Tyler, *Nature*, 2009, **459**, 113-117.
207. E. Kaiser, Colescot.Rl, Bossinge.Cd and P. I. Cook, *Anal. Biochem.*, 1970, **34**, 595-598.
208. N. J. Greenfield, *Anal. Biochem.*, 1996, **235**, 1-10.
209. P. Luo and R. L. Baldwin, *Biochemistry (Mosc).* 1997, **36**, 8413-8421.
210. A. Jasanoff and A. R. Fersht, *Biochemistry (Mosc).* 1994, **33**, 2129-2135.
211. A. Cammers-Goodwin, T. J. Allen, S. L. Oslick, K. F. McClure, J. H. Lee and D. S. Kemp, *J. Am. Chem. Soc.*, 1996, **118**, 3082-3090.
212. A. Velázquez-Campoy, H. Ohtaka, A. Nezami, S. Muzammil and E. Freire, *Curr. Protoc. Cell Biol.*, 2001.
213. G. Clore and A. Gronenborn, *Science*, 1991, **252**, 1390-1399.
214. T. Vojkovsky, *Pept. Res.*, 1995, **8**, 236-237.