Design and Synthesis of Novel Proteomimetic Scaffolds for the Inhibition of Protein-Protein Interactions

Natasha Susan Murphy

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

University of Leeds
School of Chemistry

September 2013

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 paragraph on 'Constrained Peptides' in 'Chapter 1' was adapted from the review article: 'Inhibition of alpha-helix mediated protein-protein interactions using designed molecules', V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson; Nat. Chem. 2013, ***5****, 161-173. The contributions of the authors are as follows: NM (candidate) wrote the corresponding section of the review and produced the related figures. AJW edited the manuscript into its present form. Other authors have equally contributed to the review (on different sections).*

The work described in Chapter 2 is adapted from the research article: 'Development of a Solid Phase Methodology for Synthesis of Aromatic Oligoamide Inhibitors of α-helix Mediated Protein-Protein Interactions', 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. The contributions of the authors are as follows: NM (candidate) was the lead author on this piece of work and wrote the original manuscript. AJW edited the manuscript into its present form. P. Prabhakaran and V. Azzarito contributed to the synthesis of several Fmoc protected monomers. J. P. Plante devised the original synthesis for the Fmoc protected monomers. X-ray crystallography was carried out by M. J. Hardie and C. A. Kilner.*

A minor contribution was made to the paper '2-O-Alkylated para-benzamide α-helix mimetics: the role of scaffold curvature'; 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. Elements of the molecular modelling carried out in 'Chapter 1' by NM (candidate) are included in this manuscript.*

*A synthetic procedure was devised by NM (candidate) in the paper entitled 'Aromatic oligoamide foldamers with a "wet edge" as inhibitors of the helix mediated p53-hDM2 protein-protein interaction', P. Prabhakaran, A. Barnard, A. I. Bartlett, N. S. Murphy, T. A. Edwards and A. J. Wilson, EurJOC, 2013, **17**, 3504–3512. The synthetic procedure is also used to synthesise a core starting material in 'Chapter 4'.*

Acknowledgements

Firstly, I'd like to thank my supervisor Prof. A. J. Wilson for accepting me on to do a great PhD project after completing a summer internship and Masters research project. I know you hoped this day would never come, but the time has finally arrived for me to spread my wings and leave the nest! Thank you to the Mary and Alice Smith Memorial Fund for funding, to all the collaborators on the project and to Professor Colin Fishwick for acting as my internal assessor and providing me with valuable feedback on all of my reports. Thank you to Dr. S. Warriner for giving me advice and poetry, and to Dr. K. Long, Dr. P. Prabhakaran and Dr. A. Bartlett for expressing proteins for me to carry out my work. Thank you also to all the technical, administrative and stores staff for making life in the department that little bit easier. Thank you to all the people that have come and gone and those who I will leave behind that I do not give a mention to; the PhD would not have been the same without you.

Past and present members of the Wilson group, you have all been amazing and I couldn't have picked a better bunch of people to share the joys and heartaches of research with: Jeff, Maria, Panchami, Alice and Anna for being our all knowing post docs; Jimmy, Fred, Adam and Mazza for helping me to settle in as a Wilsonian newbie; George for dancing and eating too much Nandos with us girls; Dave for being Dave and letting me play with his hair; Kelly for being one of the nicest people ever; George for letting me pick his brain during my write up; Sílvia, I wish you'd got here earlier so we could enjoy more fun times together. Thanks also go to Alun for all the gossips and Kat for joining me in working at the weekends because we're too cool. Thank you to Qi and Jing for showering me in oriental food and gifts and to Andy for being my ally. Thanks to the boys for allowing us to have quadruple dates, you know who you are. Thanks to Nic for listening to me for hours talk about nothing and for always making a plan happen, I can't wait for our next holiday! Finally, my lovely, amazing homies Kérya and Vale, I really wouldn't have made it through without you. Thank you for all the great memories, and helping me through the hard times and just generally being there to support each other. Now we are making our way back out into the real world (on our own) and I wish you all the best and happiness in your lives. Good luck also to all of you in the middle of your PhD...the end will come.

I would also like to thank my family; Maxine for doing what sisters do and for giving me the best Niece in the world. She is so beautiful, clever and funny and I can't wait for her to become a scientist like her favourite Auntie. Brian, you've helped me so much over the last few years so thank you for that and for making my mum so happy and taking care of her while I'm not there. And of course, thank you to Sam for being amazing and

supporting me in so many ways. You have been my rock and thank you for getting me through to finally handing this thesis in and ready to start my new adventure. We will make it to South America!

So I dedicate this thesis to you Sam and to my Mummy. I can't even begin to say how much you've done for me over the years, you took me to Leeds when I was 18 to start my university life and ferried me to and from ever since. Who knew I'd be here for that long? You made all of this possible and for that I am truly grateful. I love you both so much and hope I've made you proud! And you Dad xxx

Abstract

Proteins are essential parts of living organisms and protein-protein interactions (PPIs) interactions mediate many essential regulatory pathways. As such, PPIs have been implicated in a number of diseased states, however, it is currently unclear how to effectively target them due to the relatively poorly defined surface at the protein interface. When PPIs are mediated by the binding of an α-helix, key interactions usually occur on non-adjacent residues, appearing on the same face of the α-helix, resulting in close interactions. The abundance of this secondary structure in proteins and its relative rigidity provide an ideal basis for the design of synthetic mimics. In this thesis, an account of the design strategies developed to address this problem is provided, and sets the work described herein in context. Previously, the Wilson group developed 3-*O*-alkylated oligobenzamide (3HABA) scaffolds as potential α-helix mimetics and suitably functionalised trimers were identified as micromolar inhibitors of the p53/*h*DM2 interaction. To understand more how to develop potent inhibitors of this interaction, a larger library was necessary. This was achieved by generating a library of 3HABA building blocks encompassing a range of natural and unnatural functionalities. Parallel to building the monomer library, development of a general solid phase methodology for deactivated anilines was essential. The course taken was to provide a solution for this challenging technical problem, and to identify the scope of the SPS methodology. Using the developed methodology, libraries of compounds targeting the p53/*h*DM2 and Mcl-1/NOXA B interactions were synthesised and their biophysical properties evaluated resulting in directions for future library development. To extend the approach to helix mediated PPIs involving more than one face, bifacial scaffolds designed at target the ER/coactivator complex are also described. This works discusses how molecular modelling, initial biophysical testing and docking studies led to second generation ligands with better *in silico* properties.

Table of Contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 6

List of Schemes

Chapter 2

Chapter 3

Chapter 4

List of Tables

Chapter 2

Chapter 4

Abbreviations and Symbols

3AHB	Methyl-3-aminohydroxybenzoate
3D (1D/2D)	Three-Dimensional
3HABA (2HABA)	3-Hydroxyaminobenzoic acid (2-Hydroxyaminobenzoic acid)
3HNBA	3-Hydroxynitrobenzoic acid
AIB	Amplified In Breast Cancer
Bak	Bcl-2 antagonist/killer
Bax	Bcl-2-associated X protein
Bcl-2	B-cell lymphoma 2
Bcl-x_L	B-cell lymphoma extra large
BH	Bcl-2 homology domain
Bid	BH3 interacting domain death agonist
Bim	Bcl-2 interacting mediator of cell death
CaM	Calmodulin
CD	Circular dichroism
Cdc42	Cell division cycle 42
CH1	Cysteine/histidine-rich 1
DAHB	methyl-3-aminodihydroxybenzoate
Dbs	Dbl's big sister
DHABA	3,6-dihydroxyaminobenzoic acid
DHNBA	3-dihydroxynitrobenzoic acid
DIAD	Diisopropyl azodicarboxylate
DMF	Dimethylformamide
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EC_{50}	Half maximal effective concentration
EDCI	1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide

ERK	Extracellular-signal-regulated kinase
Fmoc-	Fluorenylmethyloxycarbonyl-
gp41	Glycoprotein 41
GRIP1	Glucocorticoid receptor-interacting polypeptide 1
HCTU	*O*-(6-Chlorobenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate
*h*DM2 / *m*DM2	Human double minute 2 / murine double minute 2
*h*DMX	Human double minute X
HEK293	Human Embryonic Kidney 293
HIF-1α	Hypoxia-inducible factor 1-alpha
HIV	Human Immunodeficiency Virus
HPLC	High performance liquid chromatography
IC_{50}	Half maximal inhibitory concentration
IL-2	Interleukin receptor
K_d	Dissociation constant
K_i	Inhibition constant
LC-MS	Liquid phase chromatography – mass spectrometry
MAML1	Mastermind-like protein 1
Mcl-1	Myeloid cell leukemia-sequence 1
NK1-NK3	Neurokinin Receptors (1-3)
NMP	*N*-methyl-2-pyrrolidone
NMR	Nuclear magnetic resonance
NOE / NOESY	Nuclear Overhauser Effect (Spectroscopy)
NOTCH1	Notch homolog 1, translocation-associated (Drosophila)
NR	Nuclear receptor
PDB	Protein data bank
PELP1	Proline, glutamic acid and leucine rich protein 1
Puma	P53 upregulated modulator of apoptosis

Ras	Rat sarcoma (guanine nucleotide–binding protein)
RMSD	Root mean squared deviation
RNA	Ribonucleic acid
RNAse	Ribonuclease
SAR	Structure activity relationship
smMLCK	smooth muscle myosin light-chain kinase
SOS	Son of Sevenless
SPPS / SPS	Solid phase (peptide) synthesis
t ½	Half life
SRC-1 / SRC-3	Steroid receptor coactivator-1 / -3
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
Tif-2	Transcriptional intermediary factor-2
VEGF	Vascular endothelial growth factor

Amino Acid Codes:

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
Ornithine	Orn	-

Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

Chapter One

Protein-Protein Interactions and Strategies for their Inhibition

Section 1.4.1 of this chapter contributed to a review article published in Nature Chemistry.[1]

1.1 Relevance of Protein-Protein Interactions

Proteins are essential parts of living organisms and participate in every process within the cell. They are responsible for a plethora of functions including catalysis, transportation, signalling and the transmission of information from DNA to RNA. Furthermore, proteins are central to the immune system, acting as vehicles for the immune response but also facilitating viral entry into cells. Due to the extensive involvement of proteins in cellular processes, continuous effort has been made to predict their functions from amino acid sequences and where possible their tertiary structure. Many of these biological functions involve protein-protein interactions (PPIs) so identification, characterisation and inhibition of PPIs is crucial for drug discovery.[2-4]

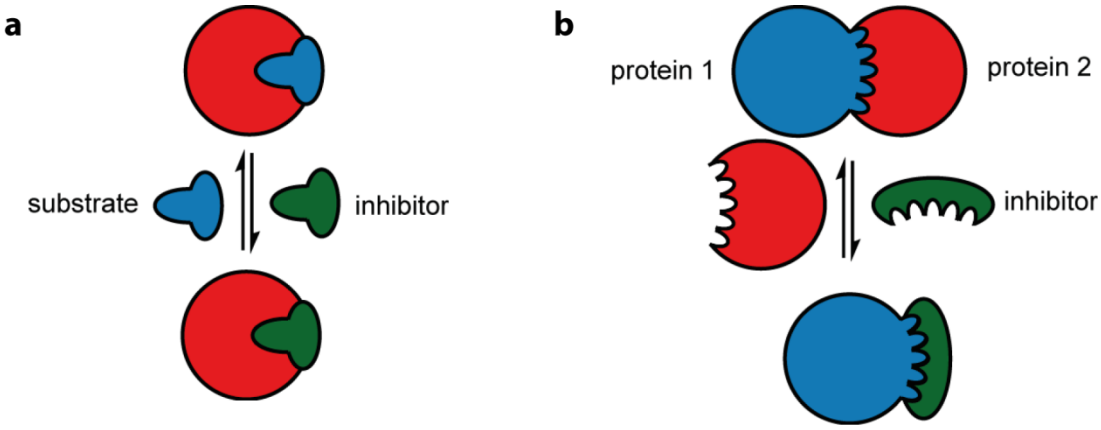


Figure 1.1 (**a**) Binding in an enzyme-substrate complex containing few strong interactions. (**b**) Binding in protein-protein complex containing many small, additive interactions.

Designing small molecule inhibitors to fit enzyme active sites has been relatively successful.[5] The catalytic area is found within a well-defined cavity or cleft within the enzyme and contains multiple recognition sites (**Fig. 1.1a**) allowing inhibitors to be designed with appropriate functionality (e.g. hydrogen bonds, π-π interactions, electrostatic forces and salt bridges). Native substrates may also present effective templates for inhibitor design making it possible to synthesise small compounds with many interactions. PPIs on the other hand were once considered "undruggable" and have not been a focus of drug development until recently. The extensive, relatively featureless surface of proteins (800-1100Å^2) with distant and varied interactions creates a daunting task for inhibitor development (**Fig. 1.1b**)[6]; designing inhibitors to complement the poorly defined hydrophobic, charged or polar domains in order to show competitive inhibition is difficult. There has been considerable interest in PPIs and a major goal is to gain a better understanding and quantification of the key features controlling these interactions. This will hopefully lead to greater success in the

prediction of protein associations and assist with the elucidation of cellular pathways and drug design.[3]

1.1.1 Structural and Thermodynamic Features of Protein-Protein Interaction

Many studies have been made to formulate a system allowing for reliable predictions about PPIs to be made. These have been hampered by many different factors including protein flexibility, the presence of even partial disorder, the existence of ensembles with distinct conformations separated by energy barriers, and the cooperativity in protein-protein association.[3] Wells and co-workers exploited a technique (alanine scanning) which involved the interfacial residues being mutated to alanine systematically to determine the change in binding free energy. Alanine scanning presented evidence that there are regions on a protein surface that distinguish it from the rest of the protein surface; the functional epitope also known as a 'hot spot'.[7] A hot spot is defined as a residue whose substitution by alanine leads to a significant drop in the binding free energy ($\Delta G \geq 2$ kcal/mol) to its protein partner.[8] Hot spots are usually found within densely packed areas in which clusters of amino acids are in contact with each other forming a network of interactions. These can be described as 'hot regions' and explain why those residues have a dominant contribution to the stability of the complex.[2,9] **Figure 1.2** demonstrates how close key binding residues can be in a hot spot and shows how a small molecule is able to interact with the protein over these regions.

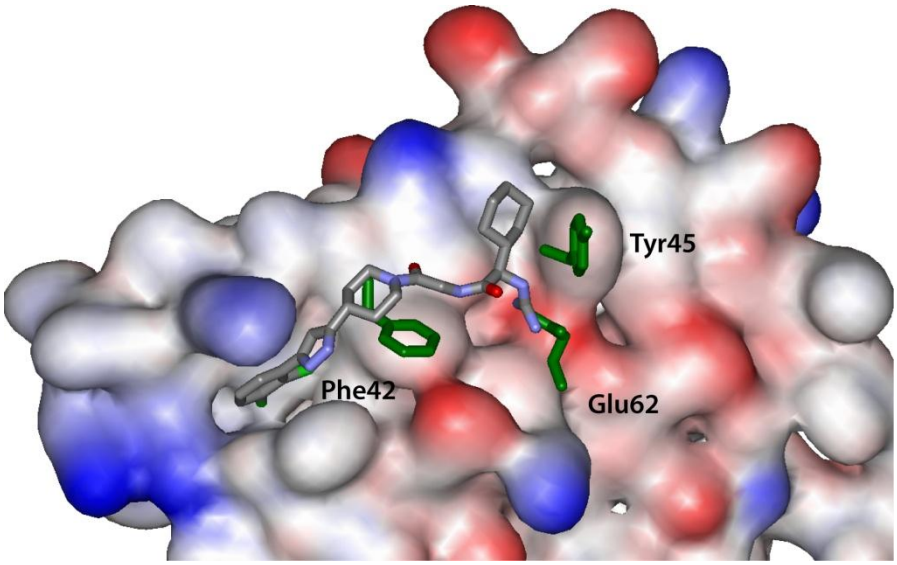


Figure 1.2 Ligand bound to hot region on the interleukin receptor, IL-2. Hot spot residues for binding to IL-2 are shown in green (PDB ID: 1PW6).[10]

1.1.2 α-Helix Mediated Protein-Protein Interactions

Over 30% of protein secondary structure is helical making it the most abundant secondary structural unit in proteins. The helix contains important shape and sequence-selective recognition motifs. The importance is highlighted by the number of proteins that bind their partner through an α-helix and it is anticipated that this interaction may be general. With 3.6 residues and a 0.54 nm rise per turn and average dihedral angles Φ and Ψ of -60° and -45° (**Fig. 1.3a**), side-chains are placed above one another every 3-4 turns. Thus helices can be considered to have three distinct faces (**Fig. 1.3b**).

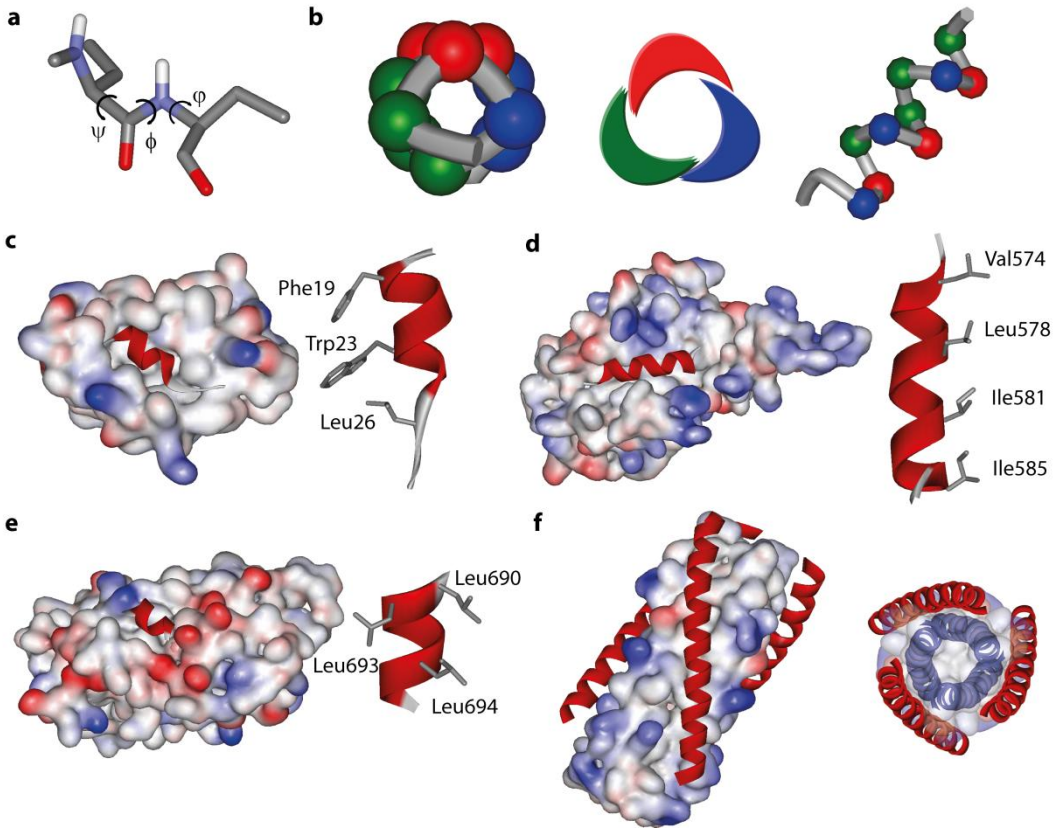


Figure 1.3 (**a**) Backbone dihedral angles in a polypeptide chain (**b**) View down the centre of a helix highlighting the three faces (left) with a cartoon representation (middle) and an alternative view of the helix (right); α-carbons represented by CPKs, with each colour corresponding to one of the three helical faces. (**c**) p53/*h*DM2 – regulates cellular stress (PDB ID: 1YCR). (**d**) Bcl-x_L-Bak regulates apoptosis (PDB ID: 1BXL). (**e**) ER/coreceptor – regulates growth and function of different tissues (PDB ID: 2QZO). (**f**) gp41 hexameric coiled-coil leads to Viral fusion (PDB ID: 1AIK).

In a number of α-helix mediated PPIs, the binding interactions occur through non-adjacent residues displayed on the same face of the α-helix; normally the *i*, *i*+3/*i*+4, *i*+7/*i*+8 (and *i*+11) residues are involved in binding resulting in close interactions. The p53/*h*DM2[11] and Bcl-2 family[12] of PPIs have such a binding motif, with primarily

hydrophobic residues making key interactions within a hydrophobic helix binding cleft (**Fig 1.3c-d**). The interaction between estrogen receptor and its coactivator proteins however, involves two faces of the helix and a charge clamp on the receptor surface helps to further stabilise the complex (**Fig. 1.3e**).[13] Other interactions may involve three faces of the helix or binding in a much shallower groove. The gp41 hexameric coiled coil assembly demonstrates even greater diversity of PPIs through the packing of three, much longer, 40 residue *N*-terminal helices with a trimeric coiled coil core consisting of *C*-terminal helices (**Fig. 1.3f**).[14] The Arora group used computational alanine scanning mutagenesis of the Protein Data Bank to identify and assess helical interfaces in PPIs. They identified a list of target PPIs amenable to disruption by helix mimetics and grouped them according to the number of helical faces involved in binding. The most common motif was identified to involve hydrophobic residues displayed on a single face of the helix at the *i*, *i* + 4 and *i* + 7 positions.[15]

1.2 PPI Targets and their biological importance

Mutation in a cell's DNA is the fundamental cause of cancer, although many mutations may occur in cells which are harmless. Mutations of DNA can be inherited or can be caused by external factors such as sunlight, viruses and certain lifestyle choices (*i.e.* smoking, drinking and diet). In healthy cells, a defence mechanism called apoptosis is initiated which is a highly conserved, specific and selective means of controlling tissue mass and shape. It prevents mutations from spreading by controlling cell death within multicellular organisms.[16] Some cancers, however, are caused by mutations in the DNA of genes such as oncogenes or tumour suppressors. These genes control when a cell is required to grow and importantly, when to stop growing. The process of apoptosis does not take place in cells in which these mutations occur and so the cancer cell survives. The cells then begin to grow uncontrollably and processes such as mitosis take place allowing the mutation to spread. The result of such mutations can lead to tumours.[17]

1.2.1 Protein 53 / Human Double Minute Two Interaction

Tumour protein 53 (TP53 or p53) plays a central role in cell cycle regulation and as such is a major tumour suppressor in humans. [18] It has been found that over 50% of human tumours contain a mutated or absent gene that codes for p53.[19] In healthy cells, the production and degradation of the p53 protein is regulated by the binding of *h*DM2 (**Fig. 1.4**), which in turn is induced by p53 in a negative feedback loop. Regulation is controlled by *h*DM2 *via* multiple mechanisms: it physically blocks the interaction between p53 and DNA thereby inhibiting its transcriptional activity, it induces nuclear export of p53, and on binding acts as a "tag" stimulating p53 degradation through its E3 ubiquitin ligase activity.

Over expression of *h*DM2 in cancer cells is shown to inactivate p53 by preventing it from exerting its apoptotic activity and allowing for cell proliferation. In about 30% of soft tissue and human osteogenic sarcomas, *h*DM2 is over expressed due to gene amplification. This indicates that *h*DM2 has an important role in the development of these tumours and so this PPI has become a major target for cancer chemotherapy.[20]

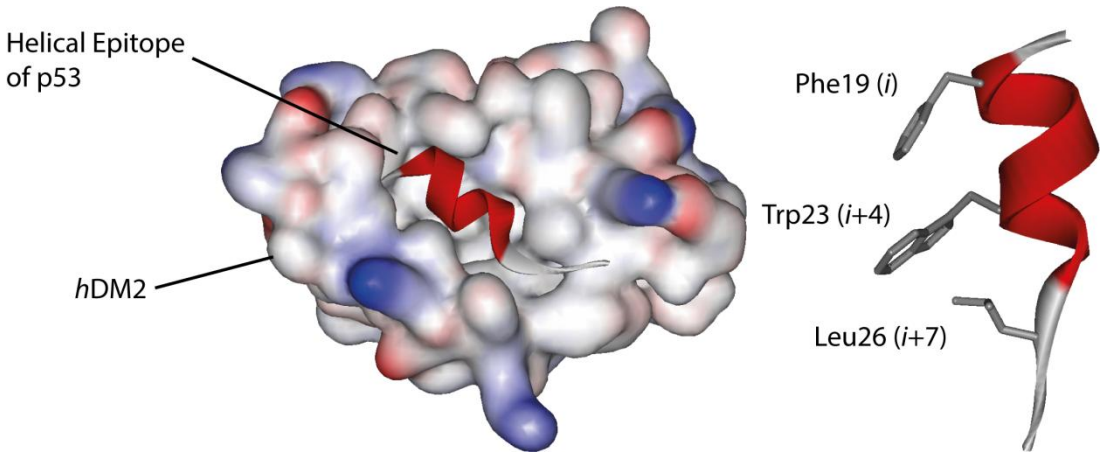


Figure 1.4 *h*DM2 and the helical binding epitope of p53 showing the key interacting residues (PDB ID: 1YCR).

The hypothesis is that by inhibiting binding interactions in the p53/*h*DM2 complex, p53 protein will be freed and this will allow it to exert its apoptotic properties. Crystal structures of the complex have revealed a hydrophobic binding pocket on *h*DM2 in which an α-helical region of p53 makes hydrophobic interactions. These contacts are made through Phe19, Trp23 and Leu26, corresponding to the *i*, *i*+4 and *i*+7 residues of the helix respectively (**Fig. 1.4**). Attaching these or similar functionalised residues to a rigid scaffold occupying the same orientation should result in an effective *h*DM2 inhibitor.[11, 20]

1.2.2 Nuclear Hormone Receptors

Steroid hormones occur naturally in the body and are essential in controlling the growth and activity of normal cells in many different tissues. The estrogen receptors (ER) and androgen receptor (AR) are ligand-activated transcription factors that belong to the nuclear hormone receptor superfamily and are responsible for mediating the physiological effects of the steroid hormone ligands. Estrogens and testosterones bind to the receptors with a high specificity and affinity and the resulting ligand/receptor dimer is able to exert its effects at both nuclear and cell membrane sites. Transcriptional control by the nuclear receptors requires interaction with coregulator complexes, either coactivators for stimulation or corepressors for inhibition of target gene expression. Interaction with the coregulator occurs after the ligand is bound, concomitant with structural changes which occur on binding of the ligand, exposing the surface required for coregulator binding. Along with the

desired regulation of normal cells, hormones have also been found to control the growth of several forms of cancer tumours. Drugs or treatments that block the effects of hormones, or lower the levels can therefore be used to treat some types of cancer.[21-23]

1.2.2.1 Estrogen Receptors

Estrogen receptors can be split up into two main groups, estrogen receptor alpha (ERα) and estrogen receptor beta (ERβ).[22] ERα and ERβ are coded by separate genes which are found on different chromosomes accordingly. Both genes can be split into six structural domains (termed domains A-F) and within these, well defined functional domains are contained. Within ERα and ERβ, homology between the domains vary: 96 % homology in the DNA binding domain (C), 53% homology between E/F domains and A, B and hinge (D) domains are not well conserved. It is within the E and F domains that the coregulatory domains are found, such as the ligand binding domain (LBD) and the ligand-dependent activation function 2 (AF-2) surface. The LBD contains sequence specific features essential for interaction with ER coactivators and corepressors.[21] Complexation between ERs and other cellular factors is required to modulate ER-mediated transcriptional activity and so selectivity between the two ERs can be achieved due to the lack of homology within the nucleotide sequence.[24] This thesis however, will focus on the ERα.

1.2.2.1.1 Estrogen Receptor α

ERα regulates the growth and function of tissues found in the female reproductive system such as breast, uterus and ovaries. It has also been identified as playing an important role in many pathological processes with 70% of breast tumours developing due to the stimulatory effect of estrogens. In gene transcription, ERα forms complexes with coactivators enabling transcriptional activation to occur. The p160 protein family is a group of ERα coactivators and consists of three members; SRC-1, Tif-2/GRIP1 and SRC-3/AIB1. Knockout studies showed that in ERα-positive breast cancer, the p160 gene is amplified indicating that these coactivators are important in ERα signalling.[24] These coactivators interact with the ERα *via* small amphipathic α-helical peptide sequences containing a common recognition motif; LXXLL (L = leucine, X = any amino acid) also known as the nuclear receptor box (NR box) (**Fig. 1.5**)[25]. The amino acids at position X vary, helping to facilitate specific NR box recognition for a particular nuclear receptor. The NR box binds to a hydrophobic cleft on the ERα AF-2 surface on the LBD controlled by the specificities of the activating ligand, 17-*β*-oestradiol (E2).[24] Studies using X-ray crystallography showed that the leucine side-chains in positions *i* and *i*+4 are projected into a hydrophobic groove whilst the side chain at the *i*+3 position projects into a hydrophobic pocket (**Fig. 1.5**).[7] The crystal structure also suggests that the peptide backbone interacts with the charged residues

that flank the binding groove on the ER further stabilising the complex. It has been proposed that these features may allow the ERα-coactivator interaction to be targeted by LXXLL motif-like inhibitors, preventing transactivation from occurring.[24]

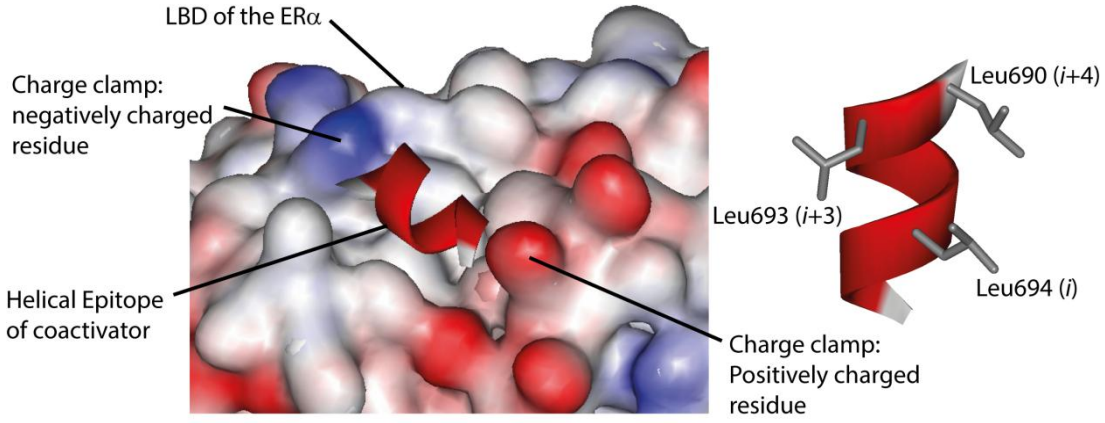


Figure 1.5 The LBD on ERα and the binding helical epitope of a nuclear receptor coactivator. The leucine residues in the LXXLL recognition motif are shown in grey (PDB ID:2QZO).

1.3 Approaches for Inhibition of α-Helix Mediated Protein-Protein Interactions

The need to modulate PPIs is increasing, and the challenges this presents have been discussed in addition to the critical features on the protein which can be exploited to overcome these challenges. The rest of this chapter will highlight major advances in developing generic approaches for PPI modulation. Several strategies for PPI inhibition have been developed and these can be grouped into three major categories. Type I (constrained peptides and β-peptides) and Type III (proteomimetics) mimetics use the native helix motif as a basis for inhibitor design and these strategies will be discussed in depth. Type II mimetics on the other hand are typically small molecule inhibitors, identified using conventional drug discovery methods such as high throughput screening (Nutlin-3a: **Fig. 1.6a**)[26] and fragment-based design (ABT-737 and ultimately ABT-263: **Fig. 1.6b**)[27]. Despite these important breakthroughs, identifying small molecule inhibitors of PPIs has proven challenging and since they are designed to selectively target one specific PPI, they have less relevance within the context of designing inhibitors and developing a general rule for PPI inhibition is much less likely; these types of inhibitors will therefore not be discussed.

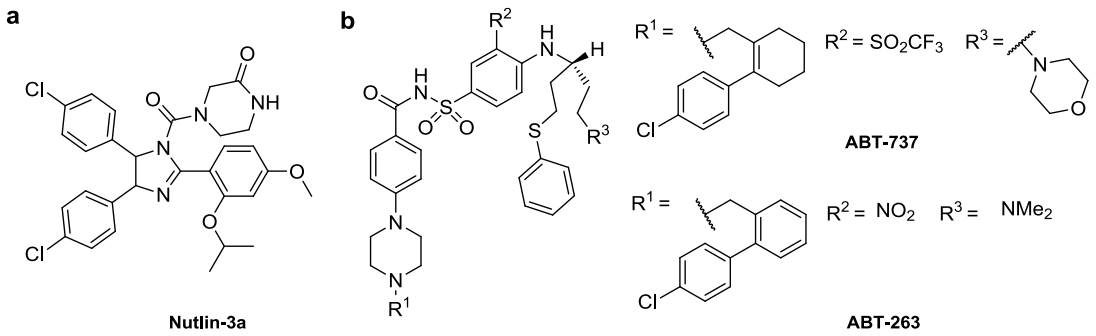


Figure 1.6 Chemical structures of (**a**) Nutlin-3a identified *via* HTS and (**b**) ABT-737 and ABT-263 idenified *via* a fragment based approach.

1.4 Type I Mimetics

Peptides are often considered therapeutically undesirable due to poor transport properties and their sensitivity to proteolytic degradation.[28] However, peptides have the advantage of providing a highly conserved and complex set of functions that cannot be completely mimicked by a small molecule, and which potentially results in diminished interference with normal biological processes.[29] The development of therapeutics composed of the helical domain central to the PPI of interest has therefore attracted due attention. Type I mimetics are short fragments of peptide which strive to replicate the local topography of the α-helical structural motif. Once removed from the stabilising environment of a protein, synthetic peptides become far less organised in solution, adopting only random conformations and the ability to bind to the partner protein is consequently retarded.[30] Several approaches are being explored to improve helicity whilst also addressing problems of proteolytic stability and other pharmacokinetic factors. These approaches can be grouped into two general categories: constrained peptides and helical foldamers.

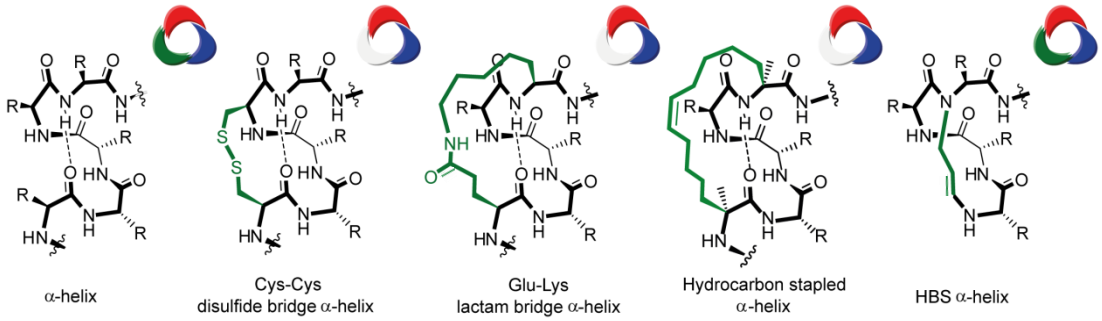


Figure 1.7 Schematics illustrating different approaches for covalent helix stabilisation (helix faces are coloured red, blue and green, with white faces indicating that the covalent stabilisation interferes with molecular recognition of that face).

1.4.1 Constrained Peptides

This class of peptide can also be categorised within helix stabilisation methods. These are methods which restrict the peptide backbone, assisting with pre-organisation of the amino acid residues and initiating helix formation. Early strategies for helix stabilisation include helix nucleating templates and metal-mediated bridges, however, covalent cyclisation affords the most successful bioactive conformations for inhibiting the desired target.[31] On a single face of the α-helix structure, lie residues at the *i*, *i*+4, *i*+7 and *i*+11 positions. A number of methods employ pairs of residues at these positions covalently linking side-chains, whilst other cyclisation strategies use main chain-main chain connections such as the hydrogen bond surrogate (HBS) method (**Fig. 1.7**).

1.4.1.1 Disulfide Bridges

One of the first examples of helix stabilisation *via* disulfides was reported by Wemmer *et al*. who synthesised hybrid peptides containing a disulfide linkage which formed the same basic 3D structure as Apamin. The hybrid peptide was effective in inducing a biological response which necessitates a helical conformation.[32] Following this, Schultz and co-workers studied the effect of stereoisomers of cysteine positioned at the *i* and *i*+7 positions. When D-Cys and L-Cys were incorporated at the *i* and *i*+7 positions respectively, intramolecular disulfide formation was achieved with little perturbation on helical conformation. L,L analogues showed a marginal increase in helicity from the unconstrained peptide and L,D analogues resembled random coils or β-sheet conformations.[33] Later studies from the Mierke and Spatola groups concurred that the D,L configuration of oxidised cysteines at the *i* and *i*+3 positions was also the best combination for targeting their desired receptors.[34,35] Schultz also demonstrated the structure temperature dependence of a peptide containing a single disulfide bridge. Unlike typical α-helices, the ends of the peptide relaxed to a random-coil conformation on heating whilst the residues within the bridge preserved partial helicity.[33]

An alternative method for constraining peptides is to use side chain-side chain lactam bridges. Spatola and co-workers compared the efficiency of disulfide and lactam constraints in stabilising peptides containing the conserved NR box LXXLL pentapeptide, with the aim of designing potent and selective inhibitors of steroid receptor-coactivator interactions.[36] A disulfide bridged nonapeptide inhibited the ERα/coactivator interaction with an inhibition constant an order of magnitude higher than the lactam bridged analogue (K_i = 25 nM vs 220 nM); confirming molecular modelling predictions. It was also considerably more potent than a comparable linear peptide with 13 residues and approximately 15 times more selective for ERα than ERβ. CD spectra of the peptide

indicated minimal helical character in an aqueous environment, however, an X-ray crystal structure (PDB ID: 1PCG) confirmed that the disulfide linked peptide does bind in the expected conformation implying receptor-induced conformational changes of the constrained coactivator peptide on binding.[36]

Following this, Spatola and co-workers carried out an extensive SAR study on disulfide bridged Selective Estrogen Receptor Modulators (SERMS) to determine what factors affect the binding affinity and selectivity. The study probed the effects of changing configuration, disulfide ring size and peptide chain flexibility *via* homocysteine and penicillamine incorporation (**Fig. 1.8**). They also manipulated the LXXLL NR box to contain a range of unnatural leucine surrogates. Several important concepts came out of the SAR study allowing them to design potent and selective inhibitors,[35] alongside replacement of the disulfide with a thioether link to produce a potent inhibitor of the ERα/co-activator interaction with an inhibition constant of 6.9 nM. The cystathionine linked peptide also demonstrated over 9 fold selectivity for ERα than ERβ.[37]

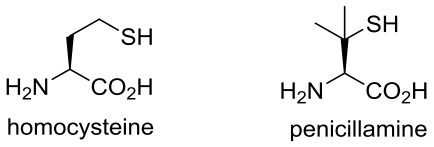


Figure 1.8 Chemical Structure of unnatural amino acids; homocysteine and penicillamine.

1.4.1.2 Lactam Bridges

The aforementioned lactam link is another approach for constraining peptides and is probably the most investigated. Early designs of lactam bridges produced biologically active cyclic peptides and since then much work has focused on side chain-side chain lactam bridges incorporated into longer peptides. The earliest demonstration following this rationale was reported by Rosenblatt *et al*.[38] A parathoid hormone related protein (PTHrP) analogue was stabilised *via* a lactam link between Lys and Asp in the *i* and *i*+4 positions respectively. The constrained peptide was 5-10 times more potent than the parent linear peptide and incorporation of the corresponding D-amino acids at either position considerably reduced potency. The authors speculated that configuration dependence may reflect steric requirements imposed by the receptor or by steric hindrance imposed on Lys and Asp by neighbouring residues. Subsequently, Geistlinger and Guy applied a similar approach for stabilisation of estrogen-binding coactivator peptides.[39]

McDowell and co-workers explored the use of constrained peptides to inhibit HIV type 1 fusion events.[40] HIV infection requires membrane fusion mediated by the *C*-terminal heptad region (CHR) of the gp41 peptide binding in an antiparallel fashion in a groove on the *N*-terminal heptad region (NHR) of gp41. Comparison of the relative activities of a linear peptide with various constrained peptides demonstrated a correlation between helicity and inhibitory potency, whilst activity of the peptide containing a restraint on the binding face was reduced entirely. All truncated peptides, however, were found to be less potent than the full length analogues.[40] The study was taken further by Kim and co-workers; they aimed to stabilise a 14 residue *C*-terminal peptide targeting the hydrophobic pocket of HIV-1 gp41 by incorporating unnatural helix favouring amino acids and chemical cross-links.[41] They found that cross-linking two Glu residues at the *i* and *i*+7 positions with an α,ω diaminoalkane group resulted in the most potent inhibitor (IC_{50} = 35 μM, K_d = 1.2 μM). An X-ray crystal structure of the cross-linked peptide bound to the HIV-1 gp41 hydrophobic pocket demonstrated that it binds with virtually the same helical conformation as the native *C*-terminal peptide (**Fig. 1.9a,b**).[41]

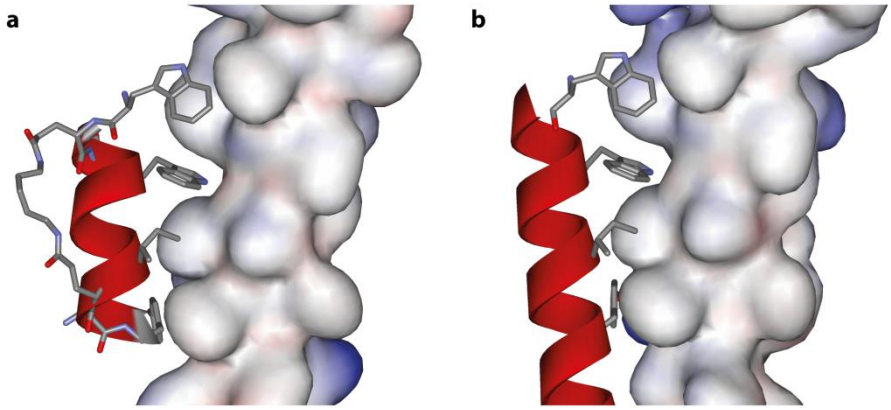


Fig. 1.9 (**a**) Crystal structure showing the binding of a cross-linked peptide bound to the hydrophobic pocket on HIV-1 gp41 (PDB ID: 1GZL) (**b**) Crystal structure showing the binding of the *C*-terminal peptide in the hydrophobic pocket HIV-1 gp41 (PDB ID: 1AIK).

Fairlie and co-workers reported an approach for stabilisation of short α-helical peptides which relies upon a Lys1→Asp5 linkage. They used a pentapeptide module to create cyclic peptides with different arrangements of linkers. This provided versatility in designing helix mimics to expose different side chains appropriate for different helix faces (**Fig. 1.10**). Biological activities of the water-soluble helices, derived from proteins with diverse functions and different receptors, were shown to be superior to linear analogues and even surpass the native ligand.[42] With this approach the group successfully mimicked the helical epitopes of (i) a quorum sensing pheromone which abolished growth of the bacteria *S. pneumonia* at sub micromolar concentrations, (ii) the F fusion protein of Respiratory Syncytial Virus showing picomolar inhibition of viral fusion, (iii) the RNA-binding viral

protein HIV-1 Rev showing nanomolar affinity for the RNA segment Rev Responsive Element and (iv) the human hormone nociceptin and induced intracellular ERK phosphorylation at picomolar concentrations (the most potent agonist identified to date).

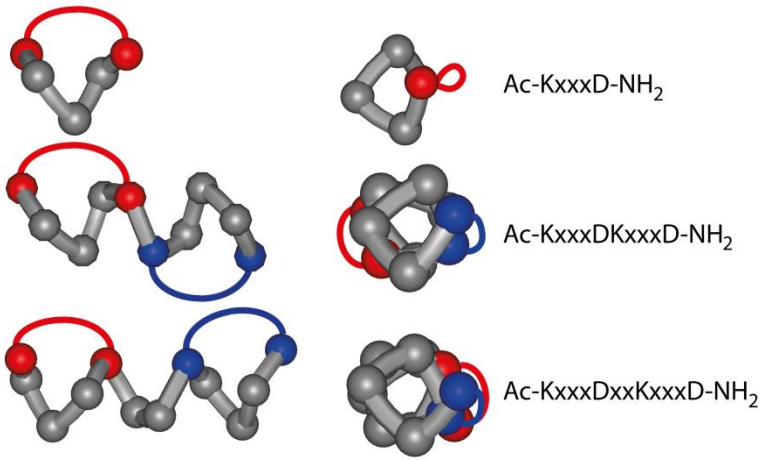


Figure 1.10 Cyclic pentapeptide modules having different distributions of interacting side chains (grey CPK = 'x' amino acids) and linking bridges (red or blue = K→D linkage)

1.4.1.3 Hydrocarbon Stapling

Whilst disulfide and lactam bridges are successful at stabilising α-helices, these naturally occurring functionalities are accordingly somewhat susceptible to degradation in cells. Grubbs and Blackwell first introduced a non-native carbon-carbon bond constraint *via* a ring-closing metathesis (RCM) reaction with *O*-allyl serine residues in an effort for enhanced biostability.[43] Verdine and co-workers elaborated on this approach and incorporated unnatural α,α-disubstituted amino acids with olefin tethers into a *C*-terminal peptide sequence of RNAse A; multiple configurations, linker lengths and positions were explored to identify ideal combinations to maximise helix content.[44] This new hydrocarbon-stapled backbone approach provided a platform for a number of significant studies in this field over the last decade. By incorporating analogous olefin bearing tethers into a 23 residue Bid BH3 peptide at the *i*, *i*+4 positions, Korsmeyer and co-workers generated peptides with a noticeable improvement in peptide α-helicity, protease resistance and *in vitro* and *in vivo* biological activity.[45] *In vivo* studies found a stabilised peptide penetrated a wide panel of leukaemia cells and selectively triggered the apoptotic pathway (IC_{50} values ranging from 1.6 - 10.2 μM). Tumour suppression and often tumour regression was also displayed in mice bearing established human leukaemia xenografts when treated with the stabilised peptide and this approach has ultimately been commercialised.[45]

Apoptosis is widely agreed to be mediated by the competitive interactions between pro- and anti-apoptotic members of the Bcl-2 family, however, how the pro-apoptotic Bax and Bak proteins trigger apoptosis remains a matter of debate.[46] The Walensky group

developed a series of ligands based on the Bid BH3 domain (referred to as "death domain" because of its ability to promote apoptosis). These ligands were stabilised by the hydrocarbon cross-linker and recapitulated the α-helical character of native death domains, directly binding to Bax and initiating Bax mediated mitochondrial apoptosis. [47] In addition, the Walensky group described an inhibitor of the Mcl-1 protein; in this instance, the most potent stapled peptide was not one derived from a known BH3 effector protein with a high affinity for Mcl-1 but rather a stapled peptide derived from Mcl-1 itself .[48]

Further targets that have been addressed using this approach included p53/*h*DM2 and ER/coactivator interactions.[49, 50] Crystal structures of stapled peptides bound to these targets have been described (**Fig. 1.11**), and in the case of ER, highlight that caution must be exercised in interpreting structure–property relationships for these ligands: the hydrocarbon linker itself can bind in the cleft of the target protein. Transcription factors have proven to be among the most difficult targets for therapeutic intervention owing to a much larger interface and the absence of a hydrophobic pocket. Bradner and co-workers successfully developed a direct-acting agonist of the oncogenic transcription factor NOTCH1 by evaluating a series of stapled peptides based on the coactivator peptide MAML1.[51] An extensive *in vitro* analysis of a 15-residue peptide stapled at the *i* and *i* + 4 positions, conclusively demonstrated binding to a NOTCH1-transcription factor complex, directly antagonising recruitment of MAML1. This stapled peptide was shown to suppress NOTCH1 signalling whilst a bioluminescent murine model exhibited a direct link between inhibition of the NOTCH pathway and anti-leukaemic activity *in vivo.*[51]

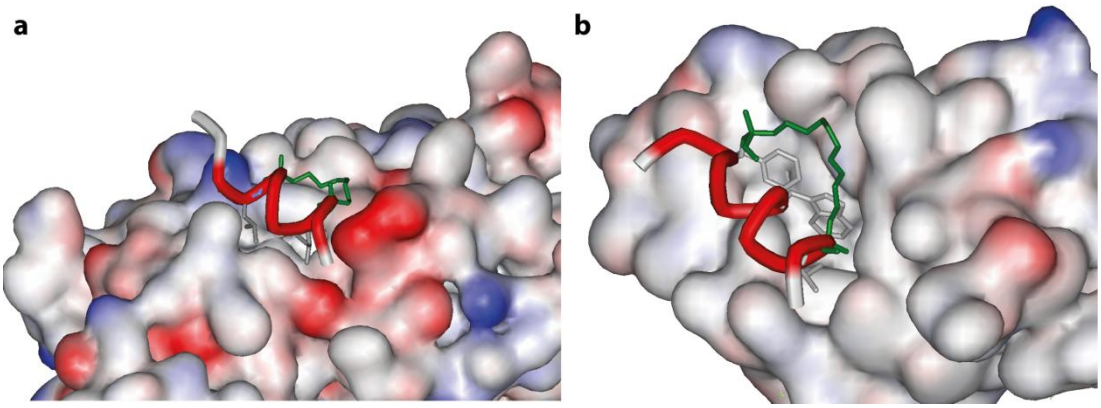


Figure 1.11 Crystal structures of stapled peptides bound to (**a**) *h*DM2, mimicking the p53/*h*DM2 interaction (PDB ID: 3B3V) and (**b**) ERα, mimicking the interaction between ERα and it's coactivators (PDB ID: 2YJA). Staples are shown in green and key binding residues shown in grey.

Finally, Walensky and co-workers recently demonstrated that hydrocarbon double-stapling structurally fortifies longer bioactive peptides, conferring protease resistance *in vitro* and *in vivo* and that oral absorption of such peptides is achievable. These peptides also displayed enhanced antiviral activity against neutralisation-resistant HIV-1 virus through inhibition of gp41 assembly when compared to singly stapled or unmodified peptides.[52] The hydrocarbon stapling approach can thus be considered as one of the main success stories in the development of designed inhibitors of PPIs thus far: Aileron Therapeutics was founded in 2005, and acquired exclusive rights from numerous institutions such as Harvard, Dana-Farber Cancer Institute, Materia and NYU to develop and commercialise a drug discovery pipeline based on stapled peptides. The company have recently completed the first-ever stapled peptide clinical trial for treating endocrine disorders, are due to start clinical trials in 2014 with Roche using an optimised p53 pathway reactivator and have recently secured $30 M in new financing to advace the pipeline of clinical candidates. [53]

1.4.1.4 Hydrogen Bonding Surrogate

The HBS motif presents a covalent linkage in lieu of a native main-chain hydrogen bond.[54, 55] The method established by Arora and co-workers employs ring-closing metathesis on olefin-bearing residues at the *i* and *i* + 4 positions. This method is attractive as recognition features on the helix surface are not encumbered by the constraining element; the cross-link is positioned on the inside of the helix, whereas side chain tethers block at least one face of the putative helix.[56] In addition to this, incorporation of the crucial residues (for lactam bridges, disulfides etc.) may remove important side chain functionalities.[56] A highly helical HBS-peptide of the BAK BH3 α-helix was shown to bind to Bcl-x_L with a K_d value of 69 nM and was 60-fold more resistant to trypsin-mediated proteolysis than the linear analogue,[39] whilst a HBS-helix derived from the *C*-peptide from gp41 was shown to bind to the *N*-terminal hydrophobic pocket and inhibit gp41-mediated cell fusion (IC_{50} = 43 μM) by inhibiting formation of the six-helix bundle.[57] In further work, a HBS-helix of p53 (K_d = 160 nM for p53/*m*DM2 interaction) was shown to target the complex with similar selectivity to that of Nutlin-3 when screened against other proteins which are known to bind helical peptides (including *h*DMX and Bcl-2 family proteins).[58] *In vitro* studies of a HBS-helix of the *C*-terminal transactivation domain of HIF-1α exhibited a K_d value of 420 nM for binding to p300 CH1 domain and down-regulated VEGF transcription.[59] Finally, the Arora group used this approach to target the Ras/SOS interaction. Based on the guanine nucleotide exchange factor SOS, they reported a cell-permeable synthetic α-helix that interferes with the Ras/SOS interaction and down-regulates Ras signalling in response to receptor tyrosine kinase activation.[60] The Arora group has more recently explored the use of a thioether linkage (teHBS; **Fig. 1.12**)) as an alternative to the all hydrocarbon tether and demonstrated

this modification is appropriate by comparison with a conventional HBS helix for inhibition of p53/*m*DM2 (K_d = 224 nM for teHBS and 232 nM for HBS).[61]

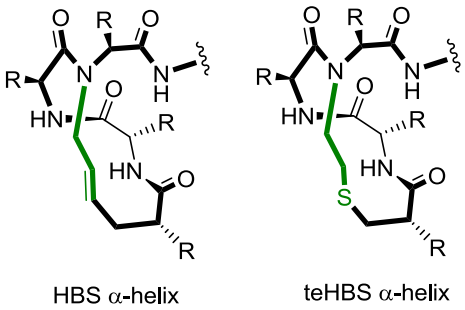


Figure 1.12 Schematics illustrating the different HBS and teHBS approaches.

1.4.1.5 Photocontrolled α-Helices

The above strategies have developed some potent peptide based inhibitors of PPIs, however, the activity of these constrained peptides cannot be controlled. Alleman and co-workers have developed a system in which protein-binding activity can be controlled by an external stimulus in a reversible manner. The approach utilises an azobenzene cross-linker which is introduced *via* cysteine residues at the *i*, *i*+4, *i*+7 or *i*+11 positions. The cross-linker undergoes *cis/trans* isomerisation on irradiation which switches the peptides between random coil-like and α-helical conformations.[62] This strategy has previously been used to control binding of α-helices to DNA but has since been pursued for photocontrollable peptides for PPI inhibition.[63]

Using this azobenzene system, Alleman and co-workers synthesised photocontrollable helices based on the BH3 domain peptides of Bak and Bid to target the anti-apoptotic Bcl-x_L.[64] Cross-links were located at the *i*, *i* + 7 and *i*, *i* + 11 positions in a Bak peptide and *i*, *i* + 4 positions in a Bid peptide. Significant helix stabilisation was observed in peptides when the cross-link was in the *cis* configuration with *i*, *i* + 4 and *i*, *i* + 7 Cys linkages (**Fig. 1.13a**), whilst linkers at *i*, *i* + 11 positions found the *trans* configuration to be helix-stabilising (**Fig. 1.13b**). Peptides in their helix-stabilised configurations displayed high affinities for Bcl-x_L with dissociation constants of 55 nM (Bid), 49 nM (Bak^{i+7}) and 21 nM (Bak^{i+11}), and in some cases cross-linked peptides showed 200-fold selectivity for binding to Bcl-x_L over helix-binding *h*DM2. Helix destabilised forms of Bak (K_d = 825 nM) and Bid (K_d = 1275 nM) peptides were found to bind more weakly to Bcl-x_L. The data demonstrates a unique ability to control activity of constrained peptides and holds potential for studying and modulating cellular function by selectively interfering with PPIs.[64]

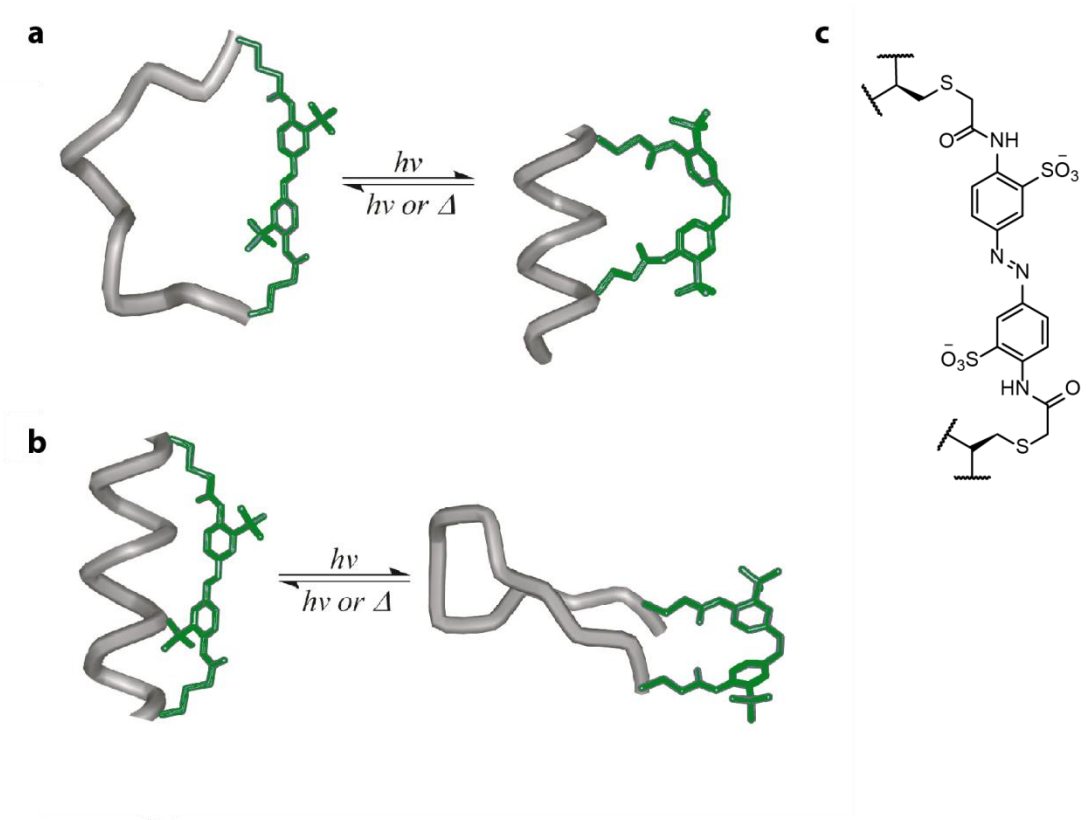


Figure 1.13 Photocontrol of peptide conformational preference with azobenzene cross-linker (green). (**a**) Helical conformation stabilised in the *cis* configuration with *i*, *i* + 4 and *i*, *i* + 7 Cys linkages. (**b**) Helical conformation stabilised in the *trans* configuration with *i*, *i* + 11 Cys linkages. (**c**) Chemical structure of the photocontrollable cross-linker (*trans* configuration).

A range of alternative cross-linkers are now being developed that confer enhanced properties and/or ease of synthesis. Two recent examples exploit biphenyl- and *meta*-xylene-derived cross-links.[65, 66] The former furnished optimised inhibitors of the Mcl-1 protein, whereas the latter provided novel inhibitors of calpain through mimicry of its proteinogenic enzyme inhibitor. Studying the literature, there are contradicting theories as to what factors will increase the inhibitory activity of short constrained peptides other than the linker itself. This includes which amino acids to cross-link and their relative positions, where in the sequence should they be and the length and flexibility of the linkers. Some groups aim to maximise helical content of short peptides for a smaller entropic penalty on binding whilst others assert that this may distort the helical conformation into an inactive form. A less defined conformation is thus considered adequate and a linker is necessary merely to reduce the number of degrees of freedom of the peptide in the unbound state. As such, there is currently not a truly generic technology for mimicking short protein helices with constrained peptides.

1.4.2 Helical Foldamers

The alternative approach for Type I mimetics are foldamers, whose conformation is designed to mimic that of the canonical α-helix. In this generic structure, it is known that intramolecular hydrogen bonding can confer stability or instability; complementary hydrogen bonding donors (amide N-H) and hydrogen bond acceptors (amide C=O) located on non-consecutive residues along the peptide (**Fig. 1.14**) confer conformational specificity and add stability to the secondary structure. In a series of modelling experiments, the Gellman group studied the ability of unnatural β- and γ-amino acids to confer such stabilisation in modified peptides.[67] From this, they postulated that absent nearest neighbour backbone hydrogen bonding, observed in β-peptide derivatives, would induce conformational specificity. This identified β-peptides as promising candidates for helical foldamers.

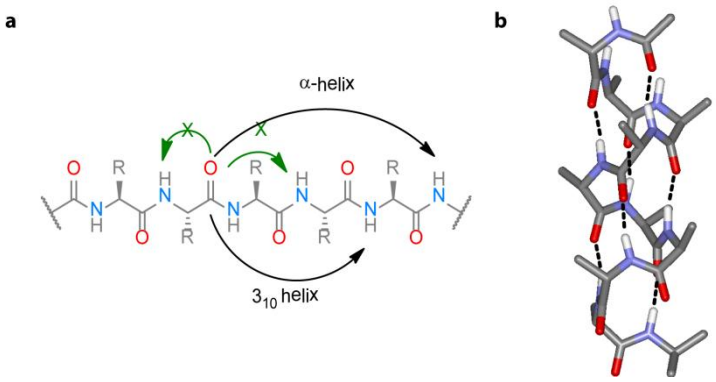


Figure 1.14 (**a**) Schematic of a polyamide backbone displaying favourable (black) and unfavourable (green) hydrogen bonding. (**b**) Intramolecular hydrogen bonding network within the α-helix.

β-peptides are generated from a simple backbone alteration to α-amino acids: addition of a methylene unit. Initially, it was thought that the extra carbon would introduce additional flexibility, reducing its tendency to adopt well defined folded states in solution. Subsequent structural studies illustrated that conformational constraints and substitutions, furnished β-peptides with an increased tendency to fold over α-peptides.[68] Furthermore, Seebach *et al* showed that this 'simple' backbone modification could impart enhanced resistance to proteolysis and more favourable pharmacokinetics to β-peptides.[69] As previously mentioned, long range backbone hydrogen bonding is inherent to helical stability and β-peptides are named after the number of backbone atoms per hydrogen bond ring; 14-helix and 12-helix being the best characterised. When designing mixed α/β-peptides systems, one has to consider a number of features: the constitutional (β^2, β^3, $\beta^{2,3}$ etc) and configurational ((R), (S)) variety of the building blocks, the position of a single β-amino acid in the peptide sequence and the spacing between β-amino acids. This substitution pattern has been shown to affect the peptide folding preferences (helices, β-sheets, hairpins etc) and has

been extensively reviewed.[70] More recently, such systems have been employed in the mimicry of biological relevant peptides and this section will focus of the success of helical foldamers in this respect.

1.4.2.1 β-peptides

The question of biological relevance of these hybrid peptides was first addressed by Seebach *et al*, who demonstrated that small seven-nine residue amphiphathic β-peptides could mimic alipoproteins by mildly inhibiting small-intestinal cholesterol absorption (compared to no activity with α-peptides).[71] They also proved more resistant to proteolysis,[71] demonstrating these 'first-generation' ligands had potential for designing biologically active β-peptides, despite having several inherent differences.[72]

The Schepartz group subsequently designed a β^3 decapeptide which had significant 14-helical stability in aqueous solution due to side chain-side chain salt bridges on one helical face and internal macrodipole stabilisation. Targeting the p53/*h*DM2 interaction, key residues from the activation domain of p53 (p53AD: Phe19, Trp23, Leu26) were strategically positioned within the peptide to align on a single face upon folding. The most potent inhibitor targeted the interaction with an affinity close to that of the native peptide (IC_{50} = 80.0 ± 3.2 μM for the β^3 peptide vs IC_{50} = 2.47 μM for p53AD).[73] Parallel SAR studies found introduction of a 6-chlorotryptophan analogue in place of the Trp side chain improved potency by 10 fold.[74-76] The Schepartz group also applied this approach to inhibit gp41 mediated cell-cell fusion with a β^3-decapeptide. Although less potent than a prescribed HIV fusion inhibitor (IC_{50} = 5.3 μM vs IC_{50} = 0.11 nM), the β^3 peptides are one third of the size, metabolically stable and available for combinatorial optimisation, showing potential for inhibitors of other systems that employ common fusion mechanisms with extended helices.[77]

1.4.2.2 Mixed α/β peptides

In addition to β-peptides, which exclusively contain β-amino acids, another canonical foldamer class is one which contains α- and β-amino acids (initially in a 1:1 alteration) and as such are termed α/β peptides. α/β peptides have been extensively characterised, displaying 4 distinct helices, each having a unique spatial arrangement of side chains (with a given sequence separation) analogous to side chain arrangement found in α-helices.[78-81] Preliminary studies from the Gellman group were made to identify foldameric ligands for the BH3-recognition cleft of Bcl-x_L. Assessing the affinity of over 200 α/β- and β-peptides with different helical conformations towards Bcl-x_L, only α/β-peptides thought to adopt a 14/15-helical secondary structure displayed significant binding.[82] Sequence-affinity relationship profiling ultimately led to the identification of a potent chimeric inhibitor (K_i =

2.2 nM). These were made up of an *N*-terminal α/β-peptide segment of Bak, fused to a *C*-terminal α-peptide segment ((α/β+α)-peptides): only 60% of the Bak peptide could be successfully mimicked by a 'continual' α/β- or β- scaffold.[83] The success of a chimeric approach suggested that mimicry of large binding epitopes may be accomplished by replacement of short segments of the epitope with distinct foldameric scaffolds, decreasing the proportion of α-amino acid residues, with the ultimate goal of metabolic stability without compromising affinity. Subsequent studies produced a crystal structure of the most potent peptide mimicking the Bim-BH3 peptide bound to Bcl-x_L (**Fig. 1.15**). This provided much insight into results from previous binding studies, highlighting the importance of subtle changes in side chain positioning.[84]

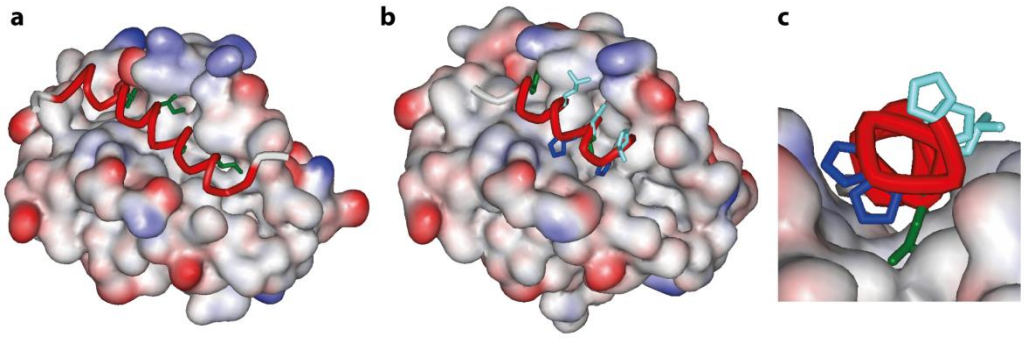


Figure 1.15 (**a**) X-ray crystal structure of pro-apoptotic Bim bound to Bcl-x_L (PDB ID: 3FDL) (**b**) X-ray crystal structure of a chimeric Bim mimetic bound to Bcl-x_L : comparison with **a** shows the foldamer to bind in the same hydrophobic cleft (PDB ID: 3FDM) (**c**) Top view of the α/β foldamer highlighting the interacting α- (green) and β-residues (dark blue) and the solvent exposed β-residues (light blue).

Traditionally, a structure-based approach was the method of choice which involved changing the native helix by incorporating constraints and various side chains to maximise helicity, whilst projecting the key side-chains in the correct orientation. Building on previous studies, the Gellman group adapted their approach by applying a 'sequence-based' design which involved replacing subsets of regularly spaced α-residues with the corresponding β^3-residues. Each replacement introduces an extra methylene unit at regular occurrences, keeping the fundamental sequence of the native helix the same. This was applied to the design of α/β-peptides of Puma (another pro-apoptotic member of the Bcl-2 family), incorporating an ααβαααβ sequence repeat along the peptide, and starting at varying positions along the peptide;[85] peptides with this repeat pattern have been previously shown to adopt a helix-like conformation by crystal structures.[86] From evaluation of seven peptides, a potent inhibitor of Bcl-x_L was identified (K_i = 1 nM) which was also found to bind to Mcl-1(K_i = 150 nM).[85] Co-crystal structures of these sequenced-based designed inhibitors show that the β-residues align on a single face of the helix producing a slight curvature to the peptide. The study revealed how subtle changes in key geometrical features (helix radii,

helix bowing etc) can greatly affect the binding affinity.[87] Subsequent studies were made to evaluate different sequences, including ααβ and αααβ repeat units. The αααβ pattern produced four low-nanomolar inhibitors of the Bcl-2 anti-apoptotic proteins based on the Bim-BH3 pro-apoptotic peptide, inducing cytochrome *c* release in wild-type mouse embryonic fibroblast (MEF) extracts.[88]

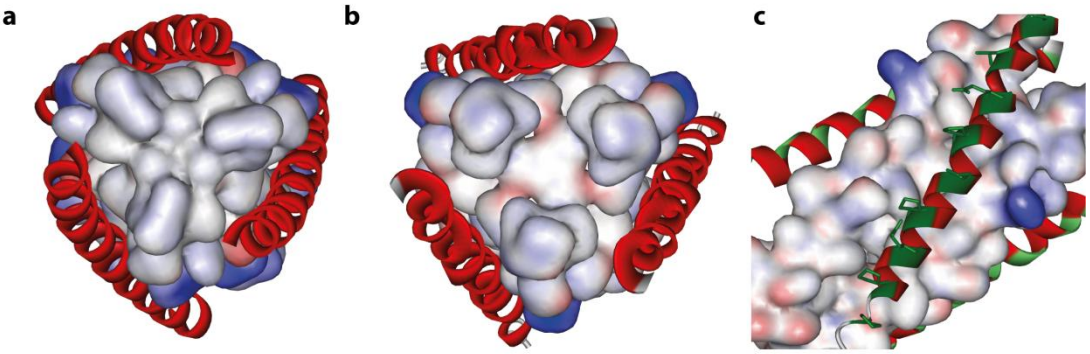


Figure 1.16 (**a**) Top view of the native gp41 six helix bundle, (**b**) Top and (**c**) lateral views of the six helix bundle formed by α/β-CHR foldamers and native α-NHR peptides. β-residues are represented in green **c** highlighting the single face positioning of the residues.

The basics of the sequence-based approach were applied to design α/β-peptide mimics of the long HIV membrane protein gp41.[89] This involved a two-step process: β^3-residue insertion into the α-sequence of the CHR domain of gp41 forming ααβαααβ repeats and then systematically changing selected β^3-residues with cyclically constrained β-residues to rigidify the backbone. As the α/β-peptide has to be considerably longer than those previously designed, the second step was necessary to overcome the entropic penalty associated with pre-organising longer oligomers. Crystallographic data showed the α/β-CHR peptides form a near identical six helix bundle with the NHR domain as the native CHR sequence (**Fig. 1.16a,b**), and cell fusion and virus-infectivity assays indicate the α/β-peptides effectively block HIV cell-fusion (K_i = 9 nM, ~380-fold improvement on the analogous acyclic α/β-peptide). Furthermore, antiviral activity was demonstrated to be comparable to that of the native α-peptide with ~280 fold enhanced resistance to proteolysis.[89] As demonstrated in **Figure 1.16c**, the β-residues in an ααβαααβ based peptide are found along a single face of the helix. Gellman *et al* exploited this feature to impart stability to the gp41 α/β-peptide mimics. Inserting β- analogues of acid/base residue pairs in appropriate positions so as not to remove key functionality, stabilisation was achieved *via* side chain ion pairing at the *i* and *i* + 3 and/or *i* and *i* + 4 positions, replacing cyclic constraints.[90]

1.5 Type III Mimetics

Although a number of amino acids are involved in constructing the highly ordered, helical motif, it is only certain residues which are responsible for the interaction with its target protein. *Proteomimetics* strive to replicate the topography of the native helix by mimicking the spatial projection of key binding residues, rather than recapitulating the entire helical motif. In doing so, molecular weight is reduced and biostability is enhanced whilst retaining the important binding features of the α-helix: or the helix pharmacophore.[91] Currently, there are numerous proteomimetics designed with a sufficiently rigid scaffold to help project side chains, analogous to those of the α-helix, with the correct spatial orientation. This is important for two reasons (i) mimicking the 3D orientation of the α-helix side chains provides correct binding conformation; residues have the correct orientation and distances required for efficient binding[92] and (ii) less entropy is paid on binding due to the rigidity of the scaffold allowing for higher affinity binding to the protein. Ideally, however, the scaffold must have some flexibility to be able to bind to the target protein and adopt the best conformation for binding.

1.5.1 Proof-of-Concept and Early Design Strategies

The template selected for one of the first attempted α-helix mimetics was a 1,1,6-trisubstituted indane.[92] It is a relatively rigid template allowing the spatial orientation of the substituents to be predicted. Molecular modelling of the indane template showed a good overlay exists between the 1- and 6- position to the Cα and Cβ bonds of the *i*, *i*+1 residues of the α-helix. The second substituent on the 1-position of the indane also corresponds to the *i*-1 residue. Willems and co-workers reported that derivatives with large hydrophobic side chains (Phe-Phe and Trp-Phe at the 1 and 6 positions) showed micromolar affinities similar to the parent dipeptides in binding to the tachykinin receptors NK1, NK2 and NK3.[92] Further work found that 1,1,6-trisubstituted indanes showed similar affinities to tachykinin receptors and other neuropeptide targets. [93] Due to the small size of indanes (**Fig. 1.17a**), they are limited to mimicking two successive amino acids. As a result, they did not prove effective at inhibiting α-helix mediated PPI's, however, they demonstrated that small molecules were able to project functionality in suitable spatial orientations for binding to a target protein and many larger scaffolds have since been designed.[4]

Development of the terphenyl α-helix originated from a biphenyl molecule. Similar to several other molecules such as allenes, alkylidenes cycloalkanes and spiranes, these structures contain a helical twist. However, when these compounds are superimposed onto a polyalanine helix, only the biphenyl skeleton complies with the requirements of an α-helical conformation - for example the disposition of the amino acid side chain and twist angle.[94] In

addition to this, the biphenyl unit is contained within 2.1 % of all reference drug molecules indicating that this structure has favourable bioactive properties. This is attributed to the flexibility, size and shape of the unit allowing it to bind to a wide variety of protein pockets.

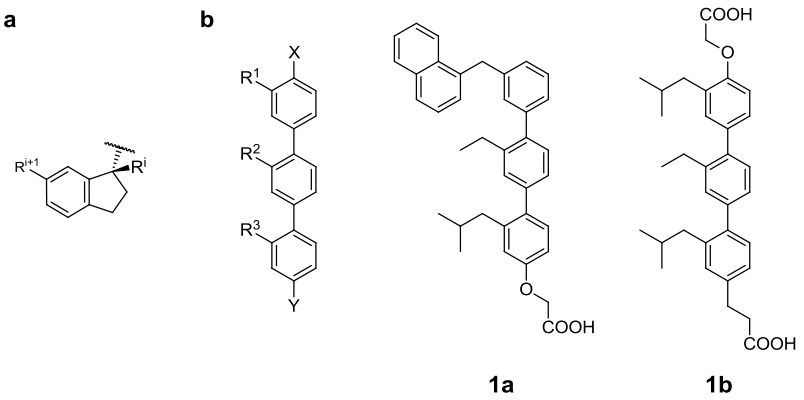


Figure 1.17 (**a**) The indane scaffold: proof-of-concept for peptide mimetics. (**b**) The terphenyl scaffold: terphenyl derivative inhibitors of the CaM/smMLCK **1a** and Bak/Bcl-x_L **1b** interactions.

Hamilton and co-workers were the first to report the use of the terphenyl backbone as an α-helix mimetic.[95] The tris-functionalised 3,2',2''-terphenyl scaffold (**Fig. 1.17b**) was an attractive template for the initial design due to the simplicity of the structure; it was also an entirely non-peptidic scaffold which could be synthesised in an iterative fashion. The aryl core assumes a staggered conformation enabling functional side chains to be projected with similar distances and angular relationships to those located in α-helices. The tris-*ortho*-substituted terphenyl was found to mimic the *i*, *i*+4, and *i*+7 residues by adopting a staggered conformation, representing two turns of an α-helix. Libraries of compounds, incorporating various functionalised side chains, were subsequently synthesised.

Initial tests using terphenyl derivatives showed them to be antagonists for the interaction between CaM and an α-helical domain on smMLCK. Derivative **1a** competes with smMLCK with an IC_{50} of 9 nM.[95] In subsequent studies, CD was used to determine the ability of terphenyl derivative **1b** to target the HIV-1 transmembrane envelope glycoprotein, gp41, and disrupt the assembly of a fusion-competent hexameric core. A dye-transfer cell fusion assay confirmed inhibition of the HIV-1 mediated fusion exhibiting an IC_{50} of 15.70 ± 1.30 μM.[96] Further studies involving fluorescence polarisation assays have shown terpheny**l** derivatives to be capable of inhibiting the p53/*h*DM2[97] and Bcl-x_L-Bak[98] PPIs. Bcl-x_L-Bak selectivity over p53/*h*DM2 was achieved by a subtle exchange of a methyl-1-naphthyl for methyl-2-naphthyl side chain, demonstrating the ability to selectively modulate different complexes.[98] Terphenyls were also shown to be active in intact cells (HEK293): inducing apoptosis by the intended mechanism by disrupting BH3-mediated interactions with Bcl-x_L.[99]

1.5.2 Second Generation Design Strategies

Difficulties in the synthesis and solubility of terphenyl due to the hydrophobic template inspired Hamilton and co-workers to develop a new scaffold based on a trispyridylamide foldamer (**Fig. 1.18a**).[100] The polyamide backbone adopts a planar conformation with all alkoxy side chains projected on the same face of the molecule (*syn*) and tilted at 45° to maximise interactions of the lone pair on the oxygen atom with the amide NH group. The observed geometry results from a bifurcated hydrogen bonding network between the amide proton and the pyridyl nitrogen and ether oxygen. Electrostatic clashes between the amide carbonyl and the pyridine nitrogen in the *anti* conformation further favours the observed *syn* geometry. Overlaying the scaffold with a polyalanine α-helix reveals a close resemblance between side chains and the *i*, *i*+4 and *i*+7 residues and a modular synthesis facilitated the generation of a small library of compounds. Fluorescence anisotropy assays demonstrated derivative **2** to inhibit the Bcl-x_L-Bak complex (K_i = 1.6 μM). [100]

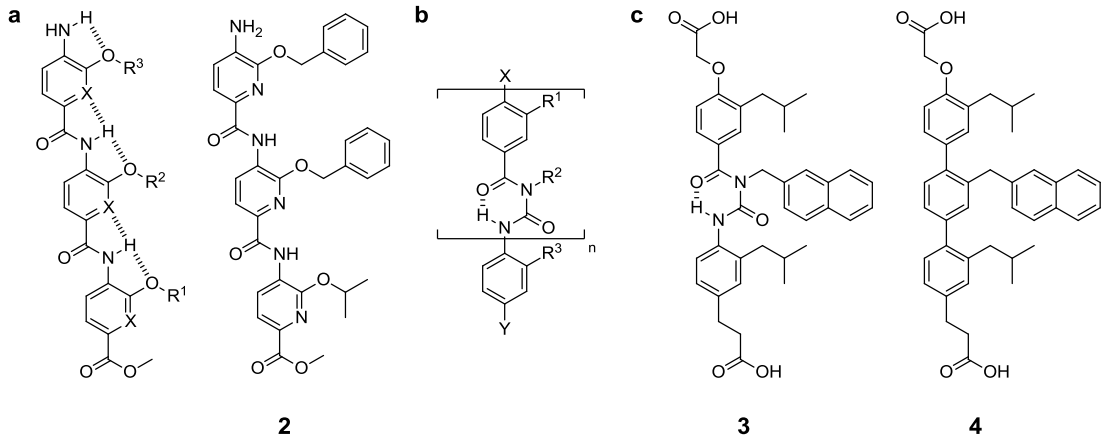


Figure 1.18 (**a**) The trispyridylamide scaffold: trispyridylamide derivative inhibitor of the Bcl-x_L-Bak complex **2**. (**b**) The benzoylurea scaffold. (**c**) Benzoylurea **3** and terphenyl **4** isosteric inhibitors of the Bcl-xL-Bak interaction.

To date, most reports on synthetic inhibitors of PPIs have focused on mimicking small structural domains, most commonly up to two or three turns of the helix. Longer α-helices, however, are frequently found in higher-ordered structures such as helix bundles, coiled coils and transmembrane domains of proteins and play a critical structural and functional role. The Hamilton group designed a benzoylurea scaffold (**Fig. 1.18b**) to mimic extended regions of an α-helix.[101] The central aromatic ring on the terphenyl scaffold is replaced with an acylurea motif: a foldameric, aromatic ring isostere. A modular synthesis provides easy incorporation of natural and unnatural amino acid side chains, installing desired recognition properties with reasonable synthetic effort. This strategy achieved significantly elongated helix mimetics with four to five benzoylurea subunits spanning

lengths of up to 37.1 Å, corresponding to helices of approximately seven turns (30 amino acids). To test the validity of the benzoylurea scaffold against α-helix mediated PPIs, derivative **3** was synthesised, an isostere of the most potent terphenyl inhibitor **4**[98] of the Bcl-xL-Bak interaction (**Fig. 1.18c**). The benzoylurea **3** exhibited comparable inhibition to that of terphenyl **4** (K_i = 2.4 μM and K_i = 114 nM respectively) demonstrating the potential for this scaffold as an α-helical mimetic.[98]

Further work reported by the Hamilton group focused on novel terephthalamide and 4,4-dicarboxamide scaffolds (**Fig. 1.19a,b**).[102, 103] This was in search for compounds with more versatile syntheses and better physical properties than those possessed by terphenyl compounds. In the terephthalamide design, two carboxamide groups have been inserted in place of flanking phenyl rings on the terphenyl helix. This restricts rotation through intramolecular hydrogen bonds between the amide NHs and the alkoxy oxygen atom, thereby influencing the position of the amino acid side chain. The dicarboxamide scaffold combines the hydrophobic core of the oligophenyl scaffold with the accessible carboxamide groups of the terephthalamides. Inhibitory activities of derivatives of both scaffolds were tested against the Bcl-x_L/Bak interaction (K_i = 0.78 μM **5a**; 1.8 μM **6**). Further studies demonstrated that treatment of human HEK 293 cells with terephthalamide derivative **5b** resulted in disruption of the same interaction in whole cells with an IC_{50} of 35 μM.[102] Despite 4,4-dicarboxamides displaying the 'required' 4 key residues, improved potency over terephthalamides was not always seen. This raises the fundamental question of what is actually required to effectively target PPIs.

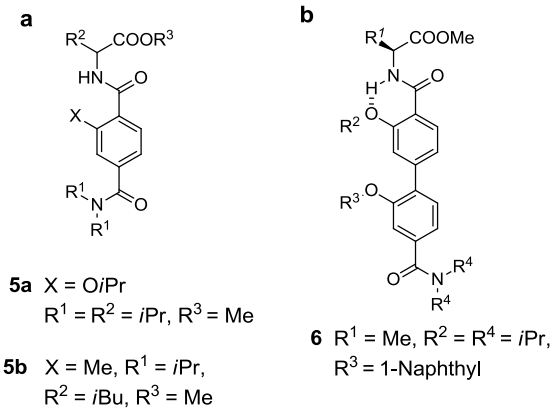


Figure 1.19 (**a**) The terephthalamide scaffold: terephthalamide derivative inhibitors of the Bcl-x_L/Bak interaction **5a** and **5b**. (**b**) The 4,4-dicarboxamide scaffold: 4,4-dicarboxamide derivative inhibitors of the Bcl-x_L/Bak interaction **6**.

1.5.3 Amphiphilic α-Helix Mimetics

Rebek and co-workers developed amphiphilic α-helix mimetics based on a pyridazine scaffold (**Fig 1.20a**).[104] Inspired by Hamilton's success with the terphenyl scaffold, they set out to develop a structurally similar compound, featuring more hydrophilic components with a synthetically easier route. The scaffold consists of a series of oxazole-pyridazine-piperazine rings, that present both a hydrophobic surface for recognition onto the protein surface and a "wet edge" which is rich in hydrogen bonding motives. The hydrophilic backbone ensures that during complexation with its target, the "wet edge" remains directed towards the solvent, and it was anticipated to exhibit increased water solubility through protonation of the basic piperazine ring at physiological pH. The binding affinity of the scaffold and similar derivatives for Bcl-x_L was determined by a fluorescence anisotropy assay. Results suggest that the presence of the cationic ammonium group may cause unfavourable interactions on binding to the receptor. A wide range of compounds based on the pyridazine scaffold were analysed, however, they do not approach the affinity demonstrated by the terphenyl compounds.[105]

With a similar objective, the 5-6-5-imidazole–phenyl–thiazole scaffold (**Fig 1.20b**) was designed by Hamilton and co-workers in which the terminal positions of the original terphenyl are replaced with more hydrophilic five-membered heterocycles.[106] The interaction between Dbs and Cdc42 regulates the resistance of cancer cells to cytotoxic therapies. Derivative **7** was designed to mimic the key binding region on Dbs, and a fluorescence assay demonstrated it to disrupt the interaction with IC_{50} of 67.0 μM.

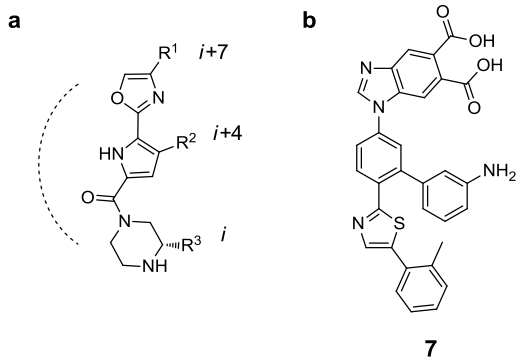


Figure 1.20 Amphiphilic helix mimetics: (**a**) the pyridazine scaffold and (**b**) a 5-6-5-imidazole–phenyl–thiazole scaffold **7** based on key binding region on Dbs.

1.5.4 Backbones with Chirality

König and co-workers reported on the synthesis of an inherently chiral 1,4-dipiperazino benzene scaffold (**Fig. 1.21a**).[107] This scaffold is another development on the terphenyl scaffold in which the relative orientation of the key side chain functionalities

remain the same as those in terphenyl-type helix mimetics, substituting the two outer phenyl rings for piperazine rings (also improving their water solubility). X-ray crystallography revealed 1,4-dipiperazino to display side chains in a similar geometrical arrangement (*i*, *i*+3 and *i*+7 residues) to an idealised α-helical structure. The subject of chirality was also addressed by the Arora group. They developed a chiral oligooxopiperizine scaffold (**Fig 1.21b**) after postulating that scaffolds containing chiral backbones may be more effective in discriminating between chiral protein pockets. Molecular modelling of the scaffold revealed it to have an inherent helical structure, confirmed by CD studies, and predicted the side chains to mimic residues at the *i*, *i*+4 and *i*+7 positions.[108] The chiral nature of these compounds will allow investigation into stereochemical aspects of protein - α-helix mimetic recognition. However, no reports on inhibition of PPIs have been described for any of these scaffolds.

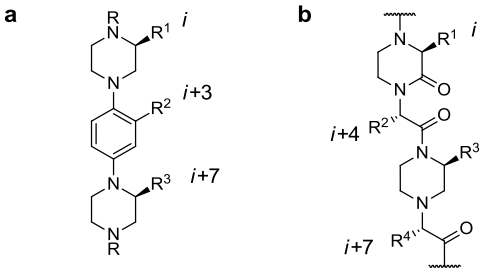


Figure 1.21 Inherently chiral scaffolds: **(a)** the 1,4-dipiperazino benzene scaffold, (**b**) the oligooxopiperizine scaffold.

1.5.5 Multi-facial Inhibitors

The design of many mimetics described thus far, have focused on displaying key functionality on a single face. However, approximately 40% of all multiprotein complexes that occur through a helix, have key residues displayed on two or more faces.[15] Such interactions are found in the ER/coregulator complexes. Careful analysis of a coactivator protein by Katzenellenbogen *et al* revealed that the three leucine residues in the LXXLL motif are on the edge of an equilateral triangle, and so one design criteria for ER inhibitors, was to have a central core with appropriate substituents attached.[109] Several scaffolds (triazene, pyrimidine, trithiane and cyclohexane) targetting this interaction were synthesised based on molecular modelling and docking studies. A series of fluorescence anisotropy competition assays found compounds from the pyrimidine (**Fig. 1.22a**) series to be the most successful: pyrimidine **8a** showed the highest affinity with a K_i of 29 μM. [109] Similar design criterion can be applied to AR/coregulator complexes. AR coactivators also bind through the consensus LXXLL sequence in addition to sequences containing multiple phenylalanine or tryptophan residues (WXXLF, FXXLF etc.).[110] Further studies from the Katzenellenbogen

group reflect the size differences in ER and AR binding pockets. Selective pyrimidine-core inhibitors were achieved by varying the size / length of hydrophobic side chains: ER/coactivator inhibition IC_{50} = 7.9 μM **8b**; no binding **8c**, AR/coactivator inhibition IC_{50} = no binding **8b**; 1.5 μM **8c**.[111]

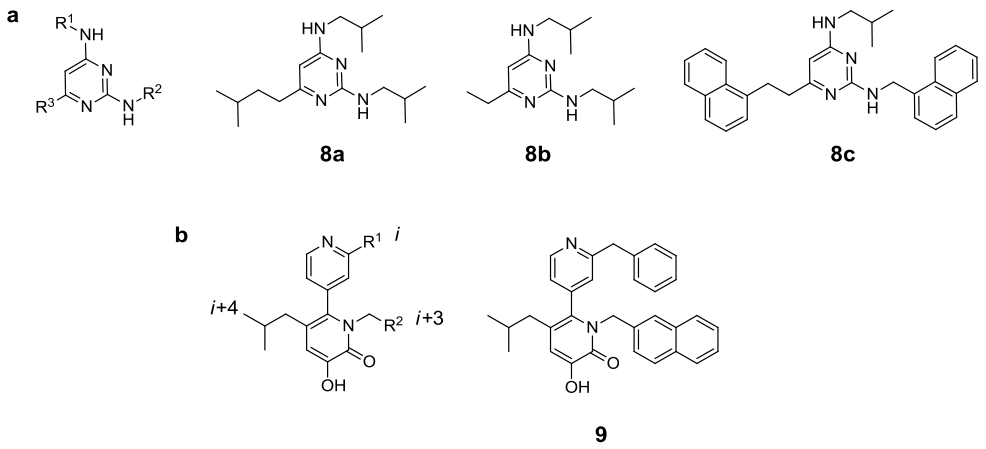


Figure 1.22 (**a**) The pyrimidine scaffold: pyrimidine derivative inhibitors of steroid hormone/coactivator interactions (**8a-c**). (**b**) The pyridylpyridone scaffold: pyridylpyridone derivative inhibitor of the ER/coactivator interaction **9**.

Designing an inhibitor of the LXXLL motif requires a molecule which mimics the distances and angular projections of the *i*, *i*+3 and *i*+4 residues on the helix (i.e. consecutive residues). Both the Willems[93] and Hamilton groups[112] have demonstrated that placing a second *ortho* substituent on a bis-heteroarly scaffold can achieve this. With this is mind, the Hamilton group designed an α-helix mimetic for inhibition of the ER/coregulator complex - the pyridylpyridone scaffold (**Fig. 1.22b**).[113] The indane-based scaffold was chosen as it has improved water solubility and bioavailability as well as a synthesis which allowed for easy introduction of substituents into the 2-pyridyl and 1,5-pyridone positions. An X-ray crystal structure of a pyridylpyridone derivative showed that it adopts a conformation that should effectively mimic the required residues (*i*, *i*+3 and *i*+4) of the GRIP1 coactivator peptide whilst in the solid state. Further investigation *via* fluorescence anisotropy found that these molecules inhibit the interaction with low micromolar inhibition constants. Pyridylpyridone **9** exhibited the best binding (K_i = 4.2 μM) which is comparable to that of the control SRC-1 NRII peptide.[113]

1.5.6 Oligobenzamide Based Inhibitors

Library generation of designed helix mimetics would allow for thorough investigation into the binding region on any given PPI. Oligobenzamides are attracting due attention as proteomimetic scaffolds due to a synthetic accessibility *via* amide bond formation, an ability to incorporate a variety of functionalities mimicking natural and unnatural moieties and a predictable conformation. After initial molecular modelling studies

on a 3-*O*-alkylated oligoamide scaffold (**Fig. 1.23**) to mimic residues on an idealised helix (*i*, *i*+4 and *i*+7 positions),[114, 115] the Wilson group demonstrated derivatives of the scaffold act as inhibitors of the p53/*h*DM2 interaction: **10** (IC_{50} = 1.0 μM) was comparable in potency to the native peptide (IC_{50} = 1.2 μM).[116] Boger and co-workers constructed a large library (>8000) of helix mimetics containing tris-, bis- and monobenzamide scaffolds: the 'missing' benzoic acid building blocks being replaced with natural amino acids. Variations of the mimetics contained aniline, nitro or boc groups at the *N*-terminus and acid or methyl ester groups at the *C*-terminus. Compounds in this study failed to exhibit notable affinity for the p53/*h*DM2,[117] however, later studies with monobenzamide derivatives (**Fig. 1.23e**) report the inhibition of gp41 assembly with micromolar affinity (K_i = 0.6-1.3 μM) and effective activity in a cell-cell assay (IC_{50} = 5-8 μM).[118] The Wilson group later modified the original scaffold, producing a 2-*O*-alkylated oligoamide scaffold. Stronger intramolecular hydrogen bonding resulted in a backbone with reduced curvature; derivatives of this series, however, exhibited comparable affinity to the regiosomeric 3-*O*-alkylated analogues.[119]

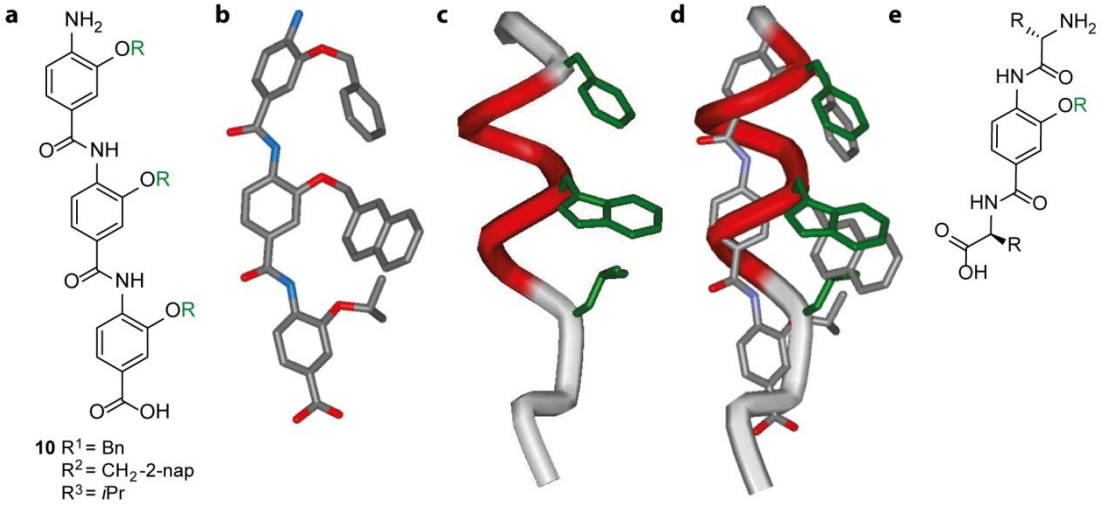


Figure 1.23 The 3-*O*-alkylated oligobenzamide scaffold: (**a**) chemical structure of the scaffold (**b**) low energy conformation of the scaffold containing suitable side chains to target the p53/*h*DM2 interaction, (**c**) p53 helix displaying the key binding residues, (**d**) overlay of scaffold with p53 demonstrating a good geometrical match, (**e**) derivative monomer scaffold targeting gp41 assembly.

After a structure-based computational design to identify p53/*h*DM2 inhibitors, the Guy group reported on a similar scaffold in which side chains are attached directly onto the ring and lacking terminal aniline / acid moities (**Fig. 1.24a**). A library of 173 compounds was synthesised (including dimer intermediates) (**Fig. 1.24b**) using solution phase parallel chemistry and a series of fluorescence polarisation assays identified several low micromolar inhibitors: most potent inhibitor **11** achieving a K_i of 12 μM.[120] A library of hybrid dimers of

this scaffold, with a pyridyl building block (**Fig 1.24c**) was constructucted by Craik and co-workers to target the dimerisation interaction of Kaposi's sarcoma-associated herpesvirus protease. A fluoregenic activity assay screening >180 compounds identified low micromolar inhibitors: **12** displaying the most potent inhibition with an IC_{50} of 3.1 μM. The Wilson group also applied an *N*-alkylated oligobenzamide scaffold (**Fig. 1.24d**) in the mimicry of α-helices based on observations made from the Rebek group. Rebek *et al* demonstrated that the intrinsic preference for the *cis* geometry can be inverted in self-assembled capsules through non-covalent interactions in the *trans* conformation and so it was hypothesised that the *N*-alkylated scaffold may adopt a similar extended conformation.[121] The initial report by Wilson and co-workers represented the first solid-phase synthesis of α-helix mimetics and showed them to inhibit the p53/*h*DM2 interaction with low micromolar affinity (IC_{50} = 2.8 - 4.1 μM).[122] Subsequent studies adapted the synthesis for a microwave assisted solid phase procedure, incorporating a range of natural and unnatural functionalities / side chains.[123]

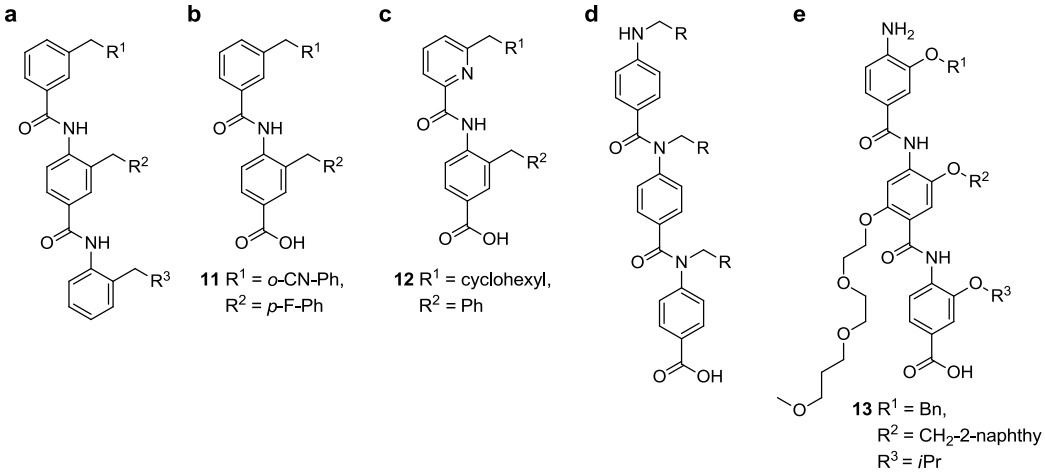


Figure 1.24 Oligobenzamide based scaffolds: (**a-c**) scaffolds alkylated directly onto the benzene ring, (**d**) *N*-alkylated oligobenzamide scaffold, (**e**) 3-*O*-alklated oligobenzamide derivative containing a wed edge.

A novel bis-benzamide scaffold was designed by Ahn and co-workers to address the amphiphilic nature of many native α-helices. The scaffold places side chains mimicking the *i* and *i*+7 residues on one face, and side chains mimicking the *i*+2 and *i*+5 residues on a second face: producing either amphiphlic or a potential bifacial mimic.[124] The Wilson group elaborated on this approach to improve the solubility of the 3-*O*-alkylated scaffold: synthesising 3,6-*O*-dialkylated building blocks displaying a hydrophobic side chain from the 3- position and a hydrophilic glycol chain from the 6- position (**Fig. 1.24e**).[125] **13** showed comparable efficiency in a fluorescence anisotropy assay, targeting the p53/*h*DM2 interaction, to the original non-functionalised analogue **10**, with an IC_{50} = 7.54 μM but with greatly improved solubility. The solubilising group was therefore shown to have little impact

on binding affinity, indicating improvements in helix mimetic properties could be achieved with this type of orthogonal functionalisation.[125]

1.6 Project Aims

Several strategies have evolved over the last 10-15 years in the field of PPI mediation. Although some of the most potent inhibitors so far have been 'small molecules', these are specifically tailored for a given PPI and are therefore considered to lack versatility. Several general features of the α-helix, however, provide a template for inhibitor design from which generic approaches might be devised to target a wide range of PPIs. Constrained peptides, foldamers and proteomimetics are classed as designed inhibitors, and although a true generic approach is still sought, each strategy displays many unique features and offers valuable knowledge into the important features of the protein-protein interface.

Previously in the Wilson group, the 3-*O*-alkylated oligobenzamide scaffold has been synthesised using a solution phase synthesis employing monomers with a variety of *O*-alkyl substituents, namely a variety of hydrophobic side chains to target the p53/*h*DM2 interaction:[114] fluorescence anisotropy assays have shown some of these to act as micromolar inhibitors.[116] The work described herein will build on this work, reporting on the development of a methodology which will allow for an efficient solid-phase synthesis of libraries of compounds based on the aforementioned scaffold. A variety of natural and unnatural amino acid side chains will be incorporated through Fmoc protected monomer building blocks to screen in fluorescence competition assays against multiple PPI targets. Further work will discuss the development of potential ER/coactivator inhibitors *via* molecular modelling and docking studies, structural characterisation and biophysical experiments on a bisbenzamide scaffold containing modified 3-*O*-alkylated building blocks.

1.7 Oligoamide Naming Protocol

The following chapters contain several scaffolds based on an *O*-alkylated benzamide structure. In order to differentiate between them with ease, including intermediate building blocks, a generic naming protocol has been devised. The following list describes the naming patterns adopted throughout and scaffolds **a**-**h** can be found in **Figure 1.25**:

- The 3-*O*-alkylated oligobenzamide scaffold **a** can be thought as having a '*3HABA*' scaffold derived from **3**-**h**ydroxy**a**mino**b**enzoic **a**cid
- The isomeric 2-*O*-alkylated oligobenzamide scaffold **b** can be thought of as having a '*2HABA*' scaffold derived from **2**-**h**ydroxy**a**mino**b**enzoic **a**cid

- The synthesis of the scaffolds can be illustrated as using '*3HABA based'* building blocks or can be more specifically referred to as:
    - '*3HNBA*' building blocks (**c**) for **3-h**ydroxy**n**itro**b**enzoic **a**cid
    - '*3AHB*' building blocks (**d**) for methyl-**3-a**mino**h**ydroxy**b**enzoate
    - Analogous naming is applied to 2HABA based building blocks
- Dimer **e** can be thought as having a '*DHABA*' scaffold derived from **3,6-di**hydroxy**a**mino**b**enzoic **a**cid
- Dimer **f** is a hybrid containing both '*3HABA*' and '*DHABA*' based building blocks
- The synthesis of the scaffolds can be illustrated as using '*DHABA based'* building blocks or can be more specifically referred to as:
    - '*DHNBA*' building blocks (**g**) for 3-**di**hydroxy**n**itro**b**enzoic **a**cid
    - '*DAHB*' building blocks (**f**) for methyl-3-**a**mino**di**hydroxy**b**enzoate
    - Hybrid oligomers can be expressed such as *DHABA:3HABA* (**e**)

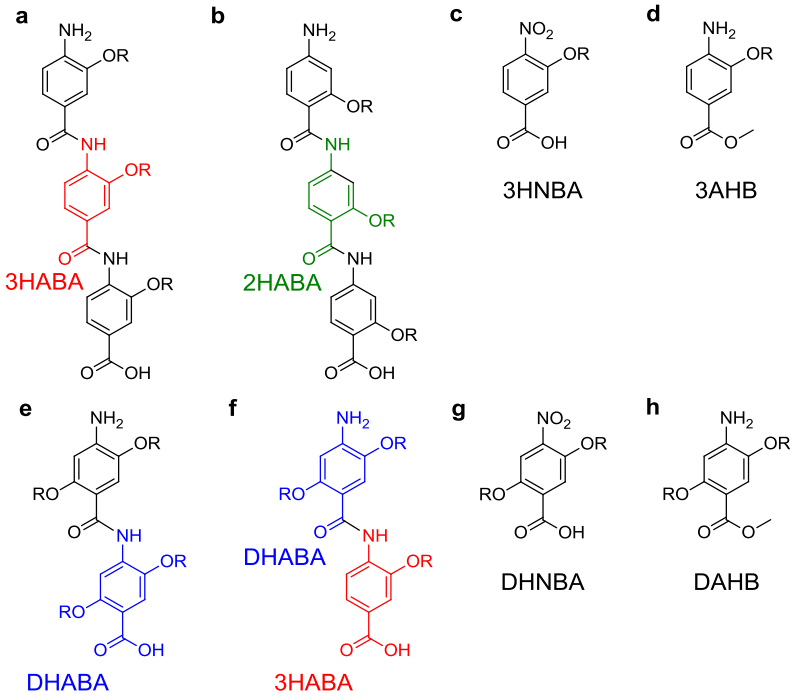


Figure 1.25 Structures of scaffolds and building blocks referred to in the naming protocol: (**a**) 3HABA based scaffold, (**b**) 2HABA based scaffold, (**c**) 3HNBA building block (3HABA based), (**d**) 3AHB building block (3HABA based), (**e**) DHABA based scaffold, (**f**) hybrid DHABA:3HABA dimer scaffold, (**g**) DHNBA building block (DHABA based), (**h**) DAHB building block (DHABA based): n.b. some acronyms are not highlighted in this thesis but should be applied in future work

1.7.1 Numbering System

To simplify the numbering and NMR assignment of our trimers, a sequential nomenclature numbering system has been devised, where each of the monomer building blocks is considered separately **(Fig. 1.26)**. Assignment of the compounds is as follows; naming proceeds from *N* to *C* terminus where each 3HABA residue is assigned a number with respect to its position on the chain and each individual residue numbered using the standard system. Side chain assignment follows a peptide nomenclature pattern in which the carbon attached to the alkoxy oxygen is assigned as Cα, and numbering of further aliphatic parts continue with Cβ, etc. In the case of aromatic side chains, the aromatic carbons are numbered CAr1, CAr2, etc. The numbering of the protons is based on the carbon numbering system. The monomer number is added as a prefix to the individual carbon / proton number for differentiation. Examples of oligomers are given below however monomer intermediates follow the same assignment. Protons in the Fmoc structure are differentiated by the prefix F; protons from the CH_2 group are numbered FHα, the neighbouring CH will be FHβ, and the aromatic protons are FH2 to FH5.

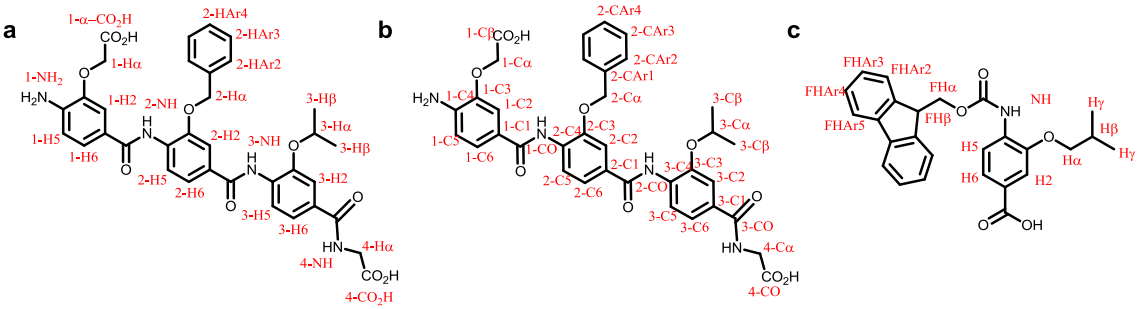


Figure 1.26 Oligomer numbering system: a) Shows proton numbering of oligomers, b) shows carbon numbering of oligomers, c) shows proton numbering of monomer building blocks.

- Naming of compounds in the experimental, using **a** (**Fig. 1.25**) as an example, will be found as follows: **NH_2-[*O*-CH_2-CO_2H-(3-HABA)]-[*O*-Bn-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H**
    - Each square bracket represents a single building block and provides information on the type of building block (3HABA) and its corresponding side chain
    - A nitro derivative will have a **NO_2-** prefix
    - A ester derivative will have a **-CO_2Me** suffix

Chapter 2

Solid Phase Methodology for Synthesis of O-Alkylated Aromatic Oligoamide Inhibitors of α-helix Mediated Protein-Protein Interactions

This chapter is adapted from the research article published in Chem. Eur. J.[126] and contributed to work in a research article published in Org. Biomol. Chem.[119]

2.1 Introduction

A major effort in modern bio-organic chemistry focuses on the design, synthesis and structural characterisation of *foldamers*:[68] non-natural oligomers that adopt well-defined secondary, tertiary and quaternary structures.[127-130] One ultimate objective of such studies is to recapitulate the functional behaviour of biomacromolecules.[131] Particular emphasis has been placed on inhibitors[60, 82, 88, 89, 132, 133] of α-helix mediated[15] protein-protein interactions[134] – an endeavour which in its own right represents a major challenge.[135, 136] The development of synthetic methodology that allows access to small-to-medium sized libraries of foldamers incorporating diverse side chains represents the cornerstone upon which such studies are pursued. In this regard, it is noteworthy that the most robust methodology exists for peptoids,[137] β-peptides[138] and more recently oligoureas;[139] templates that have seen the most significant use in a biological context.[82, 140] The Wilson group[114, 116, 119, 122, 123] and others[117, 124, 141-143] have recently reported on the use of aromatic oligoamides[130] as potential α-helix mimetics.[144, 145] Rather than topographical mimicry of the α-helix (as is the case for β,[75] α/β[82, 88, 89] and other foldamers[132]), these compounds mimic an α-helix by presenting key side chains from a rod-like template in a spatial orientation that matches that of the α-helix (**Figure 2.1**).[146]

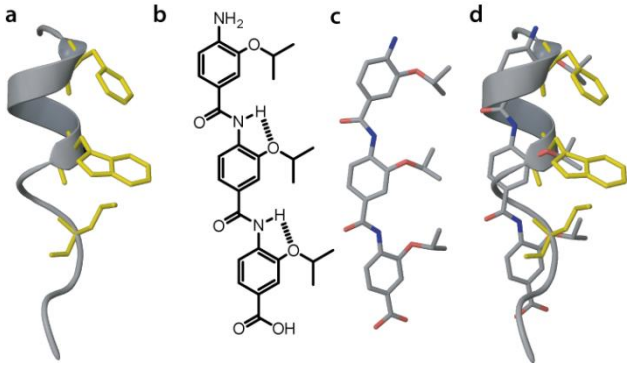


Figure 2.1 (**a**) Idealised α-helix (taken from PDB ID: 1YCR) with *i*, *i* + 4 and *i* + 7 side chains highlighted. (**b**) Chemical structure of 3HABA helix mimetic. (**c**) Minimised structure of a helix mimetic with $R^1 = R^2 = R^3 = i$Pr. (**d**) Idealised α-helix superimposed onto minimised aromatic oligoamide.

The 3-*O*-alkylated oligoamide (3HABA) scaffold (**Fig. 2.1b,c**) was evaluated for α–helix mimicry by comparison with the p53 transactivation domain (PDB ID: 1YCR) in which three side chains – Phe19, Trp23 and Leu26 – are shown to play a key role at the p53/*h*DM2 interface.[119] Of the structures within 1.5 kJ mol^{-1} of the lowest energy conformation, all were observed to position side chains on one face, indicating the desired conformation should be accessible. The RMSD was calculated on the agreement between the

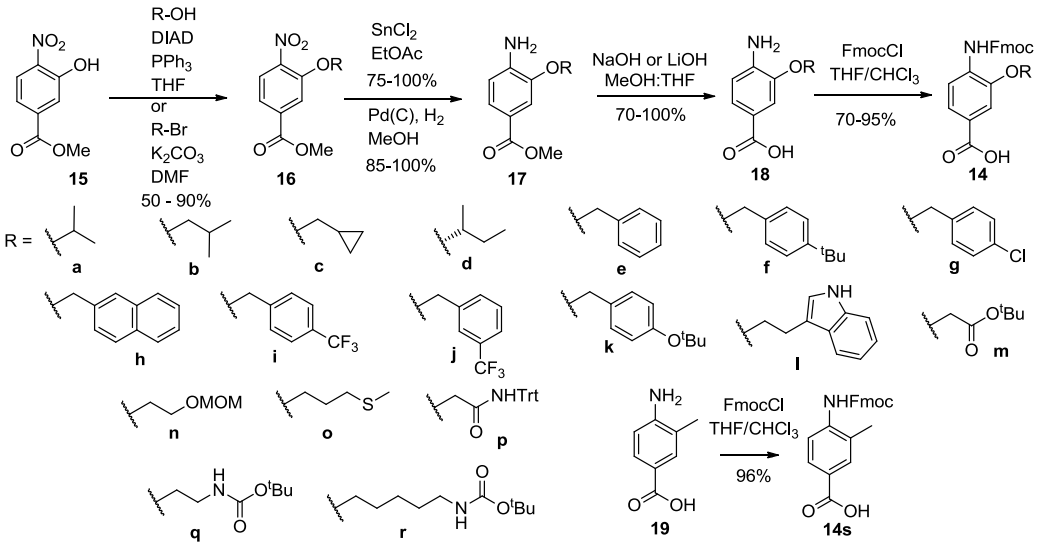
α carbon of the peptide and the alkoxy oxygen of the scaffold. The scaffold aligns well in parallel to the dipole moment of the peptide and in an anti-parallel orientation which further validates the hypothesis of a rod like α-helix mimetic. RMSD values averaged at 0.4951 and 0.4953 for the parallel (**Fig. 2.1d**) and anti-parallel orientations respectively.

Although solution methods for assembly of very large[147] and long aromatic oligoamides[148] have been described, a significant advance in this area would be the ready availability of solid phase methods tolerant to a diverse array of side chains; this would facilitate library generation and ease of purification. Other than the Wilson groups' reports on *N*-alklyated aromatic oligoamides,[122, 123] only a limited number of reports have been described on the synthesis of benzanilides[149, 150] and related aromatic oligoamides[151, 152] that meet the criteria outlined above. During development of the methodology, the Ahn group published an alternative strategy to synthesise 3-*O*-alkylated trisbenzamides.[150] This followed an *N*→*C* chain elongation to address the reduced reactivity of the aniline and each cycle followed a sequence of four iterative reactions. This method is not compatible with other methods for solid phase peptide synthesis (SPPS), therefore addition of amino acids or incorporation of oligobenzamides into peptides may be problematic. In addition to this, the linear nature of the synthesis (a total of 9 reactions once resin bound) affords relatively poor yielding trimers: averaging at approximately 30 % yield. For these reasons, continuing with development of the methodology was deemed synthetically vital; this chapter describes the development of such a method that can be used for synthesis of 3-*O*-alkylated aromatic oligobenzamides using standard Fmoc solid phase strategies and iterative coupling and deprotection steps. Using microwave irradiation, trimers can be assembled on a solid-support in 2.5-4 hr in sufficient purity for screening purposes – a significant improvement on any previously published method. The methodology is tolerant to a large and diverse collection of monomers and amenable to synthesis of longer oligomers. The approach and observations in developing it should have wide applicability for the synthesis of aromatic oligoamide foldamers in general.

2.2 Results and Discussion

In developing the approach it was desirable to avoid use of novel protecting group chemistries and so fundamental features were sought that are compatible with standard Fmoc solid phase strategies such as the use of Fmoc as a semi-permanent protecting group and permanent acid labile protecting groups on the side chains. On this basis, a four step synthesis of a broad array of monomers **14a-r** was developed (**Scheme 2.1**) exploiting either alkylation of the intermediate phenol at the diversification point using alkyl halides or alcohols under *Mitsunobu* conditions. As is shown, a full array of peptide based side chains

covering the entirety of functionality found in native peptide side chains is accessible (with the exception of Cys, Arg and His) whilst several non-natural side chains and chiral side chains can also be incorporated. An Fmoc protected monomer mimicking Gly was also synthesised; the standard monomer design positions the phenolic oxygen as the atom mimicking the α-position of the amino acid within the helix and therefore this represents a poor mimic of Gly and would require protection during synthesis. Commercially available 3-methyl-4-aminobenzoic acid **19** was thus protected with Fmoc to furnish the Gly mimic **14s**.



Scheme 2.1 Synthesis of Fmoc protected monomers for SPPS.

There are several noteworthy points as follows: (i) for benzylic side chains it was necessary to use tin (II) chloride for nitro group reduction as opposed to palladium on charcoal, so as to avoid cleavage of the side chain,[114] (ii) for the hydrolysis step, care may be required to avoid cleavage of the side chain, requiring use of lithium hydroxide and mild conditions (e.g. room temperature). Cleavage of the side chain by elimination of the phenol can occur under forcing conditions – a feature which prevented us from obtaining a monomer mimicking histidine. Additionally, for side chains possessing an electron withdrawing group γ to the phenol, elimination is promoted during monomer synthesis, hydrolysis and coupling (see below) (**Fig. 2.2a**). Finally, for the *tert*-butyl ester side chain, deprotection of the *tert*-butyl group using sodium hydroxide was observed presumably *via* a ketene intermediate (**Fig. 2.2b**).

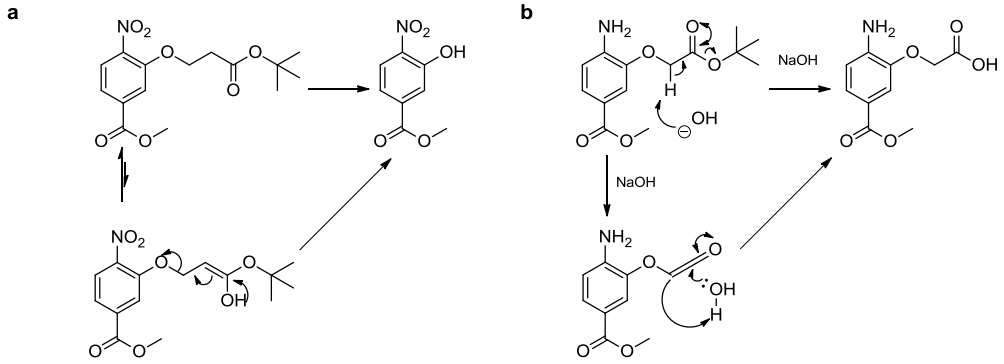
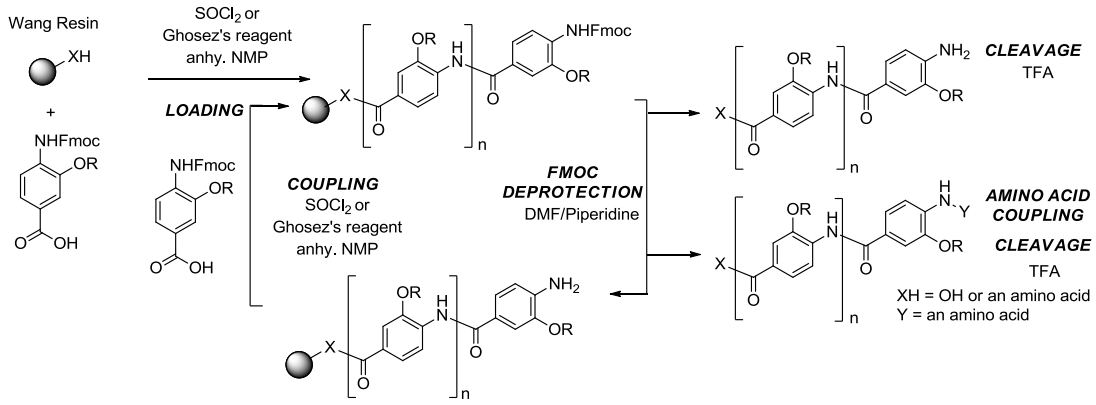


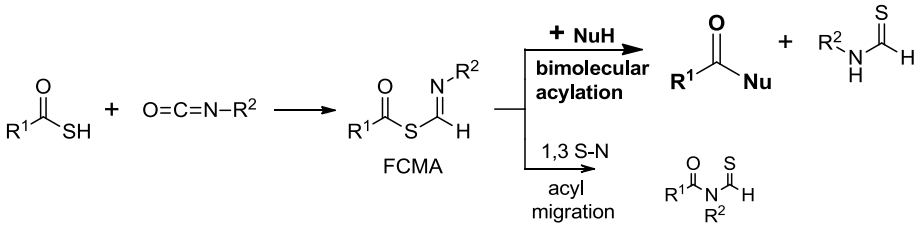
Figure 2.2 Examples of (**a**) Elimination in a side chain containing an electron withdrawing group γ to the phenol (**b**) deprotection of tert-butyl group.



Scheme 2.2 Solid phase synthesis protocol.

A general outline of the solid phase synthesis method is illustrated in **Scheme 2.2**. In terms of developing this methodology, the amide bond forming reaction is challenging as the substrate is a deactivated aniline. Acid labile Wang resin was selected for these studies and resin loading was achieved using thionyl chloride or Ghosez's reagent, however, standard coupling reagents such as HCTU and EDCI do suffice, either directly to the resin or to Gly loaded resin. A series of screening experiments for anilide formation was then attempted using the isopropyl monomer **14a**, chloroform as solvent and microwave assistance (using a CEM™ peptide synthesiser) to identify suitable coupling reagents. Of these, only those forming acid chlorides proved successful (i.e. thinoyl chloride, dichlorotriphenylphosphorane and Ghosez's reagent). This was not entirely surprising given that prior studies,[114, 116, 122] in the absence of microwave, indicated that strongly activated acids (e.g. acid chlorides) would be necessary to mediate formation of the amide bond. Due to the acidic nature of these reagents and the acid sensitive nature of the resin and possible protecting groups, several strategies to make activated monomers such as activation with a pentafluorophenol or an *N*-hydroxysuccinimide group on the carboxylic acid were also

investigated, however, both failed to induce coupling. Another method inspired by work from the Danishefsky group involved converting the acid to a thioester[153] (**Scheme. 2.3**). Addition of an isocyanate should form a very strong bimolecular acylating agent (formimidate carboxylate mixed anhydride (FCMA) intermediate) facilitating anilide formation. Treatment of monomers **14a** + **14e** with Lawesson's reagent resulted in highly unstable thioacids and were poor yielding for subsequent amide hydrolysis.



Scheme 2.3 Scheme showing reaction of a thioacid with an isocyanate.

Using the successful acylating agents, optimisation of the method then proceeded by attempting oligomer synthesis and broadening the monomer set. Unfortunately, the majority of monomers in the set were found to be poorly soluble in chloroform and so testing was carried out in DMF. Using *in situ* formation of the acid chloride from Ghosez's reagent and microwave irradiation no anilide formation was observed. Similarly, pre-activation or isolation of the acid chloride followed by microwave assistance was unable to effect the anilide formation. An explanation for these results was obtained from LC-MS analysis of the reaction mixture which revealed capping of the immobilised aniline by both DMF and Ghosez's reagent to give a stable amidine. This capping reaction which is observed even where the acid chloride is used directly, indicates that the solvent reversibly reacts with the acid chloride to generate the *Vilsmeier* intermediate which can then cap the aniline. This behaviour is not observed for synthesis of *N*-alkylated aromatic oligoamides[123] – one explanation is that capping of an *N*-alklyated aniline results in an unstable intermediate which cannot lose a proton to form the amidine (**Fig. 2.3**). With these results in hand, a solvent screen was performed to identify polar aprotic solvents which would not lead to such side-reactions.

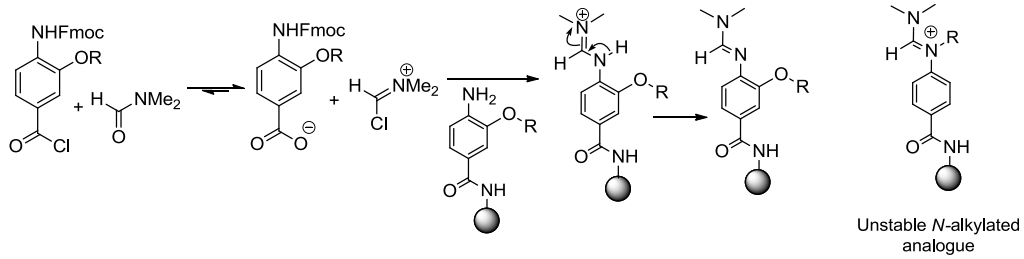


Figure 2.3 Mechanism for capping of anilines during SPS *via Vilsmeier* intermediates.

Table 2.1 Table of synthesised oligomers. R^1, R^2 and R^3 (lower-case letters) are compound **14** analogues and R^4 (upper-case letters) are amino acids: Trimer **20** is made up from **14a**–**14a**–**14a**–Gly. ' synthesised using in situ activation of monomer with thionyl chloride " synthesised using in situ activation of monomer with Ghosez's reagent.

Trimer	R^1	R^2	R^3	R^4	Final Purity (%)	Yield (%) Precipitate	HPLC
20	a	a	a	G	95	73	-
21	e	e	e	G	95	92	-
22	e	h	a	G	90	71	-
23	a	h	e	G	90	82	-
24	e	g	a	G	95	78	-
25	g	j	a	G	99	99	-
26	e	i	a	G	99	86	-
27'	a	d	a	G	99	73	35
28'	a	n	a	G	99	-	19
29	a	s	a	G	99	64	32
30'	a	f	a	G	99	79	21
31	a	k	a	G	-	-	-
32	a	l	a	G	-	-	-
33'	a	m	a	G	99	69	22
34''	a	o	a	G	99	-	22
35	a	p	a	G	90	69	32
36'	a	q	a	G	99	-	28
37'	a	r	a	G	99	-	17
38	a	a	a	-	-	-	-

Oligomer	Sequence	Final Purity (%)	Yield Precipitate	HPLC
39	a b d (I)	99	-	35
40	G a a a (G)	-	-	-
41	a c b b (G)	95	91	-
42	a c c a(G)	95	93	-
43	a a a a a a (G)	95	87	-

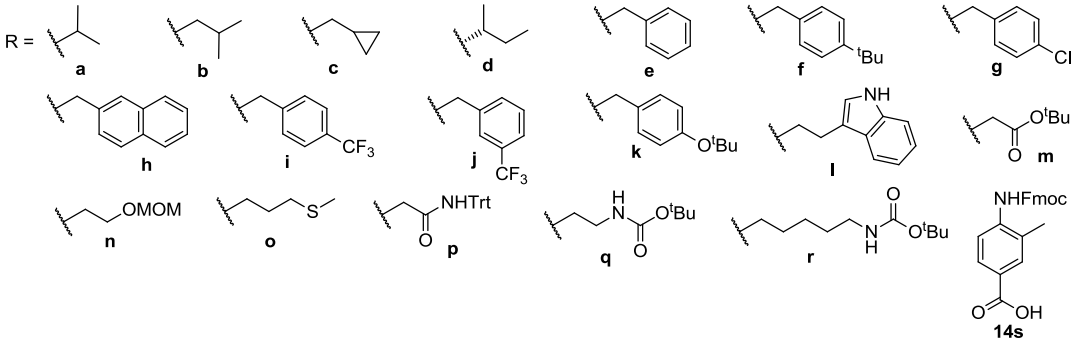


Figure 2.4 Chemical structure of side chains incorporated into trimers using the SPS procedure.

The solvent screen identified NMP as a suitable solvent with which to perform solid-phase coupling to give the aromatic benzamides. After further screening and optimisation, it was established that direct use of the acid chlorides obtained from thionyl chloride or pre-activation using Ghosez's reagent prior to coupling in the microwave

synthesiser could affect coupling in high yield. It is important to note that the highest coupling conversion obtained using dichlorotriphenylphosphorane was 20% and a significant impurity of mass 557 was obtained in both DMF and NMP and so work using this reagent was discontinued. The use of acid chlorides was preferable for the majority of alkyl/ aryl Fmoc protected monomers **14** as these could be precipitated and stored for at least one month with no decomposition. For the more highly functionalised monomers which tended not to precipitate upon reaction with thionyl chloride, it was preferable to use the *in situ* method (these highly functionalised monomers also tended to be less stable as acid chlorides). For direct addition of acid chlorides, a single cycle of coupling at 50°C for 30 minutes in the absence of base was sufficient to achieve high conversion, however, for longer oligomers, double couplings were used. For Fmoc removal, no special optimisations were required and 20% piperidine in NMP was sufficient. Care was required with the global deprotection reaction which was performed off-line from the synthesiser; certain side chains (see below) were found to be susceptible to cleavage *via* elimination with the indole side chain a notable example, thus this stage of the procedure requires careful monitoring.

With these observations and optimisations established, the versatility of the method was demonstrated by synthesising a sufficient number of trimers **20**-**37** (**Table 1**) so as to demonstrate that each monomer in the set could couple and be coupled to. In addition, it was also illustrated that it is possible to couple directly to the resin **38** and amino acids other than Gly could be appended to the *C*-terminus **39** (through use of different amino acid loaded Wang resins) and to the *N*-terminus **40**. The only problematic monomer was **14k** with the resulting oligomer undergoing cleavage of the benzylic phenol under the standard deprotection conditions required to cleave the phenolic *tert*-butyl protecting group. In addition, whilst reasonable coupling was observed with monomer **14l**, isolation and characterisation of the resulting trimer was not possible. Finally, the versatility and power of the method was also illustrated though synthesis of longer oligomers **41**-**43** (up to a hexamer). This foldamer was obtained in 10 hr using double couplings and the NMR spectrum is shown in **Figure 2.5** for the product obtained direct from the resin. This spectrum is typical of the spectral data that is obtained direct from resin cleavage and indicates that the oligomers are obtained in sufficient purity for preliminary screening. In a number of cases, this was not the case, however, cleaner material can be obtained by preparative HPLC.

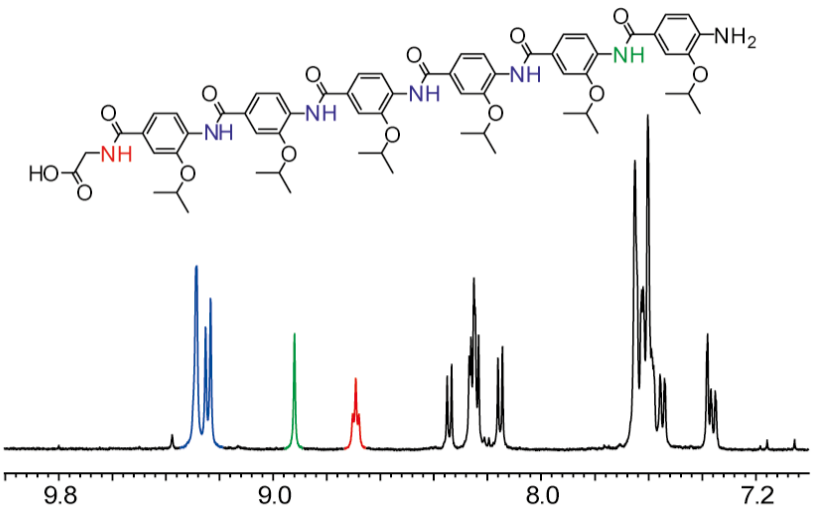


Figure 2.5 ^{1}H NMR spectrum (500 MHz, DMSO-d_6, 60 ^{o}C) of hexamer **43**.

2.2.1 Conformational Studies

2.2.1.1 Molecular Modelling

As previously discussed in *Chapter 1*, some PPIs involve much longer helices: for example the gp41 hexameric coiled coil interaction. To illustrate the potential for longer oligomers to act as mimics of extended helices, molecular modelling was performed on the hexamer **43** as is illustrated in **Figure 2.6**. To do this a conformational search was carried out on hexamer **43** and the structure was minimised by employing a full *Monte Carlo* search in the software Macromodel® using the MMFFs (Merk Molecular Force Fields) method. 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. The results revealed the lowest energy conformation was the extended structure with all six side chains lying on the same face; a conformation displaying an alternative arrangement of side chains, however, has a relative potential energy of +3.2 kJ mol^{-1} demonstrating a variety of rotamers are accessible. Using a crystal structure of gp41[14] (PDB ID: 1AIK), a series of superpositions was taken from the hexamer using different combinations of side chains (eg side chains R^1, R^2 + R^3 or R^5, R^4 + R^3) and the extended helix using different combinations of residues (e.g. *i*, *i* + 3 and *i* + 7 or *i*, *i* + 4 and *i* + 8) and at varying positions on the helix (e.g. towards the *N* or *C* terminus). From the relatively small set sampled in comparison to the available combinations, RMSD values ranging from 0.421-0.788 Å was achieved when superimposing 3 atom pairs consisting of the oxygen of the alkoxy group and the α carbon of the amino acids.

The superposition of the lowest energy conformation of the hexamer using the alkoxy oxygen from rings 2, 3, and 4 with residues at *i*, *i* + 3 and *i* + 7 positions respectively (residues Thr569, Leu566 and Gln562) is shown in **Figure 2.6a**. This demonstrates side chains from rings 1, 2, 3, 4 and 5 are orientated in a very similar fashion to residues at the *i* - 4, *i*, *i* + 3, *i* + 7 and *i* + 10 positions. Molecular modelling studies infer having different arrangements of side chains is thermodynamically viable. With this in mind, the Ar-CO bond on ring 5 has been rotated and the overlays in **Figure 2.6b-d** show side chain 6 is found to occupy the same space as residues in the *i* + 14, *i* + 15, and *i* + 16 positions. This demonstrates that such oligomers could find use in the inhibition of more extended α-helix mediated PPIs.

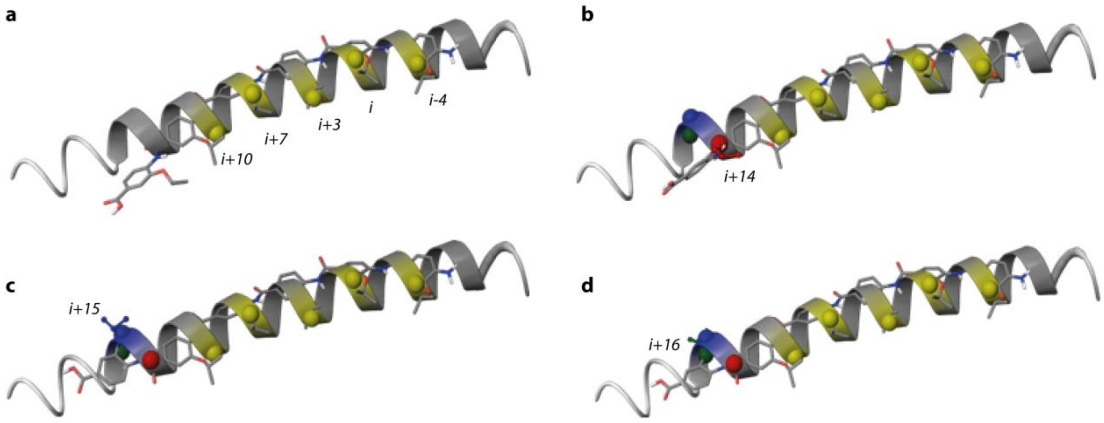
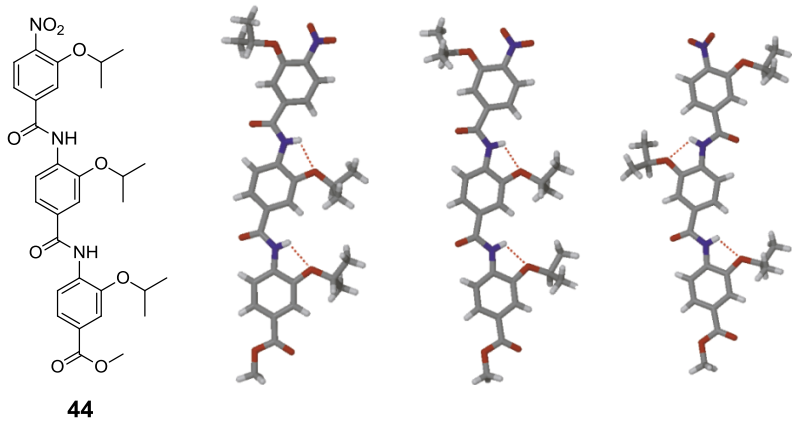


Figure 2.6 Figure showing superposition of hexamer **43** with gp41extended helix (PDB ID: 1AIK). (**a**) Low energy conformation of hexamer. (**b-d**) Demonstrating how rotation of side chain from ring 6 allows mimicry of residues at the *i* + 14, *i* + 15, and *i* + 16 positions.

2.2.1.2 X-Ray Crystal Studies

Finally several crystal structures of a representative trimer **44** (described previously)[114, 116] were obtained (Dr. P. Prabhakaran synthesised and obtained the crystals of **44**, Prof. M. Hardie and C. Kilner acquired and solved the crystal structures) comprising isopropyl monomers and with a *C*-terminal methyl ester and *N*-terminal nitro group (**Fig 2.7a-c**). These correlate with previously published analysis of the conformational preference of these oligomers[114, 154, 155] i.e. that they adopt a rod-like conformation with free rotation around the Ar-CO axes and rotation around the Ar-NH axes restricted through S(5) intramolecular hydrogen-bonding. These crystal structures are extremely significant for the project as they demonstrate that side chains can adopt both *syn* and *anti* orientations with respect to one another, whilst variations along the backbone permit the side chains to project in subtly different orientations. As most key binding residues occur on the same face of the

helix, this ability for the oligomers to rotate and project functionality in the correct orientation is a vital feature of these compounds to maximise interactions for inhibition.



44

Figure 2.7 Solid-state structures of trimer **44** (left) Structure *I* (middle) Structure *II* and (right) structure *III*.

2.3 Conclusion

A robust method for synthesis of aromatic oligoamides containing a diverse array of natural and non-natural amino acids side chains using a microwave assisted automated peptide synthesiser has been developed. A four step monomer synthesis allows generation of Fmoc protected building blocks for SPS with trimers accessible in 2.5-4 hr in sufficient purity for screening after precipitation (alkyl or aromatic side chains) or after HPLC (side chains with acid labile protecting groups). These foldamers represent excellent templates to act as mimetics of the α-helix and hence as inhibitors of protein-protein interactions. The method represents a powerful tool with which to obtain PPI inhibitors by sequence based design and for library generation to screen against unknown targets.

Chapter 3

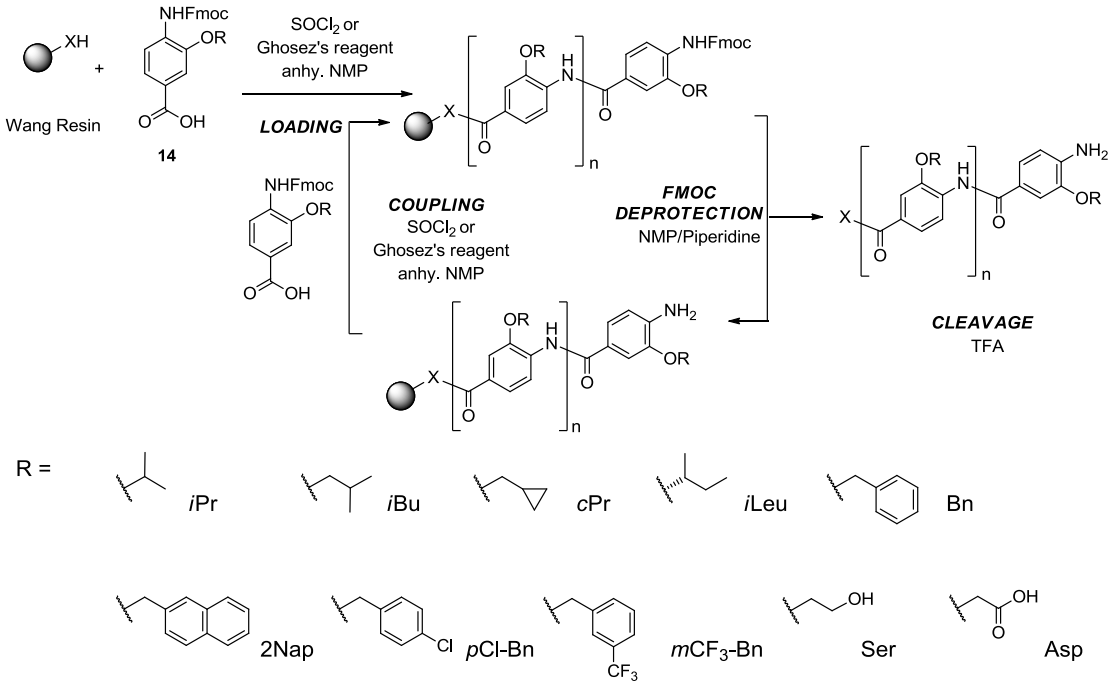
Biophysical Evaluation of

3-O-Alkylated Oligobenzamides

3.1 Addressing the Hypothesis

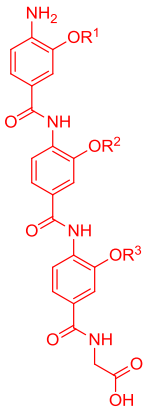
As demonstrated in *Chapter 1*, the modulation of PPIs represents a current and challenging area for drug discovery: the use of designed molecules (Type I and Type III mimetics) to target helix mediated PPIs was discussed. Small molecule modulators identified *via* traditional drug discovery routes (HTS and fragment based approach) were omitted since many PPIs share similar topographical features,[15] and a generic approach to target such complexes is an ultimate goal. The conclusion of this would be a common scaffold that contains intrinsic diversification points such that minor changes result in selective inhibitors for a number of PPI targets.

Discussed in *Chapter 2* was the 3-*O*-alkylated oligobenzamide scaffold (3HABA) developed in the Wilson group. The supposition behind the scaffold was that if a rigid rod is formed and suitable functionality is presented with a comparable spatial orientation to side chains of an α-helix[114], selective inhibitors of helix mediated PPIs could be generated.[116] Adapting the original solution phase synthesis, *Chapter 2* reports the rapid generation of a small library of oligobenzamides using a modular solid phase synthesis developed on a microwave assisted peptide synthesiser. The library focused on identifying the scope of the methodology in terms of functionality and chain length, whilst also designing potential inhibitors of the p53/*h*DM2 interaction.



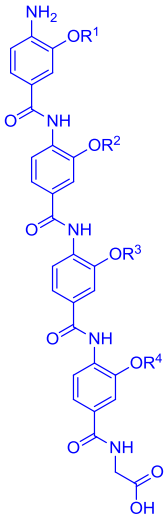
Scheme 3.1 SPS protocol for synthesis of 3-*O*-alkylated oligobenzamides. 'X' can be 'O' or 'any amino acid'.

Using the SPS method described in *Chapter 2*, a small library of 16 oligomers (**Fig. 3.1**) was successfully synthesised: designed with potentially good and bad inhibitors of the p53/*h*DM2 interaction to further test the validity of the approach. This built on previous work targeting this interaction and initiated new studies to assess the potential of the 3HABA scaffold to disrupt the Mcl-1/NOXA B interaction. Accordingly, the majority contained aromatic and aliphatic functionality; however, other functionality was included in several trimers. The iterative procedure is outlined in **Scheme 3.1** and trimers were obtained with >90% purity after precipitation (by UV chromatography and NMR spectra). After HPLC purification of trimers obtained using the *in situ* method (Method C) in *Chapter 2*, insufficient material was recovered to characterise the oligomers and also carry out screening studies, therefore no binding data was obtained for this larger set of oligomers containing a wider range of functionality. The following work will discuss how fluorescence anisotropy data is collected and processed in order to determine inhibition constants, and will examine the initial screening results from each interaction and the development of the library in response to these results.



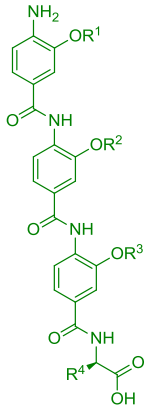
Set A

20 R^1 = *i*Pr, R^2 = *i*Pr, R^3 = *i*Pr
22 R^1 = Bn, R^2 = 2Nap, R^3 = *i*Pr
23 R^1 = *i*Pr, R^2 = 2Nap, R^3 = Bn
24 R^1 = Bn, R^2 = *p*Cl-Bn, R^3 = *i*Pr
25 R^1 = *p*Cl-Bn, R^2 = *m*CF_3-Bn, R^3 = *i*Pr
27 R^1 = *i*Pr, R^2 = *i*Leu, R^3 = *i*Pr
28 R^1 = *i*Pr, R^2 = Ser, R^3 = *i*Pr
45 R^1 = *p*Cl-Bn, R^2 = 2Nap, R^3 = *i*Pr
46 R^1 = Bn, R^2 = *m*CF_3-Bn, R^3 = *i*Pr
47 R^1 = Bn, R^2 = *i*Bu, R^3 = *i*Bu



Set B

41 R^1 = *i*Pr, R^2 = *c*Pr, R^3 = *i*Bu, R^4 = *i*Bu
42 R^1 = *i*Pr, R^2 = *c*Pr, R^3 = cPr R^4 = *i*Pr



Set C

48 R^1 = Asp, R^2 = *i*Bu, R^3 = *i*leu R^4 = Val
49 R^1 = Asp, R^2 = 2Nap, R^3 = *i*Bu, R^4 = Val
50 R^1 = Bn, R^2 = 2Nap, R^3 = *i*Pr, R^4 = Asp
51 R^1 = Bn, R^2 = 2Nap, R^3 = *i*Pr, R^4 = Lys
R^4 = amino acid

Figure 3.1 3HABA based oligomer library for screening against the p53/*h*DM2 and Mcl-1/NOXA B interaction. Set 1 (**20**-**25**, **27**) and 2 (**41**, **42**) from *Chapter 2* and Set 1 (**45**-**47**) and 3 (**48**-**51**) are newly synthesised.

3.1.1 SAR of oligomers targeting p53/*h*DM2

p53 binds with its negative regulator *via* key hydrophobic residues on the p53 activation domain, into a hydrophobic cleft on the *h*DM2 LBD. More specifically the Phe19, Trp23 and Leu26 residues at the *i*, *i*+4 and *i*+7 positions (**Fig. 3.2a**) .[11] In a previous study from the group,[116] trimer **10** (**Fig 3.2b**) containing similar substituents to the key residues on the p53 helix, was identified as the most potent 3HABA proteomimetic in a fluorescence anisotropy competition assay targeting p53/*h*DM2. Using the same side chains, an equivalent trimer **22** (with a *C*-terminal Gly) was synthesised to use as a reference to carry out an SAR study for targeting the p53/*h*DM2 interaction. In this context, oligomers were screened to determine the following:

22, **23**	Obtain a direct comparison with previously tested oligomers to study the effect of Gly and to determine if position of the aniline / acid effects the binding mode
24, **25**, **45**, **46**	To determine how halogenated aromatic groups effect binding
20, **27**, **47**	To determine how substituting aromatic groups for alkyl groups effects binding
41, **42**	To determine if larger oligomers are tolerated by the protein
28, **48**, **49**	To determine how hydrophilic functionality effects binding
50, **51**	To determine if addition of functionalised amino acids affects solubility / binding

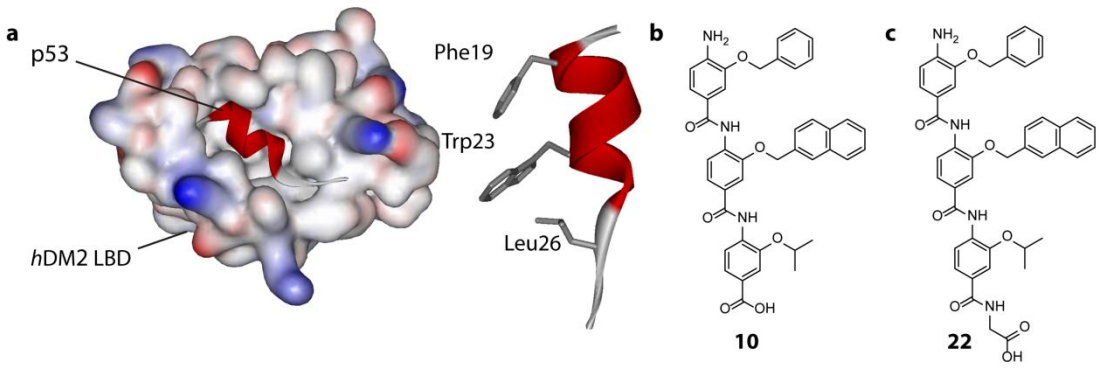


Figure 3.2 (**a**) p53/*h*DM2– regulates cellular stress (PDB ID: 1YCR): p53 binds with Phe19, Trp23 and Leu26 hydrophobic residues. (**b**) Most potent 3HABA inhibitor **10** identified from previous studies containing hydrophobic residues: R^1 = Bn, R^2 = 2Nap and R^3 = *i*Pr. (**c**) New *C*-terminal Gly analogue **22** as reference for SAR study.

3.1.2 SAR of oligomers targeting Mcl-1/NOXA B

Mcl-1 is one of the many anti-apoptotic members of the Bcl-2 family and binds to NOXA B (BH3-only pro-apoptotic member). The involvement of this interaction in

oncogenesis is still under investigation; however, it has been identified as an important therapeutic target in studies involving ABT-737 (**Fig. 1.6b**).[156] The helical binding motif is shown in **Figure 3.3a** and involves key interactions from the Glu74, Leu78, Ile81 and Val85 residues. Oligomer **48** (**Fig. 3.3b**) was synthesised as an intended inhibitor of Mcl-1/NOXA B, containing similar substituents to the key residues on NOXA B: R^1 = Asp, R^2 = *i*Bu, R^3 = *i*Leu and R^4 = Val. An additional 5 proteomimetics were selected for a preliminary SAR study, targeting Mcl-1/NOXA B, on the following rationale:

48, **49**	Oligomers containing similar functionality NOXA B: potential good mimics
47	Has similar functionality to the BID peptide which also binds to Mcl-1: potential good mimic
24	Contains large aromatic groups which may be too large for the binding cleft: potential poor mimic
27, **41**	Oligomers lacking a Glu mimic but containing other similar functionality to NOXA B: potential good mimics

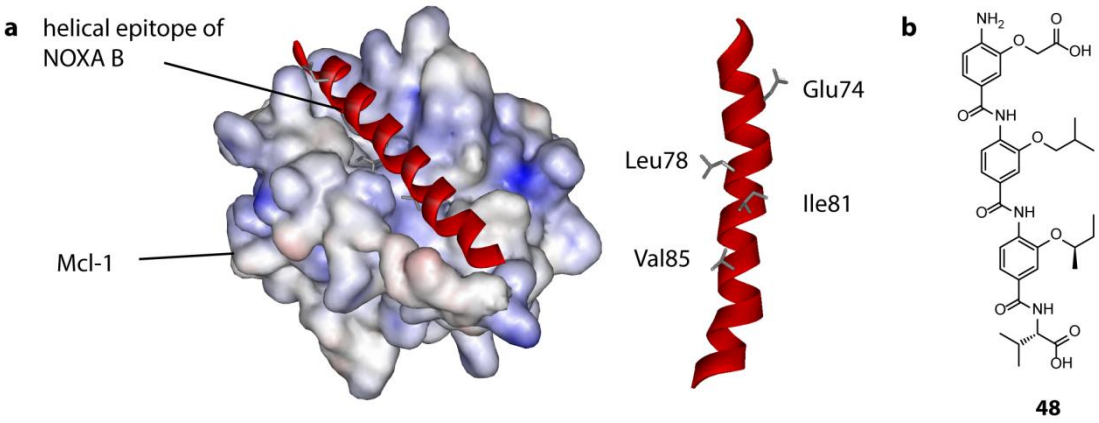


Figure 3.3 (**a**) Binding of NOXA B to its partner protein Mcl-1. Key binding residues on NOXA B positioned at the *i*, *i*+4, *i*+7 and *i*+11 residues: Glu74, Leu78, Ile81 and Val85 respectively (PDB ID: 2JM6). (**b**) Designed potential inhibitor of Mcl-1/NOXA B interaction containing appropriately functionalised side chains: **48** R^1 = Asp, R^2 = *i*Bu, R^3 = *i*Leu and R^4 = Val.

3.2 Fluorescence Anisotropy Assays

3.2.1 Fluorescence Anisotropy

Anisotropy can be defined as a property that has directional dependence, and in this context, the directional dependence of the fluorescence property of a fluorophore will be exploited. When a fluorophore is irradiated, a photon of a specific energy is absorbed and the fluorophore gets excited to a higher energy state. After the fluorescence lifetime, a lower energy photon is released (with heat), returning the fluorophore back to its original energy

state. This excitation and relaxation involves redistribution of electrons and requires excitation by a photon having a specific polarisation with respect to the axis of the molecule. In a mixure of randomly oriented molecules, only molecules oriented within a particular range of angles to the applied polarisation are excited. These molecules will, accordingly, emit with a specific polarisation relating to the orientation of the molecule. The size of a fluorophore determines the rate at which the molecule rotates: smaller entities rotate faster than larger entities. Therefore, if the fluorophore can freely rotate, changing the orientation of the molecule before relaxation and hence emission of a photon, the degree of polarisation of the emitted light will be reduced and the anisotropy will be low. Conversely, if the fluorophore is large and does not rotate before emission, the degree of polarisation of the emitted light will be much higher and, hence, the anisotropy will be high.

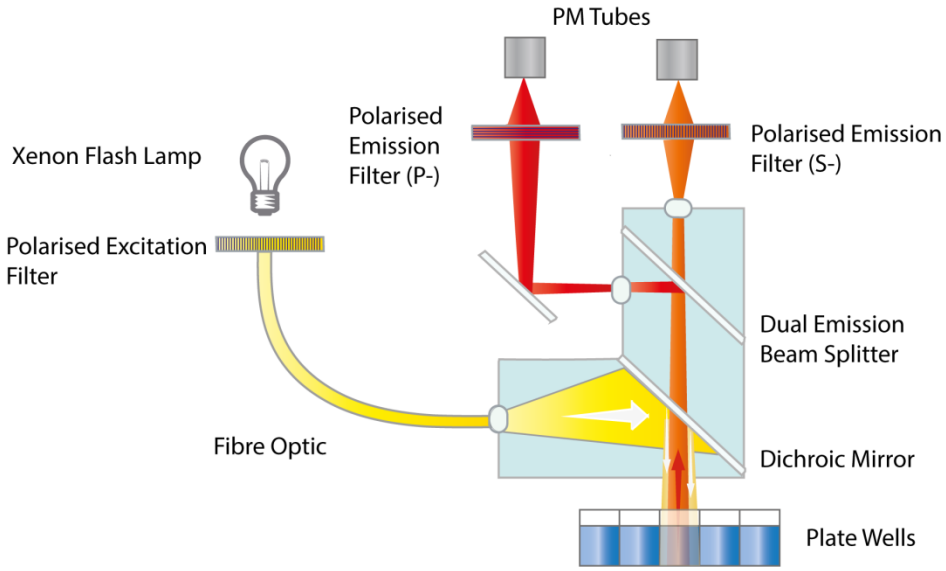


Figure 3.4 Schematic showing the mechanics of the plate reader to collect fluorescence readings.

3.2.2 Collecting and Processing Fluorescence Data

To collect fluorescence data for any given interaction, data was collected using an EnVision plate reader. **Figure 3.4** gives a schematic of this process: Broadband light is passed through a polarised excitation filter, resulting in plane polarised light with an appropriate wavelength band to the fluorophore. As the excitation light hits the dichroic mirror, it is reflected onto the sample and absorbed by correctly orientated fluorophores. The same mirror forms part of the emission channel by allowing lower energy emission wavelengths to pass through whilst blocking the excitation light, reducing saturation at the detector. The emission light then hits the dual emission beam splitter, directing half through a parallel (S-) polarising emission filter and directing half through a perpendicular (P-) polarising emission filter (S- and P- with respect to the polarised excitation filter). The few

photons which pass through the instrument are passed through a photomultiplier (PM) tube which amplifies the light signal into a voltage, which is read by the computer. Readings from S- and P- channels can be taken simultaneously, minimising artifacts such as photobleaching and sample movement, therefore G (instrument and assay dependent factor) is normalised to 1. The data is then used to calculate the average anisotropy or fraction of ligand bound, plotted against [protein] or [proteomimetic] and fitted to a logistic model to extract K_d, IC_{50} or EC_{50} values (See Appendix I).

3.2.3 Idealistic Binding Measurements

The fundamentals of the assay rely on the change in rotation time of the fluorophore to provide binding constants such as K_d, K_i (or IC_{50}s to be explained later). **Figure 3.5a** gives a graphical representation of the assay and is explained further in sections *3.2.3.1* and *3.2.3.2*.

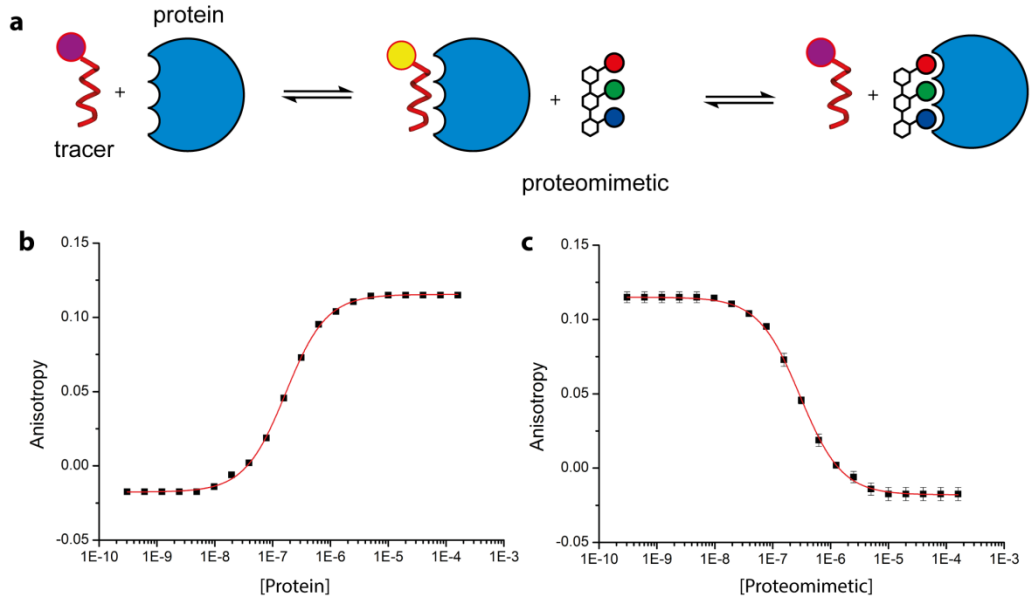


Figure 3.5 (**a**) A graphical representation of the fluorescence anisotropy assay. Left equilibrium (protein titration) allows for calculation of the K_d and right equilibrium (competition assay) allows for calculation of the K_i (or IC_{50}). (**b**) Example of a sigmoidal curve obtained after a protein titration. (**c**) Example of a sigmoidal curve obtained after a competition assay.

3.2.3.1 Protein Titration: Dissociation Constant

On the left of the first equilibrium (**Fig. 3.5a**), there is a peptide tracer (the fluorophore) and the much larger protein. When the tracer is in an unbound state, the molecule has a fast rotation due to its small size: this will therefore exhibit low anisotropy. However, when the tracer binds to the protein (right of the first equilibrium), a larger

complex is formed displaying much slower rotation: this will therefore exhibit high anisotropy. This first equilibrium provides the rate of dissociation (K_d) of the complex and is determined by carrying out a serial dilution of the protein, whilst keeping the concentration of the tracer constant. If tested over the correct range, a sigmoidal curve should be observed (**Fig. 3.5b**) displaying low anisotropy at low protein concentrations and high anisotropy at high protein concentrations.

3.2.3.2 The Competition Assay: Inhibition Constant

On the left of the second equilibrium (**Fig. 3.5a**), there is a large fluorophore complex exhibiting high anisotropy. On addition of a ligand, the tracer may be displaced becoming free to rotate: resulting in a decrease in anisotropy. This second equilibrium provides the K_i (IC_{50}) of a potential inhibitor ligand, and is determined by carrying out a serial dilution of the ligand, whilst keeping the concentration of protein and tracer constant. If tested over the correct range, a sigmoidal curve should be observed (**Fig. 3.5c**) displaying low anisotropy at higher ligand concentrations and high anisotropy at lower ligand concentrations.

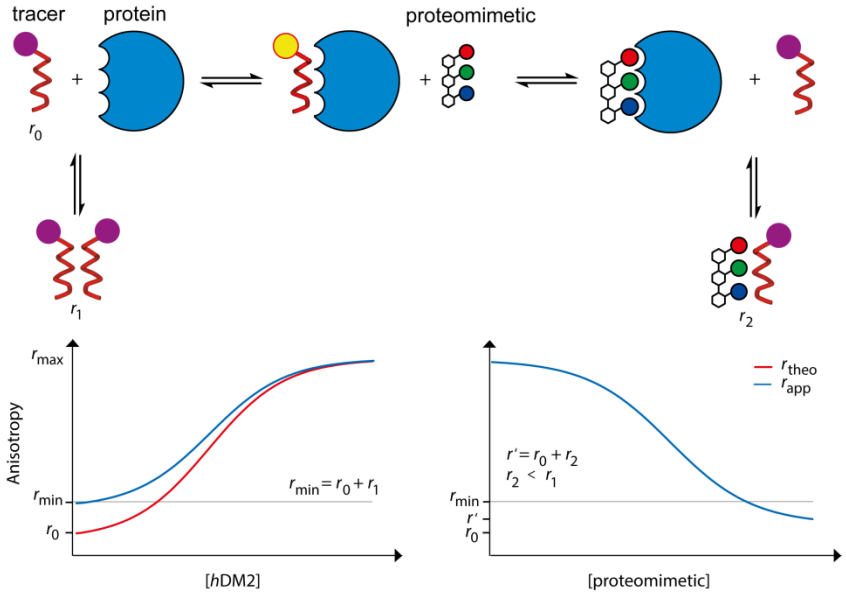


Figure 3.6 The multiple equilibria involved in the p53/*h*DM2 competition assay preventing K_i determination: Aggregation of the tracer contributes to the anisotropy during protein titration (r_1) and interaction of the tracer with the proteomimetic contributes to the anisotropy during the competition assay (r_2).

3.2.3.3 Assay Equilibria

Previous studies from the group revealed a complex set of equilibria that prevent the calculation of the inhibition constant, K_i.[116] They proposed that instead of a simple two step competition (**Fig. 3.5a**), it is more likely that the fluorophore self aggregates and also

interacts with competitor ligands. The anticipated equilibria can be explained through analysis of assay data and are shown in **Figure 3.6**

The minimum anisotropy, r_{min}, should correspond to a solution where the tracer is completely free to tumble, hence it should be equivalent to the theoretical value, r_0. R_{min} was determined during the protein titration and was actually found to be greater than r_0: caused by aggregation of the tracer, r_1. On displacement of the tracer molecule by the competitor, the tracer can interact with the competitor as [competitor] > [tracer]. The tracer:competitor complex (r_2) may be much smaller than the tracer:tracer complex and would therefore have a smaller contribution to the anisotropy $r_2 < r_1$. This can be seen from the does-response curves which have lower anisotropy values than the observed r_{min} at the beginning of the protein titration: hence why calculation of the IC_{50} (or EC_{50}) is necessary.

3.3 Biophysical Data for p53/*h*DM2 Assays

The use of fluorescence anisotropy to determine binding constants has been well established within the Wilson group with respect to the p53/*h*DM2 interaction (**Fig.3.3a**)[116, 119, 122, 125] and as such, this was the first interaction the library was tested against. Although extensively employed, it was important to repeat certain standard experiments to ensure the assay was reproducible in my hands. Following previous studies, a fluorescein tag was used as the fluorophore on the p53 activation domain (Flu-$p53_{15\text{-}31}$ purchased: Peptide Science Research Ltd.) tracer molecule and a His-tag construct of *h*DM2 was used for the protein (His-hDM$2_{17\text{-}126}$ L33E molecular cloning and expression: Dr. K. Long).

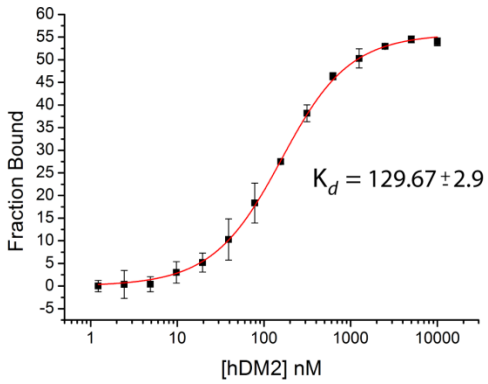


Figure 3.7 Protein titration curve of the p53/*h*DM2 interaction (K_d = 129.67 nM).

3.3.1 *h*DM2 Protein Titration

The K_d of the interaction was determined having a [*h*DM2] of 10 µM - 0.15 nM and a constant [p53*] of 54.5 nM. The plate was allowed to incubate for 1 hour and results from this experiment can be found in **Figure 3.7**: a K_d value of 129.67 ± 2.9 nM (see Appendix II)

was obtained which is comparable to data previously obtained in the group (K_d = 75.38 ± 4.2[116] and 164.4 ± 10.8 nM).

3.3.2 Competition Assays

3.3.2.1 p53 Displacement Assay

The first control experiment required was the competition assay involving displacement of labelled p53 (p53*) by unlabelled p53. In addition to understanding if the assay was reproducible in my hands, three parallel assays were made with varying DMSO concentrations (0%, 5% and 10% in the first well) to determine whether DMSO had a significant effect on the results. High DMSO concentrations were likely to be necessary due to solubility of the 3HABA oligomers in the aqueous buffer. The [p53] ranged from 50 µM – 34 nM and the [p53*] was constant at 54.5 nM across the plate. The plates were allowed to incubate for 1 hour each and results from this set of experiments are shown in **Figure 3.5**: consistent IC_{50} values were obtained across the three DMSO concentrations and were comparable with values obtained previously in the group (1.49-2.15 µM vs 1.2 µM[116]).

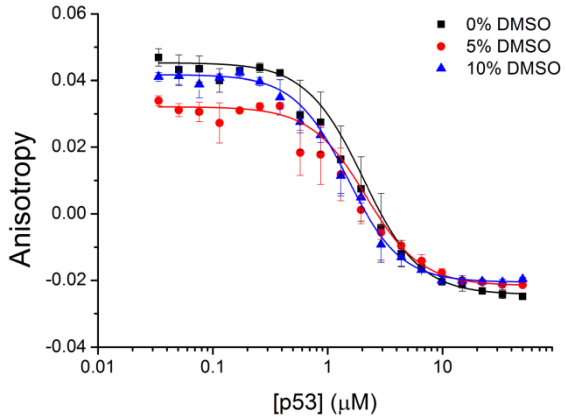


Figure 3.8 p53 displacement curves at 0%, 5% and 10% DMSO concentrations: IC_{50} = 2.04 ± 0.15 µM, 2.15 ± 0.16 µM and 1.49 ± 0.07 µM respectively.

3.3.2.2 Nutlin Competition Assay

Nutlin-3a (**Fig. 1.6a**) is known to bind to *h*DM2 and is one of the most potent inhibitors of the p53/*h*DM2 interaction identified to date.[26] A second positive control test using this inhibitor in the competition assay was accordingly carried out to ensure comparable results. The [Nutlin-3a] in the first well was at 10 µM in 10% DMSO and the [p53*] was constant at 54.5 nM across the plate. The plates were allowed to incubate and time course studies, taking readings at 30 min intervals for 3 hours, were made. IC_{50} values were consistent within this incubation period and the does-response curve for Nutlin-3a is

shown in **Figure 3.9**. In my hands, Nutlin-3a demonstrated an IC_{50} value of 24.00 ± 01.0 nM which is comparable to the literature value (90 nM).[26]

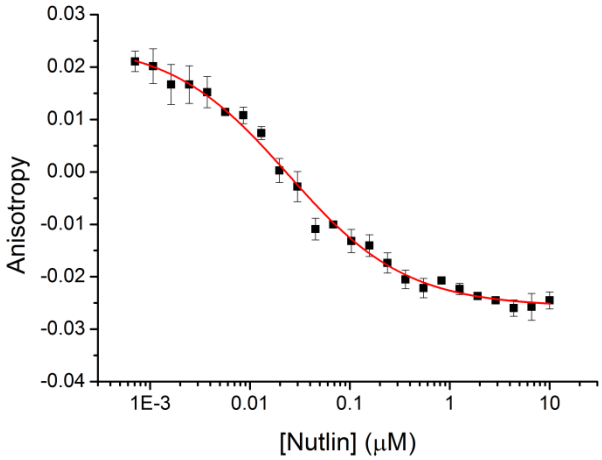


Figure 3.9 Nutlin-3a dose-response curve against the p53/*h*DM2 interaction: IC_{50} = 24.00 ± 01.0 nM.

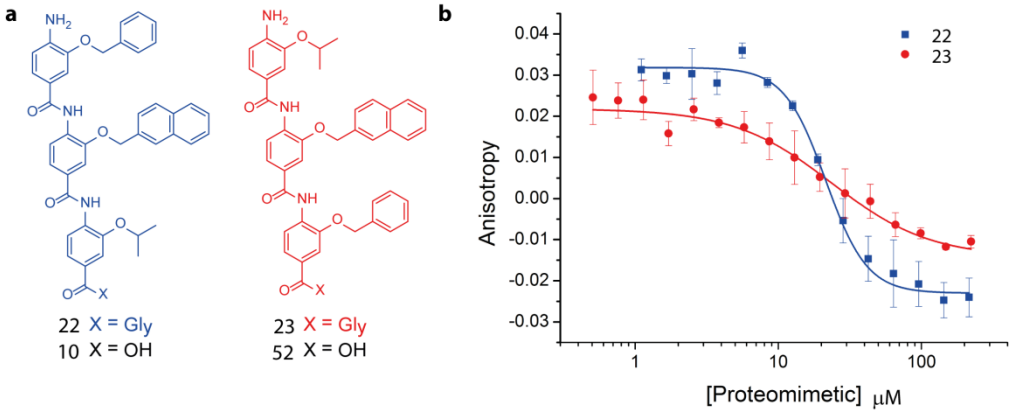


Figure 3.10 (**a**) Oligomers **22**, **23**, **10** and **52** mimicking the Phe19, Trp23 and Leu26 residues of p53. (**b**) Dose-response curves for oligomers **22** and **23**: IC_{50} = 21.43 ± 1.22 μM and 23.65 ± 4.51 μM respectively.

3.3.3 3HABA Library Screening

In previous studies, the 3HABA oligomers were synthesised in a solution-phase iterative process and oligomers were comprised solely of 3HABA based building blocks. In the SPS procedure, synthesis on a Gly loaded Wang resin achieved the best loading and so oligomers have a *C*-terminal Gly residue. To determine the effect this had on binding, **22** and **23** were synthesised as analogues of potent inhibitors from previous studies (**10** and **52**).[119] Binding curves for **22** and **23** are shown in **Figure 3.10** and exhibit comparable IC_{50} values of 21.43 ± 1.22 and 23.65 ± 4.51 respectively: suggesting the position of the aniline / acid moieties on the scaffold may not influence the orientation the inhibitor binds with. However, the shape or steepness of the curves are quite different and this feature is

expressed as the Hill coefficient which describes the cooperativity of a ligand binding to a larger protein. In this instance it gives an indication if other non-specific interactions are occurring between the proteomimetic and tracer, thereby having an effect on binding. The Hill coefficients for **22** and **23** are 1.27 ± 0.28 and 3.00 ± 0.46 respectively, suggesting that **22** binds noncooperatively and oligomer **23** has more negative cooperative binding *via* interference of the tracer:protein complex by formation of a proteomimetic:tracer complex. The previous binding studies determined **10** and **52** to display IC_{50} values of 6.35 ± 0.30 μM and 4.15 ± 0.20 μM respectively,[119] indicating the glycine may have a negative effect on the potency of the 3HABA scaffold which may be attributed to the poorer solubility of the Gly analogues.

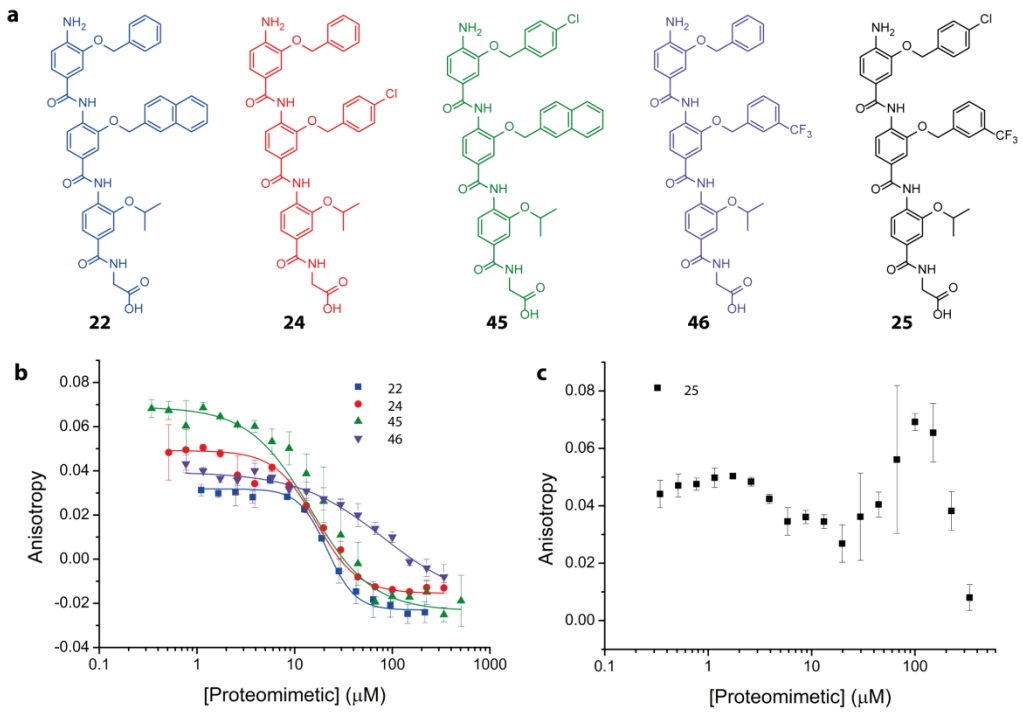


Figure 3.11 (**a**) Reference oligomer **22** and trimers **24**, **45**, **46** and **25** varying aromatic groups at the R^1 and R^2 positions. (**b**) Comparison of dose-response curves between **22** and oligomers **24**, **45**, **46** and **25**: IC_{50} = 16.51 ± 0.71 μM **25**; 14.93 ± 2.30 μM **45**; 80.71 ± 6.91 μM **46**. (**c**) Competition assay data for oligomer **25**: Data points unable to fit to a logistic model.

Oligomers **24**, **45**, **46** and **25** were synthesised to see how halogenated aromatic groups as Phe19 and Trp23 substitutes may influence binding. Dose-response curves from competition assays from this set are shown in **Figure 3.11b-c**. Using trimer **22** as the reference molecule (R^1 = Bn, R^2 = 2Nap, R^3 = *i*Pr), the following results and conclusions were made. A *p*Cl-Bn side chain replaced R^2 in trimer **24** and the R^1 in trimer **45**. In both of these trimers, a slight improvement in binding is observed (IC_{50} = 16.51 ± 0.71 μM **24**, IC_{50} = 14.93 ± 2.30 **45**) suggesting that insertion of a *p*Cl-Bn group could be beneficial. A *m*CF_3-Bn side chain replaced R^2 in trimer **46** and a dramatically higher IC_{50} is observed (80.71 ±

6.91 μM). R^1 and R^2 were replaced with *p*Cl-Bn and *m*CF_3-Bn respectively in trimer **25** and a reproducible does-reponse curve is shown in **Figure 3.12c** (see Appendix II for additional data). The data obtained is not reasonable to fit and provided no evidence for binding of this compound. In light of these results, it may be reasoned that the *m*CF_3-Bn group may not be favourable for binding.

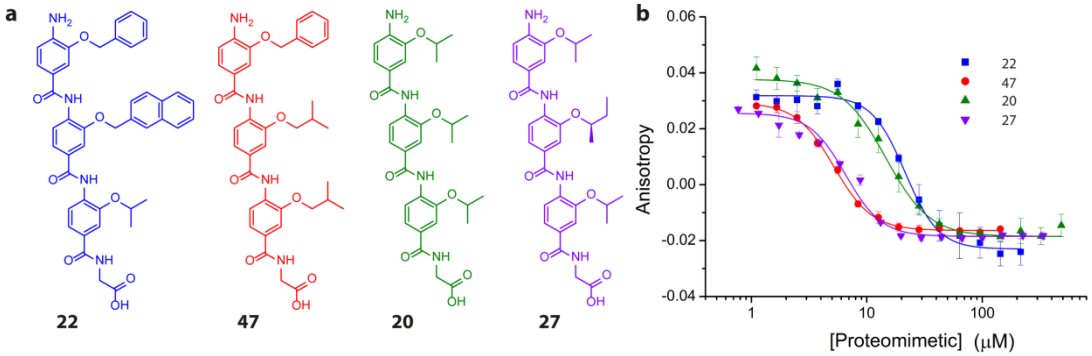


Figure 3.12 (**a**) Reference oligomer **22** and trimers **47, 20** and **27** with aromatics substituted for alkyl groups at the R^1 and/or R^2 positions. (**b**) Comparison of the dose-response curve of **22** against those for oligomers **47, 20** and **27**: IC_{50} = 5.19 ± 0.17 μM **47**, 14.61 ± 0.97 μM **20**, 6.56 ± 0.20 μM **27**.

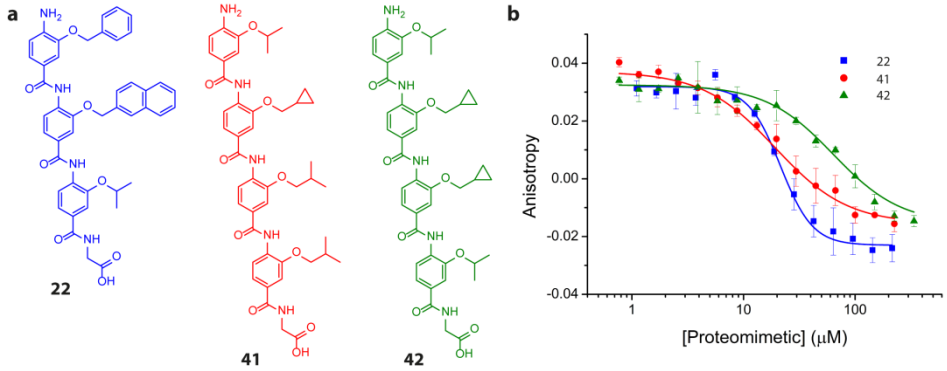


Figure 3.13 (**a**) Reference oligomer **22** and tetramers **41** and **42**. (**b**) Comparison of the dose-response curve of **22** against those for tetramers **41** and **42**: IC_{50} = 25.27 ± 4.81 μM **41**, 63.93 ± 4.94 μM **42**

Trimers **47, 20** and **27** (**Fig. 3.12a**) were modified to assess how inserting alkyl groups at the R^1 and R^2 positions might affect binding. Dose-response curves from competitions assays for this set are shown in **Figure 3.12b**. With respect to **22**, trimer **47** has an *i*Bu group replacing the bulky naphthyl group at the R^2 position. Surprisingly, removal of one aromatic group dramatically improves the IC_{50} value (5.19 ± 0.17 μM **47**). Trimer **20** has both aromatic side chains replaced with *i*Pr groups and again, an increase in binding affinity was observed and found to be reproducible (IC_{50} = 14.61 ± 0.97 μM and 10.03 ± 1.19 μM). A second trimer with all alkyl side chains (**27**) obtained a surprising low IC_{50} (6.56 ± 0.20 μM). Although it is not clear why improved binding is observed, suggestions for this could

include increased self aggregation of dimers containing aromatic side chains from hydrophobic interactions (such as π-π stacking) resulting in fewer molecules available for binding, or simply the 2Nap side chain could be a poor Trp mimic.

Tetramers **41** and **42** (**Fig. 3.13a**), containing only alkyl side chains, were also tested against this interaction. Dose-response curves for these oligomers are found in **Figure 3.13b** and IC_{50} values were determined to be 25.27 ± 4.81 μM and 63.93 ± 4.94 μM respectively. Although a great deal of information can not be obtained from these results, seemingly minor changes (R^3 *i*Bu→*c*Pr and R^4 *i*Bu→ *i*Pr) have resulted in a change in potency. This highlights the need for large libraries of inhibitors containing side chains with many simple variations in order to fully and effectively probe the *h*DM2 binding site.

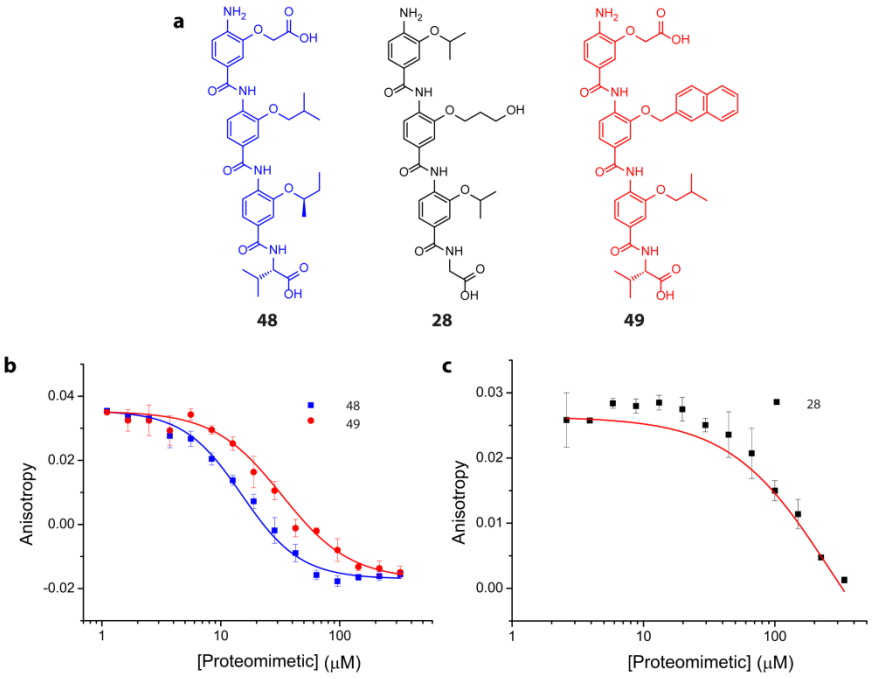


Figure 3.14 (**a**) Oligomers **48, 28** and **49** containing functionalised side chains. Dose-response curves for (**b**) acid functionalised oligomers **48** and **49** exhibiting IC_{50} values of 14.99 ± 0.89 μM and 32.68 ± 0.01 μM respectively, (**c**) alcohol functionalised **28** exhibiting an IC_{50} value of > 250 μM.

Oligomers **48** and **28** (**Fig. 3.14a**), containing a Asp and Ser analogue respectively, were screened against this target to deduce what effect the corresponding functionality may have on binding. Dose-response curves from competition assays for these oligomers are found in **Figure 3.14b**,**c**. It was pleasing to see that the acid functionalised oligomer **48** demonstrated an increased potency compared with trimer **22** (IC_{50} = 14.99 ± 0.89 μM). This corresponds to results obtained from the Hamilton group, who observed increased binding for a terphenyl scaffold when flanked with carboxylic acid groups.[157] Further studies have

shown that favourable electrostatic interactions can also be made with a nearby His residue.[158] This positive result led to the synthesis of trimer **49** (**Fig. 3.14a**), containing an *i*Bu→2Nap substitution. Despite a hypothesis that this may create more favourable hydrophobic contacts, a decrease in potency was observed (IC_{50} = 32.68 ± 0.01 µM) further suggesting 2Nap could be a poor Trp mimic. Incorporation of an alcohol moiety (**28**) led to a dramatic decrease in potency (IC_{50} = >250 µM) suggesting unfavourable interactions or simply poor affinity with the binding site.

From data obtained in the previous screening studies, it was evident that including certain hydrogen bonding moieties could help increase their potency. Unclear from **48** and **49**, however, is at what position the functionality is most likely to make electrostatic contacts. Oligomers **50** and **51** (**Fig. 3.15a**) were subsequently synthesised using an appropriately loaded wang resin as a simple method to incorporate acid and base moieties onto the trimers, and to compare results directly with trimer **22**. Dose-response curves from competition assays for these oligomers are found in **Figure 3.15b**. Trimer **50** containing a *C*-terminal Asp exhibited improved inhibition with an IC_{50} value of 9.54 ± 0.54 µM. This could indicate possible favourable interactions from incorporating acid functionality, however, the change in the shape of the curve suggests there could be some form of cooperative binding occurring. Trimer **51** with a *C*-terminal Lys exhibited much poorer inhibition (the bottom plateau was fixed at -0.02 to determine the IC_{50} at > 250 µM). Capping with an *N*-terminal Asp would be the appropriate next step to better understand what position the functionality is best positioned, however, the methodology to couple functionalised amino acids onto trimers is not yet developed.

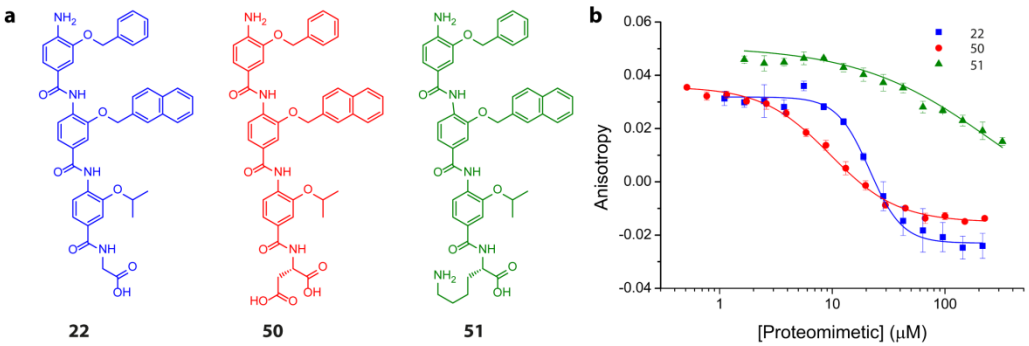


Figure 3.15 (**a**) *C*-terminally functionalised analogues of **22**, Asp **50** and Lys **51**, (**b**) Comparison in the dose-response curves of oligomers **22**, **50** and **51**: IC_{50} = 21.43 ± 1.22 µM, 9.54 ± 0.54 and > 250 µM respectively.

3.4 Biophysical Data for Mcl-1/NOXA B assays

A fluorescence anisotropy (FA) assay for this interaction has recently been developed in the group. This has provided optimised conditions and allowed the repetition of the protein titration and peptide displacements assays to generate reproducible results in my hands. A fluorescein tag was used as the fluorophore on the NOXA B (FITC-Noxa $B_{68\text{-}87}$ synthesis: Dr. P. Prabhakaran) tracer molecule and Mcl-$1_{172\text{-}327}$ (molecular cloning and expression: Dr. A. Bartlett) was used for the protein.

3.4.1 Mcl-1 Titration

Determining the K_d of this interaction was carried out having a [Mcl-1] range from 5 µM – 0.15 nM and a constant [NOXA B*] of 50 nM. The plate was allowed to incubate for 1.5 hours and results from this experiment can be found in **Figure 3.16**. A K_d value of 148.19 ± 55.97 nM (see Appendix II) was obtained which is comparable to data previously obtained in the group (K_d = 32.46 ± 1.14 nM).

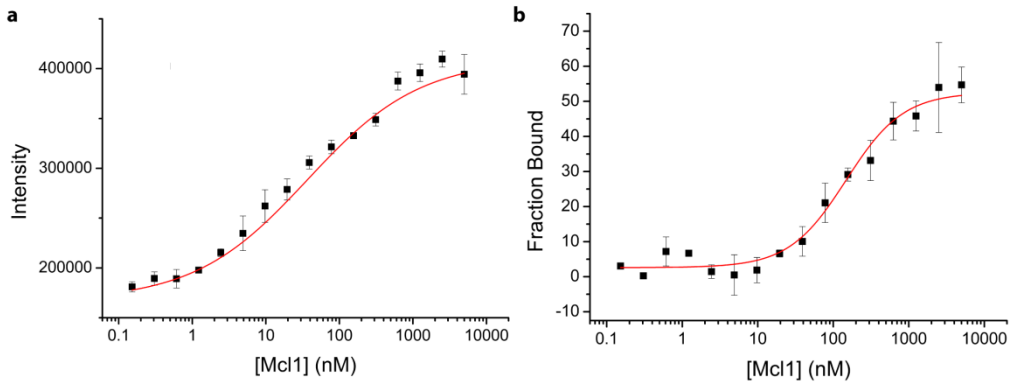


Figure 3.16 (**a**) Average intensity fitted to a logistic model to calculate λ (λ = 2.20). (**b**) Protein titration curve of the Mcl-1/NOXA B interaction (K_d = 148.19 ±55.97 nM).

3.4.2 NOXA B Displacement Assay

As a positive control, it is imperative to perform the unlabelled peptide displacement assay to ensure binding data is reproducible, and comparable with previously obtained data. In this study, assays were carried out with 0% and 10% DMSO to understand if DMSO has an effect on binding. The [NOXA B] ranged from 15 µM – 34 nM and [NOXA B*] was constant at 50 nM across the plate. The plates were allowed to incubate and time course studies, taking readings at 45 minute intervals over 3 hours, were made: IC_{50} values were consistent within this incubation period. Dose-response curves from this set of experiments are shown in **Figure 3.17**: comparable IC_{50} values were obtained across the two DMSO

concentrations and were comparable with previously obtained results (IC_{50} = 707.42 ± 104 nM **0 %** and 533.14 ± 57 nM **10 %** vs 704 ± 35 nM **10 %**)

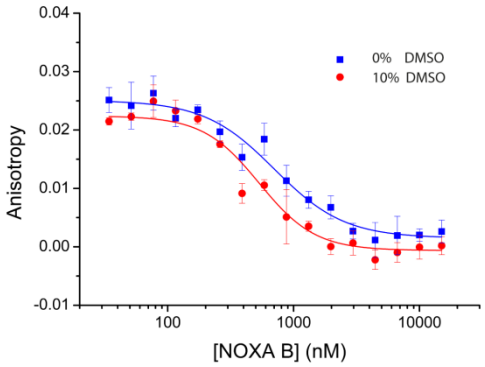


Figure 3.17 NOXA B displacement curves at 0% and 10% DMSO concentrations; IC_{50} = 707.42 ± 104.11 nM and 533.14 ± 57.10 nM respectively.

3.4.3 3HABA Screening Library

As this assay is in its infancy, there are no previous studies using 3HABA oligomers against this interaction and, hence, nothing to make comparisons against. A small selection of oligomers (**Fig. 3.18**) seen previously screened against p53/*h*DM2 were chosen to assess the potential of the 3HABA scaffold in inhibiting this new interaction. As the fluorescence intensity changes during the protein titration (see Appendix I), EC_{50} values have been calculated by plotting fraction of ligand bound against [proteomimetic] (see Appendix II).

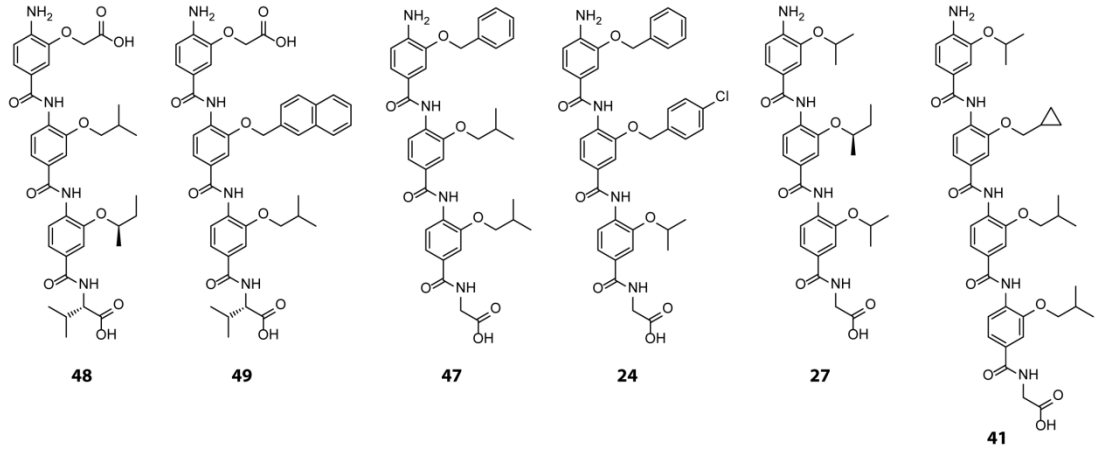


Figure 3.18 Mcl-1/NOXA B 3HABA screening library.

Oligomer **48** was designed to mimic the key binding residues of NOXA B. Synthesised on a Val loaded Wang resin, it also contains potential Ile, Leu and Glu mimics (*i*Leu, *i*Bu and Asp monomers (**14**) respectively). Oligomer **49** has a very similar composition, however, with a large aromatic group at the R^2 position which may give an indication to how large the binding pocket is. Unfortunately, as can be seen from the binding

curves for **48** and **49** (**Fig. 3.19a**,**b**), the data is fairly scattered at low [proteomimetic] with large error bars in places. This scattering may be due to aggregation of tracer molecules or proteins and / or precipitation of the protein by the proteomimetics. This is representative of competition assays from oligomers **47**, **24** and **41** and data is not reasonable to fit.

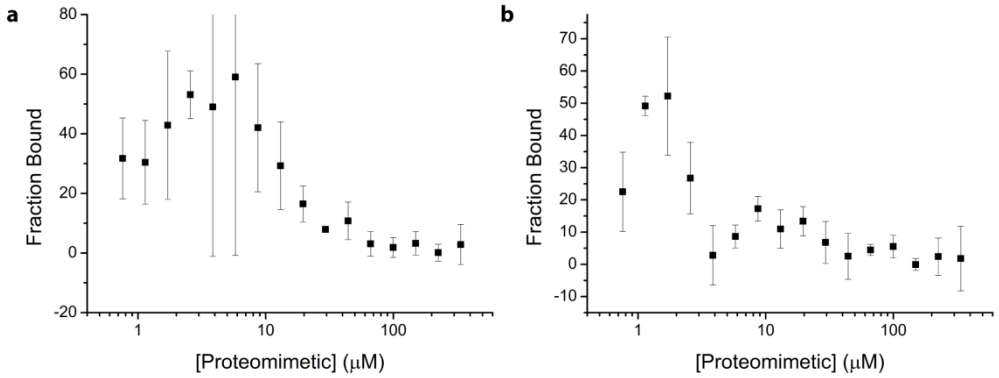


Figure 3.19 Dose-response curves for (**a**) **48**: Fitting not possible (**b**) **49**: Fitting not possible.

The most promising dose-response curve was obtained from trimer **27** shown in **Figure 3.20a**. This oligomer contains only three alkyl side chains and demonstrated an EC_{50} value of 4.66 μM, although there is still some fluctuation in this binding data. The dose-response curves from the Mcl-1/NOXA B assay is directly compared with that from the p53/*h*DM2 assay, shown in **Figure 3.20b**. For both interactions, **27** exhibits relatively good binding indicating poor selectivity for these interactions with this oligomer. As no other binding data was obtained for the Mcl-1/NOXA B interaction, the question of the selectivity of the scaffold has not yet been addressed.

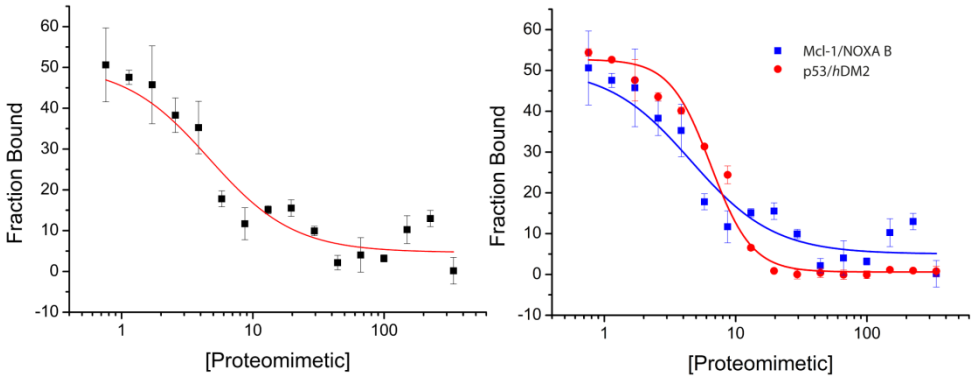


Figure 3.20 (**a**) Dose-response curve for **27**; EC_{50} = 4.66 ± 0.89 μM. (**b**) Comparison of the dose-response curves of **27** against the Mcl-1/NOXA B and p53/*h*DM2 interactions (EC_{50} = 6.56 ± 0.20 μM).

Although having deduced an EC_{50} value for the Mcl-1/NOXA B screen, the data as a whole is not particularly good. The poor binding data is not unique to the 3HABA scaffold as other screening from within the group using this assay has produced similar quality data.

Although the full library was not screened against this interaction, it is clear that the assay is not particularly robust and further screening was stopped.

3.5 Conclusion

In summary, a small library of 16 compounds was assembled to screen against the p53/*h*DM2 interaction and 6 compounds against the Mcl-1/NOXA B interaction. These were synthesised using a microwave assisted SPS methodology reported in *Chapter 2*. Although the methodology is robust with respect to alkyl and aromatic side chains, coupling of functionalised side chains is less efficient and requires purification by HPLC. The yield from this is low and prevented screening of the full library of compounds. Protein titration and peptide displacement assays were carried out for both PPIs before screening of the libraries commenced. Nice binding curves were obtained and several low μM inhibitors were identified when screening against the p53/*h*DM2 interaction (most potent inhibitor is **47**; IC_{50} = 5.19 μM). Data from the Mcl-1/NOXA B screen, however, was too poor to extract reliable binding constants with trimer **27** producing the only data reasonable to fit (EC_{50} = 4.66 μM).

Although library screening of the p53/hDM2 interaction was more successful than the Mcl-1/NOXA B screen, a much larger library would be necessary to make any real conclusions about what produces potent or poor inhibitors. However, some trends in binding affinities were identified and this will lead the way for a second library generation for the p53/*h*DM2 interaction. For example, incorporation of *p*Cl-Bn and Asp mimetics at varying positions. Both these residues apparently improved the potency of the proteomimetic and a position-activity relationship study of such side chains could provide important results. The introduction of natural amino acids as flanking residues on more oligomers could help improve solubility in addition to making key binding interactions. Furthermore, screening a wider range of functionalities with varying chain lengths and including chiral side chains will all add to understanding the restrictions imposed by the protein binding cleft. Future work would, hence, involve building a more diverse library and improve the SPS methodology for the incorporation of functionalised monomers and amino acids.

Chapter 4

Design, Synthesis and Testing of a Novel Bifacial Inhibitor

A synthetic procedure developed in *Section 4.2.1* contributed to work reported in a research article published in EurJOC.[125]

4.1 Introduction

In the development of most conventional α-helix mimetics described to date, a frequent feature or constraint was the need to design inhibitors which mimic side chains occurring on one face of the helix (i.e. at the *i*, *i*+3 (*i*+4), *i*+7 (*i*+8) positions and so on). A recent computational study carried out by the Arora group revealed that 62% of known multiprotein complexes feature a helix at the interface further highlighting the importance of α-helices in PPIs.[15] In this 480 strong set, 60% interact *via* residues on a single face whilst a third contain key binding residues on two faces and approximately 10% require all three faces for interaction with the target protein. Using carefully designed stabilised helices or helical foldamers provides a solid foundation for multifacial inhibitors[1] but due consideration should be given to designing multifacial proteomimetic inhibitors. As discussed in *Chapter 1*, there are a limited number of scaffolds which have been designed to address this challenge.[114, 155-158] On development of work carried out in *Chapters 2* and *3*, it became apparent that potential bifacial mimics could be accessed through modification of the original 3HABA scaffold. A suitable test system on which to develop the new scaffold was needed and the work described in this chapter centres on that goal.

An important group of interactions are the nuclear receptors and their coregulating peptides. This chapter has a focus on the estrogen receptor (ER), a transcription factor which regulates the growth and function of tissues found in the female reproductive system.[22] It has also been associated with a variety of diseases including cancer, cardiovascular disease, obesity and osteoporosis, highlighting the importance of this target for therapeutic intervention.[159] Approximately 70% of breast tumours develop due to the stimulatory effect of estrogens and so anti-estrogen therapies, such as tamoxifen, were developed.[160] These drugs work in a competitive manner with the estrogen ligand, inducing agonistic as well as antagonistic effects depending on the tissue.[161] In breast tissue, tamoxifen binds to the ER preventing the binding of estrogen. This has an antagonistic effect by inducing a change in the shape of the ER, shielding the coactivator binding region and in turn recruiting corepressor proteins; gene transcription and hence growth is halted.[162] Initially these have a significant effect on the reduction of breast cancer; however, in more than 80% of treated women, tamoxifen resistance develops.[163, 164] In the endometrium, tamoxifen acts as a partial agonist thereby increasing the risk of endometrial cancer in some women;[165] an alternative therapy to anti-estrogen drugs is thus needed, which may be found in the ER/coactivator PPI.

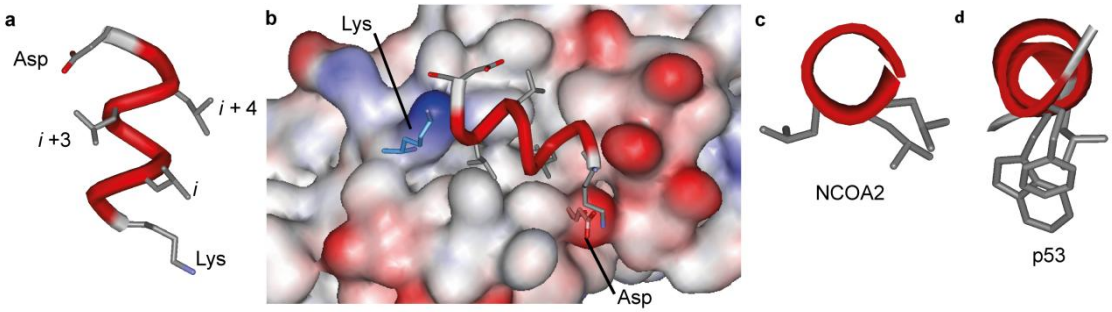


Figure 4.1 (**a**) Helical epitope of NCOA2 showing key leucine residues at the *i*, *i* + 3 and *i* + 4 positions. Also displays Asp and Lys residues that interact with the charge clamp (PDB ID: 2QZO). (**b**) Close up view of how key hydrophobic and charged residues on the NCOA2 interact with the ERα LBD surface (PDB ID: 2QZO). (**c**) Top side view of NCOA2 showing key binding residues lying on two faces (PDB ID: 2QZO). (**d**) Top side view of p53 helix showing key binding residues lying on one face (PDB ID: 1YCR).

Coactivators (p160 protein family) interact with the ERα activation function 2 (AF-2) surface on the LBD *via* small amphipathic α-helical peptide sequences containing a common hydrophobic recognition motif; LXXLL (L = leucine, X = any amino acid) also known as the nuclear receptor box (NR box) (**Fig. 4.1a**). The "X" residues in the NR box are not conserved and may provide a means to design selective inhibitors. X-ray crystal structures also show evidence of the peptide interacting with a charge clamp on the ER surface which could be utilised to further stabilise the complex (**Fig. 4.1b**).[166] Unlike key residues in the p53/*h*DM2 or Bcl-x_L/Bak complexes which are found on a single face of the helix, these key leucine residues are found on two faces as demonstrated in the nuclear receptor coactivator 2 (NCOA2) peptide (**Fig. 4.1c-d**). It has been proposed that these features may allow the ERα-coactivator interaction to be targeted by LXXLL motif-like inhibitors. Expanding on that, designing bifacial proteomimetics to project appropriate functionality in similar orientations to residues at the *i*, *i*+3 and *i*+4 positions is desired.

4.2 Design

Design of the coactivator mimetics builds on previously established work from the group. *Chapters 2* and *3* expand on this work, developing a methodology and screening of a 3HABA based scaffold reproducing the *i* and *i* + 4 (and *i* + 7) residues. The Ahn group had also previously reported the synthesis of a novel amphiphilic variant of this scaffold in which they used an Elbs persulfate oxidation on a related 2-hydroxy-4-nitrobenzoic acid (2HNBA) starting material (**53**) to generate a *para*-dihydroxylated compound (**54**).[167] It was reasoned that using a combination of the original 3-*O*-alkylated (**56** or **17**) and new 3,6-*O*-dialkylated (**55** or **57**) building blocks, residues at the *i*, *i*+3 and *i*+4 positions could be mimicked; similar to those found in the NR box. To illustrate this, we assembled the

3HABA and DHABA building blocks to form new dimer scaffolds **58** and **59**; appropriate functionalisation produced dimers **60** and **61**. These structures differ with respect to the position of the DHABA subunit; in **60** the lower, acid termini contains the DHABA building block (3HABA:DHABA) whilst in **61** the upper, amine termini contains the DHABA building block (DHABA:3HABA).

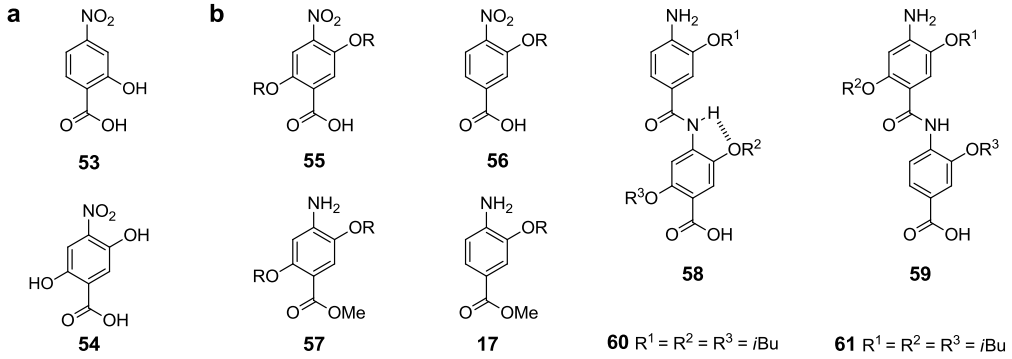


Figure 4.2 (**a**) 2HNBA starting material **53** and new DHABA building block **54**. (b) New DHABA based building blocks **55** (DHNBA) and **57** (DAHB) can be combined with original 3HABA based building blocks **56** (3HNBA) and **17** (3AHB) blocks to generate new scaffolds **58** and **59**; building blocks can be suitably functionalised to form potential ERα inhibitors **60** and **61**.

Molecular modelling was carried out using Maestro as an interface for a number of applications developed by Schrödinger.[168] Initially a conformational search of dimers **60** and **61** was run in which the structures were minimised by employing a *Monte Carlo* search in Macromodel® using the MMFFs (Merk Molecular Force Fields) method. Results of this experiment showed that in **60**, the 3-*O*-alkoxy groups are *syn* whereas in **61**, the 3-*O*-alkoxy groups are *anti* driven by a bifurcated H-bonding system (**Fig. 4.3**). Consequently, two adjacent side chains are displayed on the same face in both dimers and an alternating pattern of side chains was only seen in higher energy structures; ≥ +3.79 kJ mol^{-1} for **60** *Anti* and ≥ +9.92 kJ mol^{-1} for **61** *Syn*. This bias towards a particular conformation is caused by restricted rotation around the Ar-NH bond due to a *S*(5)-intramolecular hydrogen bond; rotation is therefore more likely to occur through the Ar-CO bond. This would also suggest that the presence of the bifurcated-like hydrogen bonding system would also restrict rotation around the Ar-CO bond in **61**. This is corroborated in the molecular modelling with the *syn* conformer of **61** having a greater relative potential energy than the *anti* conformer of **60**. The molecular modelling also infers that the *S*(5)-intramolecular hydrogen bond (2-NH-2-O3) in this system is more stable than the *S*(6)-intramolecular hydrogen bond (2-NH-1-O6) which could be attributable to steric clashes between the two isobutyl chains.

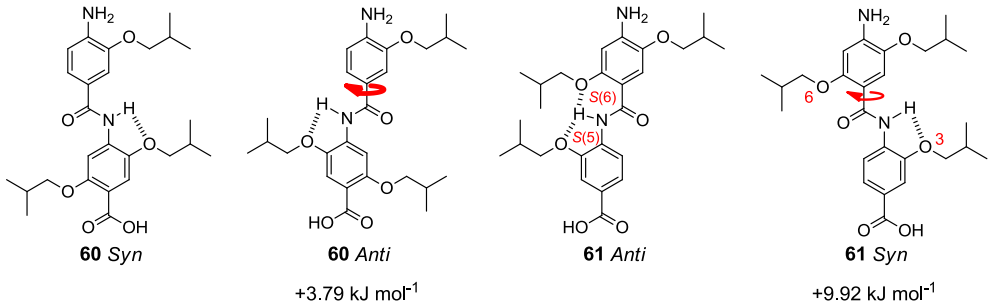


Figure 4.3 Low energy conformations of **60** (*Syn*) and **61** (*Anti*) and possible high energy conformations of **60** (*Anti*) and **61** (*Syn*). S(5) and S(6) hydrogen bonds are highlighted in **61** *Anti*.

The peptide moiety has several double-bonded resonance forms which endow peptides with permanent electric dipole moments. These can line up in secondary structures such as the α-helix, producing particularly large net dipoles which induce strong electric fields and can have an influence on its structure and function.[169] In addition to the dipole moment, NCOA2 contains charged residues which interact with the charge clamp on the ER surface. These residues match up with the dipole moment in that a lysine and an aspartic acid residue is found at the *N*- and *C*-terminus respectively. Dimers **60** and **61** thus have the potential to match the dipole moment of the coactivator peptide whilst also presenting appropriate functionality towards the charge clamp. To determine how accurately the inhibitor designs mimic projection of hydrophobic functionality in the native coactivator peptides, the RMSD values were calculated with respect to the NCOA2 peptide (PDB ID: 2QZO). To calculate the RMSD, three atom pairs were superimposed consisting of the alkoxy oxygen and the α carbon of the leucine residues at the *i*, *i* + 3 and *i* + 4 positions. The orientation of the dimer in relation to the coactivator was also altered. **Figure 4.4** shows the superposition of the lowest energy conformations; **60** lies slightly across the helix and against the dipole moment of the peptide and **61** lies with a similar orientation but in line with the dipole moment of the peptide. The RMSD values for the lowest 1.5 kJ mol^{-1} energy conformations of **60** and **61** range from 1.62-1.70 and 0.94-1.07 Å respectively indicating that **61** is predicted to be a better mimic of NCOA2. Several important features, such as being of reasonable likeness to the native coactivator peptide and containing functionality which may interact with the charge clamp, provided a good starting point for ligand generation and efforts were directed to developing a synthesis. In addition to these features, a robust *O*-alkylation already established within the group would allow incorporation of a wide variety of side-chains to enhance probing of the binding surface.

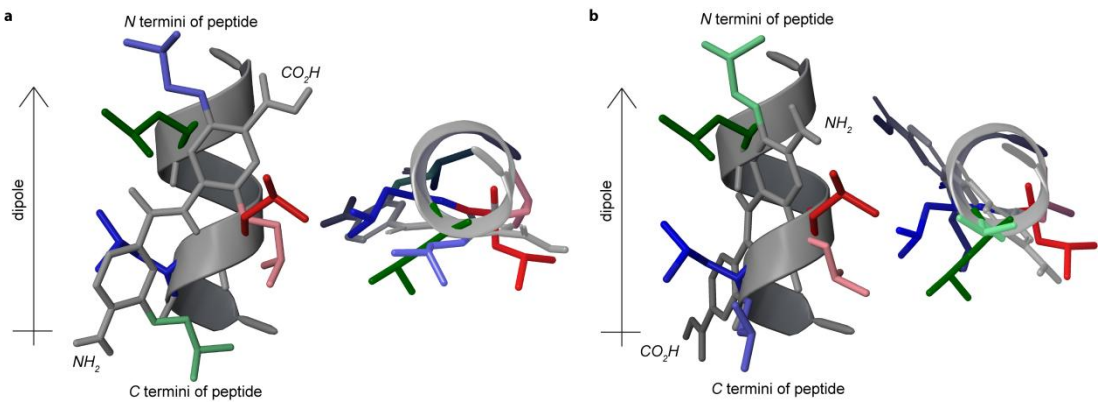
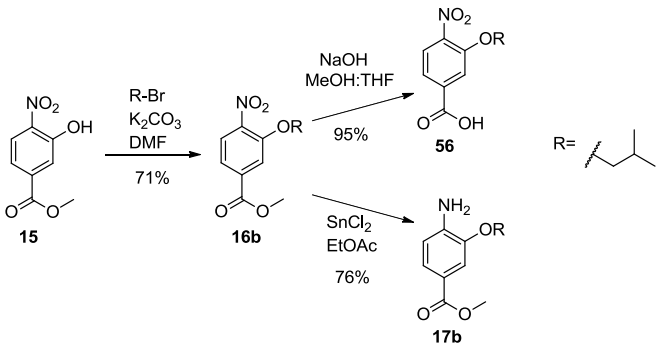


Figure 4.4 (**a**) Superposition of NCOA2 peptide and the low energy conformation of **60** (**b**) Superposition of NCOA2 peptide and the low energy conformation of **61**; (PDB ID: 2QZO).

4.2.1 Synthesis

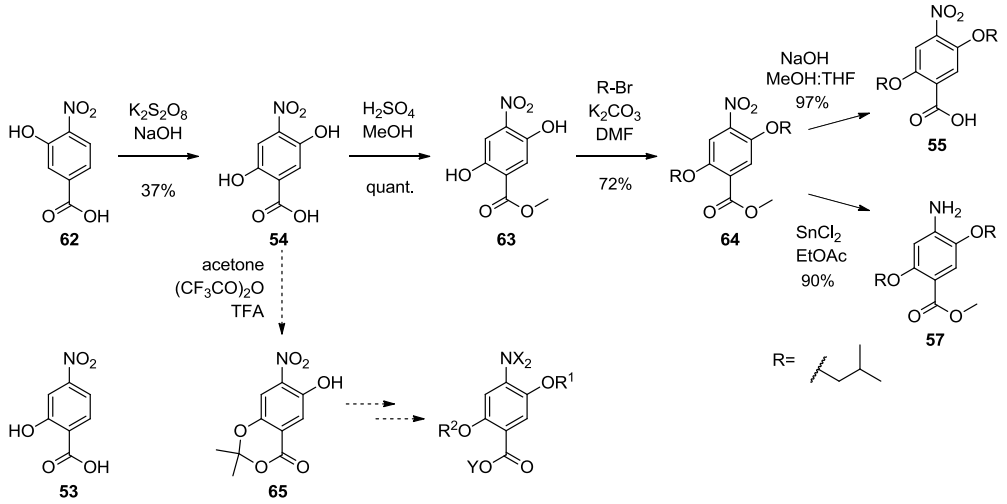
The synthesis of these compounds was achieved using two building blocks, the 3HABA based building blocks (**Scheme 4.1**) and the DHABA based building blocks (**Scheme 4.2**). The monoalkylated synthesis used methods previously developed in the group[114]. Simple alkylation using 1-bromo-2-methylpropane (or an appropriate alkyl halide) generates the alkylated product in good yield. The product from this reaction was then either reduced using tin (II) chloride (to get 3AHB intermediates) or hydrolysed using sodium hydroxide chloride (to get 3HNBA intermediates). These three simple reactions generate the two building blocks needed for the two separate scaffolds.



Scheme 4.1 Synthesis of monoalkylated 3HABA based building blocks: 3HNBA **56** and 3AHB **18**).

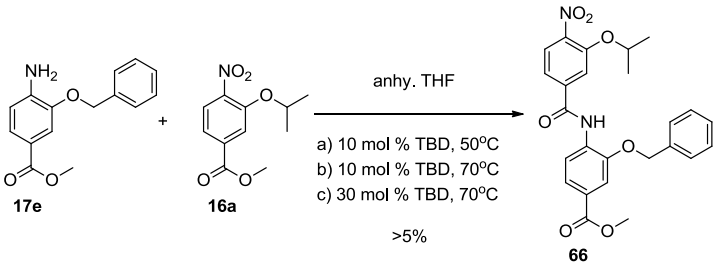
The synthesis of the dialkylated building blocks commenced with a dihydroxylation reaction following a procedure described by the Ahn group[167] starting from 2HNBA **53**. Reproducing this reaction using a cheaper alternative starting material 3-hydroxy-4-nitrobenzoic acid **62** however, results in much easier isolation of the product (**54**). If a homoalkylated product is desired, the acid needs protecting in an esterification reaction and is then alkylated with the corresponding alkyl halide. An important point to note in terms of

targeting alternative PPIs, if a bifacial mimetic is required to display different functionalities, this can be achieved *via* ketal formation between the carboxylate and 2-hydroxy group (**63**) This differentiates between the two hydroxyl groups and further alkylation and deprotection steps can deliver appropriate hetero-dialklated DHABA based building blocks.[124, 125]



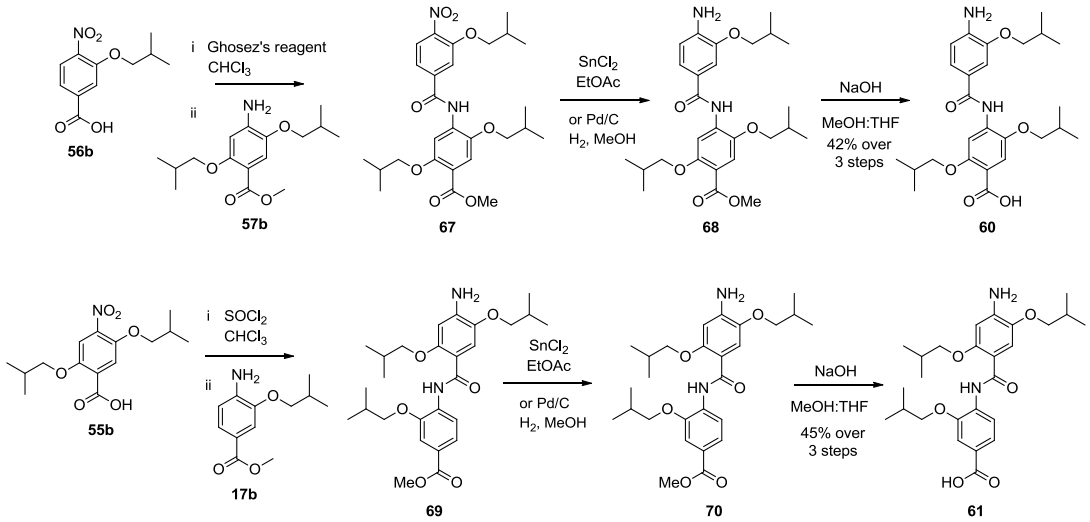
Scheme 4.2 Synthesis of DHABA based building blocks.

The Mioskowski[170] group previously reported aminolysis of esters using the catalyst TBD (1,5,7-triazabicyclo[4.4.0]dec-5-ene); one example reported contained an aniline. Successful aminolysis to **66** (**Scheme 4.3**) would mean that the final hydrolysis step forming **56** and **55** is not necessary. Although the amine was not protected, it was hypothesised that the nitro ester would be more reactive towards aminolysis and so self reaction of **18** should not be seen. Several attempts at this reaction including increasing the temperature and catalyst loading were tried however analysis by LC-MS showed that conversion to the dimer occurred with less than 5% efficiency under all conditions. Due to the very poor reactivity this route was discontinued.



Scheme 4.3 Attempted aminolysis conditions.

Akin to work carried out in *Chapter 2*, identifying if stable acyl chlorides of **55b** and **56b** could be isolated was important. A stable acyl chloride of **55b** was isolated and this carried through coupling, reduction and hydrolysis reactions to form dimer **61**, however, the acyl chloride from **56b** was too unstable to isolate. *In situ* acyl chloride formation with thionyl chloride (and building blocks **56b** and **57b**) in a microwave reactor was then attempted. Varying temperature, length of reaction and equivalents of thionyl chloride failed to result in significant conversion to **67**. Ghosez's reagent was subsequently attempted for use as the acylating agent, leading to successful formation of intermediate **67**. Further reduction and hydrolysis reactions resulted in the desired amino acid dimer **60** (**Scheme 4.4).**



Scheme 4.4 Synthesis of dimers **60** and **61**.

In addition to these dimers, the library was expanded to incorporate a variety of compounds which should have varied inhibitory potential (**Fig. 4.5**). For example, **71** which reproduces the *i*, *i* +4 and *i*+ 7 side chains was included as it was anticipated that it may be too large to fit in the binding pocket and should thus be a poor inhibitor. Both **72** and **73** contain insufficient side chains to mimic the LXXLL motif and should exhibit poor inhibition. **72** however may be a poorer inhibitor than **73** due to a masked acid group which would prevent electrostatic contacts with the charge clamp.[166] **74** may show good or even better inhibition than the original dimers as it may better probe the shape complementarity of the binding cleft due to the larger benzyl side chain. **71** and **72** were synthesised following an analogous solution phase synthesis to those shown in **Scheme 4.4**.[114] Compounds **73** and **74** were synthesised using modified syntheses with Fmoc protected building blocks as shown in **Scheme 4.5**.

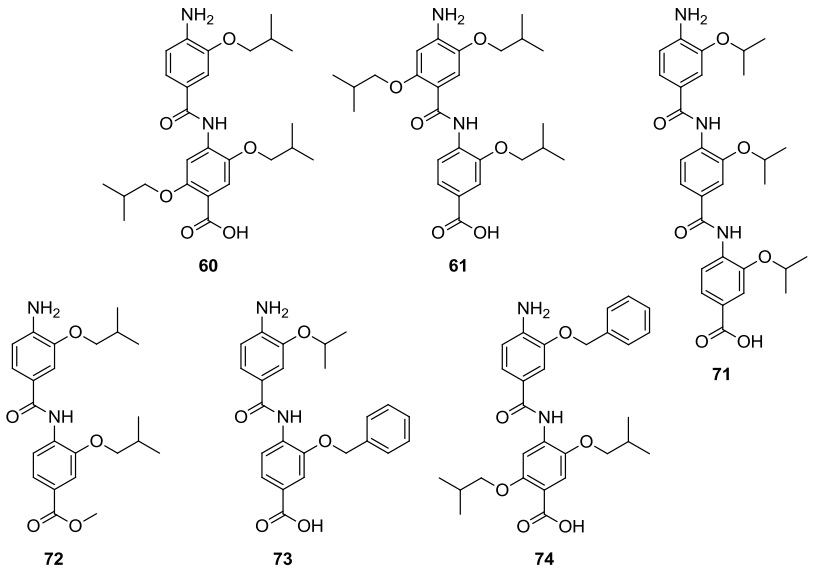
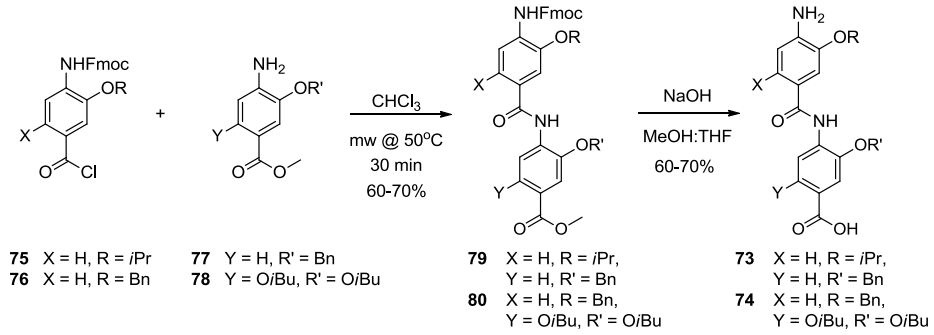


Figure 4.5 A series of compounds designed to investigate inhibition of the ERα/coactivator interaction.



Scheme 4.5 Synthetic outline of dimer formation using Fmoc protected building blocks.

4.2.2 Conformational Analyses

The new dimers were submitted to thorough analyses by NMR spectroscopy in order to determine their conformational preferences in solution; allowing for comparison with previous molecular modelling studies. After assigning all proton peaks using a combination of 1D and 2D spectroscopic techniques, NOESY spectroscopy was used to determine the conformation through interactions between protons through space. The amide proton is key in helping to determine the conformation, and the strength of the NOE may help in understanding how populated the conformations are relative to each other.

Dimer **60** was shown to rotate around the Ar-CO, indicated by strong NOEs between protons *b* and *c* with amide proton *d*; the interaction between *d* and *b*, however, seems to be marginally stronger (**Fig. 4.6**). This is in contradiction to the low energy conformational search as it suggests that both conformations are possible, but the *anti* conformation is potentially more favourable when free in solution. Rotation around the Ar-NH (S5) bond is more restricted. There is an apparent weak interaction between protons *e*

and *d* at 30 mM concentration which disappears in more dilute solutions; however, the cross peaks are also much weaker in the more dilute solutions suggesting this may not be particularly significant interaction. Minimal or absent rotation around the Ar-NH bond is consistent with molecular modelling studies. Dimer **61** was shown to have restricted rotation around both the Ar-CO (S6) and Ar-NH (S5) bonds (**Fig. 4.7**). NOEs between protons *b* and *d* with amide proton *c* are not seen at any concentration, suggesting an *anti* conformation, which is in agreement with molecular modelling studies.

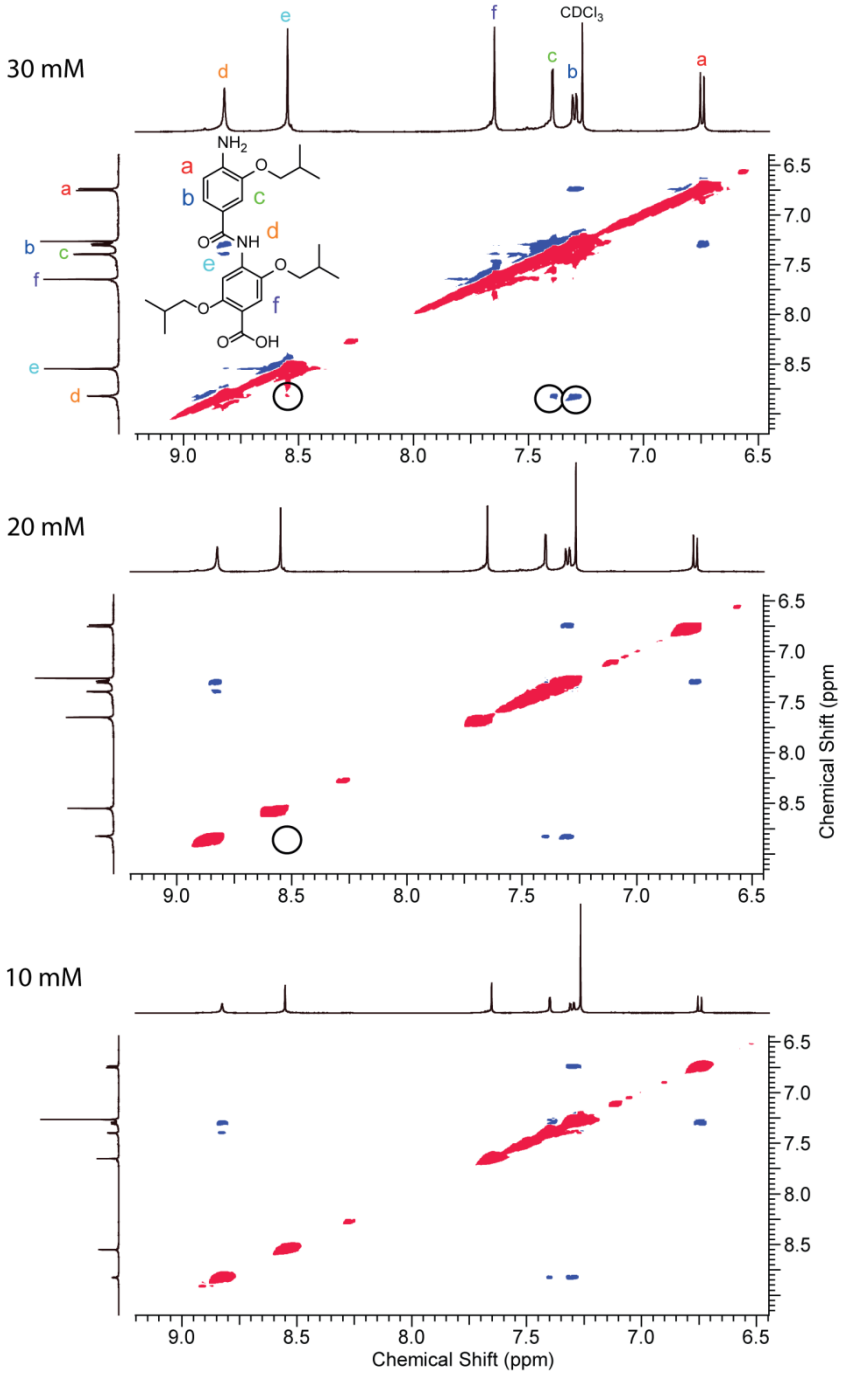


Figure 4.6 NOESY (500 MHz, $CDCl_3$) spectra of dimer **60** at 30 mM, 20 mM and 10 mM concentrations. The 30 mM spectra shows the structure and ^{1}H proton assignments and black circles highlighting any relevent NOEs.

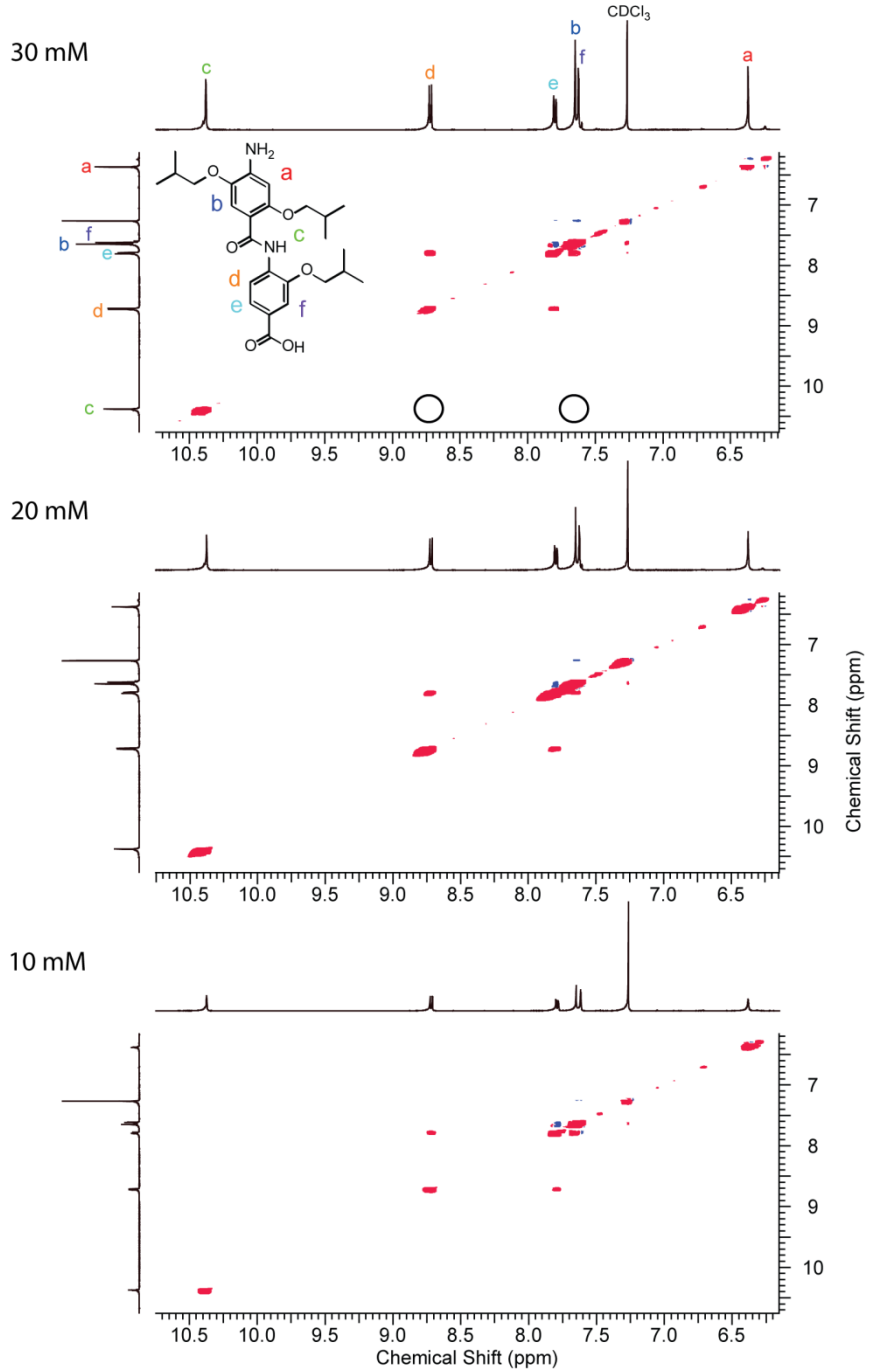


Figure 4.7 NOESY (500 MHz, $CDCl_3$) spectra of dimer **61** at 30 mM, 20 mM and 10 mM concentrations. The 30 mM spectra shows the structure and ^1H proton assignments and black circles highlighting any relevent NOEs.

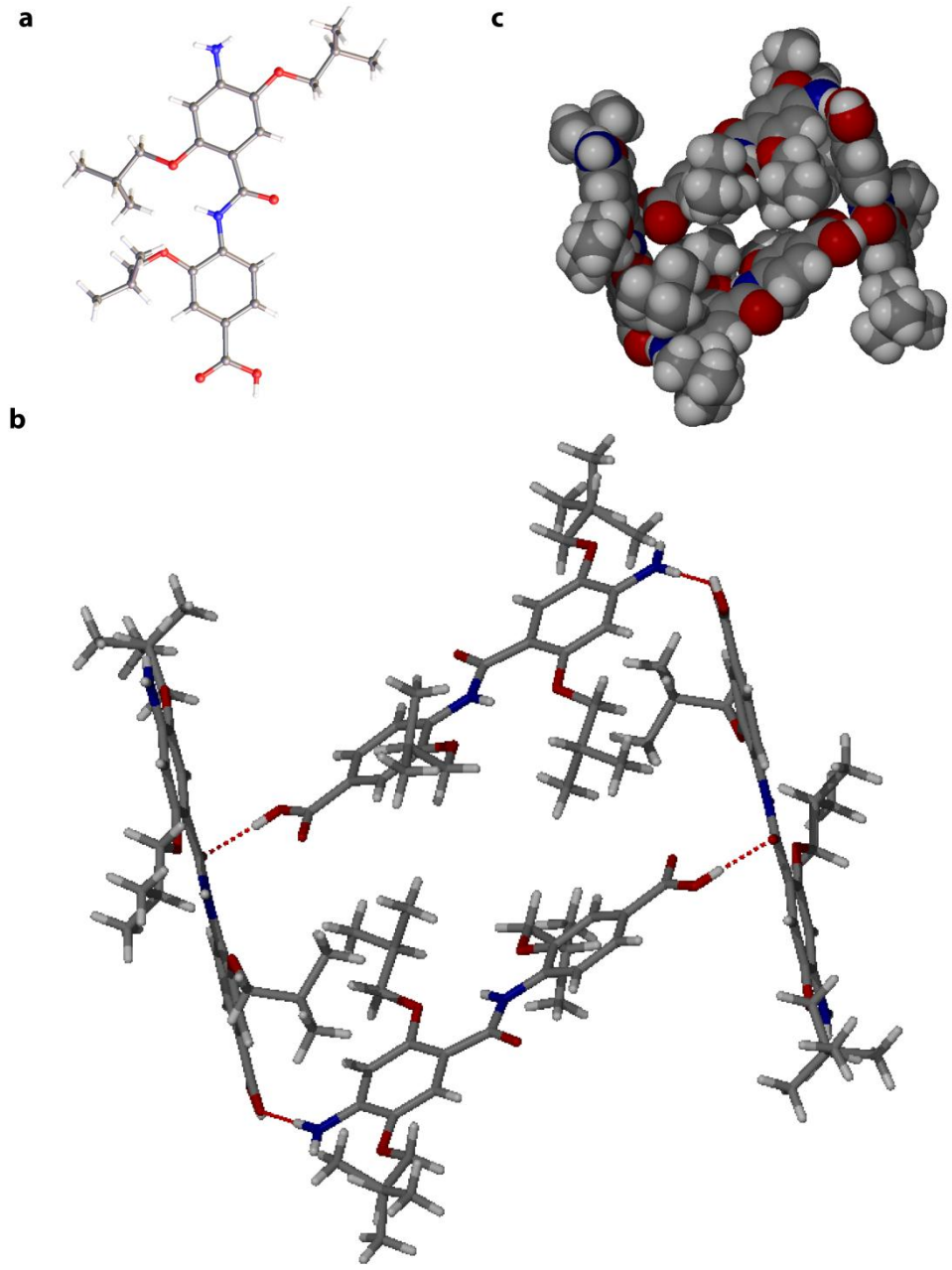


Figure 4.8 (**a**) Single X-ray crystal structure of **61**. (**b**) Cyclic tetramer packing diagram of **61** (stick) showing: aniline (NH) and carboxylic acid hydrogen bonding (CO) and carboxylic acid (OH) and amide moieties (CO) hydrogen bonding (**c**) Cyclic tetramer packing diagram of **61** (CPK) illustrating hydrophobic packing.

Initially, NOESY spectra were obtained in DMSO which is more appropriate to the aqueous conditions experienced during biological testing. Several peaks were indistinguishable in this solvent and so a thorough conformational analysis was not possible, however, comparable results were obtained where possible (See section *6.8*). In addition to solution phase conformational analyses, an X-ray crystal structure of **61** was obtained. Corroborating molecular modelling and NMR studies, **Figure 4.8a** shows the *anti* orientation, with respect to the 3-*O*-alkoxy groups, and confirms a bifurcated like hydrogen

bonding system.[125] The crystal packing of the dimer was also studied and found to be substantially different from crystal packing seen from other 3-HABA based oligomers.[171] Typically packing is seen through side chain-side chain hydrophobic interactions, however, dimer **61** packs with hydrogen bonding interactions between the aniline (NH) and carboxylic acid moieties (CO) and carboxylic acid (OH) and amide moieties (CO) on adjacent molecules. Within the chain packing structure, this forms cyclic tetramers with the hydrophobic side chains aggregating in the centre of the ring. This hydrophobic packing most likely contributes, in addition to the hydrogen bonding, to the observed packing structure (**Fig. 4.8b-c**).

4.2.2.1 H/D exchange

Unlike previous work from the group which has seen isolated *S*(5) and *S*(6) hydrogen bonding systems (as seen in **60**), dimer **61** has a bifurcated *S*(5) / *S*(6) hydrogen bonding system. Interested in quantifying what extent this new hydrogen bonding motif might affect the conformation, *H*/*D* exchange studies were performed. The experiment was performed on dimers **60** and **61** and data from a 2-*O*-alklated dimer **81** is included for direct comparison of an isolated *S*(6) system (**Fig. 4.9**).[171]

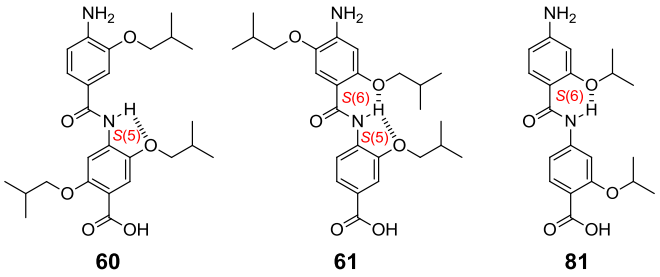


Figure 4.9 Structures of dimers containing *S*(5), *S*(6) and *S*(5)/*S*(6) hydrogen bonding.

The relative rates of a hydrogen/deuterium exchange in a ^{1}H NMR study are affected by a number of parameters. Whilst an acidic proton is anticipated to exchange more rapidly, sterically hindered and strongly hydrogen bonded atoms are expected to have a slower rate of exchange.[172] A 10% $CD_3OD/CDCl_3$ system was employed to ensure pseudo first order kinetics (**Fig. 4.10**) and the kinetic parameters for the compounds are given in **Table 1**.

The half-life for the *H*/*D* exchange studies revealed that the amide proton in **81** *S*(6) exchanges an order of magnitude more slowly than the amide proton of **60** *S*(5) and the amide proton of **61** *S*(5)/*S*(6) exchanges an order of magnitude more slowly than **81** *S*(6) and two orders of magnitude more slowly than **60** *S*(5). Comparable with previous studies, these results suggests a more stable hydrogen bond for the six-membered system relative to the five-membered analogue, and this stability is dramatically enhanced in the bifurcated system.

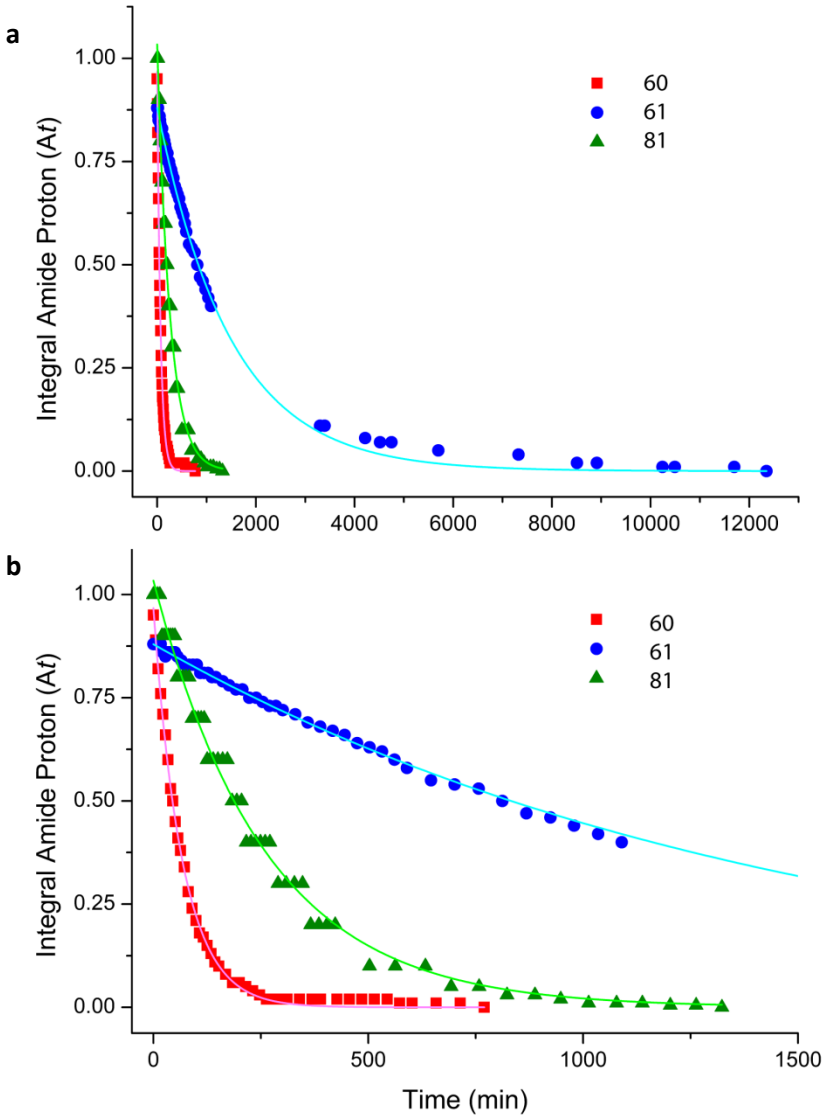


Figure 4.10 (**a**) *H*/*D* exchange kinetics for dimers **60**, **61** and **81**. (**b**) Expansion of results in **a**, better demonstrating the difference between *S*(5) and *S*(6) hydrogen bonding systems.

Table 4.1 Results of kinetic results from *H/D* exchange studies. 'δH of amide proton measure in $CDCl_3$.

Compound	δ NH'	$k_{H/D}$ (min^{-1})	$t_{1/2}$ (min)
60 S(5)	8.26	0.01485 ± 0.00017	46.7 ± 0.5
61 S(5)/ S(6)	10.38	6.7857×10^{-4} ± 0.0000093	1021.5 ± 14
81 S(6)	10.17	0.00387 ± 0.000077	179.1 ± 3.6

4.2.3 Biological Testing

Following synthesis, the compounds were sent for testing by collaborators at the Technische Univeriteit Eindhoven. This was carried out by a fluorescence polarisation competition assay which involved competing off a fluorescently labelled reference peptide based on coactivator SRC1 Box-2 bound to the ERβ LBD by one of the inhibitors. More specifically the ERβ-LBD concentration was constant at 400 nM, estradiol constant at 5 µM, and the fluorescently labeled SRC1Box2 constant at 100 nM. The inhibitor compounds were added in different concentrations across a dilution series (1 mM – 1 nM) however there were problems with solubility for compound **60**, **72** and **74**: measurements could not be performed higher than 50 µM for those compounds (**Fig. 4.11**). The results were slightly disappointing in that no compound came close to disrupting the interaction as successfully as an unlabelled SRC1 Box-2 peptide. Some of the compounds are starting to inhibit in the high µM-range but full curves to determine IC_{50} and K_i were not reached.

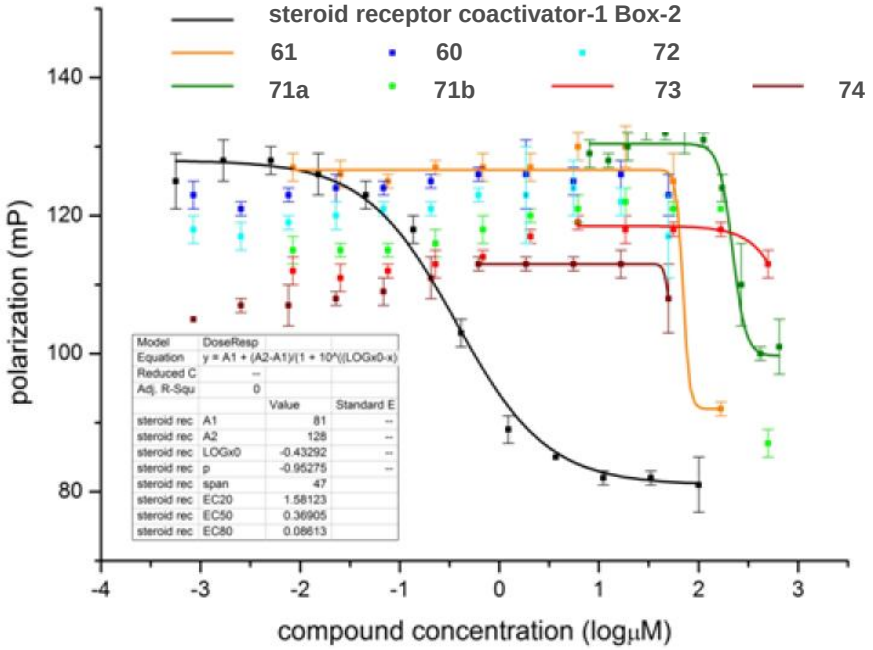


Figure 4.11 Preliminary results for biological testing using a fluorescence polarisation competition assay; **a** and **b** signify that different conditions were used during testing to improve solubility.

4.3 Docking

Interested to see how the inhibitors **60** and **61** might bind, a series of docking experiments was carried out. The LBD of ERα (PDB ID: 3ERD) was prepared for docking using the Prep Wiz function in Maestro.[155] Once the protein was refined, Glide was then used to generate a grid for docking. As the structure is dimeric, only 1 constituent monomer was used within the docking grid (**Fig. 4.12a-b**). Taking into account the size limitations of the grid and the fact that the molecule must lie entirely within it, the position where the coactivator binds was centred in the grid, however, much thought went into choosing an optimal position whilst minimising any bias. **Figure 4.12c** shows the docking grid (orange) surrounding the LBD at different angles. The protein and grid can then be used for all subsequent docking experiments.

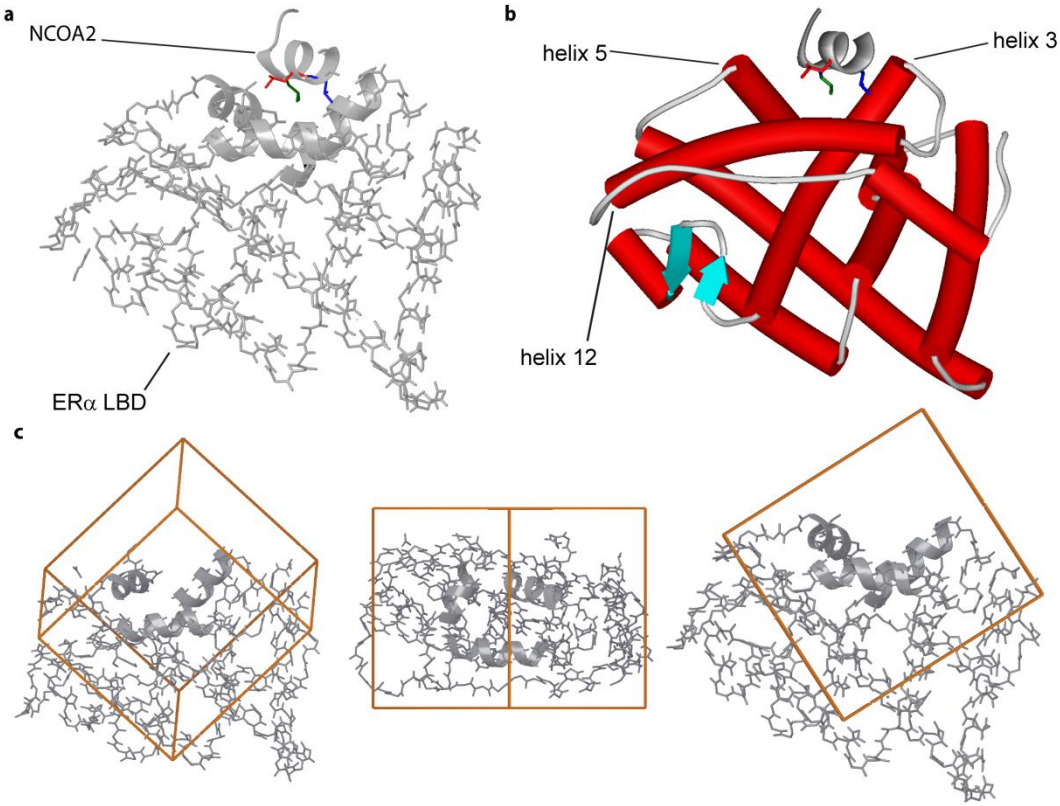


Figure 4.12 (**a**) Figure showing the binding interaction in one constituent monomer of the dimer: the larger structure is the ERα LBD and the smaller helical peptide is the NCOA2. Ribbons on the ERα LBD identify the binding region. (**b**) Figure showing a cartoon representation of **a**. This clearly shows the 3 helices (3, 5 and 12) important in binding. (**c**) Figure showing the grid (orange box) used in the docking experiments from different angles.

The next step is ligand preparation, carried out using LigPrep, which is a program developed by Schrödinger that converts two dimensional structures to 3D. This was used in conjunction with Epik to prepare the dimers and several other proposed ERα inhibitors for

docking; Epik is used to predict the ionisation states of each ligand and the energetic penalties associated with them. Having prepared the "ligands", Glide was then used to dock the results into the prepared protein. Glide is a docking program, by Schrödinger, which uses a unique scoring function and docking protocol to estimate protein-ligand binding affinities. The score or "Glide GScore" is an empirical number calculated by the summation of a number of energetic terms. As a rule of thumb, the more negative the GScore the better. This, however, is explicitly used as a guideline as in practice, compounds with high negative scores do not always interact as well as those with lower negative scores. The predicted binding scores and energies can be used as a guide to assess a ligand's ability to bind to a specified region on the protein surface.

4.3.1 Docking Discussion

In the case of most helix mediated protein-protein interfaces, the helix binding cleft is relatively flat and featureless.[173] This makes generating "acceptable" docking scores (-5 or below) less likely than when compared with docking in an enzyme active site as most of the binding energy results from hydrophobic and lipophilic interactions.[174] Along with other nuclear receptors, ligand binding on the ERα surface occurs through hydrophobic contacts but is reinforced by a charge clamp usually seen between a glutamic acid on helix 12 and lysine on helix 3[166]. Hence, designing inhibitors to exploit these electrostatic interactions presents a valuable contribution.

The original aim of these docking experiments was to gain insight into how the dimers might bind to the ERα surface, in terms of both hydrophobic and electrostatic interactions. The docking studies may have also allowed for a comparison with other compounds designed, or compounds which have been shown, to inhibit the ERα-coactivator interaction. It should have also helped in determining what other features could be exploited in second generation ligands. Initially, it was conjectured that the docking experiment would have an influence on, or be able to change, the conformation of the compounds taking into account the protein surface and would thus, produce the same or similar docking scores. During preliminary docking experiments, however, it was noticed that the GScore value assigned to a particular molecule differed quite substantially. This was an important observation and in order to obtain more conclusive docking results, the lowest 1.5 kJ mol^{-1} energy conformations generated in low energy conformational searches for **60** and **61** were processed by LigPrep and docked by Glide. Several observations were made during this process.

Using **60** as an example, the lowest 1.5 kJ mol^{-1} energy conformations had the basic *syn* (with hydrogen bonding) configuration of **a**. After being processed by LigPrep, this was

modified generating a total of 4 main conformations (**Fig. 4.13a-d**) of which 2 are energetically unlikely based on results from the low energy conformational search and evidenced by NMR analysis. Seemingly, the new conformations assigned after LigPrep did not influence the docking conformation as many changed for a second time on docking; docked poses also bore all 4 conformations. This variation in docking poses per molecule indicates that docking merely the lowest energy conformation is not sufficient and docking a larger number of conformations results in realistic conformations and poses. In future studies, one might consider rigid docking of a low energy conformation or using a core or reference ligand which will restrict the docking within a given RMSD.

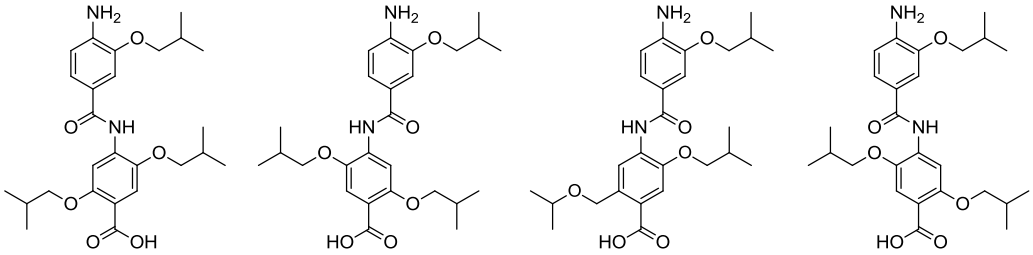


Figure 4.13 Docking conformations of **60**; **a** is the low energy conformation, **b** is higher energy conformation and **c** and **d** are highly unlikely docking poses.

On analysis of **60**, it was seen that 13 out of 89 docked poses had the ligand in its low energy conformation (**Fig. 4.13a**) with GScores ranging from -2.753 to -0.251 (Full GScore range from -3.168 to +0.754). This docking pose displays a good overlay with the native coactivator (**Fig. 4.14a**) but against the dipole moment of the coactivator peptide, akin to RMSD calculations. 16 out of 39 poses possessed the low energy conformation for **61**. The best docked pose retains the low energy *anti* conformation and has a GScore of -3.707 and values range from -3.707 to -0.921 (**Fig. 4.14b**). Although receiving a better GScore, the docking pose of **61** exhibits poorer mimicking of the coactivator peptide; only two of the side chains overlay nicely with the leucine residues and this also lies against the dipole moment of the coactivator. **Figure 4.14c** shows how **60** and **61** lie with respect to one another, in an almost mirror image like fashion. A number of equivalent docking studies were carried out with several ERα inhibitors to gain wider knowledge of available contact points on the receptor surface. α-Helix mimetic scaffolds pyridyl-pyridone[175] **82** and pyrimidine[109] **83** and HTS identified inhibitor, TPBM [176] **84** were studied (**Fig. 4.15**). The docking poses were scrutinised and realistic docking poses were identified.

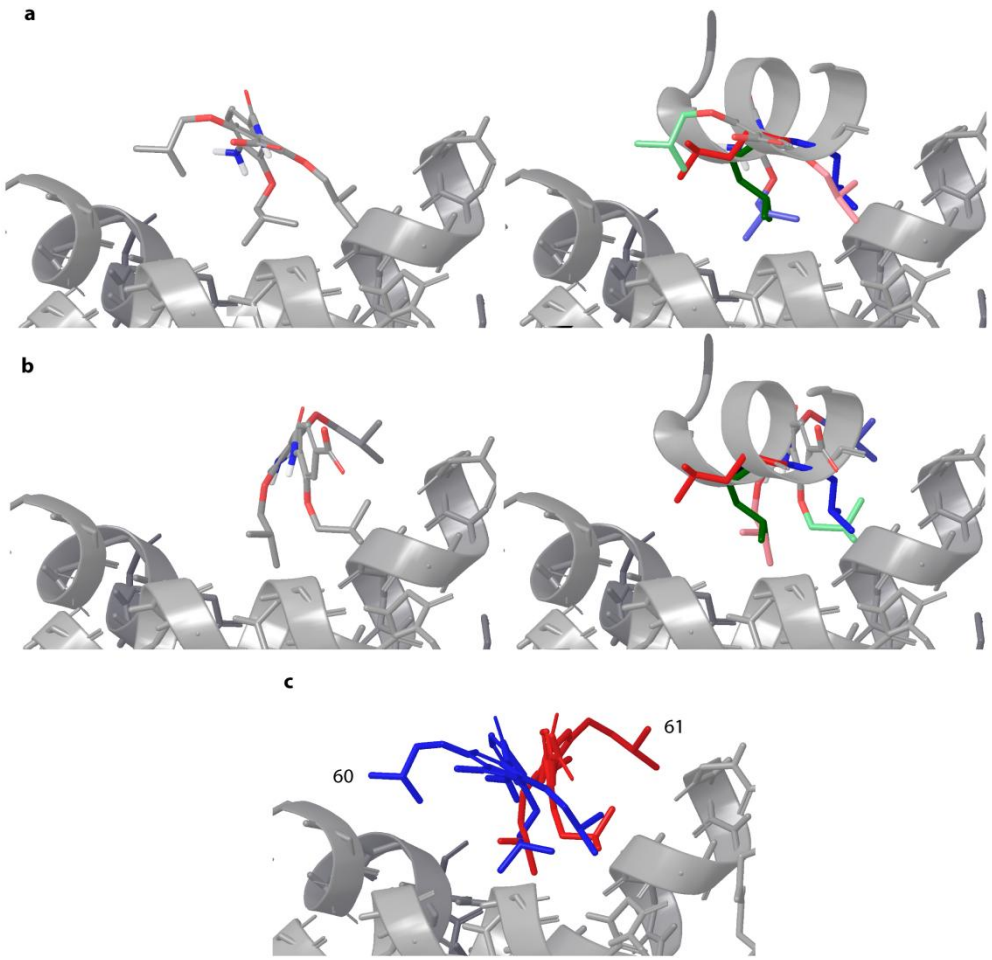
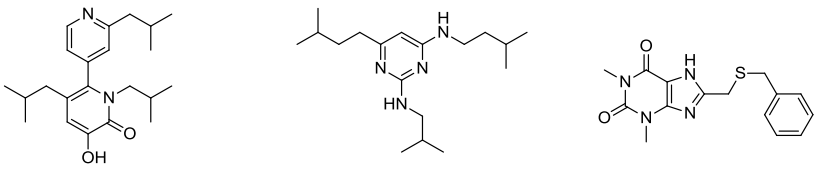


Figure 4.14 Docking poses of best scoring low energy conformations (**a**) *left*: **60**, *right*: overlay of coactivator peptide and **60**, (**b**) left: **61**, *right*: overlay of coactivator peptide and **61**. (**c**) overlay of **60** and **61**; (PDB ID:3ERD).



82 **83** **84**

Figure 4.15 ER/coactivator interaction inhibitors: pyridyl-pyridone **82**, pyrimidine **83** and TPBM **84**.

The XP visualisation tool breaks down the GScore and shows how the value is generated in energetic terms. These terms are summarised in **Table 4.2**. The XP visualisation tool also shows the docking pose including any hydrogen bonds and their lengths (Summarised in **Table 4.3**). The GScores of the dimers, most similar in value to that of the pyridyl-pyridone scaffold **82** (-2.615), are not usually considered successful docking scores but are more typical of protein-protein interfaces.[174] In contrast, GScores from **83** and **84** are found within the "acceptable" region and verging on docking scores most associated with enzyme active

site docking. Breakdown of the docking score and consideration of how the compounds dock provides reasoning for this.

Table 4.2 Summary of energetic terms (XP docking) from compounds **60**, **61**, **82-84**; ' rewards," penalties.

Best docked conformations of compounds					
Energetic Term	**60**	**61**	**82**	**83**	**84**
GScore	-2.753	-3.707	-2.615	-7.516	-4.860
Lipophilic pair term'	-2.2	-2.9	-2.1	-2.1	-2.1
Hydrophobic enclosure'	-0.2	-0.2	0.0	-0.2	0.0
Hydrophobically packed Hbond'	0.0	0.0	0.0	0.0	-0.8
Hbond pair term'	-0.9	-0.9	-1.0	-2.0	-1.7
Electrostatic rewards'	-0.3	-0.5	-0.1	-3.4	-0.7
Low MW'	0.0	0.0	0.0	-0.5	-0.4
Exposed hydrophobic ligand groups"	+0.4	+0.4	+1.7	+0.3	+0.5
Rotatable bond"	+0.3	+0.3	+0.2	+0.3	+0.5
Epik State penalty"	0.0	0.0	0.0	+1.6	+0.1

Table 4.3 Summary of potential hydrogen bonding between inhibitors **60**, **61**, **82-84** and ERα LBD.

Compound	H-Bond	Amino Acid (AA)	Length (Å)	AA Number
60	1	Lys	1.777	362
61	1	Lys	1.686	362
82	1	Lys	1.892	362
	2	Lys	2.132	362
83	1	Glu	1.490	542
	2	Glu	2.026	542
	3	Glu	1.726	380
84	1	Gln (CO)	1.796	375
	2	Gln (NH_2)	2.025	375
	3	Lys	2.084	362

Comparing the electrostatic terms and predicted hydrogen bonding of **60** and **61**, as demonstrated in **Figure 4.16a**, one can see that both make a single hydrogen bonding contact with the receptor surface (same *Hbond pair term*). The length of the Hbond in **61** is shorter and therefore stronger than that in **60**; this correlates with a slightly higher *electrostatic reward*. **83** and **84** have much more favourable electrostatic terms; the *hydrogen bond pair term* and *electrostatic rewards* for pyrimidine are particularly noteworthy with a total reward of -5.4 versus -1.4 for **83** and **61** respectively. The uncharacteristic docking scores for the ERα are owing to compounds making a number of electrostatic interactions with the receptor surface. It is evident that there are other residues on the surface that the inhibitors can interact with in addition to the residues found in the charge clamp. Both **83** and **84** benefit from more extensive hydrogen bonding interactions with Gln (375) and Glu (542). **Figure 4.16b** identifies all residues on the receptor surface which form hydrogen bonds with docked compounds **60**, **61**, **82-84**. Both dimers can successfully make hydrogen bonding interactions with the lysine in the charge clamp and in several low scoring poses, interactions with both elements in the charge clamp are possible. In these cases, poor placement of side chains result in hydrophobic penalties making hydrogen bonding with both elements of the charge clamp less likely. This demonstrates a fine balance between projecting the key hydrophobic functionality as well as key polar functionality with a correct orientation to maximise interactions with the hydrophobic cleft and charge clamp respectively.

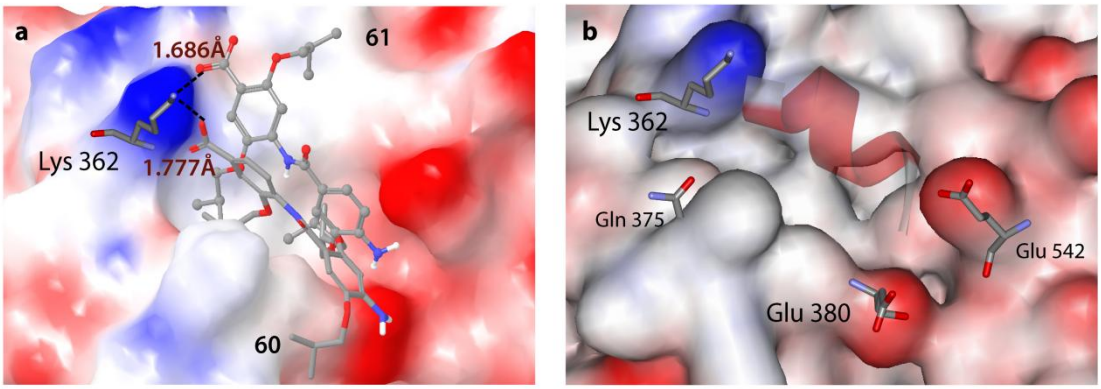


Figure 4.16 (**a**) Figure showing how **60** (stick) and **61** (ball and stick) dock on the ERα LBD including the hydrogen bonds with lysine. (**b**) Figure showing charge clamp residues Lys 362 and Glu 380 and other functionalised surface residues available for hydrogen bonding (Gln 375 and Glu 542).

4.3.2 Redesign

Several modifications to the scaffold have been envisioned to produce a second generation of inhibitors. The first set of modifications has been designed to better probe the hydrophobic binding surface. This involves generating a variety of new building blocks, which when mixed with new and existing building blocks can produce a wide range of new scaffolds. The various building blocks are shown in **Figure 4.17**. Dimers **60** and **61**, synthesised using 3HABA (**85**) and DHABA (**86**), can be modified by changing the position of side chain in the monoalkylated building block from the 3-*O*- to the 2-*O*- alkoxy position to produce **87** (2-HABA). Modification to the central amino acid character of the building blocks can produce para-dianiline (**88** and **89)** and para-diacid (**90** and **91**) constructs. Incorporation of a monoalkylated building block affords greater variety, as demonstrated in positioning of the alkyl chain at the 2- or 3- position. Removing a side chain may help to increase solubility by having less hydrophobic groups, however molecular modelling on each variant would be needed to decide which compounds to proceed with. Combining the new para-dianiline and para-diacid building blocks together, as well as with the original "amino acid" building blocks, could furnish up to 20 scaffolds.

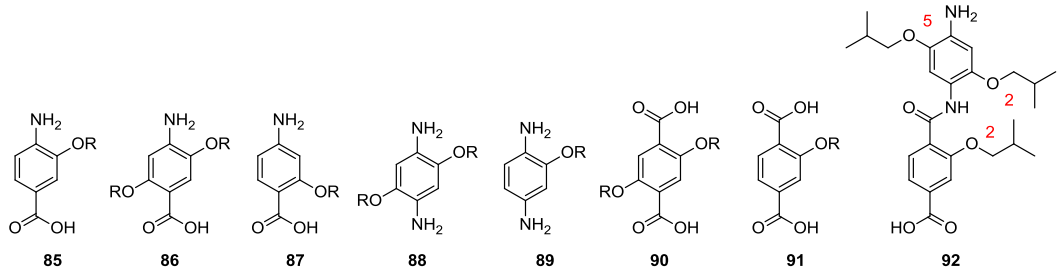


Figure 4.17 Potential building blocks for second generation scaffolds.

Scaffold **92** was submitted to preliminary molecular modelling studies to understand how these modifications may change the inhibitory potential. A low energy conformational search was carried out as discussed previously, and it was found that the 2-*O*-alkoxy groups were displayed *syn* to one another, as demonstrated in **Figure 4.17**. The RMSD values of the lowest 1.5 kJ mol^{-1} energy conformations, with respect to the NCOA2 peptide, were calculated. These values ranged from 0.59-0.62 Å, which is significantly better than the RMSD values for **60** and **61**. To illustrate this, **Figure 4.18a** shows a very good correlation between the projection of the side-chains in **92** and the NCOA2, which was not seen in earlier superpositions. Carried out using the Maestro interface as above, **92** was docked into the ERα LBD using Glide. Further demonstrating the inhibitory potential of **92**, **Figure 4.18b** shows an overlay of docking results with the native peptide. One can clearly see that the projection of the alkoxy side chains mimics the leucine side chains extremely well,

suggesting that these new scaffolds could present better inhibitors of the ERα /coactivator interaction. In addition to presenting a good mimetic, the dipole moment of **92** is in line with the dipole moment of the peptide and docking studies suggest the amine and acid moieties are positioned appropriately to make contacts with the charge clamp. In order to synthesise these new compounds, a new strategy containing unfamiliar chemistry would have needed to be developed. This was not progressed due to time constraints but has been taken on by another member of the group.

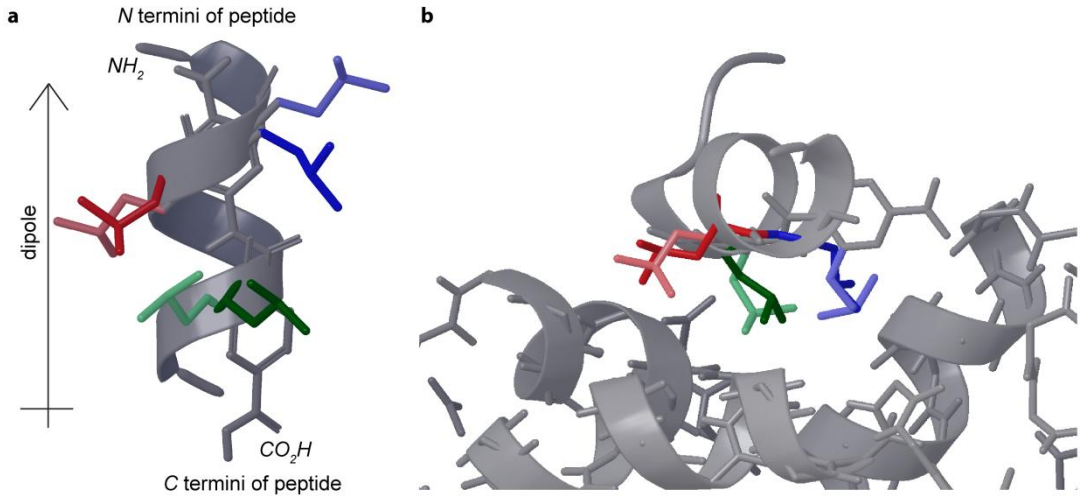


Figure 4.18 (**a**) Superposition of NCOA2 peptide and the low energy conformation of **92**. (**b**) Overlay of docked **92** and NCOA2 (PDB ID:3ERD).

A second set of potential modifications is to maximise electrostatic interactions with charged residues. This would involve putting residues such as aspartic acid and lysine onto the acid and / or amine termini of a dimer, demonstrated in **Figure 4.19** with **60** derivatives. With knowledge from initial molecular modelling studies, attaching a lysine onto the *N*-termini could have dramatic improvements on docking scores. This would essentially extend the molecule taking advantage of a newly positioned amine group to make hydrogen bonding interactions with the Glu of the charge clamp or another Glu residue situated nearby. The Lys amine is also protonated at physiological pH, unlike the scaffold aniline, which would also allow for stronger interactions. In addition to making more favourable interactions, addition of such residues may also increase the solubility of the compounds. As mentioned in the biological testing, this is an important issue to be addressed if good inhibition is to be achieved.

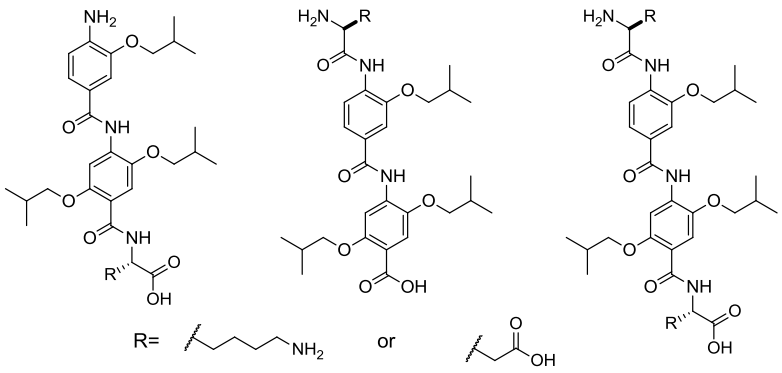
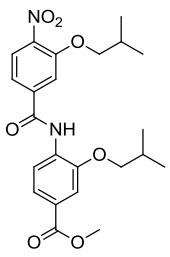


Figure 4.19 Second generation modifications to: make or improve interactions with the charge clamp; increase solubility of dimers.

4.4 Conclusion and Future Work

To conclude, two new scaffolds have been designed, modelled, synthesised and tested. Due to poor solubility and biological data, an extensive docking study was carried out leading to a second generation of ligands. Significant knowledge has been gained in the docking studies which will make replication of docking experiments for new scaffolds rapid.

Unlike other steroid receptors the androgen receptor (AR) utilises multiple mechanisms to activate gene transcription. The AR is able to bind coregulators through a canonical AR AF-2 domain *via* a restricted set of LXXLL coregulators, however, it preferentially binds to phenylalanine-rich motifs exhibiting a FXXLF (FXXFF) sequence in the AF-1 domain.[177] Due to the almost identical projections of these side chains, it may be possible to develop inhibitors relevant to the AR/coactivator interaction by changing the side chain of the already existing scaffolds. A recent study by Raj and co-workers[178] demonstrated that a simple nitro / ester 3HABA based intermediate **93** (**Fig.19**), successfully targets the AR/PELP1 (LXXLL) interaction with an IC_{50} of 40 nM, indicating the scaffolds could have potential inhibitory effects with alternative PPIs. However, in our own hands, similar scaffolds have failed to demonstrate selectivity or exhibit nanomolar inhibition of any helix mediated PPI



93

Figure 4.20 Nitro / ester 3HABA based derivative **93** shown to inhibit the AR/PELP1 interaction.

Chapter 5

Thesis Summary and Future Directions

5.1 Thesis Summary and Future Directions

As almost every biological process in the cell involves interactions between proteins, it is liable that a mutation during replication or malfunction in a biological process will eventually occur that results in the onset of disease: for example mutations in oncogenes can result in cancers, and viruses such as HIV contain enzymes which disrupt regular cellular activity. Secondary structures within proteins help characterise the interface between proteins and as the α-helix is the most abundant secondary structure in proteins, it is reasonable that this structure mediates a number of PPIs. Several helix mediated PPIs have been implicated in the development of the aforementioned diseases and have accordingly become an important focus of scientific research. Traditional small molecule therapeutics have resulted in some of the most potent inhibitors of PPIs, however, tailoring inhibitors for each PPI is not attractive as a general strategy for their inhibition. Other approaches have sought to reproduce the helical structure by stabilisation of the helical peptide (constrained peptides and foldamers) or by recapitulating the key binding residues of the helix on a core non-peptidic scaffold (proteomimetics). Both strategies have produced a variety of scaffolds, disrupting many distinct PPIs, however, a true generic scaffold is still desired.

The Wilson group developed a 3-*O*-alkylated oligobenzamide scaffold which replicated the spatial projection of residues at the *i*, *i*+3 (*i*+4) and *i*+7 (*i*+8) positions: classified within the proteomimtic strategy. When screened against the well characterised p53/*h*DM2 interaction, several low micromolar inhibitors were identified. In order to better understand what necessitates strong binding, a library of compounds, containing varied functionalities, to carry out SAR studies is a favourable approach. In this context, a SPS is described herein which discusses the synthesis of a range of Fmoc protected 3HABA building blocks, amenable to standard Fmoc SPS procedures, and the development of an automated, microwave assisted SPS methodology. In the study, various techniques to activate monomers with different side chains were developed. The most successful was stable acyl chloride formation of monomers with alkyl and aromatic side chains. Oligomers synthesised using the acyl chlorides were obtained in high yields (70-99 %) and in over 95 % purity in most cases simply after cleavage from the resin and precipitation. Monomers with acid labile protecting group generally required *in situ* acyl chloride formation due to them being highly unstable, requiring purification by HPLC. In future studies, methods for their purification will need to be addressed to obtain higher yields. Oligomer assembly was compatible with a range of natural and unnatural functionalities and a hexamer was successfully synthesised and characterised in high purity. Molecular modelling studies demonstrated the potential for longer oligomers to mimic extended helices such as those found in the gp41 hexameric coiled coil assembly. Ultimately, a rhobust SPS method for oligomer formation of deactiviated anilines was developed and is an important advance in the field.[126] There are currently no other SPS procedures for the scaffold which have as wide substrate

tolerance and can produce oligomers in high yields or purities sufficient for biological testing[150] and as such, the methodology was published in a major chemistry journal.

After assessing the scope of the methodology, an SAR oligomer library was constructed. This was largely focused on designing potential inhibitors of the model p53/*h*DM2 interaction, and screening was achieved *via* fluorescence anisotropy competition assays against this PPI. The SAR study produced some promising data, with good dose-response curves and some apparent trends in reactivity across the series. Two significant areas for future development are a larger SAR or position-AR study for the *p*Cl-Bn side chain and for the incorporation of acid containing side chains through 3HABA monomers of natural amino acids (Asp or Glu): trimers containing these functionalities exhibited improved potency compared to the most potent inhibitor in the original study.[116] The positive results in acid containing oligomers support earlier studies that inhibitors for this interaction could benefit from containing more functionality than simply mimicking the Phe19, Typ23 and Leu26 hydrophobic residues.[135, 158] Future studies should focus on further methodology development to incorporate such amino acids on the *N*-terminal position on trimers: this will not only improve future SAR studies but will also lead the way for incorporation of oligomers into proteins. A handful of oligomers were designed/identified as potential inhibitors of the Mcl-1/NOXA B interaction and a further set of competition assays were carried out against this PPI. Most data from this study was not reasonable to fit, however, a low micromolar inhibitor of this interaction was identified. Without other binding data, this has little implication and much larger SAR studies will be necessary to better understand binding requirements of this PPI. Future SAR studies can also help to determine the selectivity of the scaffold with respect to a variety of PPIs as competition assays for other helix mediated PPIs are currently being developed in the Wilson group. Further biophysical studies might include ITC which would help characterise these interactions, proteolytic stability and cell penetration studies of oligomers and obtaining crystal structures of oligomers bound to the proteins.

Over 40 % of helix mediated PPIs involve key binding residues located on more than one face, however, there are very few scaffolds in the literature designed to mimic more than one face of the putative helix.[108, 112, 118, 124] Extending our approach to mimic multiple faces, the 3-*O*-oligobenzamide scaffold was modified, resulting in two novel bifacial dimer scaffolds as inhibitors of the ERα/coactivator interaction. The development of the scaffold is discussed from initial molecular modelling, synthesis, structural characterisation and biophysical studies from our collaborators. Subsequent computational docking studies were carried out to develop more potent inhibitors and this led to second generation ligands with improved *in silico* properties. Although the project is in its infancy, the scaffolds have shown potential as bifacial mimics and there are still many features which can undergo optimisation. In addition to the *i*Bu side chains discussed, the new scaffolds can be synthesised incorporating a wide range of functionality to carry out more involved SAR studies. Docking studies also identified numerous charged amino acid residues surrounding the binding cleft,

including the charge clamp Asp and Lys residues, and future work should involve incorporation of acid and base moieties to exploit these favourable electrostatic interactions. It is anticipated that biophysical evaluation of these second generation compounds will identify potent inhibitors of the ERα/coactivator interaction and modification of side chains could also result in inhibitors of other steroid receptor/coactivator complexes (such as the AR/coactivator interaction).

On a wider scale, in addition to the 3-*O*-alkylated scaffold and the bifacial scaffolds investigated in this thesis, two other scaffolds (2-*O*-alkylated, *N*-alkylated) are undergoing biophysical evaluation as potential helix mimetics within the group. Biophyscial testing of hybrid oligomers constituting the 3-*O*, 2-*O*-, *N*-alkylated scaffolds and natural amino acids is currently under investigation and it is envisoned that a thorough SAR study involving these and possible incorporation of DHABA building blocks, creating bifacial mimics, could lead to inhibitors of helix mediated PPIs with improved biophysical properties (such as selectivity, potency and proteolytic stability).

Chapter 6

Experimental Section

General Experimental Points

All reagents were obtained from commercial sources and used without further purification unless otherwise stated. All solvents used were HPLC grade. Dry solvents were bought from Aldrich. Analytical TLC was performed using 0.2 mm silica gel 60 F_{254} pre-coated aluminium sheets (Merck) and visualised using UV irradiation or, in the case of amine intermediates, by staining with a ninhydrin solution. Flash column chromatography was carried out on silica gel 60 (35 to 70 micron particles, FluoroChem). Solvent ratios are described where appropriate. Solvents were removed under reduced pressure using a rotary evaporator at diaphragm pump pressure. Samples were freed of remaining traces of solvents under high vacuum. ^{1}H and ^{13}C NMR spectra were measured on a Bruker DPX300 or a Bruker Avance 500 spectrometer using an internal deuterium lock. Chemical shifts are reported in parts per million (ppm) downfield from TMS in δ units and coupling constants are given in hertz (Hz). Coupling constants are reported to the nearest 0.1 Hz. TMS is defined as 0 ppm for ^{1}H NMR spectra and the centre line of the triplet of $CDCl_3$ was defined as 77.10 ppm for ^{13}C NMR spectra. When describing ^{1}H NMR data the following abbreviations are used; s = singlet, d = doublet, t = triplet, q = quartet, m = multiple, br. = broad, app. = apparent. Melting points were determined using a *Reichert Austria melting point microscope* and are uncorrected. Microanalyses were obtained on a Carlo Erba Elemental Analyser MOD 1106 instrument, found composition is reported to the nearest 0.05%. Infrared spectra were recorded on a Perkin-Elmer FTIR spectrometer and samples analysed as solids (unless otherwise stated). Mass spectra (HRMS) were recorded in-house using a Micromass GCT Premier, using electron impact ionisation (EI) or a Bruker Daltonics micrOTOF, using electron spray ionisation (ES). LC-MS experiments were run on a Bruker Daltronics esquire™ series spectrometer, samples ionised by electrospray and analysed by a quadrupole ion trap mass spectrometer. All experiments were run through za C18 column on an acetonitrile/water gradient (typically 0-100% acetonitrile over 3 minutes). Intermediates **16a, 16e, 16h, 17a** and **17e** have previously been described.[114, 116, 155]

6.1 General Procedures

6.1.1 Monomer Synthesis

Procedure A (RBr Alkylation) To a stirred solution of methyl-3-hydroxy-4-nitrobenzoate **15** (1 eq.) and potassium carbonate (3 eq.) in dimethylformamide (20 mL / g) is added RBr (1.2 eq.) and the resulting mixture stirred at 50 ºC overnight under a nitrogen atmosphere. The resultant red liquid is allowed to cool and poured into water (40 mL / g) and extracted with ethyl acetate. The combined organic fractions are thoroughly washed with water and further washed by brine, dried (magnesium sulfate), filtered and evaporated to dryness.

Procedure B (*Mitsunobu*) A stirred solution containing methyl-3-hydroxy-4-nitrobenzoate **15** (1 eq.), ROH (1.1 eq.) and triphenylphosphine (1.5 eq.) in anhydrous tetrahydrofuran (30 mL / g) is cooled to 0 °C. Diisopropyl azodicarboxylate (1.5 eq.) is added and the resulting solution allowed to warm to room temperature and left stirring overnight under a nitrogen atmosphere. Organic solvents are removed under reduced pressure and the product is purified *via* column chromatography.

Procedure C (Tin Reduction) To a stirred solution containing either i) nitro/ester or ii) nitro/acid (1 eq.) in ethyl acetate (20 mL / g) tin(II) chloride dihydrate (6 eq.) is added and the resulting mixture stirred at 50 ºC overnight, under a nitrogen atmosphere (with a calcium chloride drying tube attached). On completion, the reaction mixture is allowed to cool and poured over ice. The pH is made slightly basic (~pH 8) by addition of a saturated sodium bicarbonate solution and the resulting basic mixture is allowed to stir for an hour. The aqueous mixture is extracted with ethyl acetate and the combined organic fractions washed thoroughly with brine, dried (magnesium sulfate), filtered and evaporated to dryness

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

Procedure E (Cobalt assisted reduction) To a solution containing either i) nitro/ester or ii) nitro/acid (1 eq.) in methanol, cobalt (II) chloride hexahydrate (3 eq.) is added and allowed to dissolve. Sodium borohydride (6 eq.) is added slowly and allowed to stir for 30 minutes. A further portion of cobalt (II) chloride hexahydrate (3 eq.) and sodium borohydride (6 eq.) is added as above. The solution is filtered through a celite pad which is washed with dichloromethane and the filtrate thoroughly washed with 1M hydrochloric acid to remove the metal. The organic solvents are washed with brine, dried (magnesium sulfate), filtered and evaporated to dryness.

Procedure F (NaOH Saponification) To a solution containing either i) amine/ester or ii) nitro/ester (1 eq.) in a 1:1 mixture of methanol: tetrahydrofuran (25 mL / g), a 10 % sodium hydroxide solution (5 mL / g) is added and the resulting mixture is allowed to stir at RT overnight. Addition of further portions of the hydroxide solution may be necessary. The organic solvents are removed under reduced pressure and water is added to dissolve the solid. The resulting solution is extracted with dichloromethane (unreacted starting material) and the aqueous layer acidified *via* the addition of hydrochloric acid (conc) to pH 4. The resulting precipitate is extracted into dichloromethane and the combined organic extracts are washed with water and further washed with brine, dried (magnesium sulfate), filtered and evaporated to dryness.

Procedure G (LiOH Saponification) To a solution containing either i) amine/ester or ii) nitro/ester (1 eq.) in a 1:1 mixture of tetrahydrofuran / water (25 mL / g), a saturated lithium hydroxide solution (1 eq.) is added and the resulting mixture is allowed to stir at RT overnight. The organic solvent is removed under reduced pressure and an additional amount of water is added. The resulting solution is extracted with dichloromethane (unreacted starting material) and the aqueous layer acidified *via* the addition of 1M potassium bisulfate solution to pH 4. The resulting precipitate is extracted into dichloromethane and the combined organic extracts are washed with water and further washed with brine, dried (magnesium sulfate), filtered and evaporated to dryness.

Procedure H (Fmoc protection) A solution of **18** (1 eq.) in anhydrous tetrahydrofuran (20 mL / g) is held at a reflux under a nitrogen atmosphere. A solution of fluorenylmethyloxycarbonyl chloride (1.5 eq.) in anhydrous tetrahydrofuran (10 mL / g) is then added dropwise and the resulting solution is stirred at reflux overnight. The reaction mixture is cooled to ambient temperature, concentrated and the resulting precipitate collected *via* filtration.

Procedure I (Fmoc protection) A solution of **18** (1 eq.) and sodium bicarbonate (3 eq.) in anhydrous tetrahydrofuran (20 mL / g) is held at a reflux under a nitrogen atmosphere. A solution of fluorenylmethyloxycarbonyl chloride (1.5 eq.) in anhydrous tetrahydrofuran (10 mL / g) is then added dropwise and the resulting solution is stirred at reflux overnight. Sodium bicarbonate is removed *via* hot filtration and the reaction mixture is allowed to cool to room temperature, concentrated under reduced pressure and the resulting precipitate collected *via* filtration.

6.1.2 Solid Phases Synthesis

Acyl Chloride Formation – Method A To a stirred solution of an Fmoc protected building block in anhydrous dichloromethane (20 mL / g), thionyl chloride (10 eq.) is added and the resulting mixture refluxed overnight. The organic solvents and thionyl chloride are removed under reduced pressure and the resulting solid re-dissolved in chloroform. Hexane is added to precipitate the acyl chloride which is collected *via* filtration and stored under an inert atmosphere.

Acyl Chloride Preactivation – Method B To a solution containing Fmoc protected monomers (**14**, 1 eq.) functionalised with acid sensitive protecting groups in NMP (2.5 mL), 0.9 eq. of Ghosez's reagent is added. The resulting mixture is stored under an inert atmosphere for 3 hours at 50 ºC before the addition to the resin and microwave treatment.

Acyl Chloride *In Situ* Formation- Method C To a solution containing Fmoc protected monomers (**14**, 1 eq.) in NMP (2.5 mL), 1 eq. of thionyl chloride is added immediately before addition to the resin and microwave treatment

General Points for Solid Phase Synthesis: Fmoc-Gly-Wang resin (0.79 mmol/g, 100-200 mesh; carrier: polystyrene, crosslinked with 1% DVB), Fmoc-Ile-Wang resin (0.59 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 Alfa Aeser and stored in a schlenk tube on molecular sieves under a nitrogen atmosphere. Acyl chlorides were synthesised as in Method A for acyl chloride formation and stored under an inert atmosphere. 1-Chloro-*N, N*, 2-trimethyl-1-propenylamine (Ghosez's reagent) was purchased from Sigma-Aldrich. 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. Manual SPS was carried out in 1.5 mL 'Extract-Clean' polypropylene reservoirs fitted with 20 mm polyethylene frits, both available from Alltech.

General Procedure for Oligomer Formation – Single Coupling Fmoc protected pre-loaded Wang resin (127 mg, 0.1 mmol, 1 eq.) is loaded onto a CEM™ microwave peptide synthesiser after being swelled for a total of 30 minutes in NMP and DCM solutions. A series of washes (3 x NMP), deprotection (2 x 20 % Piperidine/NMP, total of 3.5 minutes at 75 °C) and further washes (5 x NMP) prepares the resin for coupling. Fmoc protected acyl chloride **X** (0.4 mmol, 4 eq.) obtained by pre-activation or prepared separately is dissolved in NMP (2.5 mL), delivered to the reaction vessel and submitted to microwave irradiation at 50 °C for 30 minutes. A final series of filtered washes of the reaction vessel (3 x NMP) finishes a coupling cycle.

General Procedure for Oligomer Formation – Double Coupling Fmoc protected pre-loaded Wang resin (127 mg, 0.1 mmol, 1 eq.) is loaded onto a CEM™ microwave peptide synthesiser after being swelled for a total of 30 minutes in NMP and DCM solutions. A series of washes (3 x NMP), deprotection (2 x 20 % Piperidine/NMP, total of 3.5 minutes at 75 °C) and further washes (5 x NMP) prepares the resin for coupling. Fmoc protected acyl chloride **X** (0.2 mmol, 2 eq.) obtained by pre-activation or prepared separately is dissolved in NMP (2.5 mL), delivered to the reaction vessel and submitted to microwave irradiation at 50 °C for 30 minutes. A second solution containing Fmoc protected acyl chloride **X** (0.2 mmol, 2 eq.) (preactivated or isolated) in NMP (2.5 mL) is delivered to

the reaction vessel and submitted to microwave power at 50 °C for 30 minutes. A final series of filtered washes of the reaction vessel (3 x NMP) finishes a coupling cycle.

General Procedure for Cleavage After the required number of cycles, a final Fmoc deprotection is carried out and then the resin is removed from the synthesiser and transferred to a reservoir for manual cleavage. The resin is washed with dichloromethane (10 x 1 mL) and cleaved with a 1.5 mL cleavage cocktail consisting of trifluoroacetic acid-dichloromethane-triisopropylsilane with varying ratios depending on the side chains. If no protecting groups are present, a simple 1:1 trifluoroacetic acid-dichloromethane mixture is sufficient without the need for a scavenger. A 95:5 dichloromethane-trifluoroacetic acid solution was used with trimers containing naphthyl / indole side chains.

6.2 Characterisation of Monomers (Chapters 2 and 3)

Monomers **14** (**d**, **k**, **o**, **q**, **r**) and their intermediates were originally synthesised by Panchami Prabhakaran (**d**, **o**, **q**, **r**) and Valeria Azzarito (**k**) and (monomers **14k** and **14o** later synthesised by myself). These were included for completeness of *Chapter 2* experimental: all monomers and intermediates were, however, characterised by me.

4-Amino-3-isopropoxybenzoic acid 18a

Procedure F; Methyl-4-amino-3-isopropyloxybenzoate **17a** (9.00 g, 43.0 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (220 mL), 10% aqueous sodium hydroxide solution (50 mL). Work-up afforded the title compound (7.56 g, 38.7 mmol, 90%) as a colourless amorphous powder; (Found C, 61.30; H, 6.75; N, 7.01%. $C_{10}H_{13}NO_3$ requires C, 61.53; H, 6.75; N, 7.18%); δ_H (300 MHz, $CDCl_3$) 1.36 (6H, d, *J* = 6.8, Hβ), 3.85 (2H, br. s, N<u>H</u>$_2$), 4.63 (1H, sept, *J* = 6.1, Hα), 6.68 (1H, d, *J* = 8.2, H5), 7.53 (1H, s, H2), 7.61 (1H, d, *J* = 8.2, H6); δ_C (75MHz, $CDCl_3$); 21.3, 70.9, 113.2, 114.5, 118.4, 124.2, 124.4, 144.1, 169.6; υ_{max}/cm^{-1} (solid state) = 3335, 2520, 1769, 1659, 1577, 1443, 1262, 1111, 976; ESI-MS found *m/z* 196.09 $[M+H]^+$;

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-isopropoxybenzoic acid 14a

Procedure H; 4-Amino-3-isopropoxybenzoic acid **18a** (7.00 g, 35.9 mmol) in tetrahydrofuran (140 mL), fluorenylmethyloxycarbonyl chloride (13.92 g, 53.8 mmol) in tetrahydrofuran (100 mL). Work up yielded the title compound (12.28 g, 29.4 mmol, 82%) as a colourless amorphous solid; (Found C, 71.80, H, 5.65; N, 3.25%. $C_{25}H_{23}NO_5$ requires C, 71.93; H, 5.55; N, 3.36%); δ_H (300 MHz, $CDCl_3$) 1.43 (6H, d, *J* = 6.9, Hβ), 4.34 (1H, t, *J* = 6.9, FHβ), 4.54 (2H, d, *J* = 6.9, FHα), 4.73 (1H, sept, *J* = 6.9, Hα), 7.34 (2H, t, *J* = 7.5, FHAr4), 7.43 (2H, t, *J* = 7.5, FHAr3), 7.55 (1H, s, H2), 7.60-7.75 (4H, m, H5, H6 + FHAr5), 7.80 (2H, d, *J* = 8.4, FHAr2), 8.16 (1H, s, N<u>H</u>); δ_C (75MHz, $CDCl_3$); 21.9, 46.9, 66.7, 71.5, 114.3, 120.5, 125.5, 125.6, 126.3, 127.5, 128.1, 132.8, 141.1, 144.0, 147.4, 153.7, 167.3; υ_{max}/cm^{-1} (solid state) = 3333, 2975, 1709, 1599, 1542, 1497, 1442, 1337, 1240, 1105, 1053, 978; ESI-HRMS found *m/z* 418.1649 $[M+H]^+$, $C_{25}H_{24}NO_5$ requires 418.1654;

Methyl 3-isobutoxy-4-nitrobenzoate 16b

Procedure A; Methyl-3-hydroxy-4-nitrobenzoate **15** (10.00 g, 50.72 mmol) and potassium carbonate (2l.0 g, 151.9 mmol) in dimethylformamide (200 mL), 1-bromo-2-methylpropane (6.62 mL, 60.9 mmol). Following work-up, the resulting solid was crystallised (methanol/ hexane) to yield the title compound (9.17 g, 36.2 mmol, 71%) as pale yellow crystals; m.p. 68.5-69.0 °C (methanol/ hexane); δ_H (300 MHz,$CDCl_3$) 1.06 (6H, d, *J* = 6.9, Hγ), 2.16 (1H, sept, *J* = 6.6, Hβ), 3.93 (2H, d, *J* = 6.3, Hα), 3.96 (3H, s, CO_2<u>Me</u>), 7.66 (1H, dd, *J* = 8.4, 1.5, H6), 7.72 (1H, d, *J* = 1.5, H2), 7.81 (1H, d, *J* =8.4, H5); δ_C (75MHz, $CDCl_3$) 17.5, 28.6, 51.2, 74.4, 113.8, 119.5, 123.6, 133.1, 140.8, 150.5, 163.7;

υ_{max}/cm^{-1} (solid state) = 2955, 1726, 1609, 1524, 1237, 750; ESI-HRMS found *m/z* 276.0851 $[M+Na]^+$, $C_{12}H_{15}NNaO_5$ requires 276.0842;

Methyl 4-amino-3-isobutoxybenzoate 17b

Procedure C; Methyl 3-isobutoxy-4-nitrobenzoate **16b** (5.00 g, 19.7 mmol), tin(II) chloride dihydrate (26.73 g, 118.5 mmol) in ethyl acetate (150 mL). Following standard work-up the resulting solid was crystallised (dichloromethane/ hexane) to yield the title compound (3.36 g, 15.1 mmol, 76%) as colourless crystalline plates; m.p. 62.3-63.5 °C (dichloromethane/ hexane); (Found C, 64.30; H, 7.65; N, 6.35%. $C_{12}H_{17}NO_3$ requires C, 64.55; H, 7.67; N, 6.27%); δ_H (300 MHz, $CDCl_3$) 1.05 (6H, d, *J* = 6.6, Hγ), 2.13 (1H, sept, *J* = 6.6, Hβ), 3.82 (2H, d, *J* = 6.3, Hα), 3.86 (3H, s, $CO_2\underline{Me}$), 4.24 (2H, s, $N\underline{H}_2$), 6.67 (1H, d, *J* = 8.4, H5), 7.43 (1H, d, *J* = 1.5, H2), 7.53 (1H, dd, *J* = 8.4, 1.5, H6); δ_C (75MHz, $CDCl_3$) 19.8, 28.7, 52.1, 75.1, 112.4, 113.5, 119.8, 124.3, 141.6, 146.0, 167.8; υ_{max}/cm^{-1} (solid state) = 3461, 3342, 2950, 1688, 1622, 1523, 1441, 1269, 1034, 766, 635; ESI-HRMS found *m/z* 224.1287 $[M+H]^+$, $C_{12}H_{18}NO_3$ requires 224.1281;

4-Amino-3-isobutoxybenzoic acid 18b

Procedure F; Methyl 4-amino-3-isobutoxybenzoate **17b** (1.80 g, 8.1 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (50 mL), 10% aqueous sodium hydroxide (17 mL). Following acidification, the precipitate was filtered, dissolved in chloroform and dried with magnesium sulfate. The solution was filtered and the organic solvents removed under reduced pressure yielding a pink solid which was crystallised (dichloromethane/ hexane) to yield the title compound (1.22 g, 5.8 mmol, 72%) as pale pink microcrystals; m.p. 118.7-119.6 °C (dichloromethane/ hexane); (Found C, 63.15; H, 7.15; N, 6.60%. $C_{11}H_{15}NO_3$ requires C, 63.14; H, 7.23; N, 6.69%); δ_H (300 MHz, $CDCl_3$) 1.08 (6H, d, *J* = 6.7, Hγ), 2.16 (1H, sept, *J* = 6.7, Hβ), 3.85 (2H, d, *J* = 6.6, Hα), 6.71 (1H, d, *J* = 8.2, H5), 7.5 (1H, d, *J* = 1.8, H2), 7.65 (1H, dd, *J* = 8.1, 1.8, H6); δ_C (75MHz, $CDCl_3$) 19.4, 28.3, 74.7, 112.4, 113.0, 118.4, 124.9, 142.1, 145.5, 172.4; υ_{max}/cm^{-1} (solid state) = 3497, 3384, 2813, 1659, 1611, 1306, 765; ESI-HRMS found *m/z* 210.1122 $[M+H]^+$, $C_{11}H_{16}NO_3$ requires 210.1125.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-isobutoxybenzoic acid 14b

Procedure H; 4-Amino-3-isobutoxybenzoic acid **18b** (3.00 g, 14.3 mmol) in tetrahydrofuran (100 mL), fluorenylmethyloxycarbonyl chloride (5.56 g, 21.5 mmol) in tetrahydrofuran (50 mL). Work up yielded the title compound (5.25 g, 12.2 mmol, 85%) as a colourless amorphous solid; (Found C, 72.10; H, 5.80; N, 3.15%. $C_{26}H_{25}NO_5$ requires C, 72.37; H, 5.84; N, 3.25%); δ_H (300 MHz, DMSO-d_6) 1.04 (6H, d, *J* = 6.7, Hγ), 2.15 (1H, sept, *J* = 6.6, Hβ), 3.84 (2H, d, *J* = 6.6, Hα), 4.26 (1H, t, *J* = 7.0, FHβ), 4.46 (2H, d, *J* = 7.0, FHα), 7.26 (2H, t, *J* = 7.5, FHAr4), 7.36 (2H, t, *J* = 7.5, FHAr3), 7.48-7.51 (2H, m, H5 + H2), 7.56 (2H, d, *J* = 7.3, FHAr5), 7.67-7.74 (3H, m, H6 + FHAr2) 8.08 (1H, br. s, N<u>H</u>); δ_C (125MHz, DMSO-d_6) 18.9, 27.5, 46.4, 66.3, 74.6, 112.3, 120.1, 120.3, 122.1, 125.1, 126.0,

127.0, 127.7, 131.3, 140.7, 143.6, 148.6, 153.3, 166.9; υ_{max}/cm^{-1} (solid state) = 3337, 2955, 1710, 1677, 1435, 1290, 736; ESI-HRMS found *m/z* 454.1615 $[M+Na]^+$, $C_{26}H_{25}NNaO_5$ requires 454.1625.

Methyl 3-(cyclopropylmethoxy)-4-nitrobenzoate 16c

Procedure A; Methyl-3-hydroxy-4-nitrobenzoate **15** (10.00 g, 50.7 mmol), potassium carbonate (2l.00 g, 151.9 mmol) in dimethylformamide (200 mL), (bromomethyl)cyclopropane (5.91 mL, 60.9 mmol). Following work-up the resultant yellow solid which was crystallised (ethyl acetate) to yield the title compound (10.34 g, 41.2 mmol, 81 %) as large pale yellow rectangular crystals; m.p. 93.7-95.1 °C (ethyl acetate); (Found C, 57.30; H, 5.20; N, 5.55%. $C_{12}H_{13}NO_5$ requires C, 57.37; H, 5.22; N, 5.58%); δ_H (300 MHz, $CDCl_3$) 0.43 (2H, m, Hγ), 0.70 (2H, m, Hγ'), 1.34 (1H, m, Hβ), 3.99 (3H, s, $CO_2\underline{Me}$), 4.07 (1H, d, *J* = 6.8, Hα) 7.70 (1H, dd, *J* = 8.2, 1.5, H2), 7.74 (1H, d, *J* = 1.6, H6), 7.83 (1H, d, *J* = 8.3, H5); δ_C (75MHz, $CDCl_3$) 3.74, 10.3, 53.2, 75.0, 116.4, 121.7, 125.6, 135.1, 143.2, 152.2, 165.7; υ_{max}/cm^{-1} (solid state) = 3111, 1726, 1607, 1522, 1307, 1247, 750; ESI-HRMS found *m/z* 274.0689 $[M+Na]^+$, $C_{12}H_{13}NaNO_5$ requires 274.0686.

Methyl 4-amino-3-(cyclopropylmethoxy)benzoate 17c

Procedure C; Methyl 3-(cyclopropylmethoxy)-4-nitrobenzoate **16c** (10.00 g, 39.8 mmol), tin(II) chloride dihydrate (53.89 g, 238.9 mmol) in ethyl acetate (150 mL). Following work-up, the resulting solid was crystallised (ethyl acetate/ hexane) to yield the title compound (5.77 g, 26.1 mmol, 65%) as colourless microcrystals; m.p. 81.5-82.4 °C (ethyl acetate/ hexane); (Found C, 65.05; H, 6.85; N, 6.35%. $C_{12}H_{15}NO_3$ requires C, 65.14; H, 6.83; N, 6.33%); δ_H (300 MHz, $CDCl_3$) 0.29 (2H, m, Hγ), 0.56 (2H, m, Hγ'), 1.22 (1H, m, Hβ), 3.78 (3H, s, $CO_2\underline{Me}$), 3.81 (2H, d, *J* = 7.1, Hα), 4.20 (2H, br. s, $N\underline{H_2}$), 6.59 (1H, d, *J* = 8.1, H5), 7.34 (1H, d, *J* = 1.7, H2), 7.46 (1H, dd, *J* = 8.1, 1.7, H6); δ_C (75MHz, $CDCl_3$) 3.6, 10.7, 52.1, 73.7, 113.0, 113.6, 119.8, 124.4, 141.8, 145.9, 167.7; υ_{max}/cm^{-1} (solid state) = 3491, 3355, 2998, 1682, 1614, 1296, 762; ESI-HRMS found *m/z* 222.1130 $[M+H]^+$, $C_{12}H_{16}NO_3$ requires 222.1125.

4-Amino-3-(cyclopropylmethoxy)benzoic acid 18c

Procedure F; Methyl 4-amino-3-(cyclopropylmethoxy)benzoate **17c** (4.64 g, 20.1 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (120 mL), 10% aqueous sodium hydroxide (30 mL). The resulting precipitate was filtered, dissolved in chloroform and dried with magnesium sulfate. The solution was filtered and the organic solvents removed under reduced pressure yielding a beige solid which was crystallised (chloroform/ methanol/ hexane) to yield the title compound (3.65 g, 17.6 mmol, 84%) as large pale orange crystals; m.p. 154.5-155.9 °C (chloroform/ methanol/ hexane); (Found C, 63.25; H, 6.25; N, 6.65%. $C_{11}H_{13}NO_3$ requires C, 63.76; H, 6.32; N, 6.76%); δ_H (300 MHz, $CDCl_3$) 0.40 (2H, m, Hγ), 0.69 (2H, m, Hγ'), 1.33 (1H, m, Hβ), 3.95 (2H, d, *J* = 7.0, Hα), 6.72 (1H, d, *J* = 7.3, H5), 7.50 (1H, d, *J* = 1.7, H2), 7.66 (1H, dd, *J* = 7.3, 1.7, H6); δ_C (75MHz, $CDCl_3$) 3.6, 10.7,

73.8, 113.3, 113.5, 118.8, 125.5, 142.7, 145.8, 172.8; υ_{max}/cm^{-1} (solid state) = 3501, 3383, 2900, 1666, 1614, 1305, 765; ESI-HRMS found *m/z* 208.0962 $[M+H]^+$, $C_{11}H_{14}NO_3$ requires 208.0968.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(cyclopropylmethoxy)benzoic acid 14c

Procedure H; 4-Amino-3-(cyclopropylmethoxy)benzoic acid **18c** (3.13 g, 15.1 mmol) in tetrahydrofuran (100 mL), fluorenylmethyloxycarbonyl chloride (5.87 g, 22.7 mmol) in chloroform (30 mL). Work up yielded the title compound (5.26 g, 12.3 mmol, 81%) as a colourless amorphous solid; (Found C, 72.55; H, 5.25; N, 2.95%. $C_{26}H_{23}NO_5$ requires C, 72.71; H, 5.40; N, 3.26%); δ_H (300 MHz, $CDCl_3$) 0.42 (2H, m, Hγ), 0.73 (2H, m, Hγ'), 1.37 (1H, m, Hβ), 3.98 (2H, d, *J* = 7.1, Hα), 4.37 (1H, t, *J* = 6.8, FHβ), 4.58 (2H, d, *J* = 6.9, FHα), 7.36 (2H, t, *J* = 7.4, FHAr4), 7.45 (2H, t, *J* = 7.4, FHAr3), 7.57 (1H, d, *J* = 1.7, H2), 7.66-7.69 (3H, m, H5 + FHAr5), 7.75 (1H, d, *J* = 8.5, H6), 7.82 (2H, d, *J* = 7.4, FHAr2); δ_C (75MHz, $CDCl_3$) 3.7, 10.6, 47.5, 67.8, 74.5, 113.0, 117.6, 120.5, 124.3, 124.5, 125.4, 127.6, 128.3, 133.0, 141.8, 144.1, 146.9, 153.5, 169.9; υ_{max}/cm^{-1} (solid state) = 3329, 2807,1707, 1675, 1541, 1250, 735; ESI-HRMS found *m/z* 430.1628 $[M+H]^+$, $C_{26}H_{24}NO_5$ requires 430.1649.

Methyl 3-(*sec*-butoxy)-4-nitrobenzoate 16d

Procedure B; Methyl-3-hydroxy-4-nitro-benzoate **15** (2.42 g, 12.28 mmol), S(+)-*sec*-butanol (1.00 g, 13.5 mmol), triphenylphosphine (4.82 g, 18.4 mmol) with diisopropyl azodicarboxylate (3.61 mL, 18.4 mmol) in tetrahydrofuran (80 mL). Work up followed by column chromatography yielded the product (2.73 g, 10.81 mmol, 88%) as a pale yellow liquid; δ_H (500 MHz, $CDCl_3$) 1.00 (3H, t, *J* = 7.4, Hγ), 1.36 (3H, d, *J* = 6.1, $CH_\alpha(C\underline{H}_3)$), 1.69-1.83 (2H, m, Hβ + Hβ'), 3.96 (3H, s, $CO_2\underline{Me}$), 4.56 (1H, m, Hα), 7.64 (1H, d, *J* = 8.3, H6), 7.73 (1H, s, H2), 7.76 (1H, d, *J* = 8.3, H5); δ_C (125MHz, $CDCl_3$) 9.5, 18.9, 29.0, 52.8, 77.7, 116.7, 120.9, 125.1, 134.4, 143.6, 151.1, 165.4, υ_{max}/cm^{-1} (solid state) = 2980, 1731, 1606, 1531, 1294, 1108, 1094; ESI-HRMS found *m/z* 276.0823$[M+Na]^+$, $C_{12}H_{15}NO_5$ requires 276.0842.

4-((((9H-Fluoren-9-yl) methoxy)carbonyl)amino)-3-(sec-butoxy)benzoic acid 14d

4-Amino-3-(*sec*-butoxy)benzoic acid **18d** was obtained from methyl 4-amino-3-(*sec*-butoxy)benzoate **17d** by *procedure F* without purification/ isolation**.** Methyl 4-amino-3-(*sec*-butoxy)benzoate **17d** was in turn obtained from methyl 3-(*sec*-butoxy)-4-nitrobenzoate **16d** by *procedure C* without purification/ isolation. *Procedure H*; 4-amino-3-(sec-butoxy)benzoic acid **17d** (2.00 g, 9.56 mmol) in tetrahydrofuran (100 mL), fluorenylmethyloxycarbonyl chloride (3.45 g, 13.4 mmol) in tetrahydrofuran (30 mL). The crude material obtained after the reaction was purified by colum chromatography to yield the title compound (3.18 g, 7.37 mmol, 77%) as an off-white amorphous solid; (Found C, 71.60; H, 5.90; N, 3.10%. $C_{26}H_{25}N_1O_5$ requires C, 72.37; H, 5.84; N, 3.25%); δ_H (500 MHz,$CDCl_3$) 1.05 (3H, t, *J* = 7.4, Hγ), 1,39 (3H, d, *J* = 6.0, $CH_\alpha C\underline{H}_3$), 1.75 (1H, m, Hβ), 1.84 (1H, m, Hβ'), 4.34 (1H, t, *J* = 7.0, FHβ), 4.48-4.53 (3H, m, Hα + FHα), 7.33 (2H, t, *J* = 7.3, FHAr4), 7.43 (2H,

t, *J* = 7.3,FHAr3), 7.56-7.63 (4H, m, H2, H5 + FHAr5), 7.74 (1H, d, *J* = 8.5, H6), 7.80 (2H, d, *J* = 7.5, FHAr2), 8.15 (1H br. s, NH); δ_C (125MHz, $CDCl_3$) 9.7, 19.3, 29.1, 47.1, 67.5, 113.7, 117.4, 120.1, 123.0, 124.1, 125.0, 127.2, 127.9, 133.7, 143.7, 145.5, 153.0, 171.7, υ_{max}/cm^{-1} (solid state) = 3427, 2960, 2565, 1747, 1690, 1595, 1535, 1483, 1416, 1345, 1301, 1213, 1189, 1058; $[\alpha]_D^{24}$ -12.8 (*c* = 1, chloroform); ESI-HRMS found *m/z* 430.1660 $[M-H]^-$, $C_{26}H_{24}NO_5$ requires 430.1654.

Methyl 4-amino-3-(benzyloxy)benzoate 17e

Procedure C; Methyl-3-benzyloxy-4-nitrobenzoate **16e** (4.00 g, 13.9 mmol), tin(II) chloride dihydrate (18.84 g, 83.6 mmol) in ethyl acetate (100 mL). Following initial work-up, the solvents were removed under reduced pressure and the resultant orange oil was passed through a bed of silica (20% diethyl ether / dichloromethane) and the solvents removed. The resultant solid was crystallised from hexane to yield the title compound (2.58 g, 10.0 mmol, 72%) as colourless square plates; m.p. 82.2-83.6 °C (hexane); (Found C, 69.75; H, 5.85; N, 5.2%. CHN requires C, 70.02; H, 5.88; N, 5.44%); δ_H (300 MHz, $CDCl_3$), 3.90 (3H, s, CO_2Me), 4.34 (2H, br s, NH_2), 5.14 (2H, s, Hα), 6.72 (1H, d, *J* = 6.6, H5), 7.30-7.50 (5H, m, HAr2, HAr3 + HAr4), 7.60-7.63 (2H, m, H6 + H2); δ_C (75 MHz, $CDCl_3$) 52.1, 70.9, 113.1, 113.7, 119.7, 124.8, 128.2, 128.6, 129.0, 137.1, 141.9, 145.7, 167.7; υ_{max}/cm^{-1} (solid state) = 3517, 3395, 2930, 1688, 1432, 1279, 1128, 796; ESI-HRMS found *m/z* 258.1117 $[M+H]^+$, $C_{15}H_{16}NO_3$ requires 258.1125.

4-Amino-3-(benzyloxy)benzoic acid 18e

Procedure F; Methyl 4-amino-3-(benzyloxy)benzoate **17e** (5.44 g, 21.1 mmol), in a 1:1 mixture of methanol: tetrahydrofuran (135 mL), 10% aqueous sodium hydroxide (40 mL). Following acidification, the precipitate was filtered, dissolved in chloroform and dried (magnesium sulphate). This solution was then filtered and the organic solvents removed under reduced pressure to yield the title compound (3.90g, 16.0 mmol, 76%) as a colourless amorphous powder; δ_H (300 MHz, MeOD-d_4) 5.17 (2H, s, Hα), 6.75 (1H, d, *J* = 4.9, H5), 7.34 (1H, t, *J* = 4.5, HAr4), 7.39-7.42 (2H, t, *J* = 4.5, HAr3), 7.49-7.53 (4H, m, H6, H2 + HAr2); δ_C (75MHz, MeOD-d_4) 70.4, 113.0, 113.1, 118.5, 124.8, 127.6, 127.9, 128.5, 137.4, 143.2, 145.3, 169.5; υ_{max}/cm^{-1} (solid state) 3353, 2920, 2256, 1690, 1619, 1522, 1442, 1252, 1146, 1024, 879; ESI-HRMS found *m/z* 244.0973 $[M+H]^+$, $C_{14}H_{14}NO_3$ requires 244.0986;

4-(((9H-Fluoren-9-yl)methoxy)carbonylamino)-3-(benzoyloxy)benzoic acid 14e

Procedure H; 4-Amino-3-(benzyloxy)benzoic acid **18d** (7.00 g, 28.8 mmol) in tetrahydrofuran (100 mL), fluorenylmethyloxycarbonyl chloride (11.17 g, 43.2 mmol) in tetrahydrofuran (100 mL). Following work up, the resulting solid was crystallised from a 1.5:1 solution of chloroform / methanol to yield the title compound (12.32 g, 26.5 mmol, 92%) as colourless microcrystals; m.p. 242.3-243.8 °C (chloroform / methanol); δ_H (500 MHz, DMSO-d_6) 4.37 (1H, t, *J* = 6.9, FHβ), 4.50 (2H, d, *J* = 6.9, FHα), 5.31 (2H, s, Hα), 7.36-7.40 (3H, m, HAr4 +FHAr4), 7.43-7.50 (4H, m, HAr3 +

FHAr3), 7.56-7.58 (3H, m, H5 + HAr2), 7.62 (1H, s, H2), 7.75-7.81 (3H, m, H6 + FHAr5) 7.96 (2H, d, *J* =7.5, FHAr2), 8.96 (1H, br. s, NH); δ_C (125MHz, DMSO-d_6); 46.9, 66.7, 70.3, 113.5, 120.5, 121.2, 122.8, 125.6, 126.5, 127.4, 127.6, 128.1, 128.8, 132.1, 137.1, 141.1, 144.0, 146.4, 148.7, 153.8, 167.2; υ_{max}/cm^{-1} (solid state) = 3332, 2891, 1671, 1597, 1501, 1435, 1346, 1217, 1104, 984, 878; ESI-HRMS found *m/z* 488.1488 $[M+Na]^+$, $C_{29}H_{23}NNaO_5$ requires 488.1468;

Methyl 3-((4-(tert-butyl)benzyl)oxy)-4-nitrobenzoate 16f

Procedure A; Methyl-3-hydroxy-4-nitro-benzoate **15** (3.40 g, 17.3 mmol) potassium carbonate (7.15 g, 51.7 mmol), in dimethlyformamide (70 mL), 4-*tert* butyl benzyl bromide (4.10 mL, 22.4 mmol). Work up yielded the title compound (5.98 g, 169.0 mmol, 98%) as a pale yellow amorphous solid; (Found C, 66.25; H, 6.35; N, 4.10%. $C_{19}H_{21}N_1O_5$ requires C, 66.46; H, 6.16; N, 4.08%); δ_H (500 MHz, $CDCl_3$) 1.33 (9H, s, C(C$\underline{H}_3$)$_3$), 3.96 (3H, s, $CO_2\underline{Me}$), 5.25 (2H, s, Hα), 7.44-7.39 (4H, m, HAr2 + HAr3), 7.73 (1H, dd, *J* = 8.4, 1.3, H6), 7.86-7.88 (2H, m, H2 + H5); δ_C (75MHz, $CDCl_3$) 31.3, 34.7, 52.8, 71.4, 116.1, 121.6, 125.3, 125.7, 127.2, 132.0, 134.8, 142.9, 151.55, 151.58, 165.2; υ_{max}/cm^{-1} (solid state) = 2963, 1725, 1608, 1538, 1439, 1372, 1291, 1247, 1111, 1087; ESI-HRMS found *m/z* 366.1315 $[M+Na]^+$, $C_{19}H_{21}NNaO_5$ requires 366.1317.

Methyl 4- amino-3-((4-(tert-butyl) benzyl)oxy)-4-nitrobenzoate 17f

Procedure C: Methyl 3-((4-(tert-butyl)benzyl)oxy)-4-nitrobenzoate **16f** (5.70 g, 16.6 mmol), tin(II) chloride dihydrate (18.60 g, 82.7 mmol) in ethyl acetate (120 mL). Work up yielded the title compound (4.59 g, 14.7 mmol, 88%) as a yellow amorphous solid; δ_H (500 MHz, $CDCl_3$) 1.35 (9H, s, C(C$\underline{H}_3$)$_3$), 3.88 (3H, s, $CO_2\underline{Me}$), 4.26 (2H, br. s, N$\underline{H}_2$), 5.09 (2H, s, Hα), 6.69 (1H, d, *J* = 8.4, H5), 7.39-7.45 (4H, m, HAr2 + HAr3), 7.57-7.59 (2H, m, H2 + H6); δ_C (125MHz, $CDCl_3$) 31.4, 34.6, 51.7, 70.4, 112.6, 113.3, 119.5, 124.3, 125.6, 127.8, 133.7, 141.4, 145.4, 151.3, 167.3; υ_{max}/cm^{-1} (solid state) = 3476, 3352, 3199, 2958, 1682, 1620, 1593, 1523, 1441, 1316, 1151, 1109, 1034; ESI-HRMS found *m/z* 336.1577 $[M+Na]^+$, $C_{19}H_{23}NNaO_5$ requires 336.1576.

4- Amino-3-((4-(tert-butyl)benzyl)oxy)-4-nitrobenzoic acid 18f

Procedure F; 4-amino-3-((4-(tert-butyl)benzyl)oxy)-4-nitrobenzoicacid **17f** (4.50 g, 14.3 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (110 mL), 10% aqueous sodium hydroxide (25 mL). Work up yielded the title compound (4.05 g, 13.5 mmol, 94%) as a beige amorphous solid; δ_H (500 MHz, $CDCl_3$) 1.35 (9H, s, C(C$\underline{H}_3$)$_3$), 5.11 (2H, s, Hα), 6.71 (1H, d, *J* = 8.0, H5), 7.39-7.45 (4H, m, HAr2 + HAr3), 7.62 (1H, d, *J* = 1.5, H2), 7.66 (1H, dd, *J* = 8.0, 1.5, H6); υ_{max}/cm^{-1} (solid state) = 3468, 3367, 2962, 1671, 1614, 1519, 1442, 1407, 1370, 1442, 1407, 1370, 1220, 1148, 1107; ESI-HRMS found *m/z* 322.1414 $[M+Na]^+$, $C_{18}H_{21}NNaO_3$ requires 322.1419.

4-((((9H-Fluoren-9-yl) methoxy)carbonyl)amino)-3-((4-(tert-butyl)benzyl)oxy)benzoic acid 14f

Procedure H; 4-amino-3-((4-(tert-butyl) benzyl)oxy)-4-nitrobenzoicacid **18f** (3.50 g, 11.7 mmol) in tetrahydrofuran (70 mL), fluorenylmethyloxycarbonyl chloride (4.53 g, 17.5 mmol) in tetrahydrofuran

(45 mL). Work up yielded the title compound (5.20 g, 10.0 mmol, 85%) as a white amorphous solid; (Found C, 75.85; H, 6.00; N, 2.60%. $C_{33}H_{31}NO_5$ requires C, 75.99; H, 5.99; N, 2.69%); δ_H (500 MHz, $CDCl_3$) 1.37 (9H, s, C(C<u>H</u>$_3$)$_3$), 4.31 (1H, t, *J* = 7.0, FHβ), 4.52 (2H, d, *J* = 7.0, FHα), 5.17 (2H, s, Hα), 7.31(2H, t, *J* = 7.5, FHAr4), 7.39-7.42 (4H, m, HAr3 + FHAr3), 7.48 (2H, d, *J* = 8.2, HAr2), 7.56 (1H, m, H5), 7.60 (2H, d, *J* = 7.6, FHAr5), 7.71 (1H, s, H2), 7.77-7.79 (3H, m, H6 + FHAr2); δ_C (75MHz, $CDCl_3$) 31.3, 34.7, 46.9, 67.4, 70.9, 112.5, 117.3, 120.1, 123.2, 124.6, 125.0, 125.8, 127.1, 127.75, 127.8, 132.7, 133.0, 141.3, 143.6, 146.3, 151.7, 152.9, 171.2; υ_{max}/cm^{-1} (solid state) = 3430, 2958, 1740, 1681, 1595, 1534, 1489, 1351, 1242, 1224, 1199, 1058; ESI-HRMS found *m/z* 520.2139 $[M-H]^-$, $C_{33}H_{30}NO_5$ requires 520.2124.

Methyl 3-((4-chlorobenzyl)oxy)-4-nitrobenzoate 16g

Procedure A; Methyl-3-hydroxy-4-nitrobenzoate **15** (10.00 g, 50.7 mmol), potassium carbonate (2l.0 g, 151.9 mmol) in dimethylformamide (200 mL), 1-(bromomethyl)-4-chlorobenzene (12.5 g, 60.8 mmol). Following work-up, the resulting solid was crystallised (dichloromethane/ methanol) to yield the title compound (14.83 g, 46.1 mmol, 91%) as pale yellow crystalline plates; m.p. 133.7-134.8°C (dichloromethane/ methanol); (Found C, 55.75; H, 3.70; N, 4.30%. $C_{15}H_{12}NO_5Cl$ requires C, 56.00; H, 3.76; N, 4.35%); δ_H (300 MHz, $CDCl_3$) 4.00 (3H, s, CO_2<u>Me</u>), 5.29 (2H, s, Hα), 7.42 (4H, m, HAr2 + HAr3), 7.76 (1H, d, *J* = 8.4, H6), 7.84 (1H, s, H2), 7.90 (1H, d, *J* = 8.4, H5); δ_C (75MHz, $CDCl_3$) 53.4, 71.0, 116.3, 122.4, 125.9, 128.9, 129.4, 133.9, 134.7, 135.3, 142.9, 151.6, 166.5; υ_{max}/cm^{-1} (solid state) =3059, 1726. 1610, 1524, 1296, 1035, 808, 485; ESI-HRMS found *m/z* 344.0306 $[M+Na]^+$, $C_{15}H_{12}ClNNaO_5$ requires 344.0296.

Methyl 4-amino-3-((4-chlorobenzyl)oxy)benzoate 17g

Procedure C; Methyl 3-((4-chlorobenzyl)oxy)-4-nitrobenzoate **16g** (5.00 g, 15.5 mmol), tin(II) chloride dihydrate (21.00 g, 93.1 mmol) in ethyl acetate (100 mL). Following work up, the resulting solid was crystallised (dichloromethane/ hexane) to yield the title compound (3.03 g, 10.4 mmol, 67%) as pale yellow microcrystals; m.p. 118.9-119.6°C (dichloromethane/ hexane); (Found C, 61.80; H, 4.75; N, 4.70%. $C_{15}H_{14}ClNO_3$ requires C,61.76; H, 4.84; N, 4.80%); δ_H (300 MHz, $CDCl_3$) 3.86 (3H, s, H CO_2<u>Me</u>), 4.25 (2H, s, N<u>H</u>$_2$), 5.09 (2H, s, Hα), 6.69 (1H, d, *J* = 8.1, H5), 7.37 (4H, m, HAr2 + HAr3), 7.52 (1H, d, *J* = 1.5, H2), 7.57 (1H, dd, *J* = 8.1+ 1.5, H6); δ_C (125MHz, $CDCl_3$) 50.1, 68.1, 111.0, 111.8, 117.9, 122.9, 127.2, 127.5, 132.4, 133.5, 139.7, 143.4, 165.5; υ_{max}/cm^{-1} (solid state) = 3502, 3380, 2944, 1690, 1445, 1292, 1104, 800, 484; ESI-HRMS found *m/z* 292.0740 $[M+H]^+$, $C_{15}H_{15}ClNO_3$ requires 292.0735.

4-Amino-3-((4-chlorobenzyl)oxy)benzoic acid 18g

Procedure F; Methyl 4-amino-3-((4-chlorobenzyl)oxy)benzoate **17g** (2.34 g, 8.02 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (60 mL), 10% aqueous sodium hydroxide (18 mL). The resulting precipitate was filtered, dissolved in chloroform and dried with magnesium sulfate. The solution was

filtered and the organic solvents removed under reduced pressure yielding the title compound (2.06 g, 7.4 mmol, 93%) as a colourless amorphous solid; (Found C, 60.60; H, 4.40; N, 4.90%. $C_{14}H_{12}ClNO_3$ requires C, 60.55; H, 4.36; N, 5.04%); δ_H (500 MHz, MeOD-d_4) 5.16 (2H, s, Hα), 6.74 (1H, d, *J* = 8.5, H5), 7.40 (2H, d, *J* = 8.5, HAr3), 7.49 (2H, d, *J* = 8.5, HAr2), 7.51-7.54 (2H, m, H6 + H2); δ_C (125MHz, MeOD-d_4) 70.6, 114.2, 114.2, 119.5, 126.0, 129.7, 130.3, 134.8, 137.4, 144.4, 146.2, 170.5; υ_{max}/cm^{-1} (solid state) = 3438, 3293, 2563, 1688, 1296, 764, 495; ESI-HRMS found *m/z* 278.0565 $[M+H]^+$, $C_{14}H_{13}ClNO_3$ requires 278.0578.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-((4-chlorobenzyl)oxy)benzoic acid 14g

Procedure H; 4-Amino-3-((4-chlorobenzyl)oxy)benzoic acid **18g** (1.92 g, 6.9 mmol) in tetrahydrofuran (100 mL), fluorenylmethyloxycarbonyl chloride (2.68 g, 10.4 mmol) in chloroform (30 mL). Work up yielded the title compound (2.93 g, 5.9 mmol, 85%) as a colourless amorphous solid; (Found C, 69.65; H, 4.40; N, 2.65%. $C_{29}H_{22}ClNO_5$ requires C, 69.67; H, 4.44; N, 2.80%); δ_H (300 MHz, DMSO-d_6) 4.32 (1H, t, *J* = 6.9, FHβ), 4.46 (2H, d, *J* = 6.9, FHα), 5.25 (2H, s, Hα), 7.32 (2H, t, *J* = 7.5, FHAr4), 7.29-7.35 (4H, m, HAr3 + FHAr3), 7.41-7.58 (4H, m, H5, H2 + HAr2), 7.69 (1H, d, *J* = 8.2, H6), 7.74 (2H, d, *J* = 7.5, FHAr5), 7.91 (2H, d, *J* = 7.5, FHAr2), 9.01 (1H, s, N<u>H</u>); δ_C (75MHz, DMSO-d_6) 46.5, 66.2, 69.0, 113.1, 120.1, 121.1, 122.5, 125.2, 126.2, 127.1, 127.7, 128.4, 129.3, 131.7, 132.4, 135.8, 140.7, 143.7, 148.3, 153.5, 166.8; υ_{max}/cm^{-1} (solid state) = 3305, 2805, 1697, 1678, 1544, 1256, 739, 484; ESI-HRMS found *m/z* 522.1065 $[M+Na]^+$, $C_{29}H_{22}ClNaNO_5$ requires 522.1079.

Methyl 4-amino-3-(naphthalen-2-ylmethoxy)benzoate 17h

Procedure C; Methyl-3-(2-napthyl)methoxy-4-nitrobenzoate **16h** (6.25 g, 18.5 mmol), tin(II) chloride dihydrate (25.00 g, 110.8 mmol) in ethyl acetate (130 mL). Work afforded the title compound (4.38 g, 14.3 mmol, 77%) as a cream amorphous solid; (Found C, 74.4; H, 5.55; N, 4.35%. $C_{19}H_{17}NO_3$ requires C, 74.25.; H, 5.58; N, 4.65%); δ_H (500 MHz, $CDCl_3$) 3.83 (3H, s, CO_2<u>Me</u>), 4.31 (2H, br s, N<u>H</u>$_2$), 5.26 (2H, s, Hα), 6.67 (1H, d, *J* = 8.1, H5), 7.45-7.60 (5H, m, HAr), 7.80-7.86 (4H, m, HAr); δ_C (75MHz, $CDCl_3$) 52.0, 71.1, 113.1, 113.7, 119.7, 124.8, 125.9, 126.6, 126.7, 127.1, 128.1, 128.4, 128.8, 133.6, 133.7, 134.5, 142.0, 156.7, 167.6; υ_{max}/cm^{-1} (solid state) = 3373, 2862, 1698, 1618, 1440, 1254, 1023; ESI-MS found *m/z* 308 $[M+H]^+$;

4-Amino-3-(naphthalen-2-ylmethoxy)benzoic acid 18h

Procedure F; Methyl 4-amino-3-(naphthalen-2-ylmethoxy)benzoate **17h** (4.00g, 13.0 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (100 mL), 10% aqueous sodium hydroxide (25 mL). Following acidification, the precipitate was filtered, dissolved in chloroform and dried with magnesium sulfate. This solution was filtered and the organic solvents removed under reduced pressure and the solid was crystallised (chloroform/ methanol) to yield the title compound (3.19 g, 10.9 mmol, 84 %) as a colourless microcrystals; m.p. 161.8-163.5 (chloroform/ methanol); δ_H (500 MHz, DMSO-d_6) 5.32

(2H, s, Hα), 6.70 (1H, d, *J* = 8.2, H5), 7.38 (1H, d, *J* = 8.2, HAr), 7.45 (1H, s, H2), 7.50-7.54 (2H, m, HAr), 7.63 (1H, d, *J* = 8.1, H6), 7.91-7.95 (3H, m, HAr), 8.04 (1H, s, HAr); δ_C (75 MHz, DMSO-d_6); 69.8, 112.9, 113.1, 117.7, 124.6, 125.8, 126.1, 126.4 126.6, 127.9, 128.1, 128.3, 132.9, 133.2, 135.2, 143.3, 144.4, 167.9; υ_{max}/cm^{-1} (solid state) = 3401, 3000, 1678, 1613, 1521, 1441, 1263, 1149, 1025, 764; ESI-HRMS found *m/z* 316. 0983 $[M+Na]^+$, $C_{18}H_{15}NNaO_3$ requires 316.0944;

4-(((9H-Fluoren-9-yl)methoxy)carbonylamino)-3-(naphthalen-2-ylmethoxy)-benzoic acid 14h

Procedure H; 4-Amino-3-(naphthalen-2-ylmethoxy)benzoic acid **18h** (3.10 g, 10.6 mmol) in tetrahydrofuran (150 mL), fluorenylmethyloxycarbonyl chloride (4.10 g, 15.9 mmol) in tetrahydrofuran (50 mL). Following work up, the resulting solid which was crystallised (chloroform/ methanol) to yield the title compound (5.22 g, 9.1 mmol, 96%) as colourless microcrystals; m.p. 201.1-202.3 °C (chloroform/ methanol); δ_H (300 MHz, DMSO-d_6) 4.31 (1H, t, *J* = 6.9, FHβ), 4.46 (2H, d, *J* = 6.9, FHα), 5.36 (2H, s, Hα), 7.29 (2H, t, *J* = 7.4, FHAr4), 7.45 (2H, t, *J* = 5.4, FHAr3), 7.50-7.56 (3H, m, HAr), 7.63 (1H, s, H2), 7.66 (1H, d, *J* = 8.4, H6), 7.72-7.74 (3H, m, ArCH), 7.85-7.96 (5H, m, ArCH), 8.04 (1H, s, H8), 8.93 (1H, s, N<u>H</u>); δ_C (75 MHz, DMSO-d_6); 46.9, 66.7, 70.4, 113.6, 120.5, 121.3, 122.9, 125.6, 125.8, 126.4, 126.5, 126.7, 127.4, 128.0, 128.1, 128.2, 128.4, 132.2, 132.9, 133.1, 132.7, 141.1, 144.0, 148.8, 153.8, 167.2; υ_{max}/cm^{-1} (solid state) = 3423, 2927, 1673, 1602, 1440, 1216, 1034, 816, 762; ESI-HRMS found *m/z* 516.1800 $[M+H]^+$, $C_{33}H_{26}NO_5$ requires 516.1805.

Methyl 4-nitro-3-(4-trifluoromethyl)benzyloxy)benzoate 16i

Procedure A; Methyl-3-hydroxy-4-nitrobenzoate **15** (1.00 g, 5.1 mmol), potassium carbonate (2.l0 g, 15.2 mmol) in dimethylformamide (20 mL), 4-(trifluoromethyl) benzyl bromide (0.94 mL, 6.0 mmol) Following initial work-up, the solvents were removed under reduced pressure, dissolved in methanol, filtered to remove starting material and left to crystallise by slow evaporation to yield the title compound (1.05g, 3.0 mmol, 61%) as pale yellow plates; mp. 97.8-100.2°C (methanol); (Found C, 54.25; H, 3.35; N, 3.80%. $C_{16}H_{12}NO_5F_3$ requires C, 54.09.; H, 3.40; N, 3.94%); δ_H (500 MHz, $CDCl_3$) 3.98 (3H, s, CO_2<u>Me</u>), 5.26 (2H, s, Hα), 7.53 (d, 2H, *J* = 8.2, HAr3), 7.60 (2H, d, *J* = 8.2, HAr2), 7.67 (1H, d, *J* = 8.2, H6), 7.74 (1H, s, H2), 7.81 (1H, d, *J* = 8.2, H5); δ_C (125 MHz, $CDCl_3$); 54.7, 72.3, 117.7, 124.0, 125.8 (q, *J* = 270), 127.4, 127.6, 129.0, 132.4 (q, *J* = 32.5), 136.8, 140.9, 144.6, 152.9, 166.9; υ_{max}/cm^{-1} (solid state) = 3429, 2955, 1923, 1723 1591, 1334, 831, 590, 493; ESI-HRMS found *m/z* 378.0569 $[M+Na]^+$, $C_{16}H_{12}F_3NNaO_3$ requires 378.0560;

Methyl 4-amino-3-(4-(trifluoromethyl)benzyloxy)benzoate 17i

Procedure C; Methyl 4-nitro-3-(4-trifluoromethyl)benzyloxy)benzoate **16i** (1.00 g, 2.9 mmol), tin(II) chloride dihydrate (3.81 g, 16.9 mmol) in ethyl acetate (30 mL). Work up resulted in the title compound (0.93 g, 1.5 mmol, 97%) as a yellow amorphous solid; δ_H (500 MHz, MeOD-d_4) 3.79 (3H, s, CO_2<u>Me</u>) 4.19 (2H, s, N<u>H</u>$_2$), 5.12 (2H, s, Hα), 6.63 (1H, d, *J* = 8.2, H5), 7.46 (1H, d, *J* = 1.4, H2),

7.49-7.52 (3H, m, H6 + HAr2), 7.59 (2H, d, *J* = 8.1, HAr3); δ_C (75 MHz, $CDCl_3$) 52.6, 70.5, 113.5, 114.4, 120.4, 125.5, 126.6, 128.6, 129.7, (q, *J* = 270), 131.0 (q, *J* = 32.3) 141.5, 142.1, 145.7, 170.0; υ_{max}/cm^{-1} (solid state) = 3490, 3380, 2927, 1740, 1704, 1615, 1262, 1066, 705, 505; ESI-HRMS found *m/z* 326.0869 $[M+H]^+$, $C_{16}H_{15}F_3NO_3$ requires 326.0999;

4-Amino-3-(4-(trifluoromethyl)benzyloxy)benzoic acid 18i

Procedure F; Methyl 4-amino-3-(4-(trifluoromethyl)benzyloxy)benzoate **17i** (0.90 g, 2.7 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (24 mL), 10% aqueous sodium hydroxide (6.0 mL). Following acidification, the resulting precipitate was filtered, dissolved in chloroform and dried with magnesium sulfate. This solution was filtered and the organic solvents removed under reduced pressure yielding a pale yellow solid which was crystallised (chloroform) yielding the title compound (0.53 g, 1.7 mmol, 62%) as fine colourless microcrystals; m.p. 192.4 – 194.1°C (chloroform); δ_H (500 MHz, MeOD-d_4) 5.14 (2H, s, Hα), 6.63 (1H, d, *J* = 8.4, H5), 7.40-7.42 (2H, m, H6 + H2), 7.58 (4H, m, HAr2, + HAr3); δ_C (75 MHz, $CDCl_3$) 70.9, 79.9, 114.6 (q, *J* = 8.0), 119.9, 126.1 (q, *J* = 269.3), 126.5, 126.8 (q, *J* = 3.8), 129.2, 131.4 (q, *J* = 32.0), 143.5, 144.8, 146.4, 170.9; υ_{max}/cm^{-1} (solid state) = 3434, 3291, 3116, 2926, 2526, 2159, 2026, 1688, 1296, 1108, 765; ESI-HRMS found *m/z* 312.0856 $[M+H]^+$, $C_{15}H_{13}F_3NO_3$ requires 312.0842;

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-((4-(trifluoromethyl)benzyl)oxy)benzoic acid 14i

Procedure H; 4-amino-3-(4-(trifluoromethyl)benzyloxy)benzoic acid **18i** (0.46 g, 1.5 mmol) in chloroform (100 mL), fluorenylmethyloxycarbonyl chloride (0.58 g, 2.2 mmol) in chloroform (50 mL). Work up yielded the title compound (0.60 g, 1.1 mmol, 76%) as a colourless amorphous solid; (Found C, 67.55; H, 4.05; N, 2.25%. $C_{30}H_{22}F_3NO_5$ requires C, 67.54; H, 4.16; N, 2.63%); δ_H (500 MHz, DMSO-d_6) 4.32 (1H, t, *J* = 6.9, FHβ), 4.46 (2H, d, *J* = 6.9, FHα), 5.37 (2H, s, Hα), 7.31 (2H, t, *J* = 7.5, FHAr4), 7.43 (2H, t, *J* = 7.5, FHAr3), 7.60 (2H, d, *J* = 10.5, H5 + H2), 7.70 (1H, d, *J* = 8.4, H6), 7.75 (6H, m, HAr2, HAr3 + FHAr5), 7.91 (2H, *J* = 6.3, FHAr2), 9.10 (1H, s, N<u>H</u>); δ_C (75MHz, DMSO-d_6) 46.7, 66.6, 69.3, 113.4, 120.4, 121.7, 122.8, 125.6, 126.7, 127.4, 128.1, 128.2, 128.5, 129.1 (q, *J* = 30), 132.1, 141.1, 142.0, 144.0, 148.7, 153.9, 167.2; υ_{max}/cm^{-1} (solid state) = 3309, 3200-2200 (br), 1697, 1543, 1256, 1114, 739; ESI-HRMS found *m/z* 534.1500 $[M+H]^+$, $C_{30}H_{23}F_3NO_5$ requires 534.1523;

Methyl 4-nitro-3-(3-trifluoromethyl)benzyloxy)benzoate 16j

Procedure A; Methyl-3-hydroxy-4-nitrobenzoate **15** (1.00 g, 5.1 mmol), potassium carbonate (2.l0 g, 15.2 mmol) in dimethylformamide (20 mL), 3-(trifluoromethyl) benzyl bromide (0.94 mL, 6.0 mmol) Following preliminary work-up, the organic solvent was removed under reduced pressure and the solid dissolved in methanol, filtered to remove 4-(trifluoromethyl) benzyl bromide starting material and left to crystallise by slow evaporation to yield the title compound (1.25 g, 3.5 mmol, 69%) as pale

yellow microcrystals; m.p. 105.9 – 107.4 °C (methanol); (Found C, 54.10; H, 3.35; N, 3.80%. $C_{16}H_{12}NO_5F_3$ requires C, 54.09; H, 3.40; N, 3.94%); δ_H (500 MHz, $CDCl_3$) 3.90 (3H, s, $CO_2\underline{Me}$), 5.26 (2H, s, Hα), 7.48 (1H, apparent t, *J* = 7.5, HAr5), 7.55 (1H, d, *J* = 7.5, HAr4), 7.63 (1H, d, *J* = 7.5, HAr6), 7.66-7.68 (2H, m, H6 + HAr2), 7.75 (1H, s, H2), 7.82 (1H, d, *J* = 8, H5); δ_C (75 MHz, $CDCl_3$) 53.3, 71.0, 116.4 (d, *J* = 10.4), 122.6 (d, *J* = 11.8), 124.3 (q, *J* = 162.4), 124.3 (d, *J* = 9.3), 125.9, 126.1, 129.8 (d, *J* = 11.7), 130.8 (d, J = 11.8), 131.5 (q, *J* = 19.4), 135.4, 136.5, 143.1, 151.5, 165.4; υ_{max}/cm^{-1} (solid state) = 3439, 2955, 1952, 1737, 1365, 1229; ESI-MS found 378.1 $[M+Na]^+$;

Methyl 4-amino-3-(3-(trifluoromethyl)benzyloxy)benzoate 17j

Procedure C; Methyl 4-nitro-3-(3-trifluoromethyl)benzyloxy)benzoate **16j** (1.00 g, 2.9 mmol), tin(II) chloride dihydrate (3.81 g, 16.9 mmol) in ethyl acetate (30 mL). Work-up yielded the title compound (0.93 g, 2.9 mmol, 97%) as a colourless amorphous solid; δ_H (300 MHz, $CDCl_3$) 3.90 (3H, s, $CO_2\underline{Me}$), 4.31 (2H, s, $N\underline{H_2}$), 5.21 (2H, s, Hα), 6.75 (1H, d, *J* = 8.1, H5), 7.54-7.62 (2H, m, H6 + HAr6), 7.64-7.69, (3H, m, HAr2, HAr4 + HAr5), 7.76 (1H, s, H2); δ_C (75 MHz, $CDCl_3$) 52.2, 70.2, 113.0, 113.9, 119.9, 124.8, 124.9, 125.1, 125.4, 125.5, 131.4, 138.0, 141.7, 145.3, 167.6; υ_{max}/cm^{-1} (solid state) = 3492, 3356, 2957, 1888, 1772, 1691, 1614, 1266, 1110, 764; ESI-HRMS found *m/z* 348.0817 $[M+Na]^+$, $C_{16}H_{14}F_3NNaO_3$ requires 348.0818;

4-Amino-3-(3-(trifluoromethyl)benzyloxy)benzoic acid 18j

Procedure F; Methyl 4-amino-3-(3-(trifluoromethyl)benzyloxy)benzoate **17j** (0.63 g, 1.9 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (20 mL), 10% aqueous sodium hydroxide (11 mL). Following acidification, the precipitate was filtered, dissolved in chloroform and dried with magnesium sulfate. This solution was filtered and the organic solvents removed under reduced pressure yielding a pale yellow solid which was crystallised (chloroform) to yield the title compound (0.55 g, 1.8 mmol, 92%) as a colourless amorphous solid; (Found C, 57.40; H, 3.90; N, 4.25%. $C_{15}H_{12}NO_3F_3$ requires C, 57.88; H, 3.89; N, 4.50%); δ_H (300 MHz, MeOD-d_4) 5.12 (2H, s, Hα), 6.63 (1H, d, *J* = 8.7, H5), 7.39-7.42 (2H, m, H6 + HAr6), 7.44-7.53 (2H, m, H2 + HAr5), 7.64-7.68 (2H, m, HAr2 + HAr4); δ_C (75 MHz, MeOD-d_4) 70.8, 114.5, 149.6, 119.8, 125.5, 126.0, 126.0 (q, *J* = 270), 126.4, 130.7, 132.3 (q, *J* = 32), 132.6, 140.3, 144.7, 146.3, 170.8; υ_{max}/cm^{-1} (solid state) = 3515, 3418, 3379, 2927, 1879, 1761, 1673, 1524, 1227, 920, 765, 576; ESI-HRMS found *m/z* 312.0835 $[M+H]^+$, $C_{15}H_{13}F_3NO_3$ requires 326.1004;

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-((3-(trifluoromethyl)benzyl)oxy)benzoic acid 14j

Procedure H; 4-Amino-3-(3-(trifluoromethyl)benzyloxy)benzoic acid **18j** (0.42 g, 1.4 mmol) in chloroform (100 mL), fluorenylmethyloxycarbonyl chloride (0.58 g, 2.2 mmol) in chloroform (50 mL). Work up yielded the title compound (0.53 g, 1.0 mmol, 73%) as a colourless amorphous solid; δ_H (300 MHz, DMSO-d_6) 4.31 (1H, t, *J* = 7.0, FHβ), 4.45 (2H, d, *J* = 7.0, FHα), 5.35 (2H, s, Hα), 7.30

(2H, t, *J* = 6.4, FHAr4), 7.42 (2H, t, *J* = 6.4, FHAr3), 7.53 (1H, dd, *J* = 8.1, 1.5, H5), 7.60 (1H, d, *J* = 1.5, H2), 7.65-7.75 (4H, m, H6, HAr5 + FHAr5), 7.85-7.92 (3H, m, HAr6 + FHAr2), 7.97 (1H, s, HAr2), 9.14 (1H, s, NH); δ_C (125MHz, DMSO-d_6) 46.5, 66.3, 69.0, 113.2, 120.1, 121.3, 122.6, 124.2 (q, *J* = 2.6), 124.6 (q, *J* = 2.6), 125.2, 126.7, 127.4, 128.1, 129.8, 131.3 (q, *J* = 29.9*)* 131.9, 132.1, 138.7, 141.1, 144.0, 148.8, 153.9, 167.2; υ_{max}/cm^{-1} (solid state) = 3347, 3200-2200 (br), 1705, 1436, 1334, 1120, 739; ESI-HRMS found *m/z* 534.1505 $[M+H]^+$, $C_{30}H_{23}F_3NO_5$ requires 534.1523;

Methyl 3-((4-(tert-butoxy)benzyl)oxy)-4-nitrobenzoate 16k

Procedure B; Methyl-3-hydroxy-4-nitrobenzoate **15** (3.50 g, 11.7 mmol), (4-(tert-butoxy) phenyl) methanol (0.96 g, 5.4 mmol) triphenylphosphine (2.12 g, 8.1 mmol) with diisopropyl azodicarboxylate (1.59 mL, 8.1 mmol) in tetrahydrofuran (100 mL). Column chromatography yielded the title compound (1.80 g, 5.0 mmol, 95%) as pale yellow solid; δ_H (500 MHz, $CDCl_3$) 1.39 (9H, s, $C(CH_3)_3$), 3.99 (3H, s, CO_2Me), 5.26 (2H, s, Hα), 7.04 (2H, d, *J* = 8.2, HAr3), 7.39 (2H, d, *J* = 8.2, HAr2), 7.73 (1H, d, *J* = 8.2, 1.6, H6), 7.86-7.88 (2H, m, H2 + H5); δ_C (75MHz, $CDCl_3$) 28.9, 52.8, 71.4, 78.7, 116.3, 121.7, 124.2, 125.3, 128.1, 129.7, 134.8, 142.9, 151. 5, 155.7, 165.2; υ_{max}/cm^{-1} (solid state) = 3309, 2980, 1725, 1508, 1237, 895, 744; ESI-HRMS found *m/z* 382.1255 $[M+Na]^+$, $C_{19}H_{21}NNa_1O_6$ requires 382.1267.

3-((4-(Tert-butoxy)benzyl)oxy)-4-nitrobenzoic acid 17k

Procedure F; Methyl 3-((4-(tert-butoxy)benzyl)oxy)-4-nitrobenzoate **16k** (2.24 g, 6.2 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (60 mL), 10% aqueous sodium hydroxide (15 mL). Work up yielded the title compound (1.96 g, 5.7 mmol, 91%) as a yellow solid; δ_H (500 MHz, MeOD-d_4)1.36 (9H, s, $C(CH_3)_3$), 5.18 (2H, s, Hα), 6.96 (2H, d, *J* = 8.5, HAr3), 7.29 (2H, d, *J* = 8.4, HAr2), 7.71 (dd, *J* = 8.5, 1.5, H6), 7.79-7.81 (2H, m, H2 + H5); δ_C (125MHz, MeOD-d_4) 29.3, 72.3, 79.8, 117.6, 123.0, 125.2, 126.0, 128.9, 129.5, 132.1, 144.6, 152.4, 156.8, 167.7; υ_{max}/cm^{-1} (solid state) = 2977, 1692, 1607, 1520, 1506, 1434, 1294, 1250, 1162, 1109; ESI-HRMS found *m/z* 344.1156 $[M-H]^-$, $C_{18}H_{18}NO_6$ requires 344.1134.

4-Amino-3-((4-(tert-butoxy)benzyl)oxy)benzoic acid 18k

Procedure E; 3-((4-(tert-butoxy)benzyl)oxy)-4-nitrobenzoic acid **17k** (0.50 g, 1.5 mmol), cobalt chloride hexahydrate (2.07 g, 8.7 mmol) and sodium borohydride (0.68 g, 174.0 mmol) in methanol (100 mL). Work up yielded the title compound (336 mg, 1.1 mmol, 74%) as an orange amorphous solid; δ_H (500 MHz, $CDCl_3$) 1.36 (9H, s, $C(CH_3)_3$), 5.00 (2H, s, Hα), 6.64 (1H, d, *J* = 8 , H5), 6.95 (2H, d, *J* = 8.5, HAr3), 7.27 (2H, d, *J* = 8.5, HAr2), 7.54 (1H, d, *J* = 1.6, H2), 7.58 (1H, dd, *J* = 8.2, 1.6, H6); δ_C (125MHz, $CDCl_3$) 28.9, 70.4, 78.7, 113.1, 113.3, 118.4, 124.1, 125.3, 128.7, 131.2, 142.1, 145.3, 155.5, 171.7; υ_{max}/cm^{-1} (solid state) = 3491, 3384, 2978, 2932, 1692, 1613, 1509, 1421; ESI-HRMS found *m/z* 338.1355 $[M+H]^+$, $C_{18}H_{21}NNaO_4$ requires 338.1363.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-((4-(tert-butoxy)benzyl)oxy)benzoic acid 14k

Procedure I; 4-amino-3-((4-(tert-butoxy)benzyl)oxy)benzoic acid **18k** (288 mg, 0.9 mmol), sodium bicarbonate (153 mg, 1.8 mmol) in tetrahydrofuran (15 mL) and fluorenylmethyloxycarbonyl chloride (354 mg, 1.4 mmol) in tetrahydrofuran (10 mL). The pure product was precipitated from a dichloromethane-hexane solution to leave a grey amorphous solid (302 mg, 0.6 mmol, 62%); δ_H (500 MHz, $CDCl_3$) 1.31 (9H, s, $C(CH_3)_3$), 4.23 (1H, t, *J* = 6.8, FHβ), 4.45 (2H, d, *J* = 6, FHα), 5.07 (2H, s, Hα), 6.99 (2H, d, *J* = 8.2, HAr2), 7.23-7.29 (4H, m, HAr3 + FHAr4), 7.35 (2H, t, *J* = 7.3, FHAr3), 7.47 (1H, br. s, H5), 7.53 (2H, d, *J* = 7.3, FHAr5), 7.62 (1H, s, H2), 7.69-7.72 (3H, m, H6 + FHAr2), 8.10 (1H, br. s, N<u>H</u>); δ_C (125MHz, $CDCl_3$) 28.89, 47.01, 67.43, 71.00, 78.81, 128.69, 112.85, 117.38, 120.07, 123.20, 124.21, 124.62, 124.97, 127.14, 127.84, 130.36, 141.34, 143.59, 146.28, 155.86, 170.40, υ_{max}/cm^{-1} (solid state) = 3421, 2973, 1742, 1674, 1594, 1538; ESI-HRMS found *m/z* 560.2030 $[M+Na]^+$, $C_{33}H_{31}NNaO_6$requires 560.2044.

3-(2-(1H-Indol-3-yl)ethoxy)-4-aminobenzoic acid 18l

Methyl 3-(2-(1H-indol-3-yl)ethoxy)-4-nitrobenzoate **16l** was obtained from methyl-3-hydroxy-4-nitrobenzoate **15** by *procedure B* without purification/isolation. Methyl 3-(2-(1H-indol-3-yl)ethoxy)-4-aminobenzoate **17l** was obtained from methyl 3-(2-(1H-indol-3-yl)ethoxy)-4-nitrobenzoate **16l** by *procedure C* without purification/isolation. *Procedure F*; Methyl 3-(2-(1H-indol-3-yl)ethoxy)-4-aminobenzoate **17l** (2.50 g, 8.1 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (65 mL), 10% aqueous sodium hydroxide (15 mL). Work up yielded the title compound (1.90 g, 6.4 mmol, 80%) as a cream amorphous solid; δ_H (500 MHz, DMSO-d_6) 3.18 (2H, t, *J* = 6.7, Hβ), 4.19 (2H, t, *J* = 6.7 Hα), 5.46 (2H, br. s, N<u>H</u>$_2$), 6.62 (1H, d, *J* = 8.1, H5), 6.98 (1H, m, IHAr3), 7.07 (1H, m, IHAr4), 7.30-7.36 (4H, m, H2, H6, IHAr5 + IC<u>H</u>NH), 7.61 (1H, d, *J* = 7.7, IHAr2), 10.87 (1H, s, ICHN<u>H</u>); (125 MHz, DMSO-d_6) 24.8, 68.4, 110.6, 111.4, 111.9, 112.1, 117.3, 118.2, 118.3, 120.9, 123.3, 124.0, 127.3, 136.1, 142.8, 144.2, 167.5; υ_{max}/cm^{-1} (solid state) =3488, 3387, 250, 1882 (br) 1667, 1614, 1447, 1274, 745; ESI-HRMS found *m/z* 319.1048 $[M+Na]^+$, $C_{17}H_{16}N_2NaO_3$ requires 319.1053.

3-(2-(1H-Indol-3-yl)ethoxy)-4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)benzoic acid 14l

Procedure H; 3-(2-(1H-indol-3-yl)ethoxy)-4-aminobenzoic acid **18l** (1.58 g, 5.33 mmol) in tetrahydrofuran (50 mL), fluorenylmethyloxycarbonyl chloride (2.07 g, 87.00 mmol) in tetrahydrofuran (20 mL). Work up yielded the title compound as a light grey amorphous solid (2.25 g, 4.3 mmol, 81%); δ_H (500 MHz, DMSO-d_6) 3.25 (2H, t, *J* = 6.9, Hβ), 4.30-4.38 (3H, m, Hα + FHβ), 4.47 (2H, d, J = 6.9, FHα), 6.97 (1H, m, IHAr3), 7.06 (1H, m, IHAr4), 7.28 (1H, d, *J* = 2.1, IC<u>H</u>NH), 7.31-7.36 (3H, m, IHAr5 + FHAr4), 7.44 (2H, t, *J* = 7.4, FHAr3), 7.49-7.54 (2H, m, H2 + H5), 7.62 (1H, d, *J* = 7.7, IHAr2), 7.74-7.77 (3H, m, H6 + FHAr5), 7.92 (2H, d, *J* = 7.5, FHAr2), 8.86 (1H, s, FN<u>H</u>), 10.92 (1H, s, ICHN<u>H</u>); (125MHz, DMSO-d_6) 24.6, 46.5, 66.4, 69.0, 110.3, 111.4, 112.4, 118.2, 118.5, 119.7, 120.2, 121.0, 122.3, 123.2, 125.3, 125.9, 127.2, 127.3, 127.8, 131.6, 136.2, 140.8,

143.7, 148.0, 153.3, 166.9; υ_{max}/cm^{-1} (solid state) = 3358, 2951, 1908, 1714, 1673, 1439,1231, 1053, 736; ESI-HRMS found *m/z* 541.1727 $[M+Na]^+$, $C_{32}H_{26}N_2NaO_5$ requires 541.1734.

Methyl 3-(2-(tert-butoxy)-2-oxoethoxy)-4-nitrobenzoate 16m

Procedure A; Methyl-3-hydroxy-4-nitrobenzoate **15** (5.00 g, 25.4 mmol), potassium carbonate (3.50 g, 25.4 mmol) in dimethylformamide (100 mL), *tert*-butyl 2-bromoacetate (4.44 mL, 30.4 mmol). Following work up, the resulting solid was crystallised (ethyl acetate) to yield the title compound (6.26 g, 20.1mmol, 79 %) as large pale yellow crystals; m.p. 73.9-74.9°C (ethyl acetate); (Found C, 54.05; H, 5.50; N, 4.50%. $C_{14}H_{17}NO_7$ requires C, 54.02; H, 5.50; N, 4.50%); δ_H (300 MHz, $CDCl_3$) 1.50 (9H, s, $C(C\underline{H}_3)_3$), 3.97 (3H, s, $CO_2\underline{Me}$), 4.74 (2H, s, Hα), 7.63 (1H, d, *J* = 1.5, H2), 7.75 (1H, dd, *J* = 8.4, 1.5, H6), 7.89 (1H, d, *J* = 8.4, H5); δ_C (75MHz, $CDCl_3$) 28.0, 52.9, 66.5, 79.9, 83.4, 115.7, 122.3, 125.6, 134.7, 150.9, 165.0, 166.3; υ_{max}/cm^{-1} (solid state) = 2981, 1730, 1533,1305, 1225,745; ESI-HRMS found *m/z* 334.0908 $[M+Na]^+$, $C_{14}H_{17}NNaO_7$ requires 334.0897.

3-(2-(Tert-butoxy)-2-oxoethoxy)-4-nitrobenzoic acid 17m

Saponification of **16m** preceded reduction to prevent hydrolysis of the ^{t}Butyl group. *Procedure G*; Methyl 3-(2-(tert-butoxy)-2-oxoethoxy)-4-nitrobenzoate **16m** (3.00 g, 9.6 mmol) in tetrahydrofuran (100 mL), lithium hydroxide (0.40 g, 9.6 mmol) in water (100 mL). Following acidification the resulting precipitate was filtered, dissolved in chloroform and dried with magnesium sulfate. The solution was filtered and the organic solvents removed under reduced pressure yielding the title compound (1.72 g, 5.8 mmol, 60%) as a white amorphous powder; (Found C, 52.55; H, 5.00; N, 4.60%. $C_{13}H_{15}NO_7$ requires C, 52.53; H, 5.09; N, 4.71%); δ_H (300 MHz, $CDCl_3$) 1.42 (9H, s, $C(C\underline{H}_3)_3$, 4.69 (2H, s, Hα), 7.62 (1H, d, *J* = 1.5, H2), 7.76 (1H, dd, *J* = 8.1, 1.5, H6), 7.83 (1H, d, *J* = 8.1, H5); δ_C (75 MHz, $CDCl_3$) 28.0, 66.5, 83.5, 116.2, 123.1, 125.7, 133.5, 150.9, 166.3, 169.4; υ_{max}/cm^{-1} (solid state) = 3118-2555, 2984, 1740, 1695, 1532, 1432, 1265, 788; ESI-HRMS found *m/z* 320.0740 $[M+Na]^+$, $C_{13}H_{15}NNaO_7$ requires 320.0741.

4-Amino-3-(2-(tert-butoxy)-2-oxoethoxy)benzoic acid 18m

Procedure D; 3-(2-(tert-butoxy)-2-oxoethoxy)-4-nitrobenzoic **17m** (2.00 g, 6.7 mmol) in methanol (20 mL), 10% palladium on charcoal (200 mg, 10 wt%) in methanol (10 mL)) and hydrogen gas. Work up yielded the title compound (1.66 g, 6.2 mmol, 81%) as a beige amorphous powder; δ_H (500 MHz, MeOD-d_4) 1.53 (9H, s, $C(C\underline{H}_3)_3$), 4.65 (2H, s, Hα), 6.74 (1H, d, *J* = 8.3, H5), 7.39 (1H, d, *J* = 1.8, H2), 7.54 (1H, dd, *J* = 8.2, 1.7, H6); δ_C (125MHz, MeOD-d_4) 28.3, 31.2, 67.5, 114.5, 116.7, 118.8, 119.4, 125.7, 126.5, 145.7, 170.4; υ_{max}/cm^{-1} (solid state) = 3523, 3427, 3123-2565, 2988, 1740, 1685, 1537, 1454, 1276, 788; ESI-HRMS found *m/z* 268.1181 $[M+H]^+$, $C_{13}H_{18}NO_5$ requires 268.1179.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(2-(tert-butoxy)-2-oxoethoxy)benzoic acid 14m

Procedure I; 4-amino-3-(2-(tert-butoxy)-2-oxoethoxy)benzoic acid **18m** (1.00 g, 3.7 mmol), sodium bicarbonate (0.35 g, 4.1 mmol) in tetrahydrofuran (50 mL), fluorenylmethyloxycarbonyl chloride (0.87 g, 3.4 mmol) in chloroform (30 mL). The resulting precipitate was removed *via* filtration to yield the title compound (1.62 g, 3.3 mmol, 98%) as a colourless amorphous solid; (Found C, 68.40; H, 5.50; N, 2.80%. $C_{28}H_{27}NO_7$ requires C, 68.70; H, 5.56; N, 2.86%); δ_H (300 MHz, $CDCl_3$) 1.54 (9H, s, C(C<u>H</u>$_3$)$_3$), 4.34 (1H, t, *J* = 7.1, FHβ), 4.55 (2H, d, *J* = 7.1, FHα), 4.70 (2H, s, Hα), 7.36 (2H, t, *J* = 7.2, FHAr4) 7.45 (2H, t, *J* = 7.2, FHAr3), 7.58 (1H, d, *J* = 1.6, H2), 7.71 (2H, d, *J* = 7.3, FHAr5), 7.80-7.86 (3H, m, H6 + FHAr2), 8.18-8.29 (2H, m, N<u>H</u> + H5); δ_C (75MHz, $CDCl_3$) 28.1, 47.0, 67.4, 67.6, 83.2, 114.6, 117.8, 120.1, 123.2, 125.2, 125.8, 127.2, 127.8, 134.2, 141.4, 143.7, 146.1, 153.1, 167.9, 171.2; υ_{max}/cm^{-1} (solid state) = 3422, 3200-2200 (br), 1740, 1682, 1533, 1182, 760; ESI-HRMS found *m/z* 512.1675 $[M+Na]^+$, $C_{28}H_{27}NNaO_7$ requires 512.1680.

Methyl 3-(2-(methoxymethoxy)ethoxy)-4-nitrobenzoate 16n

Procedure A; Methyl-3-hydroxy-4-nitrobenzoate **15** (3.00 g, 15.2 mmol), potassium carbonate (3.15 g, 22.8 mmol), in dimethylformamide (60 mL), 2-(methoxymethoxy)ethanol (1.97 mL, 16.7 mmol). Work up afforded the title compound (3.65 g, 12.8 mmol, 89 %) as a yellow glassy solid; (Found C, 50.80; H, 5.21; N, 4.75%. $C_{19}H_{21}N_1O_5$ requires C, 50.33; H, 5.30; N, 4.91%); δ_H (500 MHz, $CDCl_3$) 3.26 (3H, s, O<u>Me</u>); 3.87-91 (5H, m, Hβ + CO_2<u>Me</u>), 4.28 (2H, t, *J* = 5.0, Hα), 4.64 (2H, s, OC<u>H</u>$_2$O), 7.63 (1H, dd, *J* = 8.5, 1.0, H6), 7.71 (1H, d, *J* = 1, H2), 7.77 (1H, d, *J* = 8.5, H5); δ_C (125MHz, $CDCl_3$) 52.8, 55.3, 65.2, 69.4, 96.6, 115.7, 121.6, 125.4, 134.8, 142.6, 151.7, 165.1; υ_{max}/cm^{-1} (solid state) = 2993, 2964, 2941, 2890, 1726, 1613, 1591, 1528, 1431, 1369, 1292, 1241, 1114, 1055, 1025; ESI-HRMS found *m/z* 308.074 $[M+Na]^+$, $C_{12}H_{15}NNaO_7$ requires 308.0746

Methyl 4-amino-3-(2-(methoxymethoxy)ethoxy)benzoate 17n

Procedure D; Methyl 3-(2-(methoxymethoxy)ethoxy)-4-nitrobenzoate **16n** (3.50g, 12.3 mmol) in methanol (70 mL), 10% palladium on charcoal (350 mg, 10 wt%) in methanol (20 mL) and hydrogen gas. Work up yielded the title compound (3.01 g, 11.8 mmol, 96%) as a colourless oil; δ_H (500 MHz, $CDCl_3$) 3.42 (3H, s, O<u>Me</u>); 3.88 (3H, s, CO_2<u>Me</u>), 3.94 (2H, t, *J* = 4.5, Hβ), 4.25 (2H, t, *J* = 4.5, Hα), 4.15 (2H, br. s, NH_2), 4.73 (2H, s, OC<u>H</u>$_2$O), 6.70 (1H, d, *J* = 8.2, H5), 7.50 (1H, d, *J* = 1.8, H2), 7.58 (1H, dd, *J* = 8.2, 1.8, H6); δ_C (125MHz, $CDCl_3$) 51.6, 55.3, 66.1, 68.2, 96.6, 113.1, 113.3, 119.4, 124.6, 141.7, 145.1, 167.2; υ_{max}/cm^{-1} (solid state) = 3481, 3366, 2949, 2887, 1704, 1619, 1521, 1442, 1294, 1264, 1217, 1152, 1108, 1037; ESI-HRMS found *m/z* 278.0999 $[M+Na]^+$, $C_{11}H_{15}NNaO_5$ requires 278.1004.

Methyl 4-amino-3-(2-(methoxymethoxy)ethoxy)benzoic acid 18n

Procedure G; Methyl 4-amino-3-(2-(methoxymethoxy)ethoxy)benzoate **17n** (3.00 g, 11.8 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (80 mL), lithium hydroxide (1.00 g, 23.3 mmol) in water (5 mL). Work up yielded the title compound (2.02 g, 8.4 mmol, 76%) as a colourless amorphous solid; δ_H (300 MHz, $CDCl_3$) 3.33 (3H, s, O<u>Me</u>), 3.86 (2H, t, *J* = 4.5, Hβ), 4.16 (2H, t, *J* = 4.5, Hα), 4.65 (2H, s, OC<u>H</u>$_2$O), 6.66 (1H, d, *J* = 8.1, H5), 7.43 (1H, d, *J* = 1.7, H2),7.52 (1H, dd, *J* = 8.1, 1.7, H6); δ_C (75MHz, $CDCl_3$) 55.3, 66.1, 68.2, 96.5, 113.3, 113.4, 118.2, 125.6, 142.5, 145.0, 172.1; υ_{max}/cm^{-1} (solid state) = 3346, 2938, 1711, 1679, 1622, 1595, 1524, 1445, 1302, 1267, 1233, 1148, 1112, 1044; ESI-HRMS found *m/z* 242.1023 $[M+H]^+$, $C_{11}H_{16}NO_5$ requires 242.1028.

4-(((9H-Fluoren-9-yl)methoxy)carbonylamino)-3-(2-(methoxymethoxy)ethoxy)benzoic acid 14n

Procedure I; Methyl 4-amino-3-(2-(methoxymethoxy)ethoxy)benzoic acid **18n** (2.20 g, 9.1 mmol), sodium bicarbonate (1.03 g, 27.4 mmol) in tetrahydrofuran (50 mL) and fluorenylmethyloxycarbonyl chloride (3.54 g, 13.7 mmol) in tetrahydrofuran (30 mL). The reaction mixture was concentrated and column chromatography yielded the title compound (2.28 g, 4.9 mmol, 54%) as a colourless amorphous solid; (Found C, 67.40, H:5.40, N: 2.85%. $C_{26}H_{25}NO_7$ requires C, 67.38, H: 5.44, N: 3.02%); δ_H (500 MHz, $CDCl_3$) 3.29 (3H, s, O<u>Me</u>), 3.88 (2H, t, *J* = 4.5, Hβ), 4.20 (2H, t, *J* = 4.5, Hα), 4.23 (1H, t, *J* = 6.6, FHβ), 4.50 (2H, d, *J* = 6.6, FHα), 4.62 (2H, s, OC<u>H</u>$_2$O), 7.26 (2H, t, *J* = 7.5, FHAr4), 7.35 (2H, t, *J* = 7.4, FHAr3), 7.56-7.58 (2H, m, H2 + H5), 7.75 (2H, d, *J* = 7.3, FHAr5), 7.86 (1H, br. s, H6) 7.93 (2H, d, *J* = 7.6, FHAr2), 8.63 (1H, s, N<u>H</u>); δ_C (125MHz, $CDCl_3$) 46.4, 54.5, 65.5, 66.1, 68.8, 95.7, 113.5, 119.2, 120.1, 122.8, 125.0, 125.6, 127.0, 127.7, 132.0, 140.7, 143.5, 147.6, 153.0, 166.7; υ_{max}/cm^{-1} (solid state) = 3299, 2945, 1709, 1682, 1597, 1539, 1499, 1415, 1300, 1246, 1114, 1088, 1043; ESI-HRMS found *m/z* 464.1673 $[M+H]^+$, $C_{26}H_{26}NO_7$ requires 464.1709.

Methyl 3-(3-(methylthio)propoxy)-4-nitrobenzoate 16o

Procedure B; Methyl-3-hydroxy-4-nitrobenzoate **15** (5.00 g, 25.3 mmol), 3-(methylthio)propan-1-ol (2.87 mL, 30.4 mmol) triphenylphosphine (9.95 g, 38.0 mmol) with diisopropyl azodicarboxylate (7.45 mL, 38.0 mmol) in tetrahydrofuran (150 mL). Column chromatography yielded the title compound (6.51 g, 22.8 mmol, 90%) as pale yellow solid; δ_H (300 MHz, $CDCl_3$) 2.13-2.21 (5H, m, Hβ, SC<u>H</u>$_3$), 2.78 (2H, t, *J* = 6.9, Hγ), 4.00 (3H, s, CO_2<u>Me</u>), 4.33 (2H, t, *J* = 5.7, Hα), 7.72 (1H, dd, *J* = 8.4, H6), 7.79 (1H, d, *J* = 1.2, H2), 7.87 (1H, d, *J* = 8.4, H5); δ_C (75MHz, $CDCl_3$) 15.5, 28.1, 30.3, 52.8, 67.8, 115.4, 121.4, 125.3, 134.8, 142.4, 151.7, 165.2; υ_{max}/cm^{-1} (solid state) = 2953, 2857, 1719, 1609, 1587, 1520, 1433, 1390, 1296, 1237, 1113; ESI-HRMS found *m/z* 308.0563 $[M+Na]^+$, $C_{12}H_{15}NNaO_5S$ requires 308.0569.

Methyl 4-amino3-(3-(methylthio)propoxy)benzoate 17o

Procedure D; Methyl 3-(3-(methylthio)propoxy)-4-nitrobenzoate **16o** (4.00 g, 12.3 mmol) in ethyl acetate (150 mL) and methanol (50 mL), 10% palladium on charcoal (400 mg, 10 wt%) and hydrogen

gas. Following work up, column chromatography yielded the title compound (2.62 g, 10.3 mmol, 73%) as a beige amorphous solid; δ_H (300 MHz, $CDCl_3$) 2.01-2.10 (5H, m, Hβ, SC<u>H</u>$_3$), 2.63 (2H, t, *J* = 7.2, Hγ), 3.79 (3H, s, CO_2<u>Me</u>), 4.09 (2H, t, *J* = 6.1, Hα), 4.17 (2H, br. s, N<u>H</u>$_2$), 6.60 (1H, d, *J* = 8.1, H5), 7.38 (1H, d, *J* = 1.8, H2), 7.47 (1H, dd, *J* = 8.1, 1.8, H6); δ_C (125MHz, $CDCl_3$) 15.5, 28.6, 30.8, 51.6, 66.7, 112.1, 113.1, 119.3, 124.1, 141.2, 145.1, 167.2; ESI-HRMS found *m/z* 256.1002 $[M+H]^+$, $C_{12}H_{18}NO_3S$ requires 256.1007.

4-Amino-3-(3-(methylthio)propoxy)benzoic acid 18o

Procedure F; Methyl 4-amino3-(3-(methylthio)propoxy)benzoate **17o** (2.50 g, 9.8 mmol) in a 1:1 mixture of methanol- tetrahydrofuran (60 mL), 10% aqueous sodium hydroxide (15 mL). Work up yielded the title compound (2.27 g, 9.4 mmol, 96%) as a cream amorphous solid; δ_H (500 MHz,$CDCl_3$) 2.12-2.17 (5H, m, Hβ, SC<u>H</u>$_3$), 2.71 (2H, t, *J* = 7.2, Hγ), 4.19 (2H, t, *J* = 6.1, Hα), 6.70 (1H, d, *J* = 8.3, H5), 7.51 (1H, d, *J* = 1.8, H2), 7.64 (1H, dd, *J* = 8.3, 1.8, H6); δ_C (125MHz, $CDCl_3$) 15.7, 28.7, 31.0, 66.9, 112.6, 113.2, 118.4, 125.2, 142.1, 145.2, 171.95; υ_{max}/cm^{-1} (solid state) = 3461, 3334, 2869, 2551, 1667, 1615, 1583, 1526, 1446, 1414, 1365, 1300, 1267, 1225, 1148, 1113, 1031; ESI-HRMS found *m/z* 264.0665 $[M+Na]^+$, $C_{11}H_{15}NNaO_3S$ requires 264.0670.

4-(((9H-Fluoren-9-yl) methoxy)carbonylamino)-3-(3-(methylthio)propoxy)benzoic acid 14o

Procedure H; 4-Amino-3-(3-(methylthio)propoxy)benzoic acid **18o** (1.70 g, 7.1 mmol), in tetrahydrofuran (50 mL) and fluorenylmethyloxycarbonyl chloride (2.72 g, 10.6 mmol) in tetrahydrofuran (30 mL). The crude material obtained after concentration was suspended in hexane and filtered (3 times). The solid obtained was then washed with methanol to get the title compound (2.94g, 6.3 mmol, 90%) as a colourless amorphous solid; (Found C, 67.45.; H, 5.35; N, 2.95; S 6.85%. $C_{26}H_{25}NO_5S$ requires C, 67.37; H, 5.44; N, 3.02; S, 6.92%); δ_H (500 MHz,$CDCl_3$) 2.19 (3H, s, SC<u>H</u>$_3$), 2.21 (2H, p, *J* = 6.4, Hβ), 2.74 (2H, t, *J* = 6.7, Hγ), 4.26 (2H, t, *J* = 6.1, Hα), 4.33 (1H, t, *J* = 6.8, FHβ), 4.56 (2H, d, *J* = 6.8, FHα), 7.35 (2H, t, *J* = 7.5, FHAr4), 7.44 (2H, t, *J* = 7.5, FHAr3), 7.61 (1H, d, *J* = 1.7, H2), 7.65 (2H, d, *J* = 7.5, FHAr5), 7.70 (1H, br. s, H6), 7.77 (1H, d, *J* = 8.1, H5), 7.80 (2H, FHAr2), 8.18 (1H, br. s, N<u>H</u>); δ_C (125MHz, $CDCl_3$) 15.9, 28.3, 31.2, 47.1, 67.4, 67.8, 112.1, 117.3, 120.1, 123.2, 124.5, 125.0, 127.2, 127.9, 133.0, 141.4, 143.7, 146.3, 153.0, 177.6; υ_{max}/cm^{-1} (solid state) = 3311, 2954, 2915, 1712, 1682, 1597, 1535, 1499, 1416, 1338, 1300, 1281, 1245, 1227, 1104, 1088, 1046, 1032; ESI-HRMS found *m/z* 486.1341 $[M+Na]^+$, $C_{26}H_{25}NNaO_5S$ requires 486.1351.

Methyl 4-nitro-3-(2-oxo-2-(tritylamino)ethoxy)benzoate 16p

Procedure A; Methyl-3-hydroxy-4-nitrobenzoate **15** (1.17 g, 5.9 mmol), potassium carbonate (2.46 g, 17.8 mmol), in dimethylformamide (40 mL), 2-bromo-N-tritylacetamide (2.25 g, 5.9 mmol). Work-up afforded the title compound (2.73 g, 5.5 mmol, 93 %) as a yellow amorphous solid; δ_H (500 MHz, $CDCl_3$) 3.90 (3H, s, CO_2<u>Me</u>), 4.60 (2H, s, Hα), 7.18-7.26 (15H, m, C(C_6<u>H</u>$_5$)$_3$), 7.66 (1H, d, *J* = 1.4,

H2), 7.72 (1H, dd, *J* = 8.5, 1.4, H6), 7.95 (1H, d, *J* = 8.5, H5), 8.06 (1H, s, NH); δ_C (125MHz, $CDCl_3$) 53.0, 68.3, 70.7, 115.6, 122.9, 126.4, 127.2, 128.1, 128.7, 135.8, 144.2, 150.2, 164.6, 165.0; υ_{max}/cm^{-1} (solid state) = 3408, 3056, 1958, 1728, 1697, 1522, 1234, 700; ESI-HRMS found *m/z* 519.1525 $[M+Na]^+$, $C_{29}H_{24}N_2NaO_6$ requires 519.1527.

Methyl 4-amino-3-(2-oxo-2-(tritylamino)ethoxy)benzoate 17p

Procedure D; Methyl 4-nitro-3-(2-oxo-2-(tritylamino)ethoxy)benzoate **16p** (2.70 g, 5.4 mmol) in ethyl acetate (60 mL), 10% palladium on charcoal (270 mg, 10 wt%) and hydrogen gas. Work up yielded the title compound (2.45 g, 5.3 mmol, 96%) as a beige solid; (Found C, 74.15; H, 5.75; N, 5.65%. $C_{29}H_{26}N_2O_4$ requires C, 74.66; H, 5.64; N, 6.00%); δ_H (500 MHz, $CDCl_3$) 3.80 (3H, s, CO_2Me), 4.45 (2H, s, Hα), 6.71 (1H, d, *J* = 8.2, H5), 7.09-7.10 (6H, m, ArH), 7.15-7.22 (9H, m, ArH), 7.45 (1H, d, *J* = 1.4, H2), 7.53-7.55 (2H, m, H6 + NH); δ_C (125MHz, $CDCl_3$) 51.9, 68.7, 70.4, 113.7, 114.6, 120.4, 125.6, 127.2, 128.1, 128.6, 140.6, 143.9, 144.3, 166.7, 166.7; υ_{max}/cm^{-1} (solid state) = 3488, 3370, 3033, 1966, 1691, 1619, 1520, 1434, 1258, 698; ESI-HRMS found *m/z* 489.1799 $[M+Na]^+$, $C_{29}H_{26}N_2NaO_6$ requires 489.1785.

4-Amino-3-(2-oxo-2-(tritylamino)ethoxy)benzoic acid 18p

Procedure F; Methyl 4-amino-3-(2-oxo-2-(tritylamino)ethoxy)benzoate **17p** (2.40 g, 5.1 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (60 mL), 10% aqueous sodium hydroxide (15 mL). Work up yielded the title compound (1.46 g, 3.2 mmol, 63%) as a colourless amorphous solid; δ_H (500 MHz, DMSO-d_6) 4.71 (2H, s, Hα), 5.53 (2H, br. s, NH_2) 6.64 (1H, d, *J* = 8.5, H5), 7.15-7.26 (15H, m, $C(C_6H_5)_3$),7.38-7.39 (2H, m, H2 + H6) 8.69 (1H, brs, NH); δ_C (125MHz, $CDCl_3$) 69.2, 71.7, 93.4, 115.0, 115.5, 125.8, 128.1, 128.7, 129.0, 130.0, 141.2, 145.6, 145.7, 170.3, 175.6; ESI-HRMS found *m/z* 451.1667 $[M-H]^-$, $C_{28}H_{23}N_2O_4$ requires 451.1663.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-(2-oxo-2-(tritylamino)ethoxy)benzoic acid 14p

Procedure I; 4-Amino-3-(2-oxo-2-(tritylamino)ethoxy)benzoic acid **18p** (1.40 g, 3.1 mmol), sodium bicarbonate (350 mg, 9.3 mmol) in tetrahydrofuran (50 mL) and fluorenylmethyloxycarbonyl chloride (1.20 g, 4.6 mmol) in tetrahydrofuran (10 mL). Work up yielded the title compound (1.52 g, 2.3 mmol, 73%) as a colourless amorphous solid; δ_H (500 MHz, DMSO-d_6) 4.29 (1H, t, *J* = 6.6, FHβ), 4.45 (2H, d, *J* = 6.6, FHα), 4.85 (2H, s, Hα), 7.13-7.14 (6H, m, ArH), 7.14-7.23 (9H, m, ArH), 7.29 (2H, t, *J* = 7.4, FHAr4), 7.41 (2H, t, *J* = 7.4, FHAr3), 7.55-7.56 (2H, m, H2 + H6), 7.73 (2H, d, *J* = 7.6, FHAr5), 7.78 (1H, br. s, H5), 7.90 (2H, d, *J* = 7.6, FHAr2), 8.83 (1H, s, NH), 9.05 (1H, s, FNH); δ_C (125MHz, DMSO-d_6) 46.7, 66.3, 67.6, 69.4, 113.4,119.8, 120.1, 123.0, 125.2, 125.7, 126.5, 127.1, 127.5, 127.7, 128.4, 132.0, 140.7, 143.6, 144.4, 147.3, 153.3, 166.8, 167.1; υ_{max}/cm^{-1} (solid state) = 3397, 3261, 3059, 1951, 1736, 1668, 1530, 1221, 1190, 1046, 739, 699; ESI-HRMS found *m/z* 697.2284 $[M+Na]^+$, $C_{43}H_{34}N_2NaO_6$ requires 697.2309.

Methyl 3-(2-((tert-butoxycarbonyl)amino)ethoxy)-4-nitrobenzoate 16q

Procedure B; Methyl-3-hydroxy-4-nitrobenzoate **15** (3.00 g, 15.2 mmol), tert-butyl (2-hydroxyethyl)carbamate (2.59 mL, 16.7 mmol), triphenylphosphine (6.00 g, 22.8 mmol) with diisopropyl azodicarboxylate (4.48 mL, 22.8 mmol) in tetrahydrofuran (90 mL). Column chromatography yielded the title compound (4.14 g, 12.2 mmol, 80%) as a pale yellow solid; (Found C, 53.00; H5.90; N, 8.20%. $C_{15}H_{20}N_2O_7$ requires C, 52.94; H, 5.92; N, 8.23%); δ_H (500 MHz, $CDCl_3$); 1.45 (9H, s, C(C<u>H</u>$_3$)), 3.61 (2H, m, Hβ), 3.96 (3H, s, CO_2<u>Me</u>), 4.24 (2H, t, *J* = 4.8, Hα), 5.09 (1H, br. s, N<u>H</u>), 7.71-7.73 (2H, m, H2 + H6), 7.87 (1H, d, *J* = 8.3, H5); δ_C (125MHz, $CDCl_3$); 28.4, 39.7, 52.9, 69.4, 79.8, 115.7, 121.8, 125.5, 135.1, 142.3, 151.6, 155.9, 165.0; υ_{max}/cm^{-1} (solid state) = 3375, 2973, 1724, 1703, 1605, 1518, 1494, 1454, 1441, 1392, 1359, 1304, 1280, 1246, 1171, 1117, 1089, 1072; ESI-HRMS found *m/z* 363.1160 $[M+Na]^+$, $C_{15}H_{20}N_2NaO_5$ requires 363.1168.

Methyl 4-amino-3-(2-((tert-butoxycarbonyl)amino)ethoxy)benzoate 17q

Procedure D; Methyl 3-(2-((tert-butoxycarbonyl)amino)ethoxy)-4-nitrobenzoate **16q** (2.00 g, 5.9 mmol) in 1:1 ethyl acetate-methanol (40 mL), 10% palladium on charcoal (200 mg, 10 wt%) and hydrogen gas. Work up yielded the title compound (1.6 g, 5.4 mmol, 91%) as a beige solid; δ_H (500 MHz, $CDCl_3$) 1.49 (9H, s, C(C<u>H</u>$_3$)), 3.59 (2H, m, Hβ), 3.86 (3H, s, CO_2<u>Me</u>), 4.12 (2H, t, *J* = 5.1, Hα), 4.28 (2H, br. s, N<u>H</u>$_2$), 4.91 (1H, br. s, N<u>H</u>), 6.68 (1H, d, *J* = 8.2, H5), 7.43 (1H, s, H2), 7.56 (1H, d, *J* = 8.2, H6); δ_C (75MHz, $CDCl_3$); 28.4, 40.1, 51.7, 67.97, 79.60, 113.2, 113.3, 119.4, 124.4, 141.2, 144.9, 156.0, 167.2; υ_{max}/cm^{-1} (solid state) = 3439, 3387, 3063, 2948, 1712, 1696, 1670, 1601, 1523, 1482, 1422, 1352, 1267, 1251, 1207, 1122, 1017; ESI-HRMS found *m/z* 333.1421 $[M+Na]^+$, $C_{15}H_{22}N_2NaO_5$ requires 333.1426.

4-Amino-3-(2-((tert-butoxycarbonyl)amino)ethoxy)benzoic acid 18q

Procedure G; Methyl 4-amino-3-(2-((tert-butoxycarbonyl)amino)ethoxy)benzoate **17q** (1.00 g, 3.2 mmol) in a 1:1 mixture of methanol- tetrahydrofuran (20 mL), lithium hydroxide (270 mg, 6.5 mmol) in water (5 mL). Work up yielded the title compound (0.92 g, 3.1 mmol, 96%) as a colourless amorphous solid; δ_H (500 MHz, MeOD-d_4) 1.47 (9H, s, C(C<u>H</u>$_3$)), 3.51 (2H, t, *J* = 5.1, Hβ), 4.05 (2H, t, *J* = 5.1, Hα), 6.71 (1H, d, *J* = 8.2, H5), 7.42 (1H, s, H2), 7.51 (1H, d, *J* = 8.2, H6); δ_C (125MHz, MeOD-d_4); 28.7, 41.0, 68.9, 80.2, 113.3, 113.9, 119.3, 125.9, 144.4, 146.4, 158.7, 170.6; υ_{max}/cm^{-1} (solid state) = 3347, 2972, 1675, 16171584, 1518, 1444, 1402, 1368, 1293, 1271, 1222, 1161, 1123, 1057; ESI-HRMS found *m/z* 297.1445 $[M+H]^+$, $C_{14}H_{21}N_2O_5$ requires 297.1450.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3(2((tertbutoxycarbonyl)amino)ethoxy)benzoic acid 14q

Procedure I; 4-Amino-3-(2-((tert-butoxycarbonyl)amino)ethoxy)benzoic acid **18q** (1.40 g, 3.1 mmol), sodium bicarbonate (1.19 g, 14.2 mmol) in tetrahydrofuran (30 mL) and fluorenylmethyloxycarbonyl chloride (1.83 g, 7.1 mmol) in tetrahydrofuran (20 mL). The reaction mixture was concentrated and

column chromatography yielded the title compound (1.83 g, 3.5 mmol, 75%) as a colourless amorphous solid; (Found C, 66.95; H, 5.80; N, 5.25%. $C_{29}H_{30}N_2O_7$ requires C, 67.17; H, 5.83; N, 5.40%); δ_H (500 MHz, DMSO-d_6) 1.39 (9H, s, C(C<u>H</u>$_3$)), 3.42 (2H, m, Hβ), 4.03 (2H, m, Hα), 4.36 (1H, t, *J* = 7.0, FHβ), 4.50 (2H, d, *J* = 7.0, FHα), 7.31-7.37 (3H, m, H2 + FHAr4), 7.42-7.45 (3H, m, H6 + FHAr3), 7.50 (1H, d, *J* = 8.4, H5), 7.78 (2H, d, *J* = 7.5, FHAr5), 7.83 (1H, br. s, N<u>H</u>), 7.93 (2H, d, *J* = 7.5, FHAr2), 9.01 (1H, s, FN<u>H</u>); δ_C (125MHz, DMSO-d_6) 27.95, 46.2, 66.2, 68.2, 77.7, 111.3, 118.6, 119.9, 122.1, 124.8, 125.1, 126.9, 127.5, 131.3, 140.5, 143.4, 147.2, 153.1, 155.6, 166.6; υ_{max}/cm^{-1} (solid state) = 3423, 3361, 2986, 2947, 1741, 1682, 1605, 1532, 1489, 1440, 1347, 1297, 1248, 1227, 1198, 1132, 1053; ESI-HRMS found *m/z* 541.1945 $[M+Na]^+$, $C_{29}H_{30}N_2NaO_7$ requires 541.1951.

3-((5-((Tert-butoxycarbonyl)amino)pentyl)oxy)-4-nitrobenzoic acid 17r

Methyl 3-((5-((tert-butoxycarbonyl)amino)pentyl)oxy)-4-nitrobenzoate **16r** was obtained from methyl-3-hydroxy-4-nitrobenzoate **15** by *procedure B* without further purification / isolation. *Procedure F*; methyl 3-((5-((tert-butoxycarbonyl)amino)pentyl)oxy)-4-nitrobenzoate **16r** (5.80 g, 15.2 mmol) in a 1:1 mixture of methanol-tetrahydrofuran (150 mL) and 10% aqueous sodium hydroxide (30 mL). Work up yielded the title compound (4.95 g, 13.4 mmol, 89% over two steps) as a cream amorphous solid; δ_H (500 MHz, $CDCl_3$) 1.45-1.56 (13H, m, Hγ, Hδ + C(C<u>H</u>$_3$)), 1.87 (2H, m, Hβ), 3.16 (2H, m, Hε), 4.16 (2H, t, *J* = 6.1, Hα), 7.72 (1H, d, *J* = 7.7, H6), 7.75 (1H, s, H2), 7.81 (1H, d, *J* = 7.7, H5); δ_C (125MHz, $CDCl_3$) 23.1, 28.4, 29.6, 40.4, 41.5, 69.7, 79.5, 115.8, 121.8, 125.2, 134.4, 142.9, 156.3, 158.3, 168.4; υ_{max}/cm^{-1} (solid state) = 3377, 2980, 2944, 1693, 1521, 1308, 1249, 1177; ESI-HRMS found *m/z* 391.1495 $[M+Na]^+$ $C_{17}H_{24}N_2NaO_7$ requires 391.1481.

4-Amino-3-((5-((tert-butoxycarbonyl)amino)pentyl)oxy)benzoic acid 18r

Procedure D; 3-((5-((tert-butoxycarbonyl)amino)pentyl)oxy)-4-nitrobenzoic acid **17r** (4.90 g, 13.4 mmol) a 1:2 mixture of ethyl acetate-methanol (90 mL), 10% palladium on charcoal (490 mg, 10 wt%) and hydrogen gas. Work up yielded the title compound (4.50 g, 13.3 mmol, 91%) as a beige amorphous solid; (Found C, 59.00; H, 7.70; N, 8.60%. $C_{17}H_{25}N_2O_7$ requires C, 60.34; H, 7.74; N, 8.28%); δ_H (500 MHz, MeOD-d_4) 1.43 (9H, s, C(C<u>H</u>$_3$)), 1.53 (4H, m, Hγ + Hδ), 1.85 (2H, m, Hβ), 3.07 (2H, t, *J* = 6.2, Hε), 4.04 (2H, t, *J* = 6.3, Hα), 6.69 (1H, d, *J* = 8.2, H5), 7.41 (1H, s, H2), 7.48 (1H, d, *J* = 8.2, H6); δ_C (125MHz, MeOD-d_4) 28.5, 28.8, 29.2, 30.1, 41.3, 69.3, 79.9, 113.3, 113.4, 119.6, 125.5, 125.6, 144.2, 146.7, 151.0, 158.6, 170.8; υ_{max}/cm^{-1} (solid state) = 3492, 3347, 2940, 1703, 1690, 1657, 1620, 1588, 1576, 1518, 1417, 1367, 1308, 1268, 1237, 1169, 1153, 1029; ESI-HRMS found *m/z* 337.1769 $[M-H]^-$ $C_{17}H_{25}N_2O_7$ requires 337.1763.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-((5-((tert-butoxycarbonyl) amino) pentyl)oxy)benzoic acid 14r

Procedure I; 4-amino-3-((5-((tert-butoxycarbonyl)amino)pentyl)oxy)benzoic acid **18r** (4.00 g, 11.8 mmol), sodium bicarbonate (2.98 g, 34.2 mmol) in tetrahydrofuran (80 mL) and fluorenylmethyloxycarbonyl chloride (4.58 g, 17.8 mmol) in tetrahydrofuran (40 mL). Precipitation of the product *via* hexane yielded the title compound (5.50 g, 9.8 mmol, 79%) as a colourless amorphous solid; δ_H (500 MHz, $CDCl_3$) 1.44 (9H, s, C(C$\underline{H}_3$)), 1.54 (2H, m, Hγ), 1.60 (2H, m, Hδ), 1.91 (2H, m, Hβ), 3.17 (2H, m, Hε), 4.13 (2H, t, *J* = 6.6, Hα), 3.34 (1H, t, *J* = 6.9, FHβ), 4.55 (3H, m, N$\underline{H}$ + FHα), 7.34 (2H, t, *J* = 7.4, FHAr4), 7.43 (2H, t, *J* = 7.5, FHAr3), 7.49 (1H, br. s, H6), 7.57 (1H, d, *J* = 1.5, H2), 7.64 (2H, d, *J* = 7.5, FHAr5), 7.75 (1H, d, *J* = 8.3, H5), 7.80 (2H, d, *J* = 7.5, FHAr2), 8.14 (1H, br. s, FN$\underline{H}$); δ_C (125MHz, $CDCl_3$) 14.2, 21.1, 23.3, 28.4, 28.7, 29.9, 47.1, 60.4, 67.4, 68.8, 94.9, 112.0, 117.3, 120.1, 124.2, 125.0, 127.2, 127.9, 141.4, 143.7, 146.4, 149.3, 153.0, 156.0, 171.2; υ_{max}/cm^{-1} (solid state) =3334, 2937, 1706, 1672, 1595, 1531, 1496, 1431, 1281, 1243, 1214, 1173, 1085, 1045; ESI-HRMS found *m/z* 559.2460 $[M-H]^-$ $C_{32}H_{35}N_2O_7$ requires 559.2450.

4-((((9H-Fluoren-9-yl)methoxy)carbonyl)amino)-3-methylbenzoic acid 14s

Procedure H; 4-Amino-3-methylbenzoic acid **19** (3.00 g, 19.9 mmol) was dissolved in tetrahydrofuran (100 mL), fluorenylmethyloxycarbonyl chloride (7.70 g, 29.8 mmol) in tetrahydrofuran (50 mL). Work up yielded the title compound (7.12 g, 19.1 mmol, 96%) as a colourless amorphous solid; (Found C, 73.90; H, 5.10; N, 3.65%. $C_{23}H_{19}NO_4$ requires C, 73.98; H, 5.13; N, 3.75%); δ_H (300 MHz, DMSO-d_6) 2.26 (3H, s, ArC$\underline{H}_3$), 4.32 (1H, t, *J* = 7.5, FHβ), 4.48 (2H, d, *J* = 7.5, FHα), 7.34 (2H, t, *J* = 7.5, FHAr4), 7.41-7.47 (3H, m, H2 + FHAr3) 7.64-7.76 (4H, m, H6, H5 + FHAr5), 7.90 (2H, d, *J* = 7.5, FHAr2), 9.21 (2H, s, N$\underline{H}$); δ_C (75MHz, DMSO-d_6) 17.8, 46.6, 65.9, 120.2, 123.3, 125.2, 126.4, 127.1, 127.4, 127.7, 130.6, 131.5, 140.6, 140.8, 143.7, 153.9, 167.0; υ_{max}/cm^{-1} (solid state) = 3271, 2831, 1701, 1685, 1528, 1253, 736; ESI-HRMS found *m/z* 374.1390 $[M+H]^+$, $C_{23}H_{20}NO_4$ requires 374.1387.

6.3 Characterisation of Oligomers (Chapter 2 and Chapter 3)

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 20

δ_H (500 MHz, DMSO-d_6) 1.32 (6H, d, *J* = 6.0, 1-Hβ), 1.35 (6H, d, *J* = 6.0, 3-Hβ), 1.39 (6H, d, *J* = 6.0, 2-Hβ), 3.93 (2H, d, *J* = 4.5, 4-Hα), 4.60 (1H, sept, *J* = 6.0, 1-Hα), 4.72 (1H, sept, *J* = 6.0, 1.9, 3-Hα), 4.79 (1H, sept, *J* = 6.0, 1.5, 2-Hα), 5.43 (2H, br. s, 1-NH_2), 6.72 (1H, d, *J* = 8.1, 1-H5), 7.32-7.35 (2H, m, 1-H2, 1-H6), 7.51-7.61 (4H, m, 2-H2, 2-H6, 3-H2, 3-H6), 8.12 (1H, app t, *J* = 8.1, 3-H5), 8.29 (1H, app t, *J* = 7.5, 2-H5), 8.81 (1H, t, *J* = 5.8, 4-N<u>H</u>), 8.94 (1H, s, 3-N<u>H</u>), 9.27 (1H, s, 2-N<u>H</u>); δ_C (125MHz, DMSO-d_6) 21.79, 21.82, 21.9, 41.2, 70.36, 71.39, 71.41 112.2, 122.6, 112.8, 119.9, 120.0, 120.3, 120.9, 121.4, 121.6, 129.2, 129.9, 131.1, 132.2, 143.24, 146.9, 147.0, 147.8, 147.9, 164.2, 164.4, 165.8, 171.3; ESI-HRMS found *m/z* 629.2558 $[M+Na]^+$, $C_{32}H_{38}N_4O_8Na$ requires 629.2582.

NH_2-[*O*-Bn-(3-HABA)]-[*O*-Bn-(3-HABA)]-[*O*-Bn-(3-HABA)]-Gly-CO_2H 21

δ_H (300 MHz, DMSO-d_6) 3.95 (2H, d, *J* = 5.6, 4-Hα), 5.14 (2H, s, 1-Hα), 5.26 (2H, s, 2-Hα), 5.29 (2H, s, 3-Hα), 6.78 (1H, d, *J* = 8.2, 1-H5) 7.26-7.64 (19H, m, ArC<u>H</u>), 7.73 (2H, m, 3-HAr2), 8.00 (1H, d, *J* = 8.3, 2-H5), 8.15 (1, d, *J* = 8.3, 3-H5), 8.88 (1H, t, *J* = 5.6, 4-N<u>H</u>), 9.20 (1H, s, 2-N<u>H</u>), 9.57 (1H, s, 3-N<u>H</u>); δ_C (125MHz, DMSO-d_6) 41.2, 69.5, 70.2, 70.3, 111.1, 111.6, 111.8, 113.2, 120.0, 120.4, 121.2, 121.6, 122.7, 127.4, 127.5, 127.6, 127.8, 127.9, 128.0, 128.4, 129.7, 130.3, 130.5, 131.4, 136.6, 136.8, 137.0, 144.7, 148.7, 149.6, 164.3, 164.5, 165.8, 171.4; ESI-HRMS found *m/z* 751.2785 $[M+H]^+$, $C_{44}H_{39}N_4O_8$ requires 751.2762.

NH_2-[*O*-Bn-(3-HABA)]-[*O*-2-CH_2-Nap-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 22

δ_H (300 MHz,DMSO-d_6) 1.35 (6H, d, *J* = 6.0, 3-Hβ), 3.95 (2H, d, *J* = 5.8, 4-Hα), 4.73 (1H, sept, *J* = 6.0, 3-Hα), 5.11 (2H, s, 1-Hα), 5.52 (2H, s, 2-Hα), 6.76 (1H, d, *J* = 8.1, 1-H5), 7.31-7.41 (3H, m, ArC<u>H</u>), 7.43-7.46 (3H, m, 1-H6 + ArC<u>H</u>), 7.50-7.56 (4H, m, 3-H6 + 1-H2 + ArC<u>H</u>), 7.60 (1H, d, *J* = 1.6, 3-H2), 7.63 (1H, dd, *J* = 8.3, 1.6, 2-H6), 7.73 (1H, dd, *J* = 8.3, 1.6, ArC<u>H</u>), 7.79 (1H, d, *J* = 1.8, 2-H2), 7.88 (2H, m, ArCH), 7.94 (1H, d, *J* = 8.3, 2-HAr2), 8.10 (2H, m, 3-H5 + 2-HAr8), 8.21 (1H, d, *J* = 8.3, 2-H5), 8.84 (1H, t, *J* = 6.0, 4-N<u>H</u>), 9.27 (1H, s, 2-N<u>H</u>), 9.34 (1H, s, 3-N<u>H</u>); δ_C (125 MHz, DMSO-d_6) 21.8, 41.2, 69.5, 70.4, 71.5, 111.2, 111.8, 112.6, 113.1, 119.9, 120.2, 121.5, 121.6, 121.8, 125.3, 125.9, 126.2, 126.4, 127.4, 127.6, 127.7, 128.1, 128.4, 129.9, 130.1, 131.1, 131.5, 132.6, 132.8, 134.4, 136.9, 144.6, 148.1, 148.9, 164.2, 164.6, 165.9, 171.4; ESI-HRMS found *m/z* 751.2774 $[M-H]^-$, $C_{44}H_{39}N_4O_8$ requires 751.2773.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-2-CH_2-Nap-(3-HABA)]-[*O*-Bn-(3-HABA)]-Gly-CO_2H 23

δ_H (500 MHz, DMSO-d_6) 1.13 (6H, d, *J* = 6.0, 1-Hβ), 3.94 (2H, d, *J* = 5.8, 4-Hα), 4.40 (1H, sept, *J* = 6.0, 1-Hα), 5.27 (2H, s, 3-Hα), 5.39 (2H, s, 2-Hα), 6.71 (1H, d, *J* = 8.1, 1-H5), 7.27 (1H, t, *J* = 7.3, 3-HAr4), 7.31 (1H, d, *J* = 1.7, 1-H2), 7.34-7.38 (3H, m, 1-H6 + 3-HAr3), 7.52-7.57 (5H, m, 3-HAr2 + ArC<u>H</u>), 7.61 (1H, dd, *J* = 8.4, 1.6, 2-H2), 7.68-7.70 (2H, m, 2-H2 + 2-H6), 7.78 (1H, d, *J* = 1.5, 3-

H2), 7.87 (1H, m, ArCH), 7.92-7.95 (2H, m, ArCH), 8.01 (1H, d, *J* = 8.1, 3-H5), 8.06 (1H, s, HAr2), 8.18 (1H, d, *J* = 8.3, 2-H5), 8.83 (1H, t, *J* = 5.8, 4-NH), 9.18 (1H, s, 2-NH), 9.54 (1H, s, 3-NH); δ_C (125 MHz, DMSO-d_6) 21.6, 41.2, 70.2, 70.4, 70.4, 111.5, 111.7, 112.5, 113.5, 120.0, 120.4, 121.4, 121.6, 121.7, 122.8, 125.5, 126.3, 126.3, 126.4, 127.4, 127.4, 127.6, 127.7, 127.9, 128.1, 128.4, 128.4, 129.8, 130.3, 130.5, 131.4, 132.6, 132.8, 134.1, 136.7, 141.9, 143.7, 148.9, 149.7, 164.4, 164.6, 165.9, 171.3; ESI-HRMS found *m/z* 751.2803 $[M-H]^-$, $C_{44}H_{39}N_4O_8$ requires 751.2773.

NH_2-[*O*-Bn-(3-HABA)]-[*O*-*p*Cl-Bn-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 24

δ_H (500 MHz, DMSO-d_6) 1.37 (6H, d, *J* = 6.0, 3-Hβ), 3.95 (2H, d, *J* = 5.8, 4-Hα), 4.74 (1H, sept, *J* = 6.0, 3-Hα), 5.16 (2H, s, 1-Hα), 5.35 (2H, s, 2-Hα), 6.76 (1H, d, *J* = 8.3, 1-H5), 7.34 (1H, t, *J* = 7.5, 1-HAr4), 7.39-7.46 (5H, m, 1-H6, 1-HAr3, 2-HAr3), 7.49-7.56 (4H, m, 1-H2, 1-HAr2, 3-H6), 7.60-7.63 (4H, m, 2-H6, 2-HAr2 + 3-H2), 7.71 (1H, d, *J* = 1.7, 2-H2), 8.10 (1H, d, *J* = 8.3, 2-H5), 8.19 (1H, d, *J* = 8.3, 3-H5), 8.83 (1H, t, *J* = 6.0, 4-NH), 9.19 (1H, s, 2-NH), 9.31 (1H, s, 3-NH); δ_C (125MHz, DMSO-d_6) 21.8, 41.2, 68.3, 69.5, 71.4, 111.2, 111.7, 112.6, 112.8, 119.9, 120.2, 121.3, 121.6, 121.7, 121.9, 127.5, 127.8, 128.4, 128.4, 129.3, 129.8, 130.1, 131.1, 131.5, 132.6, 135.8, 137.0, 144.5, 148.1, 148.7, 164.2, 164.6, 165.8, 171.4; ESI-HRMS found *m/z* 735.2243 $[M-H]^-$, $C_{40}H_{36}ClNO_6$ requires 735.2227.

NH_2-[*O*-*p*Cl-Bn-(3-HABA)]-[*O*-*m*CF_3-Bn-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 25

δ_H (500MHz, DMSO-d_6) 1.34 (6H, d, *J* = 6.2, 3-Hβ), 3.93 (2H, d, *J* = 5.8, 4-Hα), 4.71 (1H, sept, *J* = 6, 3-Hα), 5.13 (2H, s, 1-Hα), 5.41 (2H, s, 2-Hα), 6.71 (1H, d, *J* = 8.1, 1-H5), 7.39 (1H, dd, *J* = 8.0, 1.5, 1-H6), 7.44 (2H, d, *J* = 8.5, 1-HAr2), 7.48 (1H, d, *J* = 1.5, 1-H2), 7.51-7.54 (3H, m, 1-HAr3, 3-H6), 7.58 (1H, d, *J* = 1.5, 3-H2), 7.59-7.63 (2H, m, 2-H6, 2-HAr5), 7.67 (1H, d, *J* = 7.0, 2-HAr6), 7.72 (1H, d, *J* = 1.5, 2-H2), 7.85 (1H, d, *J* = 7.0, 2-HAr4), 8.00 (1H, s, HAr2), 8.07 (1H, d, *J* = 8.3, 3-H5), 8.14 (1H, d, *J* = 8.3, 2-H5), 8.81 (1H, t, *J* = 5.8, 4-NH), 9.27 (1H, s, 2-NH), 9.31 (1H, s, 3-NH); δ_C (125MHz, DMSO-d_6) 22.3, 41.8, 69.2, 69.8, 72.0, 112.0, 112.2, 113.2, 113.3, 120.7, 121.4, 121.7, 122.1, 122.4, 122.6, 124.4, 124.7 (J = 272.4), 125.1, 128.9, 129.8, 129.8 (*J* = 32), 130.0, 130.6, 130.6, 131.6, 131.8, 132.0, 132.8, 136.6, 138.8, 142.3, 144.9, 148.6, 149.5, 164.7, 165.1, 166.4, 171.9; ESI-HRMS found *m/z* 803.2133$[M-H]^-$, $C_{41}H_{35}ClF_3N_4O_8$ requires 803.2101.

NH_2-[*O*-Bn-(3-HABA)]-[*O*-*p*CF_3-Bn-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 26

δ_H (500MHz, DMSO-d_6) 1.33 (6H, d, *J* = 6, 3-Hβ), 3.93 (2H, d, *J* = 4.9, 4-Hα), 4.70 (1H, sept, *J* = 6.2, 3-Hα), 5.14 (2H, s, 1-Hα), 5.45 (2H, s, 2-Hα), 5.51 (2H, br. s, 1-NH_2), 6.71 (1H, d, *J* = 8.1, 1-H5), 7.32 (1H, t, *J* = 7, 1-HAr4), 7.36-7.41 (4H, m, 1-H2, 1-H6, 1-HAr2), 7.47-7.53 (3H, m, 1-HAr3, 3-H6), 7.58 (1H, s, 3-H2), 7.61 (1H, dd, *J* = 8.4, 1.6, 2-H6), 7.69 (1H, s, 2-H2), 7.73 (2H, d, *J* =8.3, 2-HAr3), 7.79 (2H, d, *J* = 8.1, 2-HAr2), 8.07 (1H, d, *J* = 8.1, 3-H5), 8.17 (1H, d, *J* = 8.3, 2-H5), 8.81 (1H, t, *J* = 4.9, 4-NH), 9.22 (1H, s, 2-NH), 9.29 (1H, s, 3-NH); δ_C (125MHz, DMSO-d_6) 21.8, 41.2, 69.2, 69.5, 71.4, 111.2, 111.6, 112.5, 112.6, 119.9, 120.2, 120.8, 121.9, 121.9, 124.2 (*J* = 272.3)

125.3, 127.4, 127.7, 128.4, 128.4 (*J* = 31.7), 129.9, 130.1, 131.1, 131.5, 137.0, 141.7, 142.4, 144.3, 148.1, 148.7, 164.2, 164.7, 165.8, 171.4; ESI-HRMS found *m/z* 771.2633 $[M+H]^+$, $C_{41}H_{38}F_3N_4O_8$ requires 771.2636.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Leu-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 27

δ_H (500 MHz, DMSO-d_6) 0.97 (3H, t, *J* = 7.5, 2-Hγ), 1.31-1.36 (15H, m, 1-Hβ, 2-$CH_\alpha(CH_3)$ + 3-Hα), 1.67-1.81 (2H, m, 2-Hβ + 2-Hβ'), 3.93 (2H, d, *J* = 5.6, 4-Hα), 4.60 (1H, m, 2-Hα), 4.72 (2H, sept, *J* = 6.0, 1-Hα, 3-Hα), 6.74 (1H, d, *J* = 8.1, 1-H5), 7.32-7.35 (2H, m, 1-H2 + 1-H6), 7.52-7.59 (4H, m, 2-H2, 2-H6, 3-H2 + 3-H6), 8.11 (1H, d, *J* = 8.3, 2-H5), 8.29 (1H, d, *J* = 8.3, 3-H5), 8.80 (1H, t, *J* = 5.7, 4-N<u>H</u>), 8.95 (1H, s, 2-N<u>H</u>), 9.27 (1H, s, 3-N<u>H</u>); δ_C (125MHz, DMSO-d_6) 9.4, 18.9, 21.86, 21.93, 28.6, 41.2, 70.4, 71.4, 76.1, 112.1, 112.6, 113.2, 119.9, 120.0, 120.3, 121.3, 121.6, 129.4, 129.9, 131.2, 132.2, 143.5, 147.2, 147.9, 164.2, 164.4, 165.8, 171.4; ESI-HRMS found *m/z* 643.2738 $[M+Na]^+$, $C_{33}H_{40}N_4NaO_8$ requires 643.2744.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-CH_2-OH-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 28

δ_H (500 MHz, DMSO-d_6) 1.31 (6H, d, *J* = 6.0, 1-Hβ) , 1.35 (6H, d, *J* = 6.0, 3-Hβ), 3.79 (2H, t, *J* = 4.6, 2-Hβ), 3.93 (2H, d, *J* = 5.8, 4-Hα), 4.21 (2H, t, *J* = 4.6, 2-Hα), 4.61 (1H, sept, *J* = 6.0, 1-Hα), 4.71 (1H, sept, *J* = 6.0, 3-Hα), 6.71 (1H, d, *J* = 8.8, 1-H5), 7.37-7.39 (2H, m, 1-H2 + 1-H6), 7.52 (1H, dd, *J* = 8.3, 1.7, 2-H6), 7.58-7.60 (2H, m, 2-H2 + 3-H6), 7.63 (1H, d, *J* = 1.9, 3-H2), 8.08 (1H, d, *J* = 8.3, 2-H5), 8.33 (1H, d, *J* = 8.3, 3-H5), 8.81 (1H, t, *J* = 5.8, 4-N<u>H</u>), 9.20 (1H, s, 2-N<u>H</u>), 9.28 (1H, s, 3-N<u>H</u>); δ_C (125MHz, DMSO-d_6) 21.8, 21.9, 41.2, 59.5, 70.4, 71.5, 71.5, 112.1, 112.6, 112.8, 112.9, 119.9, 120.0, 120.5, 121.0, 121.6, 121.7, 129.2, 130.0, 131.2, 143.2, 143.3, 148,0, 164.2, 164.8, 165.8, 171.4; ESI-HRMS found *m/z* 631.2378 $[M+Na]^+$, $C_{31}H_{36}N_4NaO_9$ requires 631.2374.

NH_2-[*O*-*i*Pr-(3-HABA)]-[3-Me-ABA]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 29

δ_H (500 MHz, DMSO-d_6) 1.3 (6H, d, *J* = 6.0, 1-Hβ), 1.36 (6H, d, *J* = 6.0, 3-Hβ), 2.32 (3H, s, 2-Hα), 3.93 (2H, d, *J* = 5.8, 4-Hα), 4.58 (1H, sept, *J* = 6.0, 1-Hα), 4.72 (1H, sept, *J* = 6.0, 3-Hα), 6.70 (1H, d, *J* = 8.3, 1-H5), 7.44-7.45 (2H, m, 1-H2 + 1-H6), 7.52 (1H, d, *J* = 8, 3-H6), 7.58-7.59 (2H, m, 2-H2 + 2-H6), 7.76 (1H, d, *J* = 7.5, 2-H5), 7.82 (1H, s, 3-H2), 8.10 (1H, d, *J* = 8, 3-H5), 8.81 (1H, t, 5.8, 4-N<u>H</u>), 9.26 (1H, s, 3-N<u>H</u>), 9.52 (1H, s, 2-N<u>H</u>); δ_C (125MHz, DMSO-d_6) 18.21, 21.8, 21.9, 22.0, 41.2, 70.5, 71.4, 112.6, 113.5, 119.9, 121.1, 121.6, 122.0, 125.0, 125.6, 129.5, 130.0, 130.7, 131.2, 132.9, 140.6, 143.0, 143.2, 148.0, 164.5, 165.1, 165.8, 171.4; ESI-HRMS found *m/z* 561.2358 $[M-H]^-$, $C_{30}H_{33}N_4O_7$ requires 561.2355.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-*p*-^{t}Bu-Bn-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 30

δ_H (500 MHz, DMSO-d_6) 1.30 (9H, s, 2-Ar-C(C<u>H</u>$_3$)), 1.38 (6H, m, 1-Hβ, 3-Hβ), 3.95 (2H, d, *J* = 6.0, 4-Hα), 4.59 (1H, sept, *J* = 6.0, 1-Hα), 4.74 (1H, sept, *J* = 6.0, 3-Hα), 5.30 (2H, s, 2-Hα), 6.73 (1H, d, *J* = 8.0, 1-H5), 7.36 (1H, d, *J* = 8, 1-H6), 7.40-7.44 (3H, m, 1-H2 + 2-HAr3), 7.51 (2H, d, *J* = 8.0, 2-HAr2), 7.55 (1H, d, *J* = 8.0, 2-H6), 7.61-7.63 (2H, m, 2-H2 + 3-H6), 7.76 (1H, s, 3-H2), 8.11 (1H, d,

J = 8.2, 2-H5), 8.21 (1H, d, *J* = 8.2, 3-H5), 8.85 (1H, t, *J* = 6.0, 4-NH), 9.14 (1H, s, 2-NH), 9.35 (1H, s, 3-NH); δ_C (125MHz, DMSO-d_6) 27.1, 27.2, 36.3, 39.5, 46.5, 75.3, 75.7, 76.7, 116.8, 117.9, 118.1, 118.3, 125.1, 125.4, 126.6, 126.6, 126.8, 127.1, 130.4, 132.6, 135.0, 135.3, 136.4, 136.8, 138.9, 148.7, 153.3, 154.0, 155.7, 169.5, 169.8, 171.1, 176.6; ESI-HRMS found *m/z* 733.3189 $[M+Na]^+$, $C_{40}H_{46}N_4NaO_8$ requires 733.3208.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-CH_2-indole-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 32

LC-MS analysis of this reaction indicated reasonable coupling of the central indole monomer to give the target trimer, however it was not possible to purify and isolate this oligomer. ESI-MS found *m/z* 708 $[M+H]^+$.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-COOH-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 33

δ_H (500 MHz, DMSO-d_6) 1.31 (6H, d, *J* = 6.0, 1-Hβ), 1.35 (6H, d, *J* = 6.0, 3-Hβ), 3.93 (2H, d, *J* = 5.8, 4-Hα), 4.61 (1H, sept, *J* = 6.0, 1-Hα), 4.71 (1H, sept, *J* = 6.0, 3-Hα), 4.91 (2H, s, 2-Hα), 5.43 (2H, br. s, 1-NH_2), 6.70 (1H, d, *J* = 8.5, 1-H5), 7.42-7.43 (2H, m, 1-H2 + 1-H6), 7.52 (1H, d, *J* = 8.0, 2-H6), 7.58 (1H, s, 2-H2), 7.61-7.66 (2H, m, 3-H2 + 3-H6), 8.06 (1H, d, *J* = 8.5, 2-H5), 8.32 (1H, d, *J* = 8.1, 3-H5), 8.80 (1H, t, *J* = 5.8, 4-NH), 9.28 (1H, s, 3-NH), 9.43 (1H, s, 2-NH); δ_C (125MHz, DMSO-d_6) 21.8, 21.9, 67.2, 70.4, 71.5, 112.6, 112.7, 112.8, 113.5, 119.8, 120.1, 120.7, 121.4, 121.4, 122.0, 129.1, 130.1, 131.1, 132.5, 143.3, 143.5, 147.8, 148.1, 164.0, 164.5, 165.8, 170.7, 171.4; ESI-HRMS found *m/z* 645.2087 $[M+Na]^+$, $C_{31}H_{34}N_4NaO_{10}$ requires 645.2167.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-CH_2-CH_2-S-CH_2-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 34

δ_H (500 MHz, DMSO-d_6) 1.31 (6H, d, *J* = 6.0, 1-Hβ), 1.35 (6H, d, *J* = 6.0, 3-Hβ), 2.05 (3H, s, 2-SCH_3), 2.10 (2H, m, 2-Hβ), 2.69 (2H, t, *J* = 7.1, 2-Hγ), 3.93 (2H, d, *J* = 5.8, 4-Hα), 4.24 (2H, t, *J* =5.5, 2-Hα), 4.60 (1H, sept, *J* = 6.0 , 1-Hα), 4.72 (1-H, sept, *J* = 6.0, 3-Hα), 6.71 (1H, d, *J* = 8.0, 1-H5), 7.35-7.42 (2H, m, 1-H2, 1-H6), 7.52 (1H, d, *J* = 8.0, 2-H6), 7.56-7.60 (3H, m, 2-H2, 3-H2, 3-H6), 8.09 (1H, d, *J* = 8.0, 2-H5), 8.20 (1H, d, *J* = 8.0, 3-H5), 8.81 (1H, t, *J* = 5.8, 4-NH), 9.06 (1H, s, 2-NH), 9.30 (1H, s, NH); δ_C (125MHz, DMSO-d_6) 14.7. 21.8, 21.9, 28.2, 30.0, 41.2, 67.2, 70.4, 71.5, 110.7, 112.6, 112.8, 112.9, 119.9, 121.0, 121.1, 121.6, 121.7, 129.6, 130.0, 131.2, 131.3, 143.1, 143.3, 148.1, 148.7, 164.2, 164.7, 165.8, 171.4; ESI-HRMS found *m/z* 675.2433 $[M+Na]^+$, $C_{33}H_{40}N_4NaO_8S$ requires 675.2459.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-$CONH_2$-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 35

δ_H (500 MHz, DMSO-d_6) 1.31 (6H, d, *J* = 6.0, 1-Hβ), 1.36 (6H, d, *J* = 6.0, 3-Hβ), 3.93 (2H, d, *J* = 5.6, 4-Hα), 4.61-4.71 (4H, m, 1-Hα, 2-Hα + 3-Hα), 6.71 (1H, d, *J* = 8.8, 1-H5), 7.46-7.47 (2H, m, 1-H2 + 1-H6), 7.52 (1H, d, *J* = 8.5, 2-H6), 7.58-7.63 (2H, m, 2-H2 + 3-H6), 7.86 (1H, s, 3-H2), 8.06 (1H, d, *J* = 8.5, 2-H5), 8.11 (1H, d, *J* = 8.5, 3-H5), 8.81 (1H, t, *J* = 5.6, 4-NH), 9.31 (1H, s, 3-NH), 9.73 (1H, s, 2-NH); δ_C (125MHz, DMSO-d_6) 21.8, 21.9, 41.2, 66.5, 68.2, 71.5, 112.7, 119.8, 121.0, 121.1, 122.0,

129.8, 130.1, 130.2, 131.0, 131.1, 132.1, 143.3, 148.2, 148.3, 164.1, 164.9, 165.8, 170.2, 171.5; ESI-HRMS found 622.2517 *m/z* $[M+H]^+$, $C_{31}H_{36}N_5O_9$ requires 622.2508.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-CH_2-NH_2-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 36

δ_H (500 MHz, DMSO-d_6) 1.31 (6H, d, *J* = 6.0, 1-Hβ), 1.35 (6H, d, *J* = 6.0, 3-Hβ), 3.37 (2H, m, 2-Hβ), 3.93 (2H, d, *J* = 5.8, 4-Hα), 4.38 (2H, t, *J* = 4.1, 2-Hα), 4.60 (1H, sept, *J* = 6.0, 1-Hα), 4.72 (1H, sept, *J* = 6.0, 3-Hα), 6.72 (1H, d, *J* = 8.1, 1-H5), 7.39-7.42 (2H, m, 1-H2 + 1-H6), 7.53 (1H, dd, *J* = 8.3, 1.8, 2-H6), 7.59-7.63 (3H, m, 2-H2, 3-H2 + 3-H6), 8.06 (1H, d, *J* = 8.3, 2-H5), 8.29 (1H, d, *J* = 8.9, 3-H5), 8.83 (1H, t, *J* = 5.8, 4-N<u>H</u>), 9.29 (1H, s, 2-N<u>H</u>), 9.32 (1H, s, 3-N<u>H</u>); δ_C (125MHz, DMSO-d_6) 20.7, 20.8, 63.9, 69.4, 70.3, 109.6, 111.6, 112.3, 118.8, 119.2, 119.9, 120.0, 120.9, 121.1, 128.3, 129.0, 129.9, 130.2, 142.1, 142.1, 146.6, 147.1, 163.0, 164.3, 164.7, 170.25; ESI-HRMS found *m/z* 608.2701 $[M+H]^+$, $C_{31}H_{38}N_5O_8$ requires 608.2715.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-CH_2-CH_2-CH_2-CH_2-NH_2-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 37

δ_H (500 MHz, DMSO-d_6) 1.31 (6H, d, *J* = 6.0, 1-Hβ), 1.35 (6H, d, *J* = 6.0, 3-Hβ), 1.51 (2H, m, 2-Hγ), 1.60 (2H, m, 2-Hδ), 1.85 (2H, m, 2-Hβ), 2.78 (2H, m, 2-Hε), 3.93 (2H, d, *J* = 5.5, 4-Hα), 4.17 (2H, t, *J* = 6.5, 2-Hα), 4.60 (1H, sept, *J* = 6.0, 1-Hα), 4.72 (1H, sept, *J* = 6.0, 3-Hα), 5.42 (2H, br. s, 1-N<u>H</u>$_2$), 6.70 (1H, d, *J* = 8.0, 1-H5), 7.35 (1H, d, *J* = 8.0, 1-H6), 7.38 (1H, s, 1-H2), 7.53 (1H, d, *J* = 8.5, 3-H6), 7.52-7.70 (3H, m, 2-H2, 2-H6 + 3-H2), 8.09 (1H, d, *J* = 8.0, 3-H5), 8.19 (1H, d, *J* = 8.5, 2-H5), 8.81 (1H, t, *J* = 5.5, 4-N<u>H</u>), 9.01 (1H, s, 2-N<u>H</u>), 9.28 (1H, s, 3-N<u>H</u>); δ_C (125MHz, DMSO-d_6) 21.8, 21.9, 22.2, 26.5, 27.9, 38.7, 41.2, 68.0, 70.5, 71.4, 110.9, 112.6, 112.9, 113.2, 119.8, 119.9, 121.1, 121.4, 121.7, 129.7, 130.0, 131.1, 131.3, 142.6, 143.5, 148.0, 148.8, 164.2, 164.6, 165.8, 171.4; ESI-HRMS found *m/z* 648.3048 $[M-H]^-$, $C_{34}H_{42}N_5O_8$ requires 648.3039.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-[*O*-*i*Leu-(3-HABA)]-Ile-CO_2H 39

δ_H (500 MHz, DMSO-d_6) 0.88 (3H, t, *J* = 7.4, 4-Hδ), 0.93-0.97 (6H, m, 3-Hγ + 4-CH_β(C<u>H</u>$_3$)), 1.04 (6H, d, *J* = 6.6, 2-Hγ), 1.25-1.31 (10H, m, 1-Hβ + 3-CH_α(C<u>H</u>$_3$) + 4-Hγ'), 1.53 (1H, m, 4-Hγ), 1.63-1.79 (2H, m, 3-Hβ + 3-Hβ'), 1.96 (1H, m, 4-Hβ), 2.15 (1H, 2-Hβ), 3.95 (2H, d, *J* = 6.6, Hα), 4.35 (1H, t, *J* = 7.7, 4-Hα), 4.53-4.64 (2H, m, 1-Hα + 3-Hα), 6.76 (1H, d, *J* = 8.1, 1-H5), 7.36 (1H, dd, *J* = 8.1, 1.5, 1-H6), 7.39 (1H, d, *J* = 1.5, 1-H2), 7.56-7.61 (4H, m, 2-H2, 2-H6, 3-H2 + 3-H6), 8.09 (1H, d, *J* = 8.1, 3-H5), 8.22 (1H, d, *J* = 8.1, 2-H5), 8.40 (1H, d, *J* = 7.7, 4-N<u>H</u>), 9.03 (1H, s, 2-N<u>H</u>), 9.30 (1H, s, 3-N<u>H</u>); δ_C (125MHz, DMSO-d_6) 9.3, 11.0, 15.7, 18.9, 19.0, 21.9, 25.2, 27.8, 28.5, 35.7, 57.2, 70.5, 74.6, 76.0, 110.6, 112.7, 123.0, 113.5, 119.9, 120.2, 120.8, 121.4, 121.5, 121.8, 129.6, 130.3, 131.1, 131.3, 142.2, 143.7, 148.1, 148.7, 158.1, 158.3, 164.2, 164.5, 166.1, 173.2; ESI-HRMS found *m/z* 713.3508 $[M+Na]^+$, $C_{38}H_{50}N_4NaO_8$ requires 713.3521.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-*c*Pr-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-Gly-CO_2H 41

δ_H (500 MHz, DMSO-d_6) 0.44 (2H, m, 2-Hγ), 0.62 (2H, m, 2-Hγ'), 1.06 (12H, m, 3-Hγ + 4-Hγ), 1.31 (6H, d, *J* = 6.0, 1-Hβ), 1.34 (1H, m, 2-Hβ), 2.16 (2H, m, 3-Hβ + 4-Hβ), 3.91-3.93 (4H, m, 3-Hα + 4-Hα), 4.04 (2H, d, *J* = 6.8, 2-Hα), 4.63 (1H, sept, *J* = 6.2, 1-Hα), 6.74 (1H, d, *J* = 8.3, 1-H5), 7.37-7.42 (2H, m, 1-H2 + 1-H6), 7.54-7.68 (6H, m, 2-H2, 2-H6, 3-H2, 3-H6, 4-H2 + 4-H6), 8.05 (1H, d, *J* = 8.3, 4-H5), 8.11 (1H, d, *J* = 8.1, 3-H5), 8.26 (1H, d, *J* = 8.1, 2-H5), 8.86 (1H, t, *J* = 5.6, 5-NH), 9.05 (1H, s, 2-NH), 9.44 (1H, s, 4-NH), 9.46 (1H, s, 3-NH); δ_C (125MHz, DMSO-d_6) 2.9, 10.0, 19.06, 19.10, 21.9, 25.1, 27.8, 27.9, 41.2, 67.0, 70.4, 73.2, 74.5, 74.6, 110.8, 110.9, 111.1, 112.7, 112.8, 119.8, 119.9, 120.2, 120.4, 120.8, 121.4, 122.3, 129.3, 130.1, 130.4, 130.6, 130.7, 131.5, 143.2, 143.4, 148.4, 149.7, 149.8, 164.2, 164.3, 164.5, 171.4; ESI-HRMS found *m/z* 846.3699 $[M+H]^+$, $C_{45}H_{53}N_5NaO_{10}$ requires 846.3685.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-CH_2-*c*Pr-(3-HABA)]-[*O*-CH_2-*c*Pr-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-Gly-CO_2H 42

δ_H (500 MHz, DMSO-d_6) 0.42 (4H, m, 2-Hγ + 3-Hγ), 0.58 (4H, m, 2-Hγ' + 3-Hγ'), 1.31 (6H, d, *J* = 6.0, 1-Hβ), 1.35 (6H, d, *J* = 6.0, 4-Hβ), 1.35 (2H, m, 2-Hβ + 3-Hβ), 3.93 (2H, d, *J* = 4.9, 5-Hα), 4.61 (1H, sept, *J* = 6.0, 1-Hα), 4.72 (1H, sept, *J* = 6.0, 4-Hα), 6.72 (1H, d, *J* = 8.1, 1-H5), 7.36 (1H, dd, *J* = 8.1, 1.8, 1-H6), 7.38 (1H, d, *J* = 1.8, 1-H2), 7.53 (1H, d, *J* = 8.3, H5), 7.53-7.61 (5H, m, 2-H2, 3-H2 + H5), 8.08 (1H, d, *J* = 8.5, 4-H5), 8.12 (1H, d, *J* = 8.3, 3-H5), 8.26 (1H, d, *J* = 8.3, 2-H5), 8.81 (1H, t, *J* = 4.9, 5-NH), 9.02 (1H, s, 2-NH), 9.32 (1H, s, 4-NH), 9.39 (1H, s, 3-NH); δ_C (125MHz, DMSO-d_6) 2.87, 2.91, 10.0, 21.8, 21.9, 41.2, 70.5, 71.4, 73.0, 73.2, 111.3, 111.6, 112.6, 112.7, 112.8, 119.9, 120.2, 120.3, 120.9, 121.4, 121.8, 122.0, 129.3, 130.1, 130.5, 130.0, 131.1. 131.6, 143.3, 148.1, 148.4, 149.5, 164.2, 164.3, 164.5, 165.8, 171.4; ESI-HRMS found *m/z* 830.3364 $[M+Na]^+$, $C_{44}H_{49}N_5NaO_{10}$ requires 830.3372.

NH_2-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 43

δ_H (500 MHz, DMSO-d_6) 1.34-1.42 (36H, m, 1-6Hβ), 3.95 (2H, d, *J* = 5.5, 4-Hα), 4.59-4.76 (6H, m, 1-6Hα), 6.78 (1H, d, *J* = 8.0, 1-H5), 7.34-7.37 (2H, m, 1-H2 + 1-H6), 7.54-7.64 (10H, m, 2-H2, 2H-6, 3-H2, 3-H6, 4-H2, 4-H6, 5-H2, 5-H6, 6-H2, 6-H6), 8.16 (1H, d, *J* = 8.5, H5), 8.24-8.27 (3H, m, H5), 8.34 (1H, d, *J* = 8.5, H5), 8.59 (1H, m, 7-NH), 8.87 (1H, s, 2-NH), 9.16-9.22 (4H, m, 3-NH, 4-NH, 5-NH, 6-NH); δ_C unable to obtain meaningful spectrum; ESI-HRMS found *m/z* 1160.4982 $[M+Na]^+$, $C_{62}H_{71}N_7NaO_{14}$ requires 1160.4951; found *m/z* 1182.4777 $[M+2Na-H]^+$, $C_{62}H_{70}N_7Na_2O_{14}$ requires 1182.4771.

NH_2-[*O*-*p*Cl-Bn-(3-HABA)]-[*O*-2-CH_2-Nap-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 45

δ_H (300 MHz,DMSO-d_6) 1.33 (6H, d, *J* = 6.2, 3-Hβ), 3.93 (2H, d, *J* = 5.8, 4-Hα), 4.70 (1H, sept, *J* = 6.2, 3-Hα), 5.06 (2H, s, 1-Hα), 5.49 (2H, s, 2-Hα), 6.69 (1H, d, *J* = 8.3, 1-H5), 7.39-7.50 (5H, m, 1-H2, 1-H6 + ArC<u>H</u>), 7.53 (1H, dd, *J* = 8.3, 1.7, 3-H6), 7.58 (1H, d, *J* = 1.7, 3-H2), 7.61 (1H, dd, *J* = 8.3, 1.7, 2-H6), 7.70 (1H, dd, *J* = 8.5, 2.8, 2-ArC<u>H</u>), 7.77 (1H, d, *J* = 1.8, 2-H2), 7.84-7.87 (2H, m, ArC<u>H</u>), 7.92 (1H, d, *J* = 8.8, ArC<u>H</u>), 8.07-8.09 (2H, m, 3-H5 + ArCH), 8.20 (1H, d, *J* = 8.3, 2-H5), 8.81 (1H, t, *J* = 5.8, 4-N<u>H</u>), 9.21 (1H, s, 2-N<u>H</u>), 9.30 (1H, s, N<u>H</u>); δ_C (125 MHz, DMSO-d_6) 30.4, 41.2, 68.6, 70.4, 71.4, 111.2, 111.8, 112.5, 119.9, 120.8, 121.4, 121.8, 124.9, 125.3, 125.9, 126.2, 126.4, 127.6, 128.1, 128.3, 129.2, 129.8, 130.0, 131.1, 131.6, 132.3, 132.6, 132.8, 134.4, 136.0, 139.2, 142.4, 144.1, 148.1, 148.8, 151.4, 164.2, 164.6, 165.8, 171.4; ESI-HRMS found *m/z* 809.2468 $[M+Na]^+$, $C_{44}H_{39}ClN_4NaO_8$ requires 809.2354.

NH_2-[*O*-Bn-(3-HABA)]-[*O*-*m*CF_3-Bn-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Gly-CO_2H 46

δ_H (300 MHz,DMSO-d_6) 1.34 (6H, d, *J* = 6.0, 3-Hβ), 3.93 (1H, d, *J* = 5.8, 4-Hα), 4.71 (1H, sept, *J* = 6.0, 3-Hα), 5.14 (2H, s, Hα), 5.42 (2H, s, Hα), 6.77 (1H, d, *J* = 8.3, 1-H5), 7.32 (1H, m, 1-HAr4), 7.37-7.41 (3H, m, 1-H6 + 1-HAr3), 7.48 (2H, d, *J* = 7.3, 1-HAr2), 7.53 (2H, m, 3-H6 + 1-H2), 7.58 (1H, d, *J* = 1.3, 3-H2), 7.60-7.63 (2H, m, 2-H6 + 2-HAr5), 7.67 (1H, m, 2-HAr4), 7.73 (1H, dd, *J* = 8.3, 1.5, 2-H2), 7.85 (1H, d, *J* = 7.7, 2-HAr6), 8.00 (1H, s, 2-HAr2), 8.07 (1H, d, *J* = 8.3, 3-H5), 8.13 (1H, d, *J* = 8.3, 2-H5), 8.81 (1H, t, *J* = 5.9, 4-N<u>H</u>), 9.31 (2H, m, 2-N<u>H</u> + 3-N<u>H</u>); δ_C (125 MHz, DMSO-d_6) 21.8, 41.2, 69.3, 69.5, 71.5, 111.5, 111.7, 112.7, 113.5, 114.1, 116.4, 119.9, 120.2, 121.4, 121.9, 122.2, 123.9, 125.3, 124.6, 127.4 (2), 127.8, 128.4 (2), 129.5 130.1, 131.2, 131.3, 131.4, 137.0, 138.3, 142.8, 148.1, 149.0, 158.1, 158.4, 164.2, 164.5, 165.8, 171.4; ESI-HRMS found *m/z* 793.2438 $[M+Na]^+$, $C_{41}H_{37}F_3N_4NaO_8$ requires 793.2456.

NH_2-[*O*-Bn-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-Gly-CO_2H 47

δ_H (300 MHz,DMSO-d_6) 1.02-1.04 (12H, m, 2-Hγ + 3-Hγ), 2.10-2.18 (2H, m, 2-Hβ + 3-Hβ), 3.90-3.95 (6H, m, 2-Hα, 3-Hα + 4-Hα), 5.18 (2H, s, 1-Hα), 6.73 (1H, d, *J* = 8.1, 1-H5), 7.32-7.42 (4H, m, 1-H6, 1-HAr3 + 1-HAr4), 7.48 (1H, d, *J* = 1.7, 1-H2), 7.51-7.60 (6H, m, 2-H2, 2-H6, 3-H2, 3-H6 +1-HAr2),8.03 (1H, d, *J* = 8.1, 3-H5), 8.21 (1H, d, *J* = 8.3, 2-H5), 8.81 (1H, t, *J* = 5.9, 4-N<u>H</u>), 9.01 (1H, s, 2-N<u>H</u>), 9.37 (1H, s, 3-N<u>H</u>); δ_C (125 MHz, DMSO-d_6) 19.08, 19.12, 19.4, 27.8, 27.8, 41.2, 69.5, 74.6, 110.5, 110.8, 110.8, 111.1, 112.7, 119.8, 119.8, 120.1, 120.7, 120.9, 121.5, 122.1, 127.4 (2), 127.8, 127.8, 128.4 (2), 129.6, 129.6, 130.1, 131.3, 137.1, 144.4, 148.6, 149.7, 164.2, 164.5, 165.8, 171.4; ESI-HRMS found *m/z* 705.2897 $[M+Na]^+$, $C_{37}H_{40}N_4NaO_8$ requires 705.2900.

NH_2-[*O*-CH_2-COOH-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-[*O*-*i*Leu-(3-HABA)]-Val-CO_2H 48

δ_H (300 MHz,DMSO-d_6) 0.94-1.02 (15H, m, 2-Hγ, 3-Hγ + 4-Hγ), 1.30 (3H, d, *J* = 6.0, 3-$CH_\alpha(CH_3)$), 1.63-1.77 (2H, m, 3-Hβ + 3-Hβ'), 2.12 (1H, m, 2-Hβ), 2.18 (1H, m, 4-Hβ), 3.93 (2H, d, *J* = 6.6, 2-Hα), 4.29 (1H, t, *J* = 7.6, 4-Hα), 4.57 (1H, m, 3-Hα), 4.66 (2H, s, 1-Hα), 7.01 (1H, d, *J* = 8.1, 1-H5),

7.52 (1H, s, 1-H2), 7.56-7.58 (5H, m, 1-H6, 2-H2, 2-H6, 3-H2 + 3-H6), 8.03 (1H, d, *J* = 8.3, 3-H5), 8.07 (1H, d, *J* = 8.1, 2-H5), 8.40 (1H, d, *J* = 8.1, 4-NH), 9.33 (1H, s, 2-NH), 9.36 (1H, s, 3-NH), 11.00 (1H, s, 1-Hα(CO_2H); (125 MHz, DMSO-d_6) 9.3, 18.9, 19.0, 19.4, 27.7, 28.5, 29.5, 58.4, 66.7, 74.6, 76.1, 111.0, 113.0, 115.2, 115.6, 119.7, 120.2, 121.7, 122.0, 122.7, 128.9, 130.4, 130.5, 130.6, 130.9, 131.1, 142.9, 148.2, 150.0, 163.9, 164.2, 164.9, 166.2, 173.2; ESI-HRMS was not obtainable.

NH_2-[*O*-CH_2-COOH-(3-HABA)]-[*O*-2-Nap-(3-HABA)]-[*O*-*i*Bu-(3-HABA)]-Val-CO_2H 49

δ_H (300 MHz,DMSO-d_6) 0.97-1.00 (12H, m, 3-Hγ + 4-Hγ), 2.08 (1H, m, 3-Hβ), 2.19 (1H, m, 4-Hβ), 3.9 (2H, d, *J* = 6.9, 3-Hα), 4.30 (1H, t, *J* = 7.5, 4-Hα), 4.65 (2H, s, 1-Hα), 5.47 (2H, s, 2-Hα), 6.99 (1H, d, *J* = 7.9, 1-H5), 7.52-7.66 (8H, m, 1-H2, 1-H6, 2-H6, 3-H2, 3-H6 + 2-ArCH), 7.79 (1H, s, 2-H2), 7.85 (1H, d, *J* = 7.1, 2-ArCH), 7.91-7.98 (3H, m, 3-H5 + 2-ArCH), 8.03 (1H, d, *J* = 8.1, 2-H5), 8.07 (1H, s, 2-HAr2), 8.40 (1H, d, *J* = 8.1, 4-NH), 9.43 (1H, s, 3-NH), 9.57 (1H, s, 2-NH), 10.98 (1H, s, 1-Hα(CO_2H); 125 MHz, DMSO-d_6) 18.9, 19.1, 19.4, 27.7, 29.6, 58.4, 66.7, 70.2, 74.76, 111.4, 112.0, 115.3, 115.6, 120.0, 120.1, 122.2, 122.3, 123.1, 125.1, 125.7, 126.1, 126.4, 127.6, 127.7, 128.1, 128.9, 130.0, 130.6, 130.8, 130.9, 131.0, 132.6, 132.8, 134.4, 142.9, 149.8, 149.9, 164.0, 164.2, 164.9, 166.2, 173.2; ESI-HRMS found *m/z* 776.3295 [M], $C_{43}H_{44}N_4O_{10}$ requires 776.3057.

NH_2-[*O*-Bn-(3-HABA)]-[*O*-2-CH_2-Nap-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Asp-CO_2H 50

δ_H (300 MHz,DMSO-d_6) 1.34 (6H, d, *J* = 6.0, 3-Hβ), 4.71 (1H, sept, *J* = 6.0, 3-Hα), 4.76 (1H, m, 4-Hα), 5.13 (2H, s, 1-Hα), 5.32 (2H, s, 2-Hα), 5.50 (2H, br. s, 4-Hβ), 6.72 (1H, d, *J* = 8.1, 1-H5), 7.31-7.60 (18H, m, 1-H2, 1-H6, 2-H6, 3-H2, 3-H6, 4-HβCO_2H + ArCH), 7.68 (1H, s, 2-H2), 8.07 (1H, d, *J* = 8.3, 3-H5), 8.18 (1H, d, *J* = 8.3, 2-H5), 8.70 (1H, d, *J* = 7.90, 4-NH), 9.15 (1H, s, 2-NH), 9.29 (1H, s, 3-NH); δ_C (125 MHz, DMSO-d_6) 21.8, 25.1, 35.9, 39.1, 67.0, 69.4, 69.5,71.5, 111.2, 111.7, 112.5, 112.9, 120.0, 120.2, 120.8, 121.5, 121.8, 121.8, 127.5 (2) ,127.8, 128 .4 (2), 128.5 (2), 129.3 (2), 129.8, 130.0, 131.2, 131.5, 132.6, 135.8, 137.0, 142.4, 142.9, 144.3, 148.1, 148.6, 164.2, 164.6, 165.4, 170.7, 171.8, 172.6;ESI-HRMS found *m/z* 833.1894 $[M+Na]^+$, $C_{46}H_{42}N_4NaO_{10}$ requires 833.2799.

NH_2-[*O*-Bn-(3-HABA)]-[*O*-2-CH_2-Nap-(3-HABA)]-[*O*-*i*Pr-(3-HABA)]-Lys-CO_2H 51

δ_H (300 MHz,DMSO-d_6) 1.33 (6H, d, *J* = 6.0, 3-Hβ), 1.40-1.47 (4H, m, 4-Hδ + 4-Hε) 1.77-1.88 (4H, m, 4-Hβ + 4-Hγ), 4.38-4.42 (1H, m, 4-Hα), 4.71 (1H, sept, *J* = 6.0, 3-Hα), 5.08 (2H, s, 1-Hα), 5.50 (2H, s, 2-Hα), 6.70 (1H, d, *J* = 8.1, 1-H5), 7.29–17.42 (5H, m, 1-H6 + ArCH), 7.47-7.50 (2H, m, ArCH), 7.56-7.72 (6H, m, 2-H6, 3-H6 + ArCH), 7.77 (1H, s, 3-H2), 7.84-7.87 (2H, m, ArCH), 7.92 (1H, d, *J* = 8.5, ArCH), 8.07-8.08 (2H, m, 2-H2 + 3-H5), 8.21 (1H, d, *J* = 8.1, 2-H5), 8.56 (1H, d, *J* = 7.9, 4-NH), 9.22 (1H, s, 2-NH), 9.32 (1H, s, 3-NH); δ_C (125 MHz, DMSO-d_6) 21.8, 22.8, 25.1, 30.1, 38.7, 52.3, 69.4, 70.4, 71.5, 111.1, 111.8, 112.5, 113.0, 120.1, 121.5, 121.7, 121.8, 125.3, 125.9, 126.2, 126.4, 127.4 (2), 127.6, 127.7, 127.7, 128.1, 128.3 (2), 129.8, 130.2, 131.2, 131.6, 132.6, 132.8, 134.4, 137.0, 142.4, 144.3, 148.1, 148.8, 157.7, 158.0, 164.2, 164.6, 165.8, 173.3; ESI-HRMS found *m/z* 824.3677 $[M+H]^+$, $C_{48}H_{50}N_5O_8$ requires 824.3659.

6.4 Molecular Modelling (Chapter 2)

Conformational search and Superposition with a gp41 extended helix

A conformational search was performed on hexamer **43**. The structure was minimised by employing a full *Monte Carlo* search in the software Macromodel® using the MMFFs (Merk Molecular Force Fields) method. 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. The results revealed the lowest energy conformation was the extended structure and all six side chains lie on the same face; a conformation displaying an alternative arrangement of side-chains however, has a relative potential energy of +3.2 kJ mol^{-1} demonstrating a variety of rotamers are accessible. Using a crystal structure of gp41 (PDB ID: 1AIK) we took a series of superpositions from our hexamer using different combinations of side chains (eg side chains 1,2 + 3 or 5,4 + 3) and the extended helix using different combinations of residues (e.g. *i*, *i* + 3 and *i* + 7 or *i*, *i* + 4 and *i* + 8) and at varying positions on the helix (e.g. towards the *N* or *C* terminus). From the relatively small set we sampled in comparison to the available combinations, we achieved RMSD (Root Mean Square Deviation) values ranging from 0.421-0.788 when superimposing 3 atom pairs consisting of the oxygen of the alkoxy group and the a carbon of the amino acids.

6.5 Biophysical Assessment of Proteomimetics (Chapter 3)

Assay Procedures

Protein Titration: The assay was carried out manually and plates prepared as follows (96 well plate);

- Make a 4x concentration solution of protein; if concentration range desired is 337 μM – 760 nM, then a 1.52 mM stock is made
- Make a 2x concentration solution of tracer; if the final concentration desired is 54.5 nM, then a 109 nM stock is made
- Add buffer (50 μL) to each well
- Add the protein stock (50 μL) to the first well (6 rows) and mix
- Take half of this solution (50 μL) and transfer to the second well
- Mix and transfer half, repeating to the last well: the final 50 μL is discarded
- Add tracer solution (50 μL) to each well (3 rows)
- Add buffer solution (50 μL) to each well (3 rows)
- The final volume in all wells should be 100 μL
- Cover and allow the plate to incubate for a set amount of time

Competition Assays: Competition assays, including peptide displacement assays, were all carried out using the same format and kept on a single plate to reduce intensity fluctuations which can arise when using different plates (observed in preliminary experiments). The assays were prepared as follows (96 well plate);

- Make a 4.5x concentration solution of ligand
- Make a 3x concentration solution of protein
- Make a 3x concentration solution of tracer
- Add buffer (50 μL) to each well
- Add the ligand stock (100 μL) to the first well (6 rows) and mix
- Take 2/3 of this solution (100 μL) and transfer to the next well
- Mix and transfer 2/3 again, repeating to the last well: the final 100 μL is discarded
- Add the tracer solution (50 μL) to each well (3 rows)
- Add buffer solution (50 μL) to each well (3 rows)
- The final volume should be 150 μL
- Cover and allow the plate to incubate for a set amount of time

6.6 Monomer and Dimer Synthesis (Chapter 4)

3-Isobutoxy-4-nitrobenzoic acid 56

Procedure F; Methyl 3-isobutoxy-4-nitrobenzoate **16b** (3.50 g, 13.8 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (90 mL), 10% aqueous sodium hydroxide (21 mL). Following acidification the precipitate was filtered, dissolved in chloroform and dried with magnesium sulfate. The solution was filtered and the organic solvents removed under reduced pressure yielding a pale yellow solid which was crystallised (chloroform / methanol) to yield the title compound (3.15 g, 13.5 mmol, 95%) as a yellow crystals; m.p. 199.8-202.1°C (chloroform / methanol); (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%); δ_H (300 MHz, DMSO-d_6) 0.98 (6H, d, *J* = 6.7, H-10), 2.03 (1H, sept, H-9), 4.01 (2H, d, *J* = 6.4, H-8), 7.63 (1H, dd, *J* = 8.3 + 1.3, H-3), 7.74 (1H, d, *J* = 1.3, H-7), 7.96 (1H, d, *J* = 8.3, H-4); δ_C (75MHz, DMSO-d_6) 19.0, 28.0, 75.5, 115.6, 121.5, 125.3, 136.0, 142.5, 151.4, 166.12; υ_{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;

2,5-Dihydroxy-4-nitrobenzoic acid[167] 54

To a stirred solution of 3-hydroxy-4-nitrobenzoic acid **62** (40.00 g, 218.4 mmol) in 2N 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 strongly acidified *via* the addition of sulphuric acid (conc) and the resulting precipitate was removed by filtration. The aqueous solution was refluxed for 1 hour. After cooling to ambient temperature the resulting precipitate was collected *via* filtration to yield the title compound (16.11 g, 80.9 mmol, 37%) as gold microcrystals; m.p. 245.1-246.2°C (water); (Found C, 42.05; H, 2.35; N, 6.80%. $C_7H_5NO_6$ requires C, 42.22; H, 2.53; N, 7.03%); δ_H (300 MHz, DMSO-d_6) 7.37 (1H, s, H5), 7.49 (1H, s, H2); δ_C (75MHz, DMSO-d_6) 112.5, 119.1, 119.7, 141.7, 143.1, 152.3, 170.11; υ_{max}/cm^{-1} (solid state) = 3533, 3400-2000, 169, 1598, 1442, 1244, 760, 627; ESI-HRMS found *m/z* 198.0049 $[M-H]^-$, $C_7H_4NO_6$ requires 198.0044.

M**ethyl 2,5-dihydroxy-4-nitrobenzoate 63**

To a stirred solution of 2,5-dihydroxy-4-nitrobenzoic acid **54** (5 g, 25.1mmol) in methanol (200 mL) was added slowly concentrated sulphuric acid (2 mL) and the resulting solution was stirred at reflux overnight. The reation mixture was allowed to cool to ambient temperature and sodium bicarbonate was added until carbon dioxide evolution ceased. The mixture was added to water (250 mL) and extracted with ethyl acetate (3 x 100 mL) and the combined organic fractions washed further by brine (100 mL). The organic solvents were removed by reduced pressure and the resulting orange solid was crystallised by chloroform to yield the title compound (5.33g, 25.0 mmol, quant. yield) as orange crystals; m.p. 135.3-137.2°C (chloroform); (Found C, 45.15; H, 3.25; N, 6.45%. $C_8H_7NO_6$ requires C, 45.08; H, 3.31; N, 6.57%);δ_H (300 MHz,$CDCl_3$) 4.02 (3H, s, $CO_2\underline{Me}$), 7.69 (1H, s, H2), 7.71 (1H, s,

H5), 9.75 (1H, s, $O^6\underline{H}$), 10.19 (1H, s, $O^3\underline{H}$); δ_C (75MHz, $CDCl_3$) 53.7, 112.8, 120.8, 121.4, 137.5, 146.7, 153.5, 168.9; υ_{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.

Methyl 2,5-diisobutoxy-4-nitrobenzoate 64

Procedure A; Methyl 2,5-dihydroxy-4-nitrobenzoate **63** (4.00 g, 18.8 mmol) and potassium carbonate (13.0 g, 93.8 mmol) in dimethylformamide (200 mL), 1-bromo-2-methylpropane (6.5 mL, 56.3 mmol). Following work-up, the title compound (4.41 g, 13.6 mmol, 72%) was yielded as a yellow oil; δ_H (500 MHz,$CDCl_3$) 1.04-1.07 (12H, m, $H\gamma^3$ + $H\gamma^6$), 2.11-2.17 (2H, m, $H\beta^3$ + $H\beta^6$), 3.79 (2H, d, *J* = 6.4, Hα), 3.85 (2H, d, *J* = 6.4, Hα), 3.94 (3H, s, $CO_2\underline{Me}$), 7.39 (1H, s, H2), 7.47 (1H, s, H5); δ_C (75MHz, $CDCl_3$) 19.0, 19.1, 28.3, 28.3, 52.5, 76.2, 76.6, 110.3, 118.0, 125.3, 141.7, 145.7, 151.6, 165.6; υ_{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 55

Procedure F; Methyl 2,5-diisobutoxy-4-nitrobenzoate **64** (2.16 g, 8.0 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (50 mL), 10% aqueous sodium hydroxide (13 mL). Following workup solvent was removed under reduced pressure to yield the title compound (2.13 g, 6.84 mmol, 97%) as an amorphous yellow solid; δ_H (500 MHz, $CDCl_3$) 1.05 (6H, d, *J* = 6.6, $H\gamma^3$), 1.12 (6H, d, *J* = 6.6, $H\gamma^6$), 2.14 (1H, m, $H\beta^3$), 2.25 (1H, m, $H\beta^6$), 3.91 (2H, d, *J* = 6.4, Hα), 4.04 (2H, d, *J* = 6.4, Hα), 7.50 (1H, s, H5), 7.90 (1H, s, H2); δ_C (125MHz, $CDCl_3$) 19.0, 19.1, 28.1, 28.2, 76.6, 77.6, 110.1, 119.5, 121.9, 142.4, 146.9, 150.3, 163.7; υ_{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 57

Procedure C; Methyl 2,5-diisobutoxy-4-nitrobenzoate **64** (2.24 g, 6.89 mmol), tin(II) chloride dihydrate (9.32 g, 41.33 mmol) in ethyl acetate (50 mL). Following workup solvent was removed under reduced pressure to yield the title compound (1.83 g, 6.2 mmol, 90%) as light brown greasy solid; δ_H (500 MHz, $CDCl_3$) 1.03-1.06 (12H, m, $H\gamma^3$ + $H\gamma^6$), 2.07-215 (2H, m, $H\beta^3$ + $H\beta^6$), 3.71 (2H, d, *J* = 6.4, Hα), 3.76 (2H, d, *J* = 6.4, Hα), 3.85 (3H, s, $CO_2\underline{Me}$), 6.29 (1H, s, H5), 7.33 (1H, s, H2); δ_C (75MHz, $CDCl_3$) 19.3, 19.4, 28.4, 28.5, 51.5, 75.2, 76.0, 100.0, 107.7, 114.8, 139.7, 142.2, 155.7, 166.8; υ_{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;

NH_2-[*O*-*i*Bu-(3-HABA)]-[O^3-*i*Bu-O^6-*i*Bu-(3-DHABA)]-CO_2H 60

To a solution containing 3-isobutoxy-4-nitrobenzoic acid **56** (250 mg, 1.12 mmol) in chloroform (10 mL), was added ghosez's reagent (145 μL, 1.08 mmol) and the resulting mixture was refluxed for 3 hours. Methyl 4-amino-2,5-diisobutoxybenzoate **57** (331 mg, 1.12 mmol) was subsequently added and

heated at 50 °C overnight. The solvents were removed under reduced pressure and the resulting mixture submitted to flash chromatograpy (Biotage: DCM/Ethyl acetate gradient) before taking intermediate **67** through to the next step without further isolation; Intermediate **68** was obtained from intermediate **67** by *procedure C* following standard work up without purification/ isolation; Following *Procedure F,* NH_2-[*O*-*i*Bu-(3-HABA)]-[O^3-*i*Bu-O^6-*i*Bu-(3-DHABA)]-CO_2H **60** was obtained and purified by biotage flash chromatography to isolate the title compound (estimated overall yield: 42%) as a beige amorphous solid; δ_H (500 MHz, DMSO-d_6) 1.00-1.04 (18H, m, 1-Hγ, 2-Hγ^3 + 2-Hγ^6), 2.01-215 (3H, m, 1-Hβ, 2-Hβ^3 + 2-Hβ^6), 3.75 (2H, d, *J* = 6.6, 2-Hα^6), 3.79 (2H, d, *J* = 6.6, 1-Hα), 3.84 (2H, d, *J* = 6.2, 2-Hα^3), 5.49 (2H, br. s, N<u>H</u>$_2$), 6.71 (1H, d, 1-H5), 7.30 (1H, d, *J* = 1.7, 1-H2), 7.32-7.34 (2H, m, 1-H6 + 2-H2), 8.07 (1H, s, 2-H5), 8.93 (1H, s, 2-N<u>H</u>); δ_C (125MHz, $CDCl_3$) 19.0, 19.1, 19.2, 27.8, 27.9, 27.9, 74.1, 75.0, 75.3, 106.3, 109.8, 112.4, 114.0, 114.4, 120.6, 212.2, 132.5, 141.6, 142.3, 144.7, 152.9, 164.5, 166.6;υ_{max}/cm^{-1} (solid state) = 3491, 3424, 3361, 2958, 1711, 1593, 1515, 1435, 1261, 1026, 751, 623; ESI-HRMS found *m/z* 473.2664 $[M+H]^+$, $C_{26}H_{37}N_2O_6$ requires 473.2646;

NH_2-[O^3-*i*Bu-O^6-*i*Bu-(3-DHABA)]-[*O*-*i*Bu-(3HABA)]-CO_2H 61

Acyl chloride formation Method A was used to isolate a stable acyl chloride from 2,5-diisobutoxy-4-nitrobenzoic acid **55** (150 mg, 0.48 mmol). The acyl chloride is dissolved in chloroform (10 mL) and methyl 4-amino-3-isobutoxybenzoate **17b** (107 mg, 0.48mmol) was subsequently added and the resulting mixture was heated at 50 °C overnight. The solvents were removed under reduced pressure and the resulting mixture submitted to flash chromatograpy (Biotage: DCM/Ethyl acetate gradient) before taking intermediate **69** through to the next step without further isolation; Intermediate **70** was obtained from intermediate **69** by *procedure C* following standard work up without purification/ isolation; Following *Procedure F,* NH_2-[O^3-*i*Bu-O^6-*i*Bu-(3-DHABA)]-[*O*-*i*Bu-(3HABA)]-CO_2H **61** was obtained and purified by biotage flash chromatography to isolate the title compound (estimated overall yield: 45%) as a beige amorphous solid which was crystallised from chloroform to give small rectangular crystals; δ_H (500 MHz, $CDCl_3$) 1.01 (6H, d, *J* = 6.6, 1-Hγ^3), 1.06-1.08 (12H, m, 1-Hγ^6 + 2-Hγ), 2.12-2.24 (3H, m, 1-Hβ^3, 1-Hβ^6 + 2-Hβ), 3.86 (2H, d, *J* = 6.4, 1-Hα^3), 3.93-3.96 (4H, m, 1-Hα^6 + 2-Hα), 6.39 (1H, s, 1-H5), 7.66-7.68 (2H, m, 1-H2 + 2-H2), 7.84 (1H, d, *J* = 8.5, 2-H6), 8.76 (1H, d, *J* = 8.5, 2-H5), 10.43 (1H, s, 2-N<u>H</u>); δ_C (125MHz, $CDCl_3$) 19.2, 19.2, 19.3, 27.8, 28.0, 28.3, 75.2, 75.6, 75.6, 99.9, 111.5, 112.6, 114.1, 119.1, 124.2, 134.3, 141.0, 141.6, 147.5, 152.7, 164.4; υ_{max}/cm^{-1} (solid state) = 3639, 3495, 3333, 2959, 1691, 1588, 1520, 1206, 1027, 768; ESI-HRMS found *m/z* 473.2647 $[M+H]^+$, $C_{26}H_{37}N_2O_6$ requires 473.2646;

NH_2-[*O*-*i*Bu-(3HABA)]-[*O*-*i*Bu-(3HABA)]-CO_2Me 72

To a solution containing 3-isobutoxy-4-nitrobenzoic acid **56** (107 mg, 0.45 mmol) and methyl 4-amino-3-isobutoxybenzoate **17b** (100 mg, 0.45mmol) in chloroform (2 mL) was added thionyl

chloride (65 μL, 0.90 mmol) and the resulting mixture was heated under microwave irradiation at 50 °C for 60 minutes. The solvents were removed under reduced pressure and the nitro/ester intermediate was taken through to the next step without purification/isolation. *Procedure C* followed by purification by biotage flash chromatography resulted in the title compound (yield over 2 steps: 54 mg, 0.13 mmol, 30%) as a cream amorphous solid; δ_H (500 MHz, $CDCl_3$) 1.07 (6H, d, *J* = 6.6, 1-Hγ), 1.12 (6H, d, *J* = 6.6, 2-Hγ), 2.14-2.24 (2H, m, 1-Hβ + 2-Hβ), 3.86 (2H, d, *J* = 6.6, 1-Hα), 3.92 (3H, s, CO_2Me), 3.94 (2H, d, *J* = 6.6, 2-Hα), 6.76 (1H, d, *J* = 8.1, 1-H5), 7.32 (1H, dd, *J* = 8.1, 1.9, 1-H6), 7.42 (1H, d, *J* = 1.9, 1-H2), 7.56 (1H, d, *J* = 1.7, 2-H2), 7.74 (1H, dd, *J* = 8.5, 1.7, 2-H6), 8.64 (1H, d, *J* = 8.5, 2-H5), 8.74 (1H, s, 2-NH); δ_C (125MHz, $CDCl_3$) 19.3, 19.4, 28.3, 28.4, 28.3, 52.0, 74.8, 75.1, 110.4, 111.5, 113.6, 118.2, 120.0, 123.5, 124.3, 124.5, 132.8, 146.4, 146.9, 165.1, 166.9; υ_{max}/cm^{-1} (solid state) = 3461, 3333, 2959, 1721, 1541, 1271, 1203, 1020, 759, 586; ESI-HRMS found *m/z* 415.2243 $[M+H]^+$, $C_{23}H_{31}N_2O_5$ requires 415.2227;

FmocNH-[O-*i*Pr-(3HABA)]-[O-Bn-(3HABA)]-CO_2Me 79

To a solution containing **75** (100 mg, 0.24 mmol) and **77** (62 mg, 0.24 mmol) in chloroform (2 mL) was added thionly chloride (34 μL, 0.90 mmol) and the resulting mixture was heated under microwave irradiation at 50 °C for 60 minutes. The solvents were removed under reduced pressure and the crude material purified *via* Biotage® flash chromatography (Snap® cartridge: hexane-dichloromethane-ethyl acetate ramp) to yield intermediate **79** (90.6 mg, 0.14 mmol, 60%) as an amorphous beige solid, identified by NMR and HRMS; δ_H (300 MHz, $CDCl_3$) 1.39 (6H, d, *J* = 6.0, 1-Hγ), 3.93 (3H, s, CO_2Me), 4.35 (1H, t, *J* = 7.0, FHβ), 4.53 (1H, t, *J* = 7.0, FHα), 4.65 (1H, sept, *J* = 6.0, 1-Hα), 5.23 (2H, s, 2-Hα), 7.26 (1H, m, 1-H6), 7.35 (2H, t, *J* = 7.5, FHAr4), 7.40-7.47 (9H, m, FHAr3, 1-H2, 1-H5, 2-HAr2, 2-HAr3 + 2-HAr4), 7.64 (2H, d, *J* = 7.3, FHAr2), 7.73 (1H, d, *J* = 1.6, 2-H2), 7.77-7.81 (3H, m, 2-H6 + FHAr5), 8.12 (1H, br. s, 1-NH) 8.67 (1H, d, *J* = 8.4, 2-H5), 8.80 (1H, s, NH); ESI-HRMS found *m/z* 657.2567 $[M+H]^+$, $C_{40}H_{37}N_2O_7$ requires 657.2595;

NH_2-[*O*-*i*Bu-(3HABA)]-[*O*-Bn-(3HABA)]-CO_2H 73

Procedure F; Intermediate **79** (85 mg, 0.13 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (3 mL), 10% aqueous sodium hydroxide (1 mL). Work up yielded the title compound (38 mg, 0.09 mmol, 70%) as a beige amorphous solid; δ_H (500 MHz, $CDCl_3$) 1.33 (6H, d, *J* = 6.0, 1-Hβ), 4.54 (1H, sept, *J* = 6.0, 1-Hα), 5.24 (2H, s, 2-Hα), 6.67 (1H, d, *J* = 7.9, 1-H5), 7.20 (1H, d, *J* = 7.9, 1-H6), 7.36 (1H, s, 1-H2), 7.40-7.50 (5H, m, 2-HAr2, 2-HAr3 + 2-HAr4), 7.76 (1H, s, 2-H2), 7.87 (1H, d, *J* = 8.3, 2-H6), 8.71 (1H, d, *J* = 8.3, 2-H5), 8.76 (1H, s, 2-NH); δ_C (125MHz, $CDCl_3$) 22.1, 70.8, 71.4, 112.2, 112.5, 113.6, 118.5, 120.3, 123.4, 123.8, 125.0, 127.9, 128.6, 128.9, 133.8, 136.0, 141.5, 144.8, 146.7, 165.1, 170.9;υ_{max}/cm^{-1} (solid state) = 3431, 3344, 2975m 1682, 1516, 1206, 764; ESI-HRMS found *m/z* 421.1775 $[M+H]^+$, $C_{24}H_{25}N_2O_5$ requires 421.1758;

FmocNH-[*O*-Bn-(3HABA)]-[O^3-*i*Bu-O^6-*i*Bu-(3-DHABA)]-CO_2Me 80

To a solution containing **76** (168 mg, 0.36 mmol) and **78** (104 mg, 0.36 mmol) in chloroform (2 mL) was added thionly chloride (53 μL, 0.72 mmol) and the resulting mixture was heated under microwave irradiation at 50 °C for 60 minutes. The solvents were removed under reduced pressure and the crude material purified *via* Biotage® flash chromatography (Snap® cartridge: hexane-dichloromethane-ethyl acetate ramp) to yield intermediate **80** (165 mg, 0.22 mmol, 62%) as an amorphous beige solid, identified by NMR and HRMS; δ_H (500 MHz, $CDCl_3$) 1.09 (6H, d, *J* = 6.6, 2-Hγ), 1.11 (6H, d, *J* = 6.8, 2-Hγ), 2.14-2.21 (2H, m, 2-Hβ^3 + 2-Hβ^6), 3.87-3.89 (4H, m, 2-Hα^3 + 2-Hα^6), 3.91 (3H, s, CO_2Me), 4.32 (1H, t, *J* = 7.1, FHβ), 4.53 (2H, d, *J* = 7.1, FHα), 5.24 (2H, s, 1-Hα), 7.33 (2H, t, *J* = 7.0, FHAr4), 7.41-7.46 (5H, m, FHAr3, 1-H5, 1-H6 + 1-Hα), 7.47-7.48 (4H, m, 1-HAr2 + 2-HAr3), 7.53 (1H, s, 2-H2), 7.61 (2H, FHAr2), 7.70 (1H, d, *J* = 1.5, 1-H2), 7.80 (2H, d, *J* = 7.8, FHAr5), 8.41 (1H, s, 2-H5), 8.85 (1H, s, 2-NH); ESI-MS found *m/z* 507.2 $[M+H]^+$;

NH_2-[*O*-Bn-(3HABA)]-[O^3-*i*Bu-O^6-*i*Bu-(3-DHABA)]-CO_2H 74

Procedure F; Intermediate **80** (150 mg, 0.20 mmol) in a 1:1 mixture of methanol : tetrahydrofuran (4 mL), 10% aqueous sodium hydroxide (1.5 mL). Work up yielded the title compound (57 mg, 0.11 mmol, 57%) as a beige amorphous solid; δ_H (500 MHz, $CDCl_3$) 1.09-1.10 (12H, m, 2-Hγ^3 + 2-Hγ^6), 2.16-2.25 (2H, m, 2-Hβ^3 + 2-Hβ^6), 3.90 (2H, d, *J* = 6.4, 2-Hα^3), 4.08 (2H, d, *J* = 6.4, 2-Hα^6), 5.15 (2H, s, 1-Hα), 6.77 (1H, d, *J* = 8.1, 1-H5), 7.37 (1H, m, 1-HAr4), 7.42 (2H, m, 1-HAr3), 7.46 (2H, m, 1-HAr2), 7.55 (1H, d, *J* = 1.9, 1-H2), 7.65 (1H, s, 2-H2), 8.54 (1H, s, 2-H5), 8.83 (1H, s, 2-NH); δ_C (125MHz, $CDCl_3$) 19.2, 19.3, 28.2, 28.3, 70.6, 75.5, 103.7, 110.7, 111.0, 113.5, 113.9, 120.5, 123.3, 127.4, 128.3, 128.7, 134.3, 136.4, 141.9, 146.0, 152.9, 165.3, 165.6; υ_{max}/cm^{-1} (solid state) = 3438, 3334, 2960, 1728, 1595, 1485, 1260, 1027, 733, 591 ESI-HRMS found *m/z* 507.2495 $[M+H]^+$, $C_{29}H_{35}N_2O_6$ requires 507.2490

6.7 Molecular Modelling (Chapter 4)

Molecular modelling was carried out using maestro as an interface for a number of applications developed by Schrödinger.

Conformational search

A low energy conformational search of dimers **60** and **61** was run in which the structures were minimised by employing a *Monte Carlo* search in Macromodel® using the MMFFs (Merk Molecular Force Fields) method 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.

Superposition with the ERα coactivator peptide

To calculate the RMSD, three atom pairs were superimposed consisting of the alkoxy oxygen of the dimers and the α-carbon of the leucine residues, at the *i*, *i* + 3 and *i* + 4 positions, of the NCOA2 peptide (PDB ID: 2QZO). The orientation of the dimer in relation to the coactivator was also altered thereby changing the direction of the dipole moment.

Docking

The LBD of ERα (PDB ID: 3ERD) was prepared for docking using the Prep Wiz function in maestro.[155] Once the protein was refined, Glide was then used to generate a grid for docking. As the structure is dimeric, only 1 constituent monomer was used within the docking grid (chain C). Epik was used to predict the ionisation states of each ligand and the energetic penalties associated with them. Having prepared the "ligands", Glide (default settings in extra precision (XP) mode) was then used to dock the results into the prepared protein. The predicted binding scores and energies can be used as a guide to assess a ligand's ability to bind to a specified region on the protein surface.

6.8 ^{1}H-^{1}H NOESY Spectra (Chapter 4)

NOESY spectra of **4.5** and **4.6** were acquired at 30 mM, 20 mM and 10 mM concentrations in $CDCl_3$ and at 30 mM in DMSO. 20 scans were taken over ~2.5 hours at ~299 K at 500 MHz. **Figure 6.1** shows relevant NOEs are present in DMSO (similar to those in $CHCl_3$: **Fig. 4.6** and **Fig. 4.7**), however, several aromatic peaks are less well resolved in DMSO than $CDCl_3$.

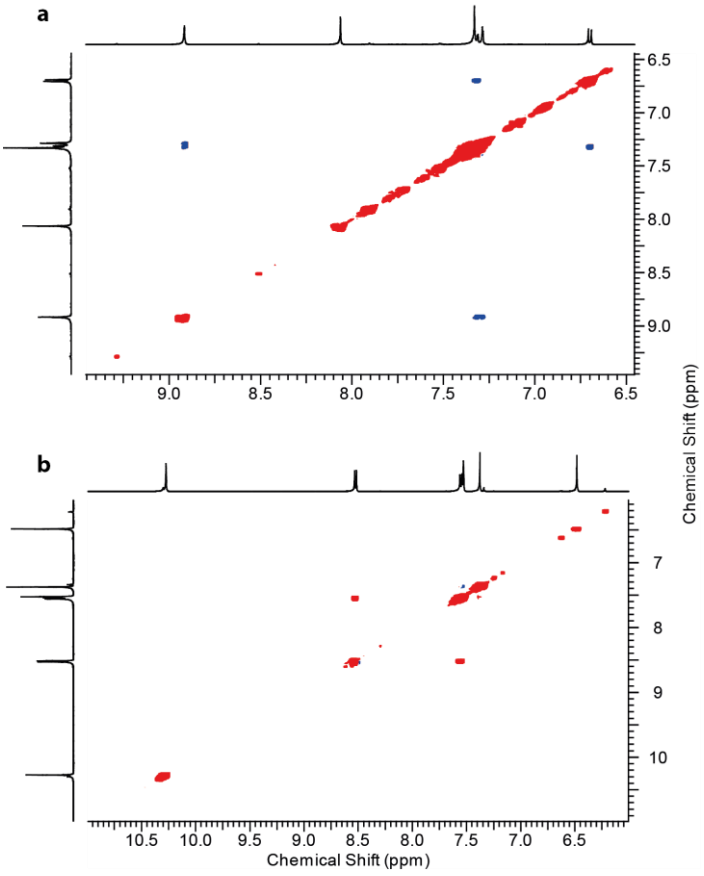


Figure 6.1 NOESY (500 MHz, DMSO-d_6) spectra of (**a**) dimer **60** at 30 mM and (**b**) dimer **61** at 30 mM.

6.9 Single Crystal X-ray Crystallographic Studies (Chapter 4)

Prismatic crystals were obtained by the slow evaporation of a solution of the compound in chloroform. A crystal of size 0.4 x 0.4 x 0.3 mm was used for data collection; θ range = 2.37≤θ≤ 26.13°, Crystals belong to Monoclinic; Space group C 2/c ; Formula = $C_{26}H_{36}N_2O_6$; Formula weight = 472.57; a = 22.224(3) Å, b = 16.806(2) Å; c = 15.8146(19) Å, β = 113.877(6)°, Volume = 5401.2(12), Z = 8, D (calculated): 1. 162 g/cm^3, μ = 0.082 mm^{-1}, Reflections collected 19204; Independent reflections 5305; Observed reflections 4102 [I >2σ(I)]; R value = 0.0445, wR_2 = 0.1195. Measurements were carried out at 120 K on a Bruker-Nonius Apex X8 diffractometer equipped with an Apex II CCD detector and using graphite monochromated Mo-Kα radiation from a FR591 rotating anode generator. The structure was solved by direct methods using SHELXLS-97 and refined using SHELXL-97. The compound crystallises in the monoclinic space group C 2/c with one molecule in the asymmetric unit. All non-hydrogen atoms were refined anisotropically. Hydrogen atoms were placed in calculated positions and refined using a riding model. All Uiso(H) values were constrained to be 1.2 times (1.5 for methyl) Ueq of the parent atom.

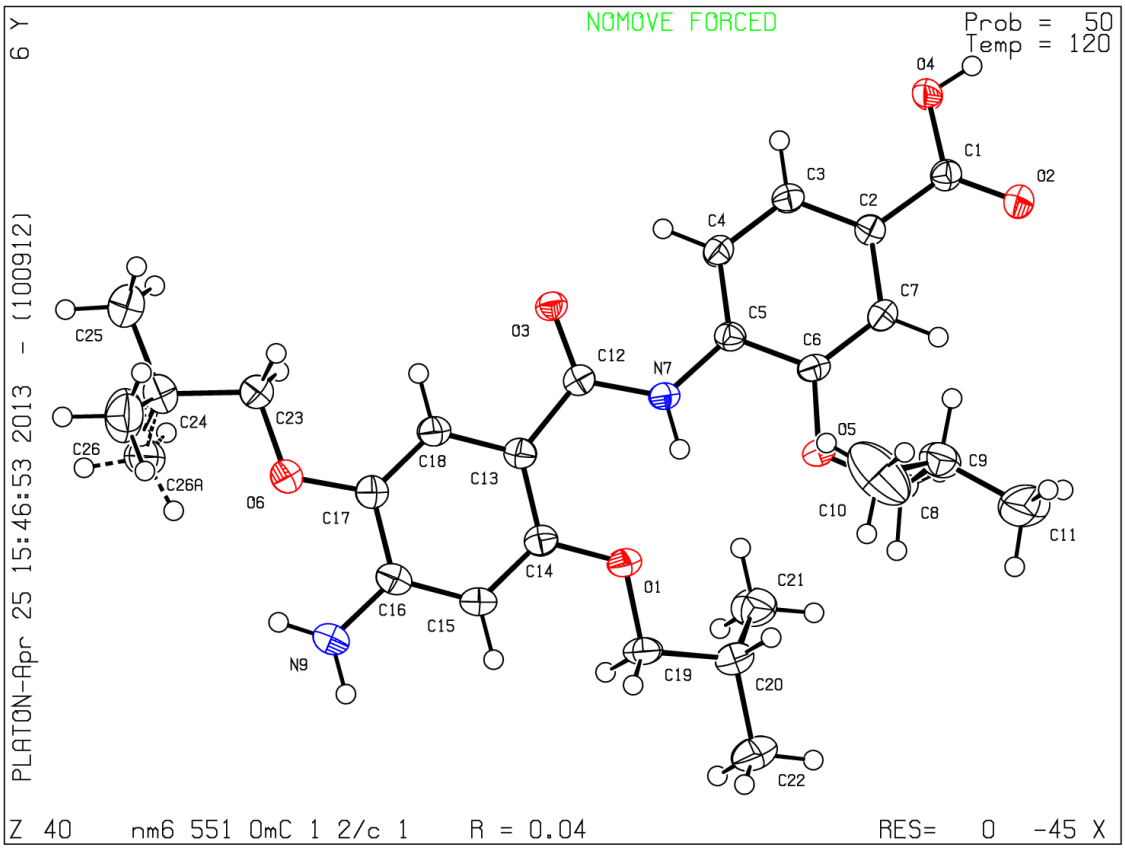


Figure 6.2 Single X-ray crystal structure data for dimer **61**.

6.10 H/D Exchange Studies (Chapter 4)

This was carried out on dimers **60** and **61** and data was taken from a previously published study for **81**[171] (Valeria Azzarito). A 10% $CD_3OD/CDCl_3$ system was used to ensure pseudo first order kinetics. An initial NMR spectrum was acquired with 450 μL of an analyte solution at 11.11 mM in $CDCl_3$ to which was added 50 μL of CD_3OD giving a final concentration of 10 mM. To minimize variability, a constant baseline correction was applied and a distinct, non-exchanging signal was used as an internal integration reference. Integration ranges for exchanging protons are user-defined and self-consistent within each experiment. Rate constants (and corresponding half-lives) were determined from the slope of a non-linear least squares fit to the graph following **Equation 6.1**.

$$A_t = A_0 e^{-kt}$$

Equation 6.1

A_t = Integral of amide proton at time t;

A_0 = Integral of amide proton at time zero (can be fixed)

K = reaction rate coefficient

The half life of the H/D exchange was determined using **Equation 6.2**.

$$t\,½ = \frac{\ln 2}{k}$$

Equation 6.2

References

1. V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161-173.
2. S. Jones and J. M. Thornton, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 13-20.
3. Z. Keskin, A. Gursoy, B. Ma and R. Nussinov, *Chem. Rev.* 2008, **108**, 1225-1244.
4. H. Yin and A. D. Hamilton, *Ange. Chem. Int. Ed.*, 2005, **44**, 4130-4163.
5. R. E. Babine and S. L. Bender, *Chem. Rev.*, 1997, **97**, 1359-1472.
6. T. Berg, *Ange. Chem. Int. Ed.* 2003, **42**, 2462-2481.
7. B. C. Cunningham and J. A. Wells, *J. of Mol. Bio.*, 1993, **234**, 554-563.
8. T. Clackson and J. A. Wells, *Science*, 1995, **267**, 383-386.
9. O. Keskin, B. Y. Ma and R. Nussinov, *J. of Mol. Bio.*, 2005, **345**, 1281-1294.
10. C. D. Thanos, M. Randal and J. A. Wells, *J. Am. Chem. Soc.*, 2003, **125**, 15280-15281.
11. P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine and N. P. Pavletich, *Science*, 1996, **274**, 948-953.
12. X. Liu, S. Dai, Y. Zhu, P. Marrack and J. W. Kappler, *Immunity*, 2003, **19**, 341-352.
13. 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.
14. D. C. Chan, D. Fass, J. M. Berger and P. S. Kim, *Cell*, 1997, **89**, 263-273.
15. B. N. Bullock, A. L. Jochim and P. S. Arora, *J. Am. Chem. Soc.*, 2011, **133**, 14220-14223.
16. M. Knowles and P. Selby, *Introduction to the cellular and molecular biology of cancer*, **Oxford:** Oxford University Press, 2005.
17. H. J. Thompson, R. Strange and P. J. Schedin, *Cancer Epidemiol. Biomarkers Prevent.*, 1992, **1**, 597-602.
18. L. Romer, C. Klein, A. Dehner, H. Kessler and J. Buchner, *Angew. Chem. Int. Ed.*, 2006, **45**, 6440-6460.
19. M. Hollstein, D. Sidransky, B. Vogelstein and C. Harris, *Science*, 1991, **253**, 49-53.
20. L. H. Chen, H. Yin, B. Farooqi, S. Sebti, A. D. Hamilton and J. D. Chen, *Mol. Cancer Ther.*, 2005, **4**, 1019-1025.
21. M. H. Herynk and S. A. W. Fuqua, *Endocr. Rev.*, 2004, **25**, 869-898.
22. 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.
23. R. C. J. Ribeiro, P. J. Kushner and J. D. Baxter, *Annu. Rev. Med.*, 1995, **46**, 443-453.
24. M. Carraz, W. Zwart, T. Phan, R. Michalides and L. Brunsveld, *Chem. Biol.*, 2009, **16**, 702-711.
25. E. H. Kong, A. C. W. Pike and R. E. Hubbard, *Biochem. Soc. Trans.*, 2003, **31**, 56-59.
26. 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.
27. 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. Xiao, X. Yang, H. Zhang, S. Fesik, S. H. Rosenberg and S. W. Elmore, *Cancer Res.*, 2008, **68**, 3421-3428.
28. S. A. Marshall, G. A. Lazar, A. J. Chirino and J. R. Desjarlais, *Drug Discov. Today*, 2003, **8**, 212-221.
29. B. Leader, Q. J. Baca and D. E. Golan, *Nature Reviews Drug Discovery*, 2008, **7**, 21-39.
30. J. M. Scholtz and R. L. Baldwin, *Annu. Rev. Biophys. Biomol. Struct.*, 1992, **21**, 95-118.
31. J. Garner and M. M. Harding, *Org. Biomol. Chem.*, 2007, **5**, 3577-3585.

32. J. H. B. Pease, R. W. Storrs and D. E. Wemmer, *Proc. Nat.l Acad. Sci. U.S.A.*, 1990, **87**, 5643-5647.
33. D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh and P. G. Schultz, *J. Am. Chem. Soc.*, 1991, **113**, 9391-9392.
34. M. Pellegrini, M. Royo, M. Chorev and D. F. Mierke, *Journal of Peptide Research*, 1997, **49**, 404-414.
35. A. K. Galande, K. S. Bramlett, J. O. Trent, T. P. Burris, J. L. Wittliff and A. F. Spatola, *ChemBioChem*, 2005, **6**, 1991-1998.
36. A. M. Leduc, J. O. Trent, J. L. Wittliff, K. S. Bramlett, S. L. Briggs, N. Y. Chirgadze, Y. Wang, T. P. Burris and A. F. Spatola, *Proc. Nat.l Acad. Sci. U.S.A.*, 2003, **100**, 11273-11278.
37. A. K. Galande, K. S. Bramlett, T. P. Burris, J. L. Wittliff and A. F. Spatola, *Journal of Peptide Research*, 2004, **63**, 297-302.
38. M. Chorev, E. Roubini, R. L. McKee, S. W. Gibbons, M. E. Goldman, M. P. Caulfield and M. Rosenblatt, *Biochemistry*, 1991, **30**, 5968-5974.
39. D. Y. Wang, W. Liao and P. S. Arora, *Angew. Chem. Int. Ed.*, 2005, **44**, 6525-6529.
40. J. K. Judice, J. Y. K. Tom, W. Huang, T. Wrin, J. Vennari, C. J. Petropoulos and R. S. McDowell, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 13426-13430.
41. S. K. Sia, P. A. Carr, A. G. Cochran, V. N. Malashkevich and P. S. Kim, *Proc. Nat.l Acad. Sci. U.S.A.*, 2002, **99**, 14664-14669.
42. R. S. Harrison, N. E. Shepherd, H. N. Hoang, G. Ruiz-Gomez, T. A. Hill, R. W. Driver, V. S. Desai, P. R. Young, G. Abbenante and D. P. Fairlie, *Proc. Nat.l Acad. Sci. U.S.A.*, 2010, **107**, 11686-11691.
43. H. E. Blackwell and R. H. Grubbs, *Angew. Chem. Int. Ed.*, 1998, **37**, 3281-3284.
44. C. E. Schafmeister, J. Po and G. L. Verdine, *J. Am. Chem. Soc.*, 2000, **122**, 5891-5892.
45. 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.
46. Peter E. Czabotar, D. Westphal, G. Dewson, S. Ma, C. Hockings, W. D. Fairlie, Erinna F. Lee, S. Yao, Adeline Y. Robin, Brian J. Smith, David C. S. Huang, Ruth M. Kluck, Jerry M. Adams and Peter M. Colman, *Cell*, 2013, **152**, 519-531.
47. L. D. Walensky, K. Pitter, J. Morash, K. J. Oh, S. Barbuto, J. Fisher, E. Smith, G. L. Verdine and S. J. Korsmeyer, *Mol.Cell*, 2006, **24**, 199-210.
48. M. L. Stewart, E. Fire, A. E. Keating and L. D. Walensky, *Nat. Chem. Biol.*, 2010, **6**, 595-601.
49. S. Baek, P. S. Kutchukian, G. L. Verdine, R. Huber, T. A. Holak, K. W. Lee and G. M. Popowicz, *J. Am. Chem. Soc.*, 2011, **134**, 103-106.
50. 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.
51. R. E. Moellering, M. Cornejo, T. N. Davis, C. Del Bianco, J. C. Aster, S. C. Blacklow, A. L. Kung, D. G. Gilliland, G. L. Verdine and J. E. Bradner, *Nature*, 2009, **462**, 182-188.
52. 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. USA*, 2010, **107**, 14093-14098.
53. S. Crunkhorn, *Nat. Rev. Drug. Discov.*, 2013, **12**, 741-741.
54. E. Cabezas and A. C. Satterhwait, *J. Am. Chem. Soc.*, 1999, **121**, 3862-3875.
55. A. Patgiri, A. L. Jochim and P. S. Arora, *Acc. Chem. Res.*, 2008, **41**, 1289-1300.
56. R. N. Chapman, G. Dimartino and P. S. Arora, *J. Am. Chem. Soc.*, 2004, **126**, 12252-12253.
57. D. Wang, M. Lu and P. S. Arora, *Angew. Chem. Int. Ed.*, 2008, **47**, 1879-1882.
58. L. K. Henchey, J. R. Porter, I. Ghosh and P. S. Arora, *ChemBioChem*, 2010, **11**, 2104-2107.

59. 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.
60. A. Patgiri, K. K. Yadav, P. S. Arora and D. Bar-Sagi, *Nat. Chem. Biol.*, 2011, **7**, 585-587.
61. A. B. Mahon and P. S. Arora, *Chem. Commun.*, 2012, **48**, 1416-1418.
62. L. Guerrero, O. S. Smart, C. J. Weston, D. C. Burns, G. A. Woolley and R. K. Allemann, *Angew. Chem. Int. Ed.,* 2005, **44**, 7778-7782.
63. H. Y. Fan, S.-A. Morgan, K. E. Brechun, Y.-Y. Chen, A. S. I. Jaikaran and G. A. Woolley, *Biochemistry*, 2011, **50**, 1226-1237.
64. S. Kneissl, E. J. Loveridge, C. Williams, M. P. Crump and R. K. Allemann, *ChemBioChem*, 2008, **9**, 3046-3054.
65. 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.
66. H. Jo, N. Meinhardt, Y. Wu, S. Kulkarni, X. Hu, K. E. Low, P. L. Davies, W. F. DeGrado and D. C. Greenbaum, *J. Am. Chem. Soc.*, 2012, **134**, 17704-17713.
67. G. P. Dado and S. H. Gellman, *J. Am. Chem. Soc.*, 1994, **116**, 1054-1062.
68. S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173-180.
69. D. Seebach, S. Abele, J. Schreiber, V. rg, B. Martinoni, A. K. Nussbaum, H. Schild, rg, H. Schulz, H. Hennecke, R. Woessner and F. Bitsch, *Chimia Int. J. Chem.*, 1998, **52**, 734-739.
70. R. P. Cheng, S. H. Gellman and W. F. DeGrado, *Chem. Rev.*, 2001, **101**, 3219-3232.
71. M. Werder, H. Hauser, S. Abele and D. Seebach, *Helvetica Chimica Acta*, 1999, **82**, 1774-1783.
72. D. Seebach, R. I. Mathad, T. Kimmerlin, Y. R. Mahajan, P. Bindschädler, M. Rueping, B. Jaun, C. Hilty and T. Etezady-Esfarjani, *Helvetica Chimica Acta*, 2005, **88**, 1969-1982.
73. J. A. Kritzer, J. D. Lear, M. E. Hodsdon and A. Schepartz, *J. Am. Chem. Soc.*, 2004, **126**, 9468-9469.
74. M. Hintersteiner, T. Kimmerlin, G. Garavel, T. Schindler, R. Bauer, N.-C. Meisner, J.-M. Seifert, V. Uhl and M. Auer, *ChemBioChem*, 2009, **10**, 994-998.
75. E. A. Harker, D. S. Daniels, D. A. Guarracino and A. Schepartz, *Bioorg. Med. Chem.*, 2009, **17**, 2038-2046.
76. J. Michel, E. A. Harker, J. Tirado-Rives, W. L. Jorgensen and A. Schepartz, *J. Am. Chem. Soc.*, 2009, **131**, 6356-6357.
77. O. M. Stephens, S. Kim, B. D. Welch, M. E. Hodsdon, M. S. Kay and A. Schepartz, *J. Am. Chem. Soc.*, 2005, **127**, 13126-12127.
78. A. Hayen, M. A. Schmitt, F. N. Ngassa, K. A. Thomasson and S. H. Gellman, *Angew. Chem. Int. Ed.*, 2004, **43**, 505-510.
79. M. A. Schmitt, S. H. Choi, I. A. Guzei and S. H. Gellman, *J. Am. Chem. Soc.*, 2005, **127**, 13130-13131.
80. S. De Pol, C. Zorn, C. D. Klein, O. Zerbe and O. Reiser, *Angew. Chem. Int. Ed.*, 2004, **43**, 511-514.
81. 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.
82. 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.
83. 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.
84. 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.

85. W. S. Horne, M. D. Boersma, M. A. Windsor and S. H. Gellman, *Angew. Chem. Int. Ed.*, 2008, **47**, 2853-2856.
86. J. K. Murray and S. H. Gellman, *Peptide Science*, 2007, **88**, 657-686.
87. 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.
88. M. D. Boersma, H. S. Haase, K. J. Peterson-Kaufman, E. F. Lee, O. B. Clarke, P. M. Colman, B. J. Smith, W. S. Horne, W. D. Fairlie and S. H. Gellman, *J. Am. Chem. Soc.*, 2011, **134**, 315-323.
89. 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. USA*, 2009, **106**, 14751-14756.
90. L. M. Johnson, D. E. Mortenson, H. G. Yun, W. S. Horne, T. J. Ketas, M. Lu, J. P. Moore and S. H. Gellman, *J. Am. Chem. Soc.*, 2012, **134**, 7317-7320.
91. A. S. Ripka and D. H. Rich, *Curr. Opin. Chem. Biol.*, 1998, **2**, 441-452.
92. D. Horwell, M. Pritchard, J. Raphy and G. Ratcliffe, *Immunopharmacology*, 1996, **33**, 68-72.
93. D. C. Horwell, W. Howson, G. S. Ratcliffe and H. M. G. Willems, *Bioorg. Med. Chem.*, 1996, **4**, 33-42.
94. E. Jacoby, *Bioorg. Med. Chem. Lett.*, 2002, **12**, 891-893.
95. B. P. Orner, J. T. Ernst and A. D. Hamilton, *J. Am. Chem. Soc.*, 2001, **123**, 5382-5383.
96. J. T. Ernst, O. Kutzki, A. K. Debnath, S. Jiang, H. Lu and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2002, **41**, 278.
97. 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.
98. 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.
99. 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.
100. J. T. Ernst, J. Becerril, H. S. Park, H. Yin and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2003, **42**, 535.
101. J. M. Rodriguez and A. D. Hamilton, *Angew. Chem. Int. Ed.*, **46**, 8614-8617.
102. 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.
103. J. M. Rodriguez, L. Nevola, N. T. Ross, G.-i. Lee and A. D. Hamilton, *ChemBioChem*, 2009, **10**, 829-833.
104. A. Volonterio, L. Moisan and J. Julius, *Org. Lett.*, 2007, **9**, 3733-3736.
105. S. M. Biros, L. Moisan, E. Mann, A. Carella, D. Zhai, J. C. Reed and J. Rebek, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 4641-4645.
106. C. G. Cummings, N. T. Ross, W. P. Katt and A. D. Hamilton, *Org. Lett.*, 2009, **11**, 25-28.
107. P. Maity and B. Konig, *Org. Lett.*, 2008, **10**, 1473-1476.
108. P. Tošovská and P. S. Arora, *Org. Lett.*, 2010, **12**, 1588-1591.
109. A. L. Rodriguez, A. Tamrazi, M. L. Collins and J. A. Katzenellenbogen, *J. Med. Chem.*, 2004, **47**, 600-611.
110. E. Hur, S. J. Pfaff, E. S. Payne, H. Gron, B. M. Buehrer and R. J. Fletterick, *PLoS Biol.*, 2004, **2**, 1303-1312.
111. J. R. Gunther, A. A. Parent and J. A. Katzenellenbogen, *ACS Chem. Biol.*, 2009, **4**, 435-440.
112. I. C. Kim and A. D. Hamilton, *Org. Lett.*, 2006, **8**, 1751-1754. *Angew. Chem. Int. Ed.*, 2007, **46**, 4471-4473.
114. J. Plante, F. Campbell, B. Malkova, C. Kilner, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2008, **6**, 138-146.

115. J.-M. Ahn and S.-Y. Han, *Tetrahedron Lett.*, 2007, **48**, 3543-3547.
116. 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.
117. A. Shaginian, L. R. Whitby, S. Hong, I. Hwang, B. Farooqi, M. Searcey, J. Chen, P. K. Vogt and D. L. Boger, *J. Am. Chem. Soc.*, 2009, **131**, 5564–5572.
118. L. R. Whitby, K. E. Boyle, L. Cai, X. Yu, M. Gochin and D. L. Boger, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 2861-2865.
119. 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.
120. F. Lu, S.-W.Chi, D.-H. Kim, K.-H. Han, I. D. Kuntz and R. K. Guy, *J. Combi. Chem.*, 2006, **8**, 315-325.
121. T. Heinz, D. M. Rudkevich and J. Rebek, *Angew. Chem. Int. Ed.*, 1999, **38**, 1136-1139.
122. F. Campbell, J. P. Plante, T. A. Edwards, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2010, **8**, 2344-2351.
123. K. Long, T. A. Edwards and A. J. Wilson, *Bioorg. Med. Chem.*, 2013, dx.doi.org/10.1016/j.bmc.2012.1009.1053.
124. S. Marimganti, M. N. Cheemala and J.-M. Ahn, *Org. Lett.*, 2009, **11**, 4418-4421.
125. P. Prabhakaran, A. Barnard, N. S. Murphy, C. A. Kilner, T. A. Edwards and A. J. Wilson, *Eur. J. Org. Chem.*, 2013, **2013**, 3504-3512.
126. N. S. Murphy, P. Prabhakaran, V. Azzarito, J. P. Plante, M. J. Hardie, C. A. Kilner, S. L. Warriner and A. J. Wilson, *Chemistry – A European Journal*, 2013, **19**, 5546-5550.
127. R. P. Cheng, S. H. Gellman and W. F. DeGrado, *Chem. Rev.*, 2001, **101**, 3219-3232.
128. D. J. Hill, M. J. Mio, R. B. Prince, T. S. Hughes and J. S. Moore, *Chem. Rev.*, 2001, **101**, 3893-4011.
129. D. Seebach, A. K. Beck and D. J. Bierbaum, *Chem. Biodivers.*, 2004, **1**, 1111-1239.
130. D.-W. Zhang, X. Zhao, J.-L. Hou and Z.-T. Li, *Chem. Rev.*, 2012.
131. C. M. Goodman, S. Choi, S. Shandler and W. F. DeGrado, *Nat. Chem. Biol.*, 2007, **3**, 252-262.
132. T. Hara, S. R. Durell, M. C. Myers and D. H. Appella, *J. Am. Chem. Soc.*, 2006, **128**, 1995-2004.
133. J. H. Lee, Q. Zhang, S. Jo, S. C. Chai, M. Oh, W. Im, H. Lu and H.-S. Lim, *J. Am. Chem. Soc.*, 2010, **133**, 676-679.
134. O. Keskin, A. Gursoy, B. Ma and R. Nussinov, *Chem. Rev.*, 2008, **108**, 1225-1244.
135. H. Yin and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2005, **44**, 4130-4163.
136. J. A. Wells and C. L. McLendon, *Nature*, 2007, **450**, 1001-1009.
137. K. Kirshenbaum, A. E. Barron, R. A. Goldsmith, P. Armand, E. K. Bradley, K. T. V. Truong, K. A. Dill, F. E. Cohen and R. N. Zuckermann, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 4303-4308.
138. J. K. Murray and S. H. Gellman, *Org. Lett.*, 2005, **7**, 1517-1520.
139. C. Douat-Casassus, K. Pulka, P. Claudon and G. Guichard, *Org. Lett.*, 2012, **14**, 3130-3133.
140. A. R. Gocke, D. G. Udugamasooriya, C. T. Archer, J. Lee and T. Kodadek, *Chem. Biol.*, 2009, **16**, 1133-1139.
141. I. Saraogi, C. D. Incarvito and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2008, **47**, 9691-9694.
142. J. M. Rodriguez, N. T. Ross, W. P. Katt, D. Dhar, G.-i. Lee and A. D. Hamilton, *ChemMedChem*, 2009, **4**, 649-656.
143. 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.
144. I. Saraogi and A. D. Hamilton, *Biochem. Soc. Trans.*, 2008, **36**, 1414.

145. C. G. Cummings and A. D. Hamilton, *Curr. Opin. Chem. Biol.*, 2010, **14**, 341-346.
146. B. P. Orner, J. T. Ernst and A. D. Hamilton, *J. Am. Chem. Soc.*, 2001, **123**, 5382-5383.
147. N. Delsuc, J.-M. Léger, S. Massip and I. Huc, *Angew. Chem. Int. Ed.*, 2007, **46**, 214-217.
148. D. Sánchez-García, B. Kauffmann, T. Kawanami, H. Ihara, M. Takafuji, M.-H. Delville and I. Huc, *J. Am. Chem. Soc.*, 2009, **131**, 8642-8648.
149. H. M. Konig, R. Abbel, D. Schollmeyer and A. F. M. Kilbinger, *Org. Lett.*, 2006, **8**, 1819-1822.
150. T.-K. Lee and J.-M. Ahn, *ACS Comb. Sci.*, 2010, **13**, 107-111.
151. N. R. Wurtz, J. M. Turner, E. E. Baird and P. B. Dervan, *Org. Lett.*, 2001, **3**, 1201-1203.
152. B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175-7185.
153. Y. Rao, X. Li and S. J. Danishefsky, *J. Am. Chem. Soc.*, 2009, **131**, 12924-12926.
154. 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.
155. S. Schrödinger Suite 2009 Protein Preparation Wizard; Epik version 2.0, LLC, New, N. York, 2009; Impact version 5.5, Schrödinger, LLC, New York, NY, 2009; Prime version and S. 2.1, LLC, New York, NY, 2009.
156. I. Azumaya, T. Okamoto, F. Imabeppu and H. Takayanai, *Tetrahedron*, 2003, **59**, 2325-2331.
157. 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.
158. Y. Rew, D. Sun, F. Gonzalez-Lopez De Turiso, M. D. Bartberger, H. P. Beck, J. Canon, A. Chen, D. Chow, J. Deignan, B. M. Fox, D. Gustin, X. Huang, M. Jiang, X. Jiao, L. Jin, F. Kayser, D. J. Kopecky, Y. Li, M.-C. Lo, A. M. Long, K. Michelsen, J. D. Oliner, T. Osgood, M. Ragains, A. Y. Saiki, S. Schneider, M. Toteva, P. Yakowec, X. Yan, Q. Ye, D. Yu, X. Zhao, J. Zhou, J. C. Medina and S. H. Olson, *J. Med. Chem.*, 2012, **55**, 4936-4954.
159. B. J. Deroo and K. S. Korach, *J. Clin. Invest.*, 2006, **116**, 561-570.
160. B. Fisher, J. P. Costantino, D. L. Wickerham, R. S. Cecchini, W. M. Cronin, A. Robidoux, T. B. Bevers, M. T. Kavanah, J. N. Atkins, R. G. Margolese, C. D. Runowicz, J. M. James, L. G. Ford and N. Wolmark, *J. Natl. Cancer Inst.*, 2005, **97**, 1652-1662.
161. L. Bergman, M. L. R. Beelen, M. P. W. Gallee, H. Hollema, J. Benraadt, F. E. van Leeuwen and A. G. Comprehensive Canc Ctr, *Lancet*, 2000, **356**, 881-887.
162. Y. Shang, X. Hu, J. DiRenzo, M. A. Lazar and M. Brown, *Cell*, 2000, **103**, 843-852.
163. R. Clarke, M. C. Liu, K. B. Bouker, Z. P. Gu, R. Y. Lee, Y. L. Zhu, T. C. Skaar, B. Gomez, K. O'Brien, Y. Wang and L. Hilakivi-Clarke, *Oncogene*, 2003, **22**, 7316-7339.
164. R. Clarke, F. Leonessa, J. N. Welch and T. C. Skaar, *Pharmacol. Rev.*, 2001, **53**, 25-71.
165. L. Bernstein, D. Deapen, J. R. Cerhan, S. M. Schwartz, J. Liff, E. McGann-Maloney, J. A. Perlman and L. Ford, *J. Natl. Cancer Inst.*, 1999, **91**, 1654-1662.
166. R. T. Nolte, G. B. Wisely, S. Westin, J. E. Cobb, M. H. Lambert, R. Kurokawa, M. G. Rosenfeld, T. M. Willson, C. K. Glass and M. V. Milburn, *Nature*, 1998, **395**, 137-143.
167. S. Marimganti, M. N. Cheemala and J. M. Ahn, *Org. Lett.*, 2009, **11**, 4418-4421.
168. Maestro, version 9.0, Schrödinger, LLC, New York, NY, 2009.
169. W. G. J. Hol, *Prog. Biophys. Mol. Biol.*, 1985, **45**, 149-195.
170. C. Sabot, K. A. Kumar, S. Meunier and C. Mioskowski, *Tetrahedron Lett.*, 2007, **48**, 3863-3866.
171. 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.

172. L. R. Steffel, T. J. Cashman, M. H. Reutershan and B. R. Linton, *J. Am. Chem. Soc.*, 2007, **129**, 12956-12957.
173. T. Edwards and A. Wilson, *Amino Acids*, 2011, **41**, 743-754.
174. J. Fernandez-Recio, M. Totrov and R. Abagyan, *J. Mol. Biol.*, 2004, **335**, 843-865.
175. J. Becerril and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2007, **46**, 4471-4473.
176. C. Mao, N. M. Patterson, M. T. Cherian, I. O. Aninye, C. Zhang, J. B. Montoya, J. Cheng, K. S. Putt, P. J. Hergenrother, E. M. Wilson, A. M. Nardulli, S. K. Nordeen and D. J. Shapiro, *J. Biol. Chem.*, 2008, **283**, 12819-12830.
177. R. Chmelar, G. Buchanan, E. F. Need, W. Tilley and N. M. Greenberg, *Int. J. Cancer*, 2007, **120**, 719-733.
178. 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**, 1923.

Appendix I

Collecting and processing fluorescence data:

To collect fluorescence data for any given interaction, assays were ran in triplicate (including 3 blank/control wells) and data was collected using an EnVision plate reader.

S- and P- channels corrected by subtracting the average polarisation value of the corresponding control wells. **Equation 1** is used to calculate the total intensity, which can then be inserted into **Equation 2** to get the anisotropy value for each individual well.

$$I = 2PG + S$$

Equation 1 Calculation for the total intensity:

$$r = (S\text{-}PG)/I$$

Equation 2 Calculation for the observed anisotropy:

I = Intensity, r = observed anisotropy, P = polarisation value for P- channel, S = polarisation value for S- channel. G = G factor,

Protein Titration

Average anisotropy values from the titration can be directly plotted and fitted using a logistic model in Origin 8.6. IC_{50} values may be extracted directly from this data, however, in order to obtain K_d, **Equation 3** is needed to calculate the fraction of ligand bound (or bound tracer: L_b). The minimum and maximum anisotropies (r_{min} and r_{max}) are identified from the data of plotted average anisotropies and λ is calculated from the change in intensity of the fluorophore in its bound and unbound state. This can be influenced by the hydrophobic environment of the protein which may vary λ away from a unitary value when the fluorophore is not influenced by its environment (λ =1). The value of L_b is then inserted into **Equation 4** ($y = L_b$) to determine the dissociation constant (k_1).

$$L_b = \frac{(r - r_{min})}{\lambda(r_{max} - r) + r - r_{min}}$$

Equation 3 Calculation for fraction of ligand bound.

$$y = \frac{\{(k_1 + [x] + [FL]) - \sqrt{\{(k_1 + [x] + [FL])^2 - 4[x][FL]\}}}{2}$$

Equation 4 Rearrangement of the equation allows for the calculation of the dissociation constant.

$L_{b} = y$ = Fraction of ligand bound, r = anisotropy, $\lambda = I_{bound}/I_{unbound}$ (intensity variation), [FL] = concentration of fluorescent ligand, $k_1 = K_d$, [x] = [Protein].

Competition Assay

In previous studies, competition assays have been unsuccessful in extracting K_i values due to a complex equilibrium between the components of the assay (discussed in section *3.2.3.3*). IC_{50} values could alternatively be calculated by plotting anisotropy values (y) against ligand concentration (x). The IC_{50} value (x_0) is extracted directly from the curve using **Equation 5**. If λ is not equal to 1 in the original protein titration, it may be appropriate for an EC_{50} value to be calculated instead. This involves fitting λ from the protein titration into **Equation 3**, with r and r_{min} anisotropy data fitted from the competition assay. EC_{50} values can then be calculated by plotting fraction of ligand bound (y) against ligand concentration (x) and fitting to a logistic model using Origin 8.6. The EC_{50} values (x_0) are extracted directly from the curve using **Equation 5**.

$$y = r_{min} + \frac{r_{max} - r_{min}}{1 + 10^{(x - \log x_0)}}$$

Equation 5 Rearrangement of equation allows for the calculation of; $IC_{50} = x_0$ when y = anisotropy and r_{min} and r_{max} are the minimum and maximum anisotropies $EC_{50} = x_0$ when y = fraction of ligand bound and r_{min} and r_{max} are the minimum and maximum ligand bound: r_{max} = [FL] (when $y = L_b$)

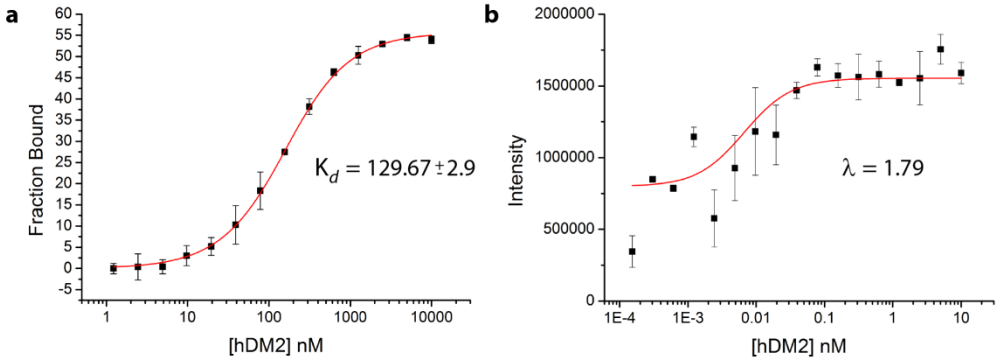
Appendix II

Assay implementation notes for p53/*h*DM2 and Mcl-1/NOXA B screen:

*h*DM2 Titration λ

In a previous protein titration experiment, λ was calculated to equal 1, however, new intensity data did not agree with this on first examination. In **Figure 3.4b**, one can see that the intensity is fairly constant at high [*h*DM2] but begins to fluctuate when at low [*h*DM2]. The average intensity was plotted against protein concentration and fitted to a logistic model to determine the intensity of fully bound and fully unbound tracer. λ was calculated to be 1.79, however, confidence in the fitting was particularly low due to scattering of the data at lower concentrations. The data from a sample of compounds was processed and the IC_{50} (y = 1) and EC_{50} (y = 1.79) values determined and variations in these values were negligible. It was decided to calculate IC_{50} values for future competition assays for the following reasons:

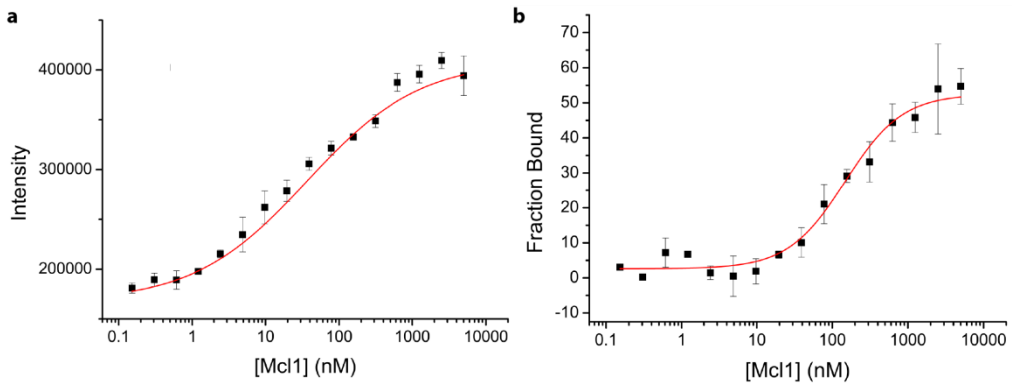
- Fluctuations in intensity readings presented uncertainty in the calculated λ
- When this λ value was used to calculate the EC_{50}, changes from the IC_{50} were negligible
- Previous studies had found λ = 1



(a) Protein titration curve of the p53/*h*DM2 interaction when λ = 1(K_d = 129.67 nM). **(b)** Average intensity fitted to a logistic model to calculate λ.

Mcl-1/NOXA B titration λ

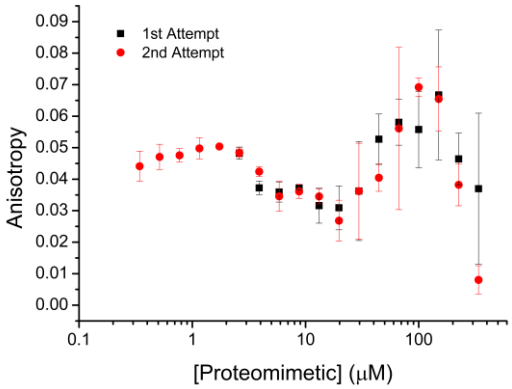
As observed previously λ ≠ 1, therefore the intensity was plotted against [protein] and fitted to a logistic model and was calculated as λ = 2.20



(**a**) Average intensity fitted to a logistic model to calculate λ (λ = 2.20). (**b**) Protein titration curve of the Mcl-1/NOXA B interaction (K_d = 148.19 ±55.97 nM) fitted to a logistic model when λ = 2.20.

Competition Assays p53/*h*DM2

- Due to poor inhibition of oligomer **25** (p53/*h*DM2 screen) the bottom plateau was not reached, however, an IC_{50} value could be obtained after fixing the bottom plateau to -0.02 which was the average r_{min} value.
- Comparison of data obtained from two competition assays for oligomer **25** against the p53/*h*DM2 interaction demonstrating a similar scattering pattern was reproducible and not due to plating error.



Competition Assays Mcl-1/NOXA B

- The fluorescence intensity changes during the protein titration. More specifically, the intensity increases with increasing [protein] indicating the environment of the protein affects the amount of fluorescence emitted by the fluorophore. This has been taken into account and EC_{50} values have been calculated by plotting fraction of ligand bound (inserting λ = 2.20) against [proteomimetic].