An Integrated Approach to the Discovery of Inhibitors of Protein-Protein Interactions

George McEwan Burslem

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

The University of Leeds
The School of Chemistry and The Astbury Centre for Structural Molecular Biology

April, 2015

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.

Portions of the work presented in chapters 2 and 4 were included in a research paper that was accepted for publication in February 2014 (G.M. Burslem, H.F. Kyle, A.L. Breeze, T.A. Edwards, A. Nelson, S.L. Warriner and A.J. Wilson, *ChemBioChem*, 2014, **15**, 1083). The contributions of the authors were as follows: G.M.B. (the candidate) performed all synthetic chemistry, peptide synthesis and modelling, H.F.K. performed all protein expression, G.M.B. and H.F.K. performed all assays jointly. G.M.B. prepared a draft of the manuscript which was then edited into its present form (attached) by all authors.

Portions of the work described chapter 2 were submitted as part of a research paper describing the requirements for HIF-1α/p300 binding (Exploration of the HIF-1α/p300 binding interface using peptide and adhiron phage display technologies to locate binding hot-spots for inhibitor development, H. F. Kyle, K. F. Wickson, J. Stott, G. M. Burslem, A. L. Breeze, D. C. Tomlinson, S. L. Warriner, A. Nelson, A. J. Wilson and T. A. Edwards, submitted). H.F.K. expressed proteins, carried out phage display experiments, performed NMR and crystallographic analysis. K.F.W. assisted with phage display experiments, J.S. performed bioinformatics analysis, G.M.B. (the candidate) prepared reagents. Binding studies were performed jointly by H.F.K. and G.M.B. A.L.B., D.C.T., S.L.W., A.N., A.J.W., and T.A.E. supervised the research. H.F.K. prepared a draft which was edited into its present form (attached) by all authors. Portions of the work performed by the candidate are included in this chapter.

Portions of the work described in chapter 3 were submitted as part of a research paper describing the synthesis of 1,3-amino alcohols (I. Colomer, O. Adeniji, G.M. Burslem, P. Craven, M.O. Rasmussen, A. Willaume, T. Kalliokoski, R. Foster, S. Marsden, A. Nelson, Bioorganic and Medicinal Chemistry, 2014, doi:10.1016/j.bmc.2015.01.058). G.M.B. and P.C. proposed the synthetic approach, I.C., O.A. and G.M.B. performed the initial synthesis validation, M.O.R and A.W. performed the final library synthesis, T.K. performed the chemoinformatics, R.F., S.M. and A.N. supervised the project. I.C. prepared a draft manuscript which was

then edited into its present from (attached) by A.N. with comments from all authors. Selected examples prepared by GMB are included in this thesis.

A review based on the synthesis of oligobenzamides as α-helix mimetics was published by the candidate in January 2014 (G.M. Burslem and A.J. Wilson, *Synlett*, 2014, **25**, 324). G.M.B. prepared a draft of the manuscript which was then edited into its present form (attached) by AJW. Portions of the original draft have been rewritten and included in chapter 4.

Acknowledgements

Firstly, I would like to thank my supervisors, Adam and Andy, for giving me the opportunity to come to Leeds and work on this project. I'm extremely grateful for all the support they've both given me as well as the level of autonomy. Thank you for all the opportunities, all the letters you've written on my behalf, letting me go to so many conferences and occasionally telling me to go home and sleep! I'm also very grateful to my co-supervisors, Stuart and Ed, for all their help, suggestions and tolerance.

I owe a huge debt of gratitude to Hannah Kyle, my partner in this project. Without her very little of this work would have been possible. I think we made a pretty good team despite the rumours, all the gin and that awful ITC room.

The work in this thesis would not have been possible without funding from both AstraZeneca and the EPSRC. Throughout the course of this project I've interacted with many people at AstraZeneca, all of whom I acknowledge gratefully. In particular, Ed Griffen and Kate Wickson for their enthusiasm and support, Paul Faulder and Martin Packer for their computational assistance and Alex Breeze for all his help and guidance.

I would also like to thank the other staff within the university who I've interacted with for all their questions, comments and assistance. Particularly, Martin Huscroft for all his help with HPLC (and for teaching me how to fix the LC-MS), Chris Empson for all his help with the plate reader and robot, Simon Barrett for keeping the NMRs running smoothly, Iain Manfield for assistance with ITC and Tanya Marinko-Covell for mass spectrometry. They should all be commended for their bravery in letting me play with their expensive equipment.

Over the time I've spent in Leeds, I'm had the pleasure of working with/alongside many people, within the Nelson/Wilson/Warriner/Edwards groups and beyond. There are too many to mention them by name but I hereby thank all of them. A few people deserve a special mention: Richard Doveston for all his help in and out of the lab (Smokestack?) and for helping me move (twice), Phil Craven for allowing me to borrow his chemical knowledge occasionally and his passionate use of the label maker and Donald "the Wonderhorse" Firth for taking me to gigs, reminding me to eat and being "actually very successful".

Finally, all my friends and family for putting up with my nonsense.

Abstract

Protein-protein interactions present challenging targets for therapeutic intervention with enormous potential for modulating biological pathways, particularly in the field of oncology. Two α-helix mediated protein-protein interactions of interest, hypoxia inducible factor-1α (HIF-1α)/p300 and eukaryotic initiation factor 4E (eIF4E) /eIF4G are introduced and their known inhibitors discussed in Chapter 1.

Initially, biophysical assays for both interactions were developed and the binding requirements between peptides derived from helix donor components (HIF-1α and eIF4G) to their protein counterparts (p300 and eIF4E respectively) were investigated. This information was used to develop competition assays capable of identifying inhibitors and provided important insight for the rational design of inhibitors.

Subsequently, a computational approach, described in Chapter 3, to inhibitor discovery was applied to both targets, using both docking and pharmacophore modelling. Several series of compounds were purchased or prepared and screened as inhibitors. The development of a synthetic route to a novel scaffold is described providing a weak small molecule inhibitor.

In parallel, a proteomimetic approach to inhibitor design was employed, using sequence based rational design, drawing on the knowledge gained in Chapter 2. By mimicking a key helical region of HIF-1α, the interaction with p300 can be disrupted, as discussed in Chapter 4. Additionally, new methods for the preparation of oligobenzamide helix mimetics were investigated allowing the preparation of challenging targets, late stage functionalization and the preparation of oligobenzamide/peptide hybrids.

Overall, this thesis provides an introduction to two therapeutically relevant interactions, provides biophysical assays for the identification of inhibitors and discloses the first biophysically characterised inhibitors of the HIF-1α/p300 interaction.

Table of Contents

List of Figures

List of Schemes

List of Tables

List of Abbreviations

4E-BP	4E-Binding Protein
AIB	Amino *iso*-butyric Acid
ARNT	Aryl hydrocarbon receptor nuclear transporter
bHLH	Basic loop-helix-loop
Boc	t-Butyl Carbamate
CBP	cAMP response element-binding protein (CREB) binding protein
CH	Cysteine/histidine-rich
CITED	CREB-binding protein/p300-interacting transactivator with ED-rich tail
CTAD	C-terminal transactivation domain
DIPEA	Diisopropylethylamine
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
eIF4E	Eukaryotic initiation factor 4E
eIF4G	Eukaryotic initiation factor 4G
FIH	Factor inhibiting HIF
FITC	Fluorescein isothiocyanate
Fmoc	9-Fluorenylmethyl
HAT	Histone acetyl transferase
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HBS	Hydrogen bond surrogate
HCTU	N,N,N′,N′-Tetramethyl-O-(6-chloro-1H-benzotriazol-1-yl)uronium hexafluorophosphate
hDM2	Human double minute 2

HIF	Hypoxia inducible factor
HPLC	High pressure liquid chromatography
HRE	Hypoxic response element
HRMS	High resolution mass spectrometry
HSQC	Heteronuclear single quantum correlation
IC_{50}	Half maximal inhibitory concentration
IRES	Internal ribosome entry site
ITC	Isothermal titration calorimetry
K_d	Dissociation constant
K_i	Inhibition constant
LC-MS	Liquid chromatography-mass spectrometry
LC-MS/MS	Liquid chromatography- mass spectrometry/mass spectrometry
m7G	7-Methylguanylate
MOM	Methoxymethyl
mRNA	Messenger ribonucleic acid
mTOR	Mammalian target of rapamycin
NMR	Nuclear magnetic resonance
ODDD	Oxygen dependant degradation domain
PAINS	Pan Assay Interference Compounds
PAS	Per-ARNT-AHR-Sim
PEG	Polyethylene glycol
PHD	Prolylhydroxylase domain
PPI	Protein-Protein Interaction
*p*TSA	*Para*-toluene sulfonic acid
pVHL	Von-Hippel Lindau protein
PyBOP	Benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate
ROCS	Rapid overlay of chemical shape
SAR	Structure activity relationship

SCX	Strong Cation Exchange
SPPS	Solid phase peptide synthesis
SUMO	Small ubiquitin-like modifier
TFA	Trifluoroacetic acid
TLC	Thin layer chromatography
VEGF	Vascular endothelial growth factor

Amino Acid Abbreviations

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

Chapter 1 – Protein-Protein Interactions: HIF-1α/p300 and eIF4E/eIF4G

The candidate confirms that the work submitted is his 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.

1.1 Protein-Protein Interactions

Historically, drug discovery has focussed on the modulation of enzymes to elicit a pharmacological response. This approach has many advantages such as a potentially well-defined binding pocket or an endogenous substrate to use as a starting point for small molecule design.[1] However protein-protein interactions (PPIs) are implicated in virtually all biological pathways[2] and are emerging as new targets for therapeutic intervention. The inhibition of protein-protein interactions represents an unaddressed challenge with significant potential in chemical biology and drug discovery.

The challenges involved in inhibiting PPIs are based around the differences from the approach used to inhibit enzymatic targets. A protein-protein interface is often seen as two large flat featureless surfaces interacting but this is seldom the case;[3] there may be key "hotspots",[4] defined or transitory binding residues or regions, which provide the majority of the affinity.[5]

1.1.1 α-Helix Mediated Protein-Protein Interactions

Of particular interest are α-helix-mediated PPIs where a key helical region of a protein, the helix donor, binds into a cleft on a partner protein, the helix acceptor (Figure 1).[6] The helix acceptor complements the well-defined structure of an α-helix and can therefore be inhibited by other small peptides with similar conformations and compatible side chains or small molecules that bind strongly into the helix acceptor cleft.
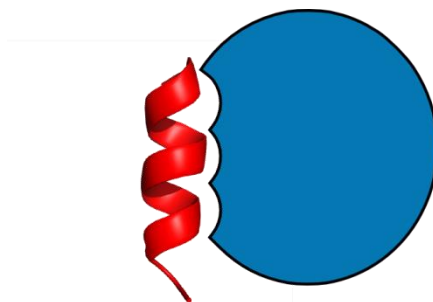
Figure 1 - Schematic representation of an α-helix mediated protein-protein interaction. The helix donor is shown in red whilst the helix acceptor is shown in blue.

It has been estimated the human interactome contains over 650,000 pairwise PPIs[7] offering many opportunities for signalling pathway modulation. Analysis of the protein structure database shows that 62% of all multi-protein complexes feature an α-helix at the interface between the proteins. Furthermore, computational alanine scanning estimates around 60% of these helix-mediated PPIs feature a proposed hotspot contained within the helical region.[8]

Due to the well conserved structure of helical protein domains, significant efforts have been focussed on the mimicry of helices to develop inhibitors.[9] These efforts have led to a wealth of technologies for inhibition of helix mediated PPIs (see Chapter 5) which have afforded potent inhibitors. This success, and the success of more traditional drug discovery approaches toward this subset of PPIs, has led to interactions involving an α-helix being considered more tractable.[9]

1.1.2 Helix-Mediated PPIs in Cancer

A wide range of helix mediated PPIs has been implicated in cancer progression.[10] The prototypical helix-mediated PPI is the interaction between p53 and hDM2 which has been widely studied allowing several potent inhibitor compounds to be discovered.[11] Many of the PPIs studied for cancer therapy are involved in the regulation of apoptosis,[12] however there exists a multitude of other PPIs relevant to cancer. Discussed below are two therapeutically-relevant α-helix-mediated PPIs which are the focus of this thesis.

1.2 Hypoxic Sensing

Homeostasis is key for health and the body has evolved a large number of sensing systems and feedback loops to allow tissue to respond to changes and return to the "normal" state. Arguably the most important factor for survival is oxygen level. Sensing hypoxia is crucial for healthy tissue growth as when growing tissue exhausts the supply of oxygenated blood, there must be a pathway through which the growth of additional vasculature is stimulated.

Hypoxia Inducible Factor-1α (HIF-1α) plays a major role in oxygen sensing within tissues. HIF-1α is constitutively expressed in cells but rapidly degraded or deactivated via oxygen-dependent mechanisms under normoxic conditions. This degradation of HIF-1α is so efficient and rapid that there is effectively no HIF-1α in cells under normoxic conditions.[13]

When HIF-1α accumulates, it translocates to the nucleus and forms a heterodimer with HIF-1β. It also forms a PPI with the promiscuous transcriptional co-factor p300 (Figure 2).[14, 15] p300 has a homolog known as CREB binding protein (CBP) which is considered largely functionally and structurally redundant.[15] The HIF-1 dimer/p300 complex binds to hypoxic response elements (HRE) on DNA and causes a plethora of downstream events via transcription mediation.[16] Perhaps most crucially, it induces expression of vascular endothelial growth factor (VEGF) which stimulates angiogenesis, thus resupplying the hypoxic region with oxygen-rich blood and relieving the hypoxic state. Another effect downstream of HRE binding is regulation

of oxygen usage, switching the cell from aerobic to anaerobic respiration by expression of glycolytic enzymes.[17]
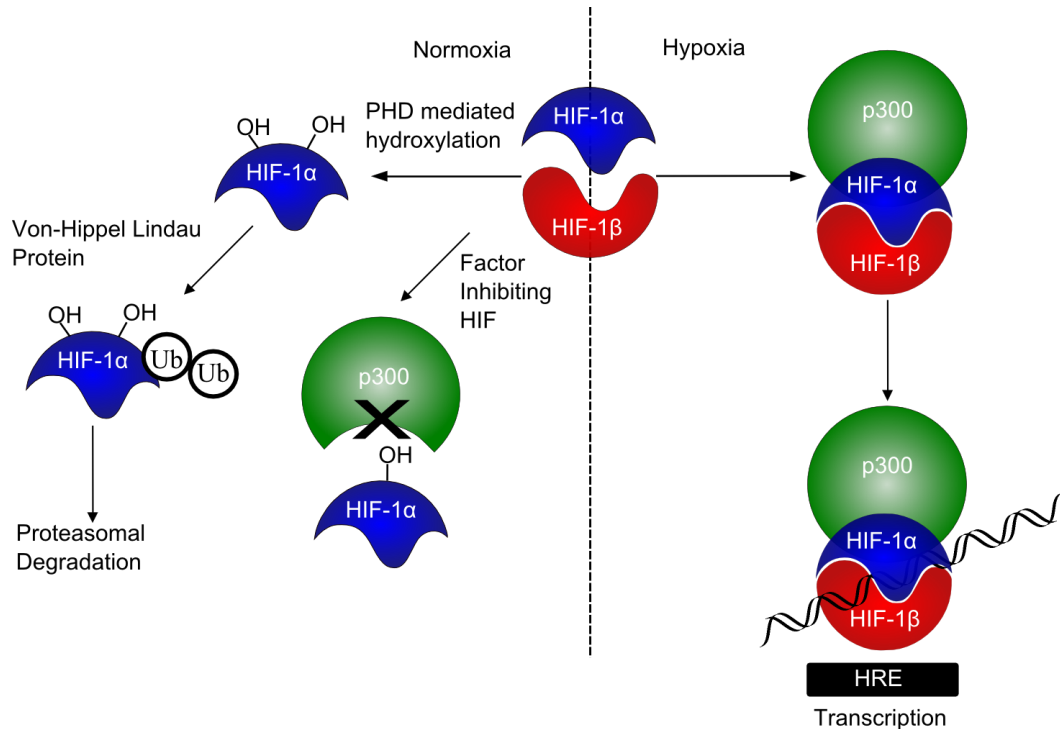
Figure 2 - HIF pathways under normoxic and hypoxic conditions. HIF-1α - Hypoxia Inducible Factor 1α, HIF-1β - Hypoxia Inducible Factor 1β, PHD – prolyl hydroxylase domain, HRE – hypoxic response element.

However, the hypoxic response pathway can be exploited by cancers. Rapidly dividing cells, forming solid tumours, will rapidly deplete the local oxygen supply and become hypoxic. The hypoxic response cascade will then resupply the tissue with oxygen via angiogenesis. The fact that the hypoxic response facilitates the growth and survival of solid tumours makes it a potential target for therapeutic intervention.[18, 19]

1.2.1 HIF Regulation

The major pathway through which HIF-1α is degraded is via prolyl hydroxylase domain (PHD) enzymes.[13, 20] PHD enzymes function via the reaction between 2-oxoglutarate and oxygen, mediated via non-haem iron, to hydroxylate two proline residues (Pro_{402} and Pro_{564}) in the oxygen dependent domains of HIF-1α (Figure 3). This PHD mediated hydroxylation of HIF-1α stimulates an interaction with the von-Hippel Lindau protein, an E3 ligase which ubiquitinates HIF-1α, targeting it for proteasomal degradation (Figure 2).[13]

Another pathway through which HIF-1α is regulated is also via oxygen-dependent hydroxylation by a protein known as Factor Inhibiting HIF (FIH).[21] Hydroxylation by FIH of an asparagine residue (Asn_{803}) on the *C*-Terminal Transactivation Domain

(CTAD) of HIF-1α prevents its association with p300.[22] Through this mechanism, FIH acts as a second oxygen sensor. Under hypoxic conditions, neither the PHDs nor FIH are active and as such the level of HIF-1α increases. Through these oxygen sensing enzymes, cells regulate their metabolism as discussed above and respond to hypoxia.

1.2.2 Structural Biology of HIF-1α/p300

As previously discussed, HIF-1 is a heterodimer of HIF-1α and HIF-1β (sometimes known as aryl hydrocarbon receptor nuclear translocator, ARNT). There are also 2 isoforms of HIF-1α, known as HIF-2α and HIF-3α. Whilst both can form dimers with HIF-1β, HIF-3α lacks the ability to bind the co-activator protein and thus is inactive.[19] When HIF-2α and HIF-3α form complexes with HIF-1β they are known as HIF-2 and HIF-3 respectively and have been reported to be expressed in different amounts in different tissues.[23]
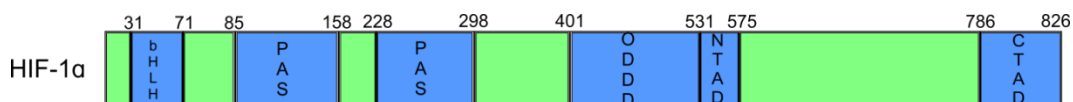
Figure 3 - Domains of HIF-1α. bHLH – basic helix-loop-helix, PAS – per ARNT-AHR-Sim, ODDD - oxygen-dependent degradation domain, NTAD – *N*-terminal transactivation domain, CTAD – *C*-terminal transactivation domain.

The dimerization occurs through a basic helix-loop-helix (bHLH) domain and 2 per-ARNT-AHR-Sim (PAS) domains on both the HIF-1α subunit and the HIF-1β subunit.[24] PAS domains are implicated in protein-protein interactions in other systems.[25] It is also thought that coiled coil co-activators play a role in the dimerization.[26] HIF-1α has several other key domains, as shown in Figure 3, the oxygen-dependent degradation domain (ODDD) which contains the proline residues which are hydroxylated under normoxia and both the *N*-terminal and *C*-terminal transactivation domains.[27] The *C*-terminal transactivation domain (CTAD) is the region that interacts with p300.

p300 is a large multidomain protein that acts as a transcriptional co-activator for many transcription factors. It also has epigenetic modulation functions via its bromodomain and HAT domains.[15] p300 has three cysteine/histidine-rich (CH) domains, CH1, CH2 and CH3, which all bind zinc. Transcription factors bind at the CH domains and this binding is usually specific to one of the CH domains. HIF-1α binds to the CH1 (sometimes known as transcriptional adapter zinc-binding (TAZ)1 domain) domain of p300.[28]

Structures of the HIF-1α CTAD complexed with p300 was determined by two groups simultaneously by solution NMR studies.[29, 30] To date, no X-ray crystal structure of the HIF-1α/p300 complex has been reported. The p300 CH1 domain

forms a rigid structure consisting of 4 helices constrained and stabilised by binding 3 zinc atoms whilst the HIF-1α CTAD is thought to be unstructured in the absence of p300. Upon binding to the CH-1 domain, HIF-1α forms three helical regions which appear to mediate the binding between the two proteins (Figure 4). The CH-1 domain of p300 is also the binding site for the CREB-binding protein/p300-interacting transactivator with ED-rich tail (CITED) family of proteins which can compete with HIF-1α.[31, 32]
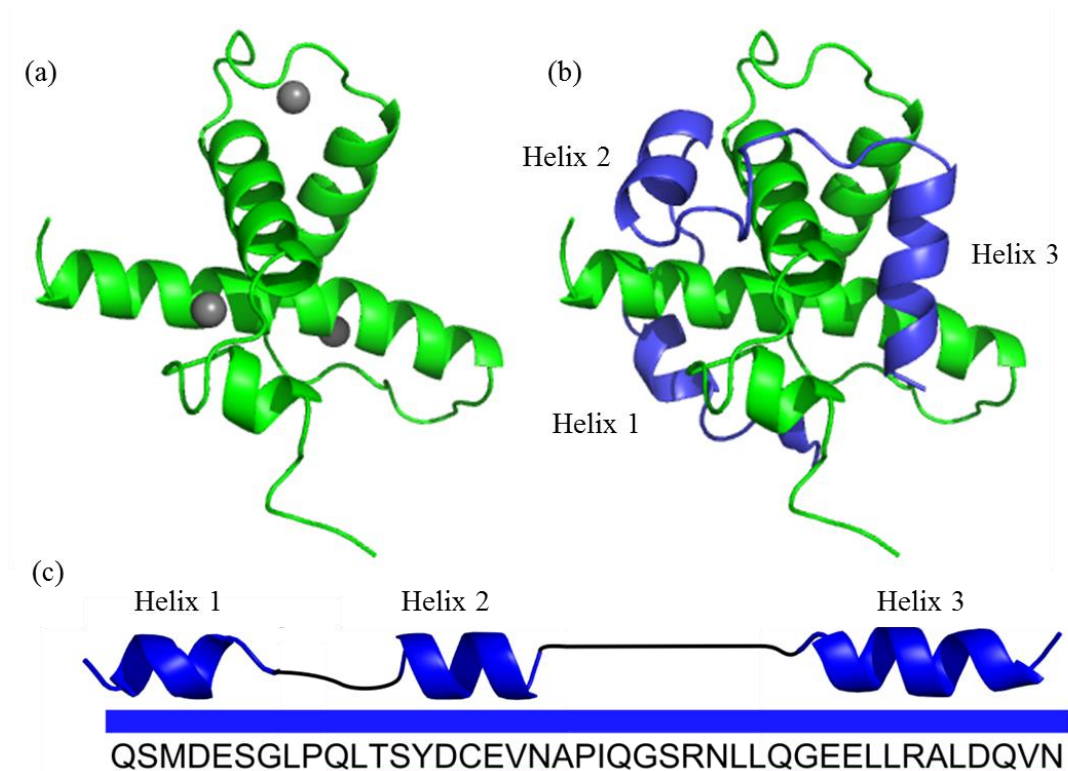
Figure 4 - NMR structures of (a) p300 CH1 domain (green) in the complex (peptide not shown) (b) Complex between HIF-1α CTAD (blue) and p300 CH1 domain (green) (zinc omitted for clarity) PDB ID: 1L8C (c) primary sequence of HIF-1α CTAD with helical regions indicated

1.2.3 HIF-1α/p300 Hotspots

Functional analysis of the binding between HIF-1 and p300 has identified several hot spot residues which are focussed on two of the helical regions, helix 2 and helix 3.[33] Throughout this thesis, the helices in the HIF-1α CTAD are numbered sequentially from *N*- to *C*-terminus as shown in Figure 4.

Helix 2 of the HIF-1α CTAD features two reported hot-spot residues. The first is the asparagine previously discussed as the target for FIH. Hydroxylation of Asn_{803} prevents a key intermolecular hydrogen bond and introduces a steric clash.[29] Also considered a hotspot residue, is Cys_{800} on helix 2.[34] A C800A mutation has been reported to reduce binding to p300.[35] It has been suggested that post-translational nitrosylation of Cys_{800} can regulate the interaction between HIF-1α and p300 but

there is conflicting evidence as to whether it increases[36] or decreases[35] binding. Both of these residues are on the same face of helix 2, binding into the surface of p300 (Figure 5).
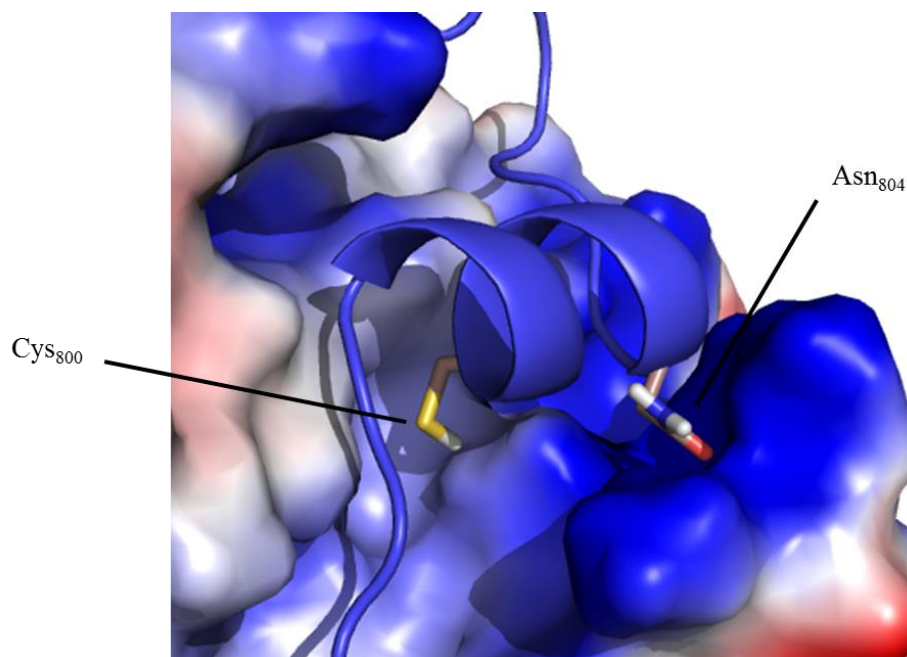
Figure 5 - Helix 2 binding site with reported hotspot residues highlighted

Truncations of the CTAD to remove helix 3 have been reported to prevent binding between HIF-1α and p300.[33] Further analysis by mutagenesis and alanine scanning has identified several important hydrophobic residue on helix 3.[34] Again, these hotspot residues align along one face of the helix in the *i, i*+4 and *i*+7 positions as shown in Figure 6.
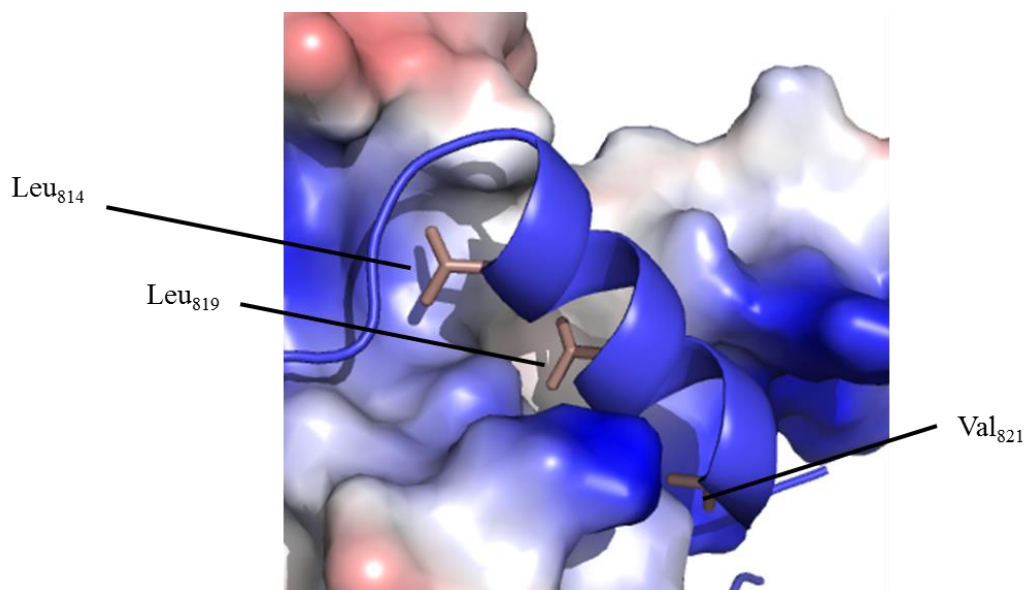
Figure 6 - Helix 3 binding site with reported hotspot residues highlighted

Whilst a large amount of work has been done to explore the requirements for binding of HIF-1α to p300 it is still not clear which of the hotspots provides the primary driving force for the interaction. Further work is required to allow targeted design of inhibitors.

1.2.4 Inhibitors of the Hypoxic Response Pathway

There are several potential points of intervention along the hypoxic response pathways, described in Figure 2, at which it has been modulated by various compounds.[19] For example, both inhibitors[37] and activators[38] of the PHD enzymes have been reported. The Crews lab has published inhibitors of the HIF-1α/pVHL interaction.[39, 40] Significantly these modulators focus on the normoxic response pathways, however under hypoxic conditions, such as in solid tumours, these pathways are not active due to their intrinsic dependence on oxygen.

Inhibitors active under hypoxic conditions must therefore focus on the interactions between proteins or between proteins and hypoxic response elements. Several families of compounds have been identified which inhibit the interaction between HIF-1α and HIF-1β. Acriflavine (Figure 7) was identified as an inhibitor of dimerization in a screen of compounds that had previously entered phase II clinical trials.[41] The Tavassoli group has used genetically encoded libraries to identify cyclic peptide inhibitors of HIF dimerization such as cyclo-CLLFVY (Figure 7).[42] Interestingly, a selective small molecule inhibitor of HIF-2 heterodimer formation has been identified that does not affect HIF-1 formation.[43]
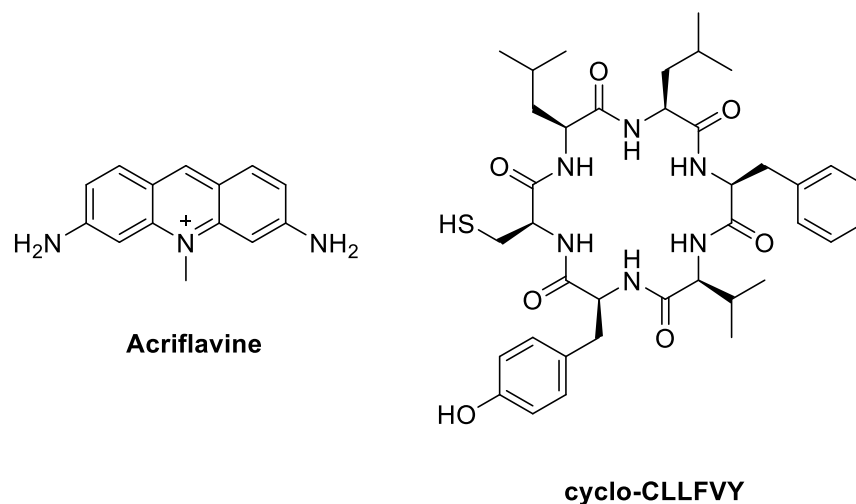
Figure 7 - Structures of HIF-1 dimerization inhibitors

Another approach to mediating the hypoxic response is to prevent transcription of hypoxic response elements. Dervan and co-workers have achieved this using rationally designed polyamides which selectively bind the HRE for VEGF.[44] Echinomycin has also been shown to prevent HIF-1α DNA interactions.[45] Inhibitors

of HIF-1β binding to coiled-coil co-activators have also been shown to decrease HIF mediated transcription under hypoxic conditions.[46]

Perhaps the main target for inhibition of the hypoxic response is the HIF-1α/p300 interaction although it has been called " a considerable challenge"[29] based on the area over which the interaction occurs. High-throughput screening has identified several compounds capable of eliciting a cellular response consistent with HIF-1α/p300 inhibition. A natural product, Chetomin (Figure 8), was identified as an inhibitor and shown to be effective at down-regulating HIF-controlled transcription as well as reducing tumour size in mouse models.[47] The epidithiodiketopiperazine (ETP) functionality contained in Chetomin has been shown to disrupt the folding of the CH-1 domain of p300, via zinc ejection, preventing interaction with HIF-1α.[48] Indeed, much simpler ETP containing compounds have been shown to be sufficient for activity.[48, 49] Additional families of compounds capable of disrupting the folding of the p300 CH1 domain by zinc ejection have been identified, including ninhydrin (Figure 8).[50] Compounds with zinc ejection-based mechanisms are likely to encounter issues with selectivity and toxicity, due to interactions with other zinc-binding proteins and the fact that p300 has multiple binding partners.[49]

Chetomin

Ninhydrin
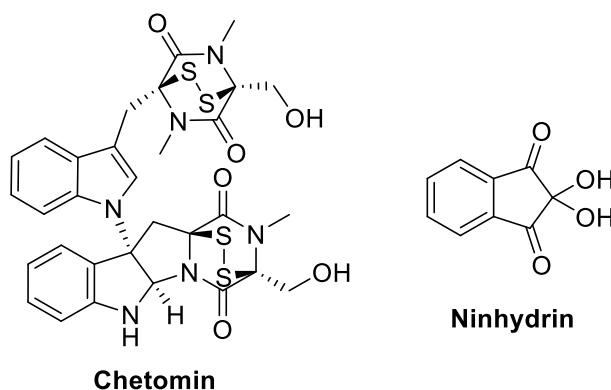
Figure 8 - Structures of compounds capable of ejecting zinc from p300

A small molecule called KCN-1 (Figure 9), reported to inhibit the HIF-1α/p300 interaction, was identified through high-throughput cell based screening of a combinatorial library.[51] KCN-1 has been shown to prevent HIF-regulated expression and reduce tumour size in animal models but the exact mode of action still remains unclear[52, 53] (See Chapter 5 for further discussion).
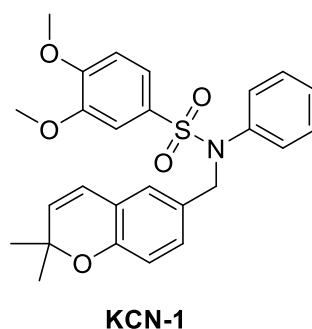
KCN-1

Figure 9 - Structure of KCN-1

The Arora group has employed hydrogen-bond surrogate (HBS) peptides to inhibit the HIF-1α/p300 interaction. HBS peptides constrain the conformation of a helix by replacing the *N*-terminal *i* to *i*+4 hydrogen-bond with a covalent link (Figure 10) thus inducing helicity. The Arora group have reported HBS peptides for both helix 2[54] and helix 3[55] regions of the HIF-1α CTAD.
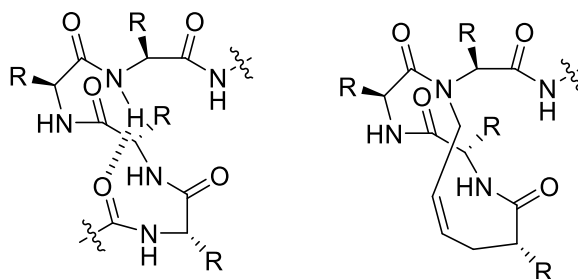
Figure 10 - Structures of *i* to *i*+4 hydrogen bond and hydrogen bond surrogate.

The work presented in Chapter 5 of this thesis describes the development of the first small molecule inhibitors **1** of HIF-1α/p300 to be characterised biophysically.[53] Subsequently, the Arora group has also published small molecule inhibitors, such as **2**, of the HIF-1α/p300 PPI based on a similar proteomimetic approach.[56, 57] Although Arora's compounds show high affinity binding to p300 and reduce HIF-regulated transcription, they have not been explicitly shown to disrupt the interaction between HIF-1 and p300.
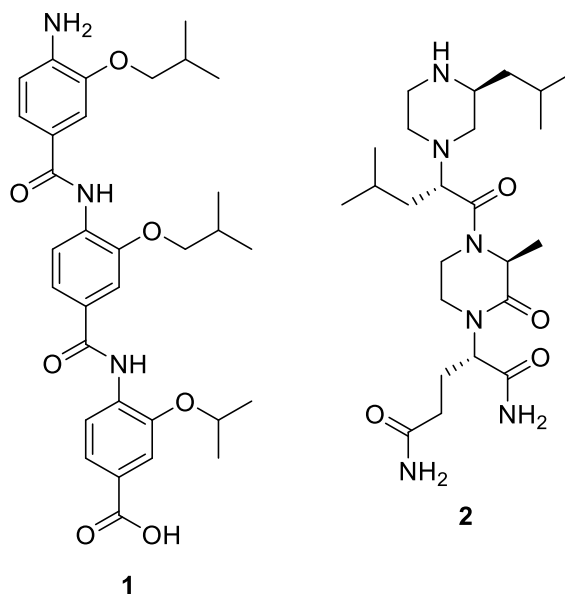
Figure 11 - Structures of proteomimetic inhibitors of HIF-1α/p300

There is still a need for potent, selective inhibitors of the interaction between the HIF-1α CTAD and the CH1 domain of p300. The inhibitors discussed above provide some evidence that the interaction is tractable and a valid target, however none of them are understood well enough or have the correct properties to enter the clinic.

1.3 Control of Translation

As discussed above, cancers can exploit cellular pathways to control protein transcription but other systems for protein level control exists, such as those controlling HIF-1α levels. Once genes are transcribed from DNA to mRNA, they must be translated into proteins. In eukaryotes, mRNA must be recruited to the ribosome for translation to occur. One way in which this occurs is the cap-dependent recruitment mediated by the eukaryotic initiation factor (eIF)-4 complex outlined in Figure 12.
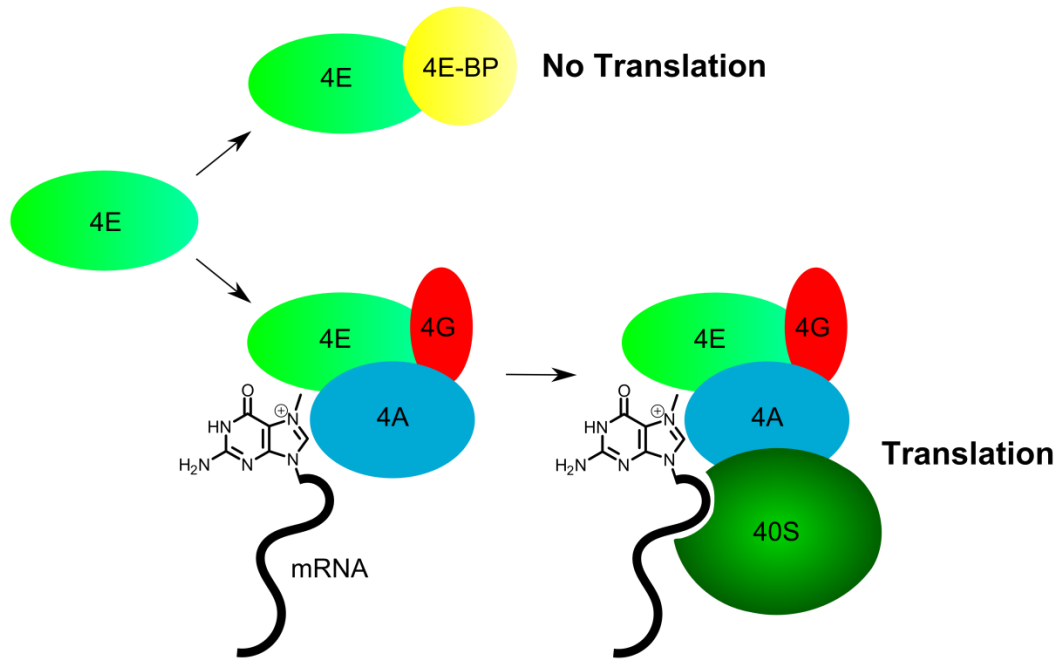
Figure 12 - Schematic representation of the eIF4 recruitment of 5'Cap-mRNA to the ribosome

The 5′-end of certain mRNAs is capped with 7-methylguanylate (m7G), the structure of which is shown in Figure 12. mRNA bearing this modification can bind to eIF4E, a subunit of the eIF4 complex which in turn forms a protein-protein interaction with the *N*-terminal domain of eIF4G.[58, 59] eIF4G is a large scaffolding protein with also binds the eIF4a subunit, which is thought to unwind mRNA. The formation of this eIF4/mRNA complex facilitates the recruitment of 5′-cap mRNA to the 40S subunit of the ribosome, initiating translation. An alternative form of cap-independent translation can occur if the mRNA features a structure known as an internal ribosome entry site (IRES) which facilitates direct recruitment to the 40S subunit.[60, 61]

eIF4E recruitment to the eIF4 complex can be inhibited by binding to a protein called 4E binding partner (4E-BP) (Figure 12). 4E-BP binds to the same site on eIF4E as eIF4G via a conserved $YX_4L\phi$ amino acid sequence (where X = any residue and ϕ = a hydrophobic residue.[62] The $YX_4L\phi$ is the same recognition element employed by eIF4G.[63] 4E-BP Phosphorylation state is crucial for the regulation of 4E-BP binding to eIF4E allowing cellular regulation of translation. When phosphorylated, 4E-BP binds with significantly lower affinity than when non-phosphorylated thus the phosphorylation event releases eIF4E to interact with eIF4G.[64] It has also been shown that 4E-BP binds to eIF4E at a distal non-competitive site which induces dissociation of eIF4G.[65]

Due to sequestration by 4E-BP and low levels of endogenous expression, eIF4E is the limiting partner for the formation of the eIF4 complex and as such regulates

which mRNAs are translated into proteins. This regulation occurs through the competition of mRNA for eIF4E; well-defined 5′ end RNA sequences are rapidly translated whilst ill-defined 5′ end RNA sequences are translated slower. This leads to regulation of the proteins expressed by cells.[66] Some key genes implicated in cancer cell survival are regulated at this level and in cancer cells the level of eIF4E is increased and thus the proteins involved in cell growth and inhibition of cell apoptosis are expressed at higher levels.[67] The mRNA strands for these oncogenes have ill-defined 5′ ends so are usually out-competed by other mRNA strands. If the eIF4E protein is overexpressed, this regulation is no longer efficient and cancer cells can survive and proliferate. However, if the interaction between the eIF4E and eIF4G proteins is inhibited, the eIF4 complex cannot form and the proteins essential for cancer cell survival cannot be produced. Thus the interaction between eIF4E and eIF4G is also a potential target for the treatment of cancer.[67]

1.3.1 Structural Biology of eIF4E/eIF4G

The eIF4E protein has been crystallised in the presence of a peptide derived from eIF4G and the m7G cap structure.[68] The X-ray crystal structure reveals that eIF4E consists of a large β-sheet and 3 long helices (Figure 13).[69] The eIF4G derived peptide, acting as a helix donor, binds on one face of the eIF4E protein whilst the m7G cap binds in a hydrophobic pocket at a distal site (Figure 13). It is unclear whether there is any cooperativity between cap and peptide binding (see Chapter 2).[70, 71] For both eIF4G and 4E-BP, the $YX_4L\phi$ eIF4E recognition domain is thought to be unstructured in the absence of eIF4E.[62]
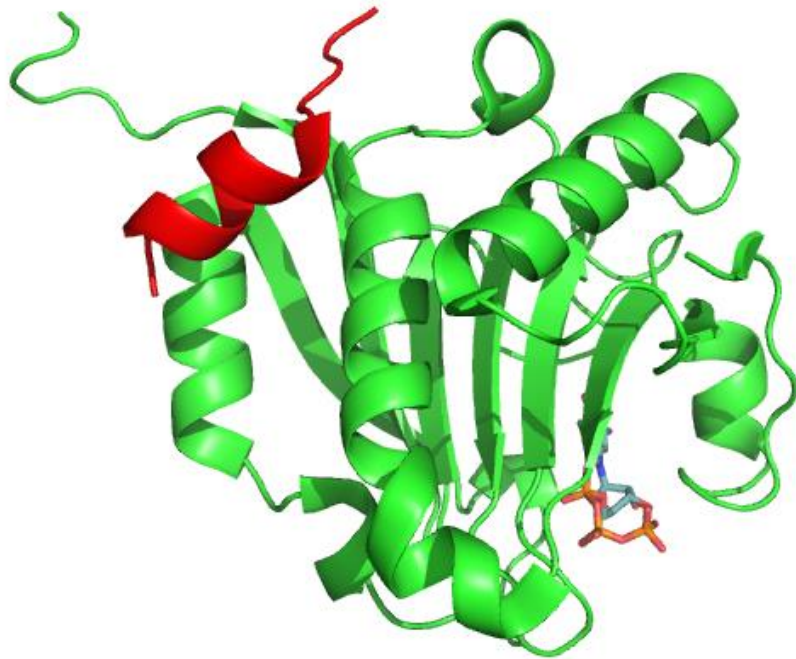
Figure 13 - Structure of eIF4E (green) bound to a peptide derived from eIF4G (red) and m7G Cap (structure coloured by element) (PDB ID: 2W97)

1.3.2 eIF4E/eIF4G Hotspots

Alanine scanning has been used to identified hot spot residues for the eIF4E/eIFG4 interaction which match the conserved sequences (Figure 14).[63] The tyrosine residue at position 416 is essential, mutation of this residue completely abrogates binding.[72] The leucine residues at positions 421 and 422 are also important but the non-conserved Leu_{422} can be exchanged for another hydrophobic residue.
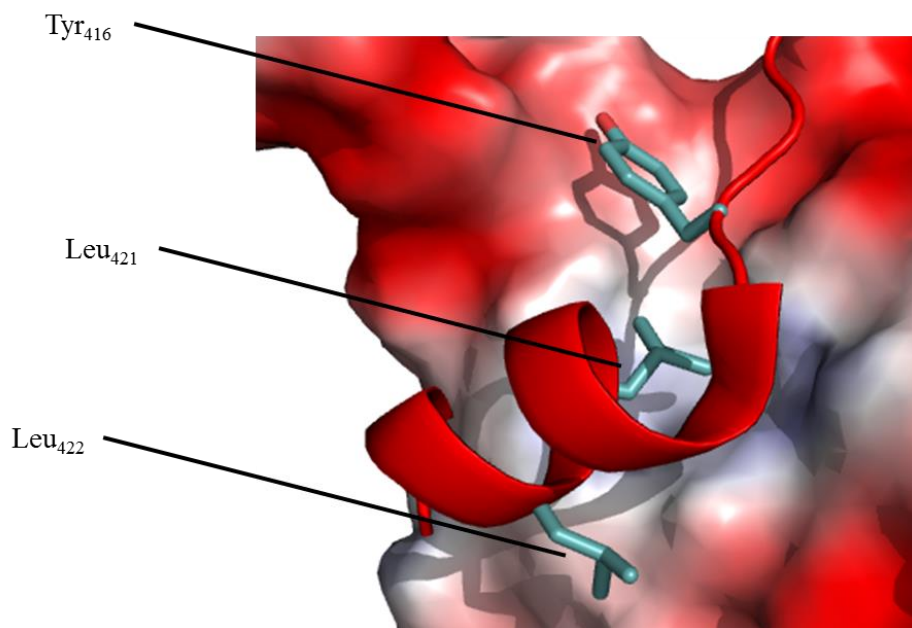
Figure 14 - Hotspot residues for the eIF4E/eIF4G interaction

1.3.3 Inhibitors of Cap-Mediated Translation

Several potential targets exist for the inhibition of cap-dependent translation and these have been exploited in several ways. For example, one of the kinases responsible for 4E-BP phosphorylation is mammalian target of rapamycin (mTOR). Inhibitors of mTOR prevent the phosphorylation of 4E-BP thus decreasing the amount of eIF4E available for eIF4 complex formation, preventing translation.[73] Many examples of mTOR inhibitors[74] such as AZD8055[75] and rapamycin[73] (Figure 15) itself have been shown to prevent 4E-BP phosphorylation and decrease cap-dependent translation.
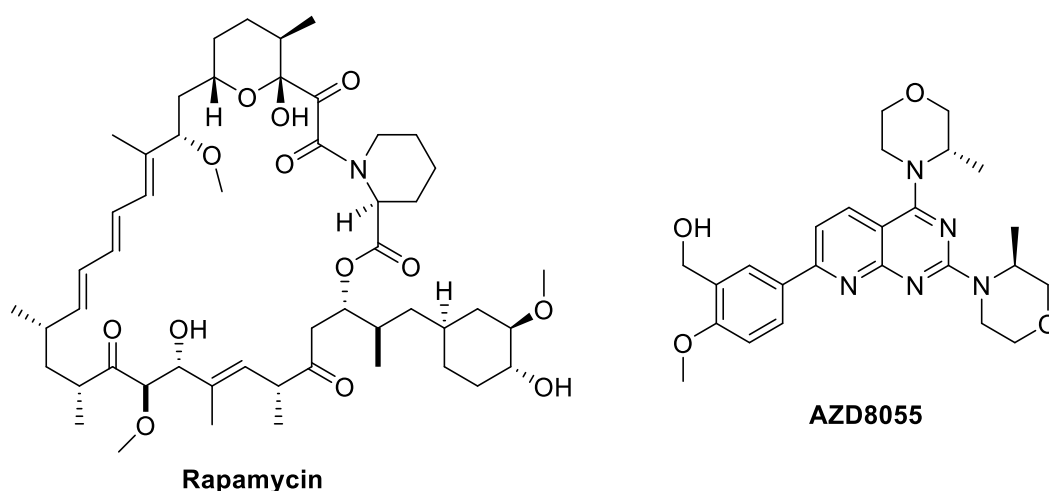
Figure 15 - Structures of mTOR inhibitors

Another potential target for the inhibition of cap dependent translation is the binding of m7G to eIF4E. Ribavirin is currently used as an anti-viral, but has been shown to bind to eIF4E and compete for 5′-cap mRNA (Figure 16).[76] It has also been shown to significantly reduce tumour size in murine models. Ribavirin mimics guanine ribonucleosides which is responsible for both its antiviral and anti-cancer activity. Figure 16 shows how Ribavirin is capable of mimicking m7G.
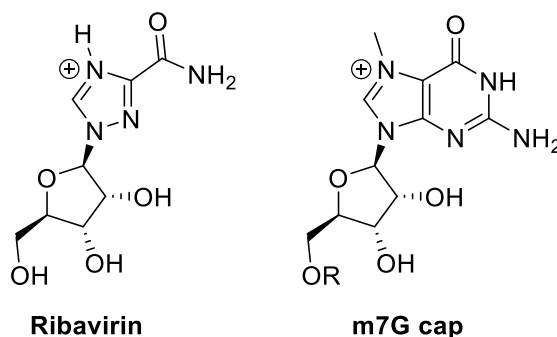
Figure 16 - Structures of compounds that bind in the eIF4E cap binding region. Ribavirin shown protonated as it is thought to be protonated at physiological pH.

Considerable effort has been expended in the development of peptide based inhibitors of the eIF4E/eIF4G interaction. Since the requirements for peptide binding is well understood, as discussed above, various modifications to the peptide sequence have been made in positions where they will be tolerated. Short peptide sequences containing the YX$_4$Lϕ sequence conjugated to cell penetrating peptides have been shown to induce apoptosis.[77] Increasing the helicity of eIF4G derived peptides by incorporating helix inducing amino *iso*-butyric acid (AIB) residues[72] or peptide stapling[78] has also been shown to increase affinity for eIF4E (Figure 17). A phage display approach was also used to identify peptides with an increased affinity for eIF4E (Figure 17).[79]

(a) Stapled peptide (b) AIB (c) Pentenyl Alanine

(d) Sequence
WT: KKRYDREFLLGF
AIB: KKRYDREFLLXF
Stapled: KKRYSRE*LL*F
Phage Display: KKRYSRDQLVAL
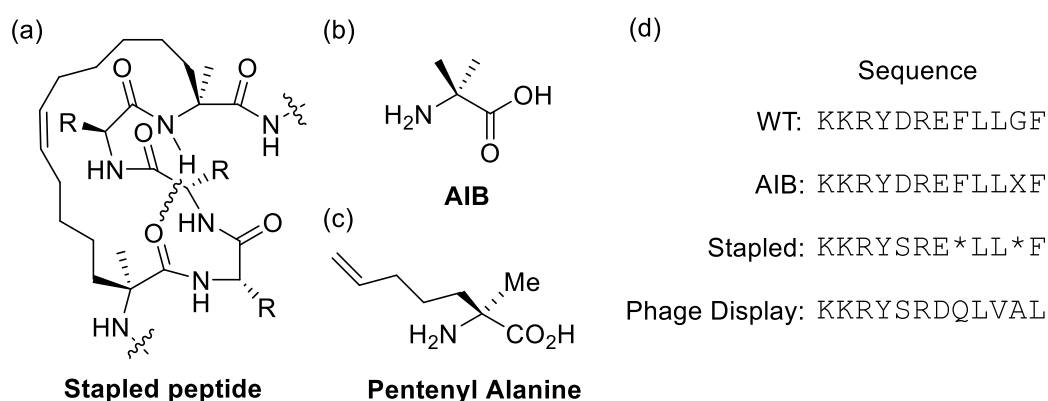
Figure 17 - Higher affinity eIF4G peptides (a) generic stapled peptide structure (b) structure of AIB (c) the structure of the amino acid incorporated for peptide stapling via metathesis (d) structures of eIF4G peptides with increased affinity. X represents AIB incorporation, * represents the stapled positions.

Finally, several groups have identified small molecules capable of inhibiting the interaction between eIF4E and eIF4G. High throughput screening (HTS) was used to identify eIF4E/eIF4G inhibitor 1 (4EGI-1) (Figure 18) which is capable of inducing the dissociation of eIF4G from eIF4E.[80] Initially, it was thought that 4EGI-1 was an orthosteric inhibitor but it was subsequently shown to act through an allosteric mechanism by X-ray crystallography.[71] Attempts have been made to optimise 4EGI-1 for affinity by increasing its rigidity.[81]

4EGI-1   4E1R-Cat
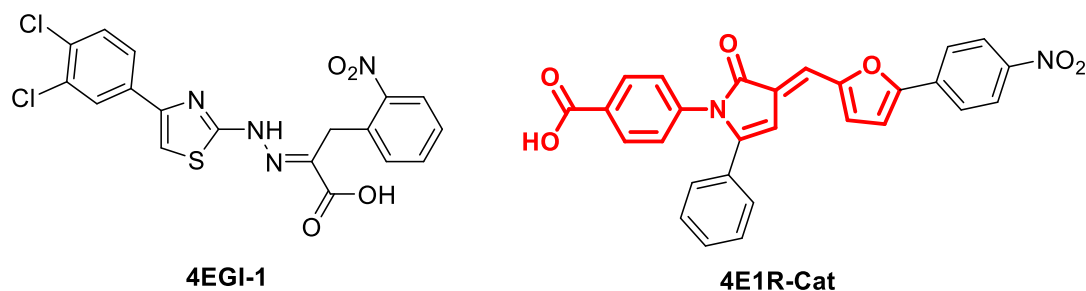
Figure 18 - Structure of small molecules inhibitors of eIF4E/eIF4G. PAINS substructure highlighted in red.

A second small molecule, 4E1R-Cat, also identified by HTS, was shown to inhibit the binding of both eIF4G and 4E-BP to eIF4E.[82] However, the central core (shown in red in Figure 18) of 4E1R-Cat has been identified as a pan assay interference[83] (PAINS) compound and 4EIR-Cat or close analogues have been identified in a number of screens against different targets.[84] Geldanamycin has also been reported to weakly inhibit the interaction between eIF4E and eIF4G.[85]

As with the HIF-1α/p300 interaction, the inhibitors discussed above provide confidence that the eIF4E/eIF4G is a relevant and tractable target for the modulation of α-helix mediated protein-protein interactions but also that further study is required to develop potent inhibitors.

1.4 Summary

This chapter briefly introduced PPIs and their role in cancer then describes two potential targets for therapeutic intervention in the field of oncology: HIF-1α/p300 which is activated as a response to a lack of oxygen and controls transcription and eIF4E/eIF4G which is activated by an overexpression of the eIF4E protein and controls translation. Initial success in the field has shown that these targets are potentially tractable and may lead to new therapeutics. The rest of this thesis will focus on enhancing the understanding of the interactions and the development of inhibitors of both the eIF4E/eIF4G and HIF-1α/p300 interactions.

1.5 Project Aims

This chapter has described two therapeutically relevant protein-protein interactions for which there is a clear need for new/improved inhibitors. The aim of this project was primarily to identify new compounds capable of inhibiting either/both of these targets.

Initially this required the development of assays capable of identifying inhibitors. As this chapter outlines, the interaction between HIF-1α and p300 occurs over a large surface area and is a challenging target for modulation, the work presented within this thesis will endeavour to further elucidate the requirements for binding and thus allow targeted inhibitor design.

Finally a variety of computational and synthetic approaches to inhibitor discovery/development were to be employed in the search for potent inhibitors of both the HIF-1α/p300 interaction and the eIF4E/eIF4G interaction.

Chapter 2 – Assay Development

The candidate confirms that the work submitted is his 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.

Portions of the work presented in this chapter were included in a research paper that was accepted for publication in February 2014 (G.M. Burslem, H.F. Kyle, A.L. Breeze, T.A. Edwards, A. Nelson, S.L. Warriner and A.J. Wilson, *ChemBioChem*, 2014, **15**, 1083). The contributions of the authors were as follows: GMB (the candidate) performed all synthetic chemistry, peptide synthesis and modelling, HFK performed all protein expression, GMB and HFK performed all assays jointly. GMB prepared a draft of the manuscript which was then edited into its present form (attached) by all authors.

Portions of the work described in this chapter were submitted as part of a research paper describing the requirements for HIF-1α/p300 binding (Exploration of the HIF-1α/p300 binding interface using peptide and adhiron phage display technologies to locate binding hot-spots for inhibitor development, H. F. Kyle, K. F. Wickson, J. Stott, G. M. Burslem, A. L. Breeze, D. C. Tomlinson, S. L. Warriner, A. Nelson, A. J. Wilson and T. A. Edwards, submitted). H.F.K. expressed proteins, carried out phage display experiments, performed NMR and crystallographic analysis. K.F.W. assisted with phage display experiments, J.S. perform bioinformatics analysis, G.M.B. (the candidate) prepared reagents. Binding studies were performed jointly by H.F.K. and G.M.B. A.L.B., D.C.T., S.L.W., A.N., A.J.W., and T.A.E. supervised the research. H.F.K. prepared a draft which was edited into its present form (attached) by all authors. Portions of the work performed by the candidate are included in this chapter.

2.1 Assay Development

To facilitate the discovery of inhibitors of the interactions of interest, namely the HIF-1α/p300 interaction and the eIF4E/eIF4G interaction, it was necessary to develop competition assays capable of monitoring the inhibition of binding between peptide (derived from HIF-1α or eIF4G) and the corresponding protein binding partner (p300 or eIF4E respectively). Fluorescence anisotropy assays were selected as the primary screen as it is a relatively high throughput biophysical assay[86] and has previously been used for the successful identification of PPI inhibitors by the Wilson group[87-90] and others[80, 91, 92]

2.1.1 The basis of fluorescence anisotropy[86]

Fluorescence anisotropy is a technique based on polarized light and the time it takes for a molecule in solution to tumble. If a fluorophore is excited with polarized light at its excitation frequency, it will emit polarised light at the emission frequency. However, when a molecule is in solution it is tumbling so that although the light emitted is still polarised with respect to the incident angle, the light emitted from a sample is no longer as polarised. The degree to which the light at the emission frequency is polarised depends on the rate of tumbling of the molecule which in turn is dependent on the size (radius of gyration) of the molecule. For example, a large molecule, such as a protein, will rotate slowly in solution compared to a small molecule, so the degree of polarised light emitted from a sample of fluorescent protein will be higher than that of a fluorescent small molecule or peptide. This can be used to monitor complexation in solution. If a labelled small molecule in solution is tumbling rapidly then the light emitted will not be polarised but should it form a complex with a larger molecule then its effective size will increase, its rate of tumbling will decrease and thus the degree of polarized light will increase as illustrated in Figure 19.
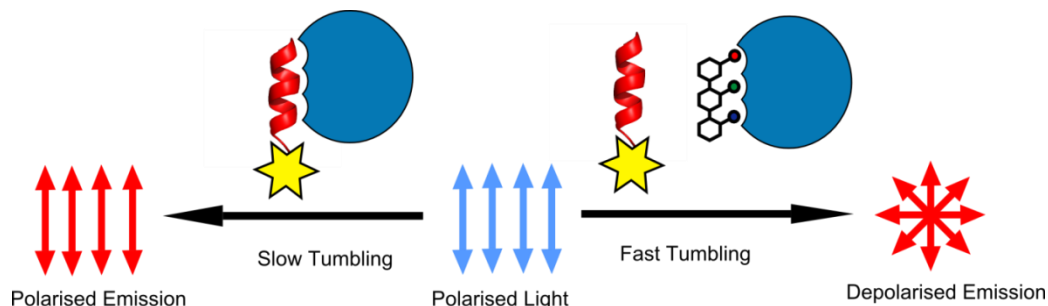
Figure 19 – Cartoon illustrating fluorescence anisotropy of peptide tracer ligand with and without competitive inhibitor present

Anisotropy is calculated by measuring the intensity of light emitted with the same polarisation as the exciting source, *s*, and the intensity of light emitted with

perpendicular polarisation to the exciting source, p. From these measured values the total intensity, i, can be calculated from equation (1). The anisotropy, r, can then be calculated using equation (2).

$$i = (2p) + s \qquad (1)$$

$$r = \frac{(s-p)}{i} \qquad (2)$$

To identify inhibitors of protein-protein interactions, anisotropy can be utilised by having a small fluorescently labelled peptide which contains key binding residues and the protein to which it binds in solution. Potential inhibitors can then be titrated into the solution and the fluorescence anisotropy measured. If the inhibitor competes away the labelled peptide from the protein then the fluorescence anisotropy should decrease because the small peptide is tumbling in solution faster than it could when complexed with the large protein. Ideally, the size difference between the peptide and protein would be large so that the change in tumbling rate would be large and the change in anisotropy would be greater.

2.2 HIF-1α/p300 Assay Development

2.2.1 Direct Binding of HIF-1α CTAD to p300

As discussed in the previous chapter, the *C*-terminal transactivation domain (CTAD) of HIF-1α interacts with the CH1 domain of p300.[30] To investigate this interaction and develop an assay capable of identifying inhibitors, the p300 CH1 domain (prepared by recombinant expression by H. Kyle) was titrated against a fluorescein labelled peptide (prepared by Proteogenix, France) derived from the sequence of the HIF-1α CTAD, shown in Figure 20, and the anisotropy measured (as described in the experimental section). An increase in the anisotropy, as shown in Figure 21, shows that as the protein concentration was increased, it bound to the fluorescent peptide and the rate of fluorophore tumbling decreased. Fitting this data to a 1:1 binding model allowed the calculation of a binding constant (K_d) of 37.4 ± 4.4 nM which is in good agreement with the literature. [29, 35, 55]

Ac-SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN-NH$_2$

FITC-Ahx-SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN-NH$_2$

Figure 20 - HIF-1α CTAD peptide sequences
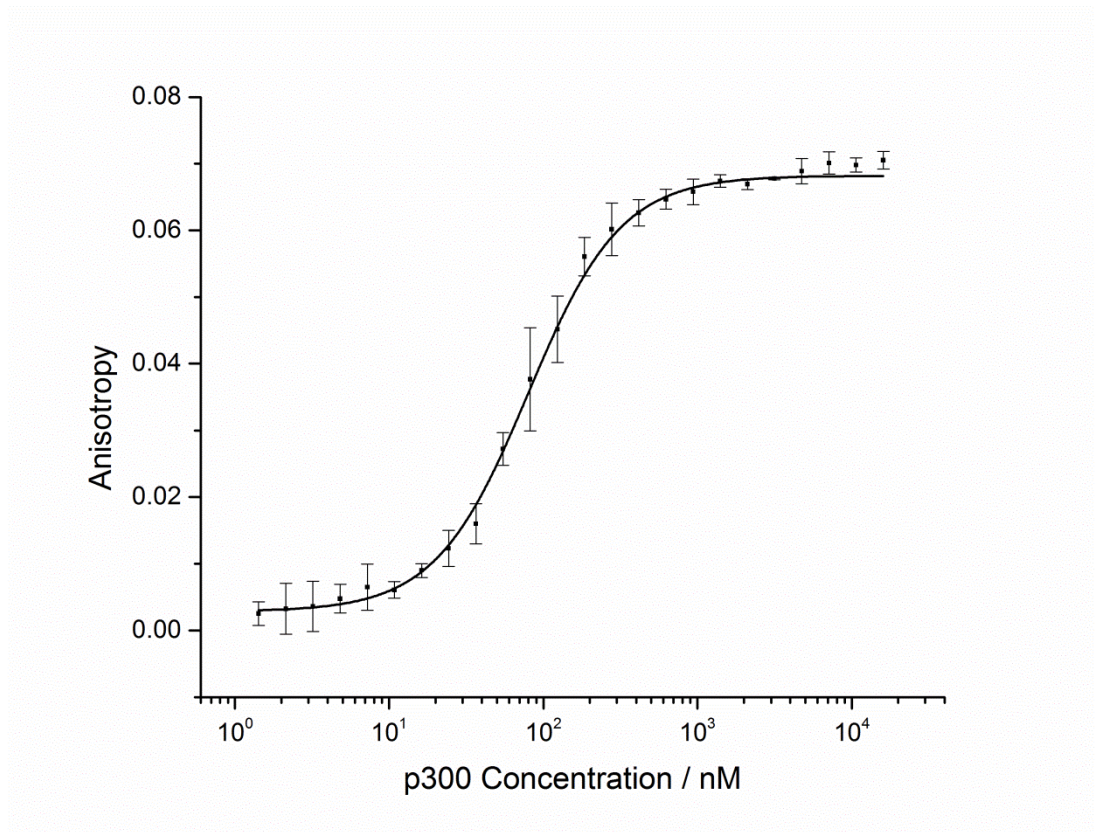
Figure 21 – Binding curve for FITC HIF-1α CTAD/p300 CH1 domain (80 nM FITC-HIF-1α CTAD, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol). Error bars represent the standard deviation of 3 repeats.

2.2.2 Time and Temperature Dependence of HIF-1α/p300 Interaction

Initially the assay plates were prepared and allowed to equilibrate for approximately 40 minutes. To ensure this was sufficient for an equilibrium to be reached, a plate was prepared and read at different time points; 40 minutes, 5 hours and 18 hours. The plot below, Figure 22, shows that no significant change was observed between 40 minutes and 5 hours. A slightly different K_d was observed after 18 hours but due to the unknown stability of the protein and the fact that incubation for 18 hours is impractical, an incubation time of 40 minutes was selected.
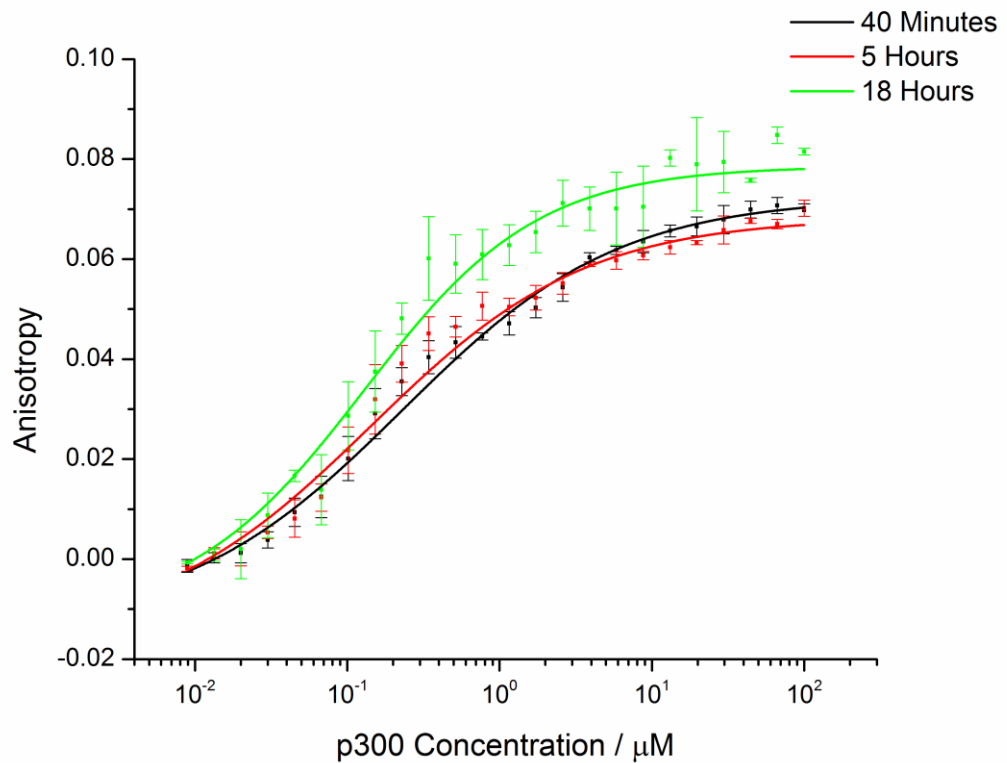
Figure 22 - Time dependence on fluorescence anisotropy assays between FITC-
HIF-1α and p300. (80 nM FITC-HIF-1α CTAD, 40 mM sodium phosphate,
100 mM NaCl, 1 mM DTT, 5% glycerol). Error bars represent the standard
deviation of 3 repeats.

A temperature screen was also conducted to assess the thermodynamic parameters
governing complex formation, a plate was prepared and measured at different
temperatures set by the plate-readers internal temperature control. The plate was
incubated within the plate reader as the temperature was increased and held at the set
temperature for a minimum of 20 minutes. A general increase in the rate of tumbling
in solution as the temperature increases manifests itself as a decrease in the
minimum and maximum anisotropy as seen in Figure 23.
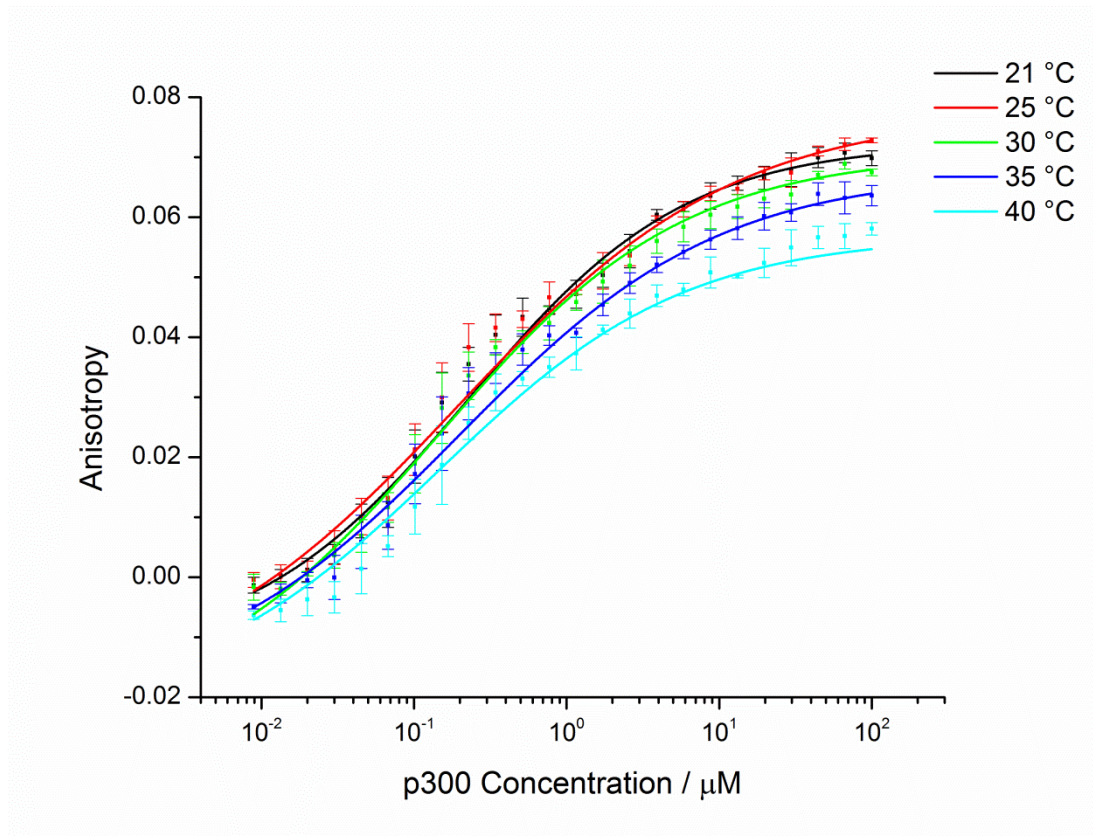
Figure 23 - Temperature dependence of HIF-1α/p300 interaction(80 nM FITC-HIF-1α CTAD, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol). Error bars represent the standard deviation of 3 repeats.

A general downwards trend was observed in the binding constant as the temperature increases. Performing a Van't Hoff analysis (a plot of $1/T$ against $\ln K_d$) on this data shows a positive gradient suggesting an exothermic interaction with an enthalpic driving force. The negative entropic contribution could be attributed to the adoption of ordered helical regions observed in NMR structures over the random coil thought to exist in the absence of p300.[30] The enthalpic/entropic contributions are corroborated by the ITC experiment discussed below.
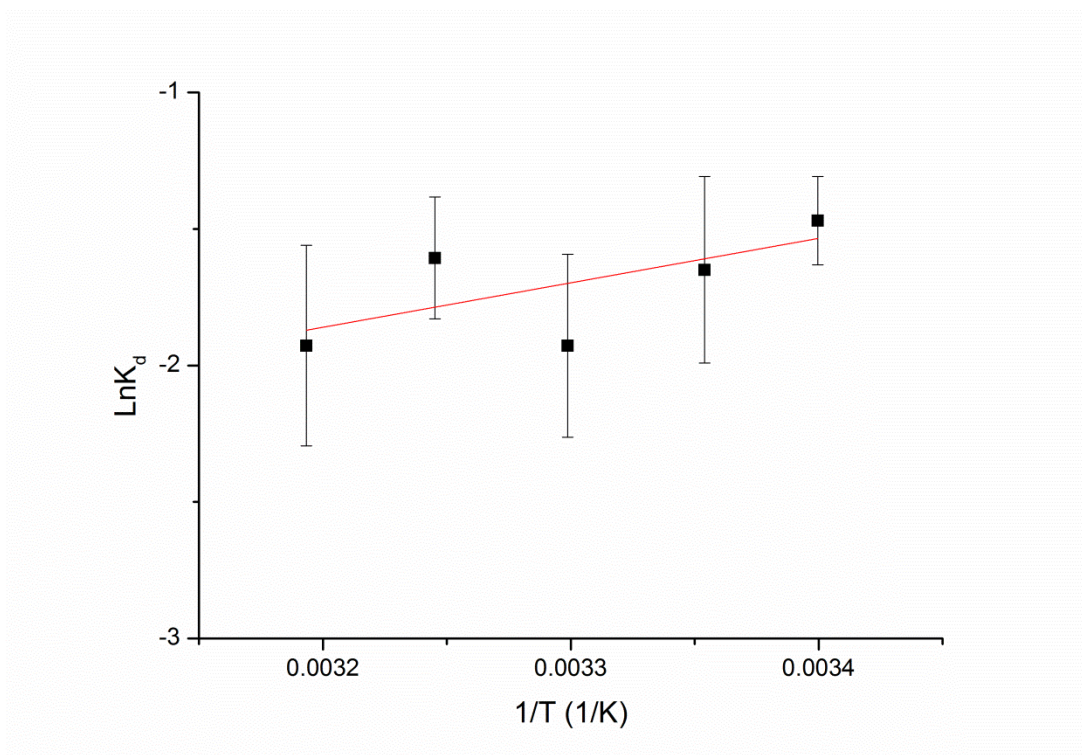
Figure 24 - Van't Hoff Analysis for the HIF-1α/p300 interaction

2.2.3 Isothermal Titration Calorimetry analysis of HIF-1α/p300

To further confirm this interaction, an isothermal titration calorimetry (ITC) experiment was conducted. ITC gives useful information about binding affinity but also provides information about the stoichiometry of an interaction and enthalpic/entropic contributions to the binding free energy. During an ITC experiment one binding partner is injected in aliquots into a temperature controlled cell containing the other binding partner. As the partners interact, a reference cell is either heated or cooled to maintain the same temperature between the two cells. The energy required to alter the reference cell temperature can be used to derive a binding isotherm. By employing this technique a K_d of 45 ± 4 nM was measured for the HIF-1α/p300 interaction as shown in Figure 25. Interestingly during these ITC experiments the delay times between injections had to be extended to allow a the signal to return to the baseline, suggesting a relatively slow on-rate.
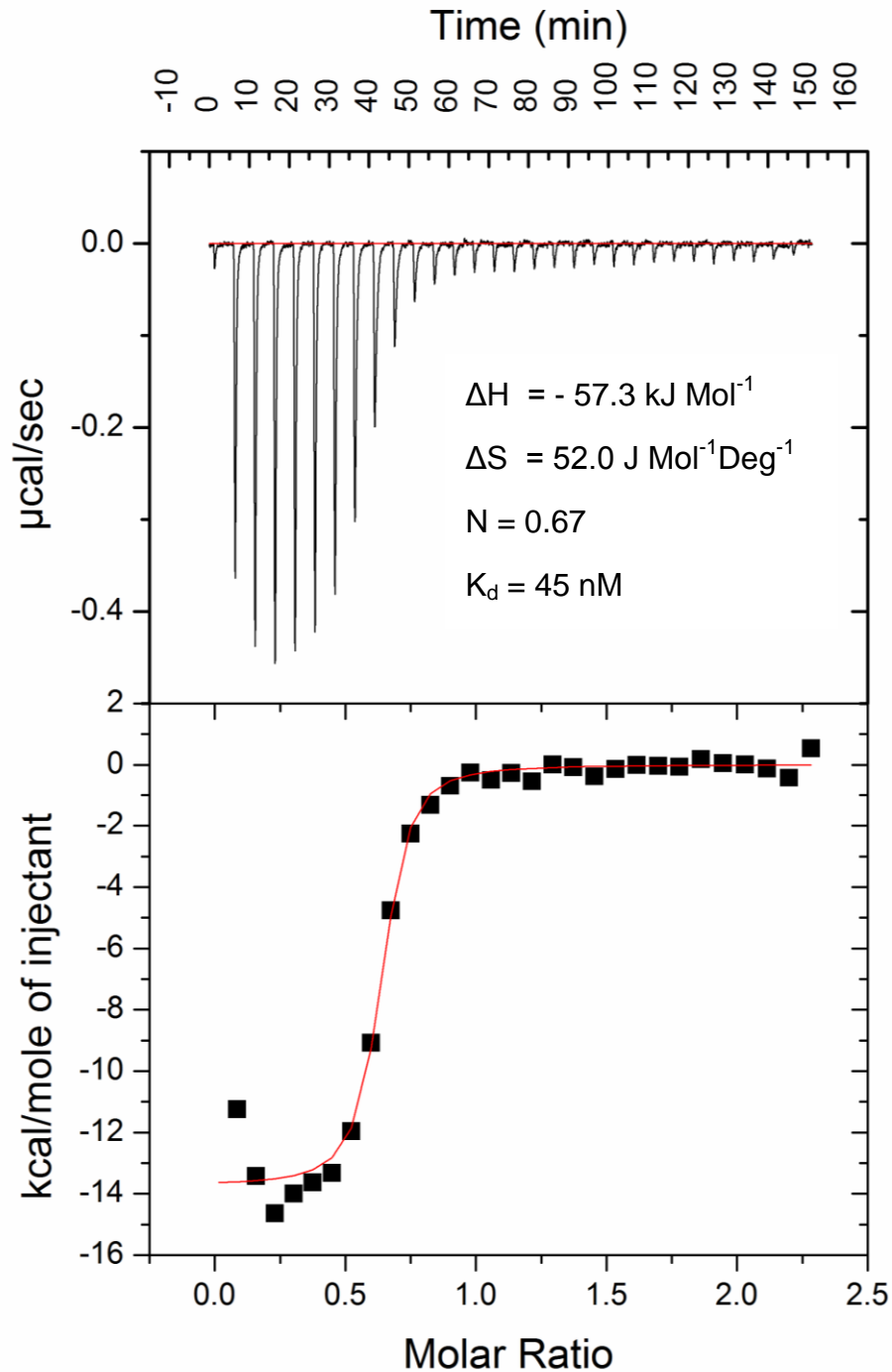
Figure 25 - Isothermal titration calorimetry analysis of the HIF-1 CTAD/p300 Interaction

The K_d measured by ITC and fluorescence anisotropy are in good agreement with each other and with literature values.[29, 35, 55] The stoichiometry, *n* of the interaction observed by ITC was 0.69, slightly less than the expected 1:1, but this can be rationalised by the possible presence of some partially unfolded/misfolded protein. The total protein concentration in the cell can be measured but the degree to which it was properly folded/functional cannot be easily assessed.

2.2.4 Competition Assay for inhibitors of HIF-1α/p300

To identify inhibitors, a competition assay was required. To this end a concentration of p300 that gave sub-maximum binding was selected and the unlabelled analogue of the HIF-1α CTAD titrated into a mixture of the p300 CH1 domain and the FITC-HIF-1α CTAD. The unlabelled peptide competes with the labelled analogue for binding to the p300, resulting in a decrease in anisotropy (Figure 26).
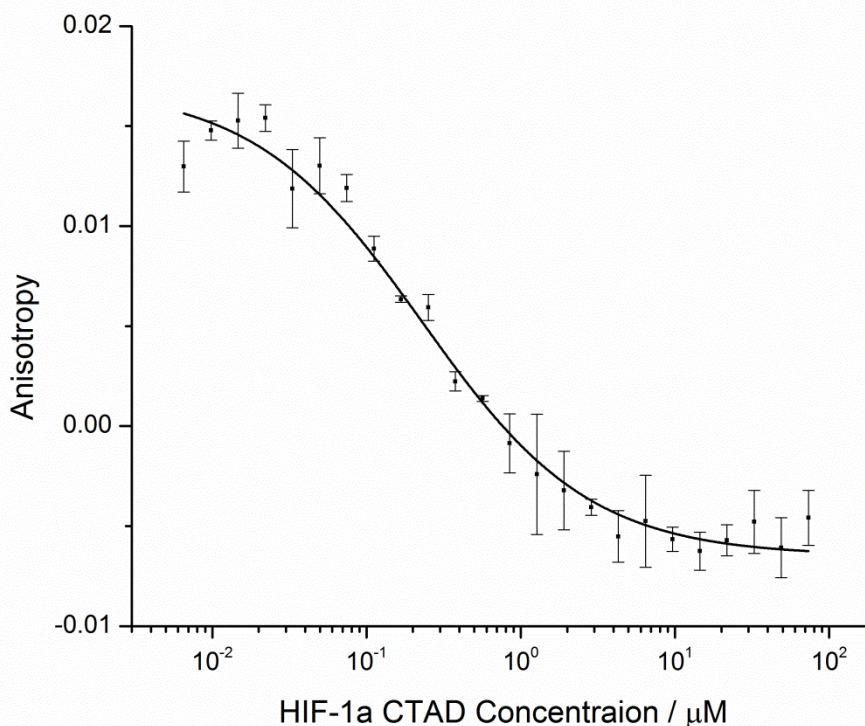
Figure 26 - Competition titration of HIF-1α CTAD vs FITC HIF-1α CTAD. (80 nM FITC-HIF-1α CTAD, 0.1 μM p300, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

The decrease in anisotropy validated the assay for the identification of inhibitors of this interaction. The IC_{50} value derived from this assay is 228 ± 25 nM which when converted to an inhibition constant (K_i) using the method of Lebeda *et al.* gives 42.2 $\pm$ 7.67 nM which is in good agreement with the direct binding measured for the same peptide sequence as above.[93] The slight discrepancy may be due to interactions made by the fluorophore in the labelled peptide.

To enable the identification of even weak inhibitors of this interaction, the screening of inhibitors at relatively high concentrations (~1 mM) was envisaged. To enable this, it was necessary to perform the assay not only in aqueous buffer solutions but also in DMSO/buffer mixtures. To ensure this would not prevent the assay from functioning a 20% volume solution of DMSO in buffer was serially diluted into a

plate containing the protein and tracer as a mock inhibitor and the anisotropy measured. No change in anisotropy or intensity was observed at this level of DMSO confirming that a moderate level of co-solvent could be used during screening.

2.2.5 HIF-1α/p300 Binding Hotspots

As discussed in chapter 1, the surface area over which the HIF-1α CTAD and p300 interact is large. By measuring the solvent exposed area of both components of the complex separately and then the solvent exposed area of the complex, the occluded surface area of approximately 4200 Å^2 can be calculated (see experimental section).[94] This is considered too large an area to be blocked by a traditional small molecule inhibitor but certain regions are often more important than others. These so called "hotspots" can be targeted with small molecules and by blocking the high affinity regions, inhibitors much smaller than the surface area of the PPI can be effective inhibitors.[95, 96]

As previously mentioned, alanine scanning has been performed on the HIF-1α CTAD to identify key hot spot residues[33] as well as random mutagenesis studies[34] however these assays were always carried out in cells. To further understand the molecular basis of the interaction and to inform rational inhibitor design a biophysical study was performed.

A previous study, focussing on post-translational modifications, reported that *N*-terminal truncations of the HIF-1α CTAD were tolerated but truncation at the *C*-terminus prevented FITC-labelled peptides from binding p300.[35] It was also demonstrated that *S*-nitrosylation of Cys800 prevented binding to p300 which is in conflict with another report.[36]

2.2.6 Comparison of Helical Regions

As described in previously, the HIF-1α/p300 interaction has been studied using fluorescence anisotropy and isothermal titration calorimetry with a peptide derived from the *C*-terminal transactivation domain (CTAD) and shown to have a binding constant of approximately 40 nM. Given the previous successes of studying helical regions as hotspots[9, 87, 97, 98] and the fact that the hotspot residues are found on helical regions of HIF-1α, it was decided to assess the relative contributions from the different helical regions of HIF-1α. To this end, peptides encompassing the helical regions were synthesized with fluorescent labels to allow direct binding measurements.
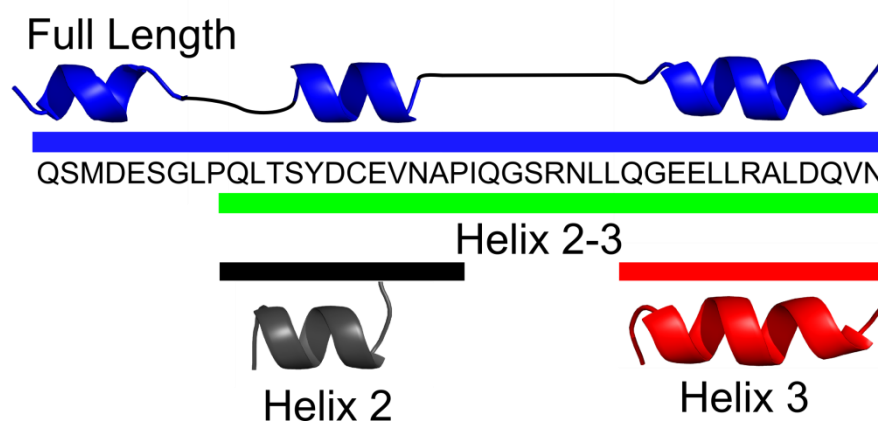
Figure 27 - Schematic representation of peptide prepared containing helical regions of HIF-1α

The hot spot residues Cys800 and Asn803 are in the *i* and *i*+4 positions of helix 2 and the hot spot residues Leu814, Leu819 and Val821 are in the *i, i*+4 and *i*+7 positions of helix 3 (see Chapter 1).

Peptides comprising the helix 2 sequence, the helix 3 sequence and the sequence of from helix 2-3 were prepared using standard Fmoc solid phase peptide synthesis (SPPS) as described in the experimental section and labelled with *N*-terminal fluorescein via an aminohexanoic acid linker. It was envisaged that since the *N*-terminal region of the HIF-1α CTAD has not previously been indicated as containing hotspots and appears as helical in only one of the two NMR structures, (PDB ID: 1L3E[29] features only 2 helices) that it would not play a significant role in binding. Truncation at the N-terminus has previously been shown to only slightly reduce activity.[33] Indeed previous work by the Arora group suggested that the helix 2 region alone was sufficient for binding.[54]

The binding of these three peptide sequences was assessed by direct fluorescence anisotropy and compared with the binding of the full length CTAD as shown in Figure 28. Perhaps most surprising is the approximately 80-fold decrease in binding between the full length CTAD and the helix 2-3 (K_d = 6.7 ± 0.5 μM) construct suggesting that the helix 1 region is important. Moreover it appeared that the helix 2 region is essentially insufficient to bind p300 in the absence of the rest of the sequence, despite containing two reported hotspot residues. The helix 3 peptide binds to p300 with a significantly lower affinity, suggesting that the previous observations regarding helix 3 hotspots were correct.
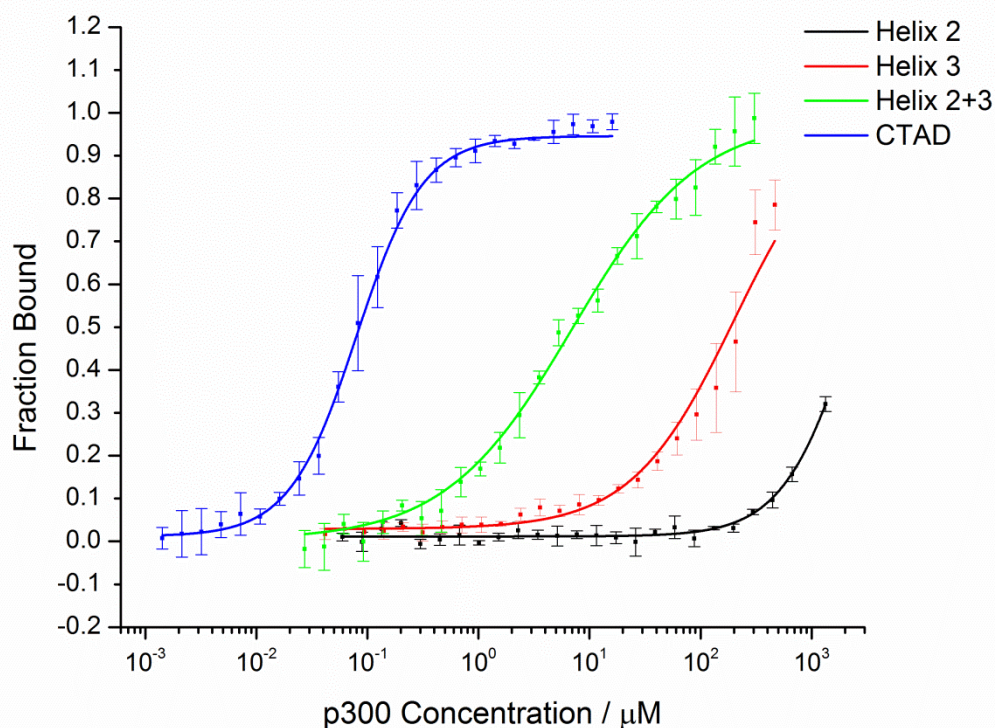
Figure 28 - Fluorescence anisotropy binding of different peptides to p300 (80 nM FITC-peptide, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol). Error bars represent the standard deviation of 3 repeats.

2.2.7 Exploration of Cooperativity

Of the two helical regions investigated here, helix 3 has the higher affinity but the disparity between the two helical regions separately compared to when linked suggested that other factors might be important. Cooperativity between the two helical regions could explain some of the observations. To explore any potential allosteric cooperativity, the fluorescently-labelled helix 3 peptide was titrated into a solution containing both the protein and a high concentration (3 mM) of the unlabelled helix 2 peptide. No increase in affinity was observed compared to that observed for binding of helix 3 in the absence of helix 2 as can be seen in Figure 29. Allosteric cooperativity by a different region of the HIF-1α CTAD cannot be ruled out but the presence of helix 2 did not appear to increase the affinity of helix 3.
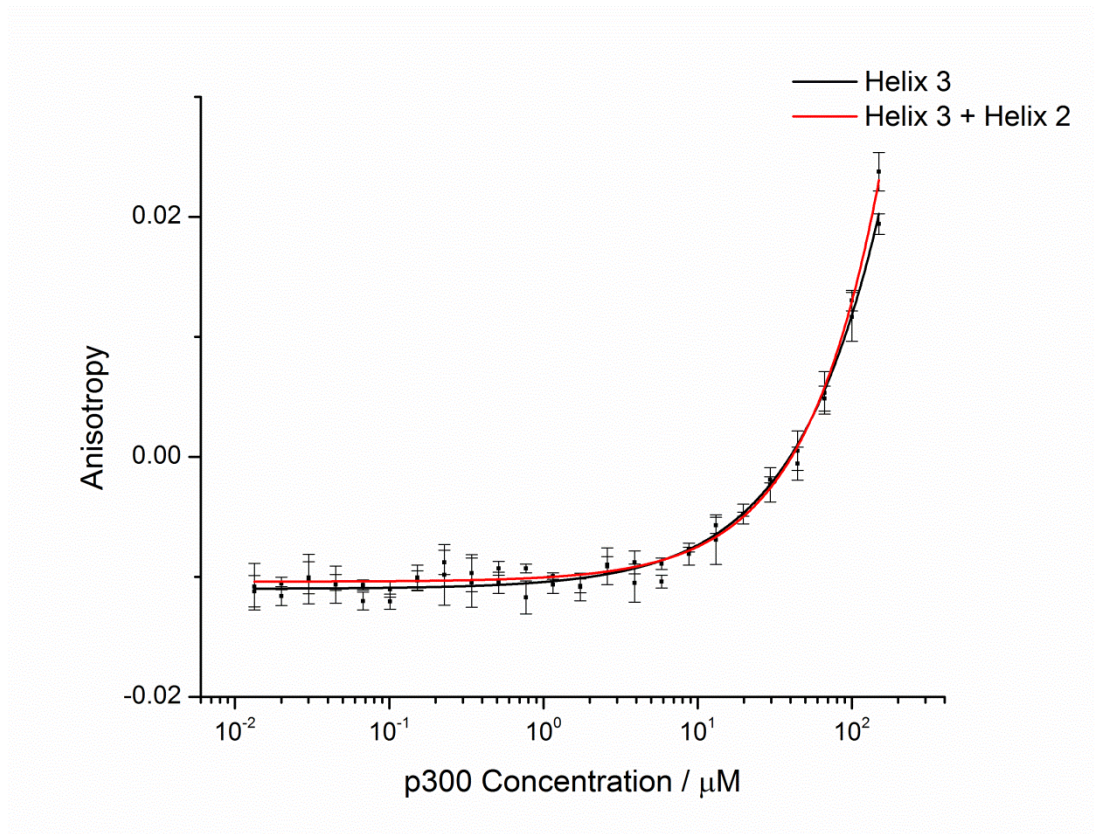
Figure 29 - Fluorescence anisotropy binding of helix 3 in the presence and absence of helix 2. (80 nM FITC-HIF-1α CTAD, 3 mM helix 2 peptide, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol). Error bars represent the standard deviation of 3 repeats.

The significant increase in binding when the helix 2 and 3 peptides are linked suggest that either a previously unidentified hotspot residue may exist in the linking region or that the chelate effect is crucial for this interaction. The chelate effect arises from an increase in effective molarity and a reduced entropic penalty for sequential binding events. Inspection of the ensemble of NMR structures suggests that this loop is relatively unstructured as shown in Figure 30. The chelate effect can significantly increase binding of two weak interactions via increasing the effective molarity.[99, 100]
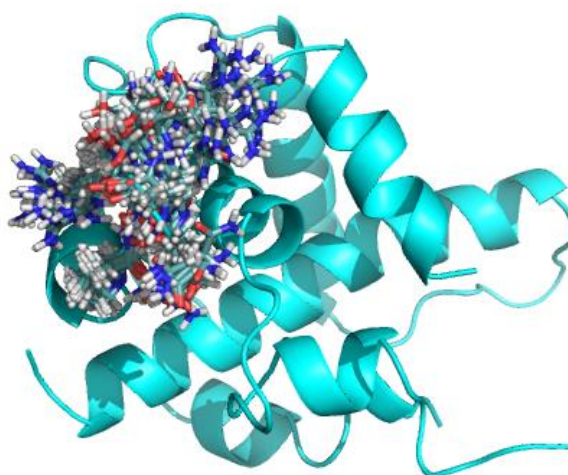
Figure 30 - Ensemble NMR structures of helix 2 to helix 3 linker region. 1
backbone representation shown for clarity.

To investigate the importance of the chelate effect on the binding between HIF-1α
and p300, a series of peptide conjugates was proposed. By linking the helical
domains of the HIF-1α CTAD with a polyethylene glycol (PEG) linker, the effect of
chelation could be explored without the linker region making specific interactions
with the protein. PEG linkers have previously been incorporated into unstructured
loop regions of proteins without perturbing binding.[101] The design of these
constructs was challenging as calculating the required length for the PEG linker is
non-trivial. Achieving the correct length is crucial, as if it is too short then both
helices could not bind simultaneously.[102] To this end, the length of the native
sequence and the linear distance between the helices was measured computationally
(Table 1). For a PEG chain to span the same distance it must undergo a random 3D
walk so longer lengths were required.[103] The commercially available monodisperse
PEG lengths chosen along with the lengths supplied by the manufacturers are shown
in Table 1.

Table 1 – Distances from helix 2-3 and PEG linker lengths.

	Length (Å)
Linear distance	21
Native linker sequence	40
PEG$_8$	40
PEG$_{12}$	53
PEG$_{24}$	95

Peptide constructs were prepared, consisting of a fluorophore (FITC), the helix 2 binding region, a PEG linker and then the helix 3 binding region as shown in Figure 31. The PEG-peptide conjugates were synthesised via standard Fmoc-SPPS utilising an Fmoc-amino-carboxy-PEG building block and *N*-terminally labelled with fluorescein to allow direct binding measurements (see Appendix I for alternative approaches).

FITC-Ahx-QLTSYDCEVAN-PEG$_8$-EELLRALDQVN-NH$_2$

FITC-Ahx-QLTSYDCEVAN-PEG$_{12}$-EELLRALDQVN-NH$_2$

FITC-Ahx-QLTSYDCEVAN-PEG$_{24}$-EELLRALDQVN-NH$_2$

Figure 31 - PEG linked peptide sequences

Titration of p300 protein against the PEG conjugates showed a significant dependency on the length of the PEG linker. As the PEG chain was lengthened the binding to p300 increased (Figure 32 and Table 2).
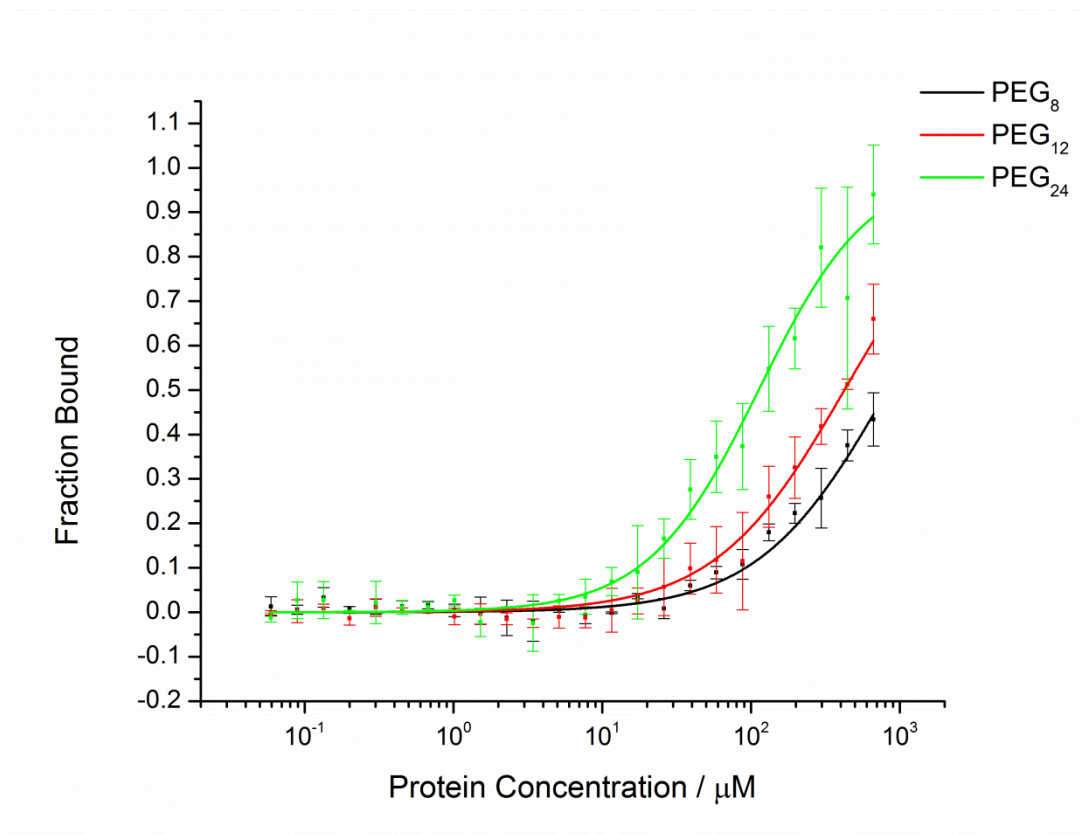
Figure 32 - Binding curves for FITC peptide PEG conjugates (80 nM FITC-Peptide, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol). Error bars represent the standard deviation of 3 repeats.

Table 2 - Binding constants for PEG linked constructs determined by fluorescence anisotropy protein titrations

Linker Length	Estimated $K_d/\mu M$
PEG_8	827 ± 99
PEG_{12}	423 ± 17
PEG_{24}	144 ± 13

The PEG_{24} construct binds tightest and is the only construct to bind with higher affinity than the helix 3 region alone (cf. 198 μM). This suggests that either the PEG length is not long enough or that interactions between the linker region and p300 are crucial. By applying a variation of the Kuhn length model, the average distance of a random walk by a polymer can be estimated. Using the published characteristics of PEG polymers[104, 105] and the self-avoiding random walk model, it can be estimated that a PEG_{72} linker would be required for an average distance of random walk of >40 Å suggesting that, despite the linear lengths of the linkers used above were sufficient, further elongation of the linker may continue to result in increased affinity.

Nevertheless, the results reported hear tentatively suggest that the chelate effect is important for the high affinity interaction between the HIF-1α CTAD and p300. Whilst an increase in affinity was not observed when both Helix 2 and 3 peptides were present but separate, an increase in affinity was observed when the same two peptides were linked by a non-peptidic linker.

2.2.8 Peptide Inhibitor Screening

Since the objective of the project was to design and prepare inhibitors of this PPI, a more complete series of peptide fragments of the HIF-1α CTAD was prepared and tested for inhibition. Despite the fact that peptide-based inhibitors are unfavourable in terms of cell permeability and instability to degradation, the identification of peptides capable of inhibiting the full length CTAD binding could inform inhibitor design. [106]

SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN   Full CTAD
SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN   Helix 1-2
SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN   Helix 2-3
SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN   Linker plus Helix 3
SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN   Helix 2 plus Linker
SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN   Linker
SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN   Helix 1; Helix 2; Helix3

Figure 33 - Sequences of peptides used to explore inhibitory ability of HIF-1α CTAD regions

Nine additional peptides, the sequences of which are shown in Figure 33, were synthesised or purchased and tested in the competition assay discussed in Chapter 2. In these competition experiments, only 2 peptides, in addition to the unlabelled CTAD, were able to inhibit to any extent (Figure 34). The peptide comprising the helices 2 and 3 (denoted helix 2-3) in agreement with the direct binding measurement described above and the peptide comprising helix 1 and 2 (denoted helix 1-2). This result again suggests that the helix 1 region may be relevant for the recognition between HIF-1α and p300.
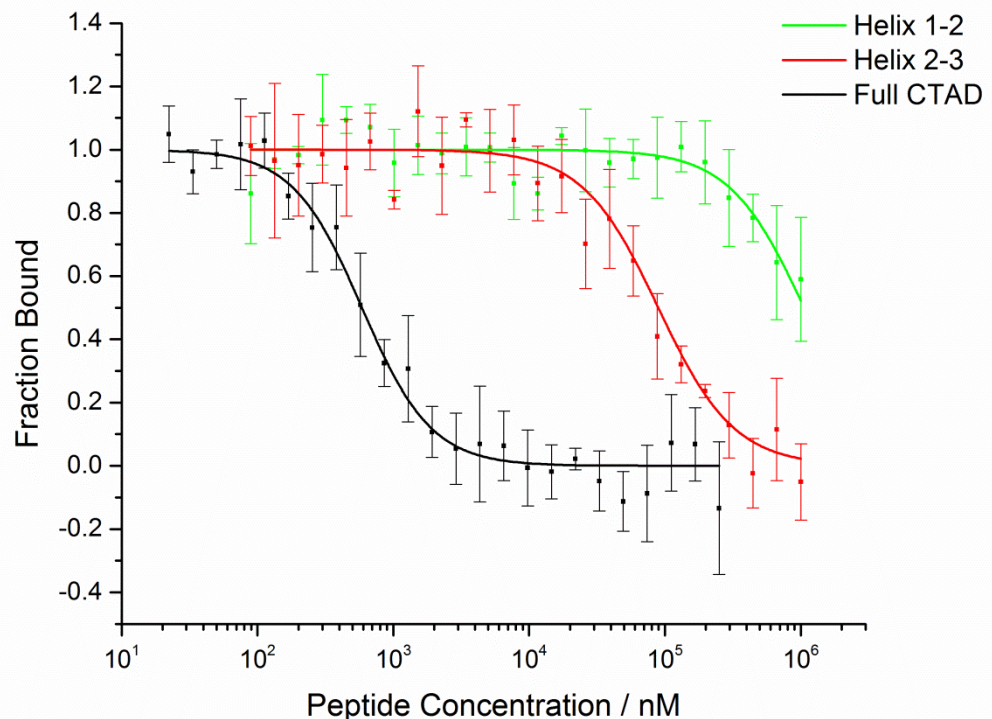
Figure 34 - Competition assay with peptide portions of HIF-1α against FITC HIF-1α CTAD. (80 nM FITC-HIF-1α CTAD, 0.1 µM p300, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

Unfortunately these competition assays did not reveal a key pharmacophore or even provide a ranking of the peptides in affinity order. Instead, suggesting that, at least for peptidic inhibitors, 2 helical regions are the minimum requirement for inhibitory activity. This is potentially due to the format of the assay which relies on competing with a particularly high affinity peptide (Kd ~ 40 nM).

2.2.9 Conclusions on HIF-1α/p300 binding requirements

Whilst the results presented in this section do not offer conclusive proof it led to the proposal of a hypothesis: that the helix 3 region of the peptide provides the majority of the affinity whilst the rest of the peptide, particularly the previously identified hotspots, impart specificity. This would be in line with previous reports carried out on much more complex biological systems. It is conceivable that mutation of the Cys_{800} or Asn_{804} residues led to decreased activity in cells due to off target binding which is not observed in the isolated recombinant system discussed here. They are also key sites for post-translational modifications as discussed in Chapter 1. Further work is clearly required in this area to more thoroughly investigate, using more sensitive biophysical techniques, the recognition features between the binding partners.

The proposed hypothesis suggests that the helix 3 binding region of p300 would be a good starting point to target inhibitors, giving them the best chance to abrogate the binding between the HIF-1α CTAD and p300.

2.3 eIF4E/eIF4G Assay Development

2.3.1 Direct binding of eIF4G peptide to eIF4E

As for the assay system developed above, a fluorescence anisotropy assay was developed for the eIF4E/eIF4G interaction. A peptide derived from the helical region of the eIF4G protein that binds to eIF4E was synthesised using conventional Fmoc solid phase peptide synthesis on 0.1 mmol scale as both the fluorescein-labelled and acetylated analogue (see experimental section).

Ac-RYDREFLLGFQ-NH$_2$

FITC-Ahx-RYDREFLLGFQ-NH$_2$

Figure 35 - eIF4G peptide sequences

Titration of increasing protein concentrations (prepared by recombinant expression by H. Kyle) into a solution of the fluorescein-labelled peptide provided a binding curve as the peptide bound to the protein and tumbled less rapidly in solution as

shown in Figure 36. Fitting of this data to a K_d model (described in experimental section) gave a K_d of 1.76 ± 0.17 µM which is in good agreement with the literature.[80, 82]
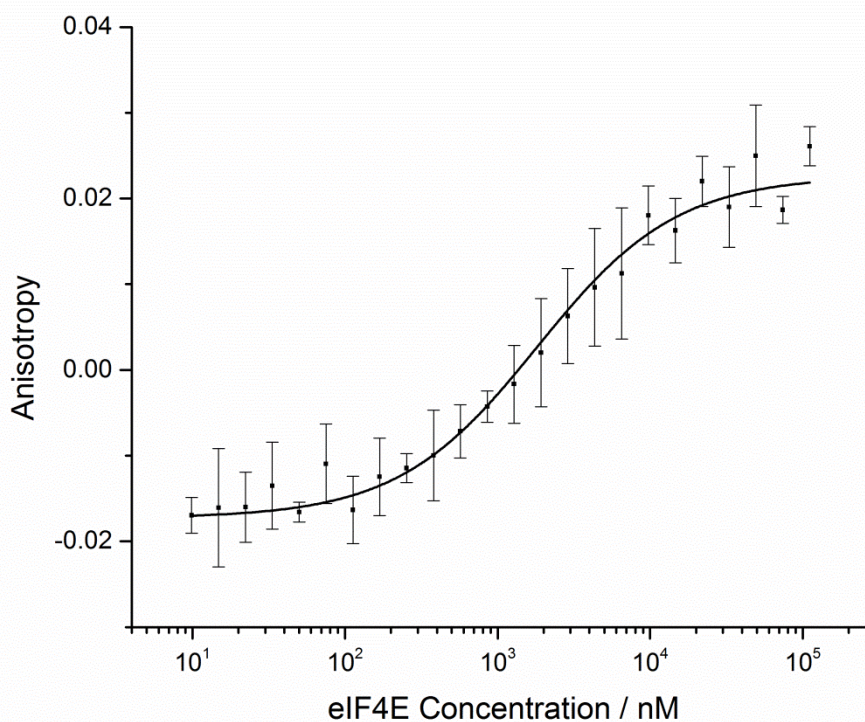
Figure 36 - Binding curve for FITC-eIF4G/eIF4E (80 nM FITC-eIF4G, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats

Since the eIF4e protein had been expressed as a small ubiquitin-like modifier (SUMO) fusion protein and the fact that a larger size difference between the bound and unbound states (high and low anisotropy respectively) gives a higher dynamic range for the assay, binding of the labelled eIF4G peptide to eIF4E-SUMO fusion protein was compared to the binding to eIF4E alone. The presence of the SUMO increases the molecular weight significantly and thus the rate of tumbling but the binding constant was unchanged as can be seen in Figure 37. Additionally, as a control, it was confirmed that no binding of the labelled eIF4G peptide to SUMO could be measured in the absence of eIF4E. This enabled easier protein preparation (no SUMO cleavage step) and gave an improved dynamic range for the assay.
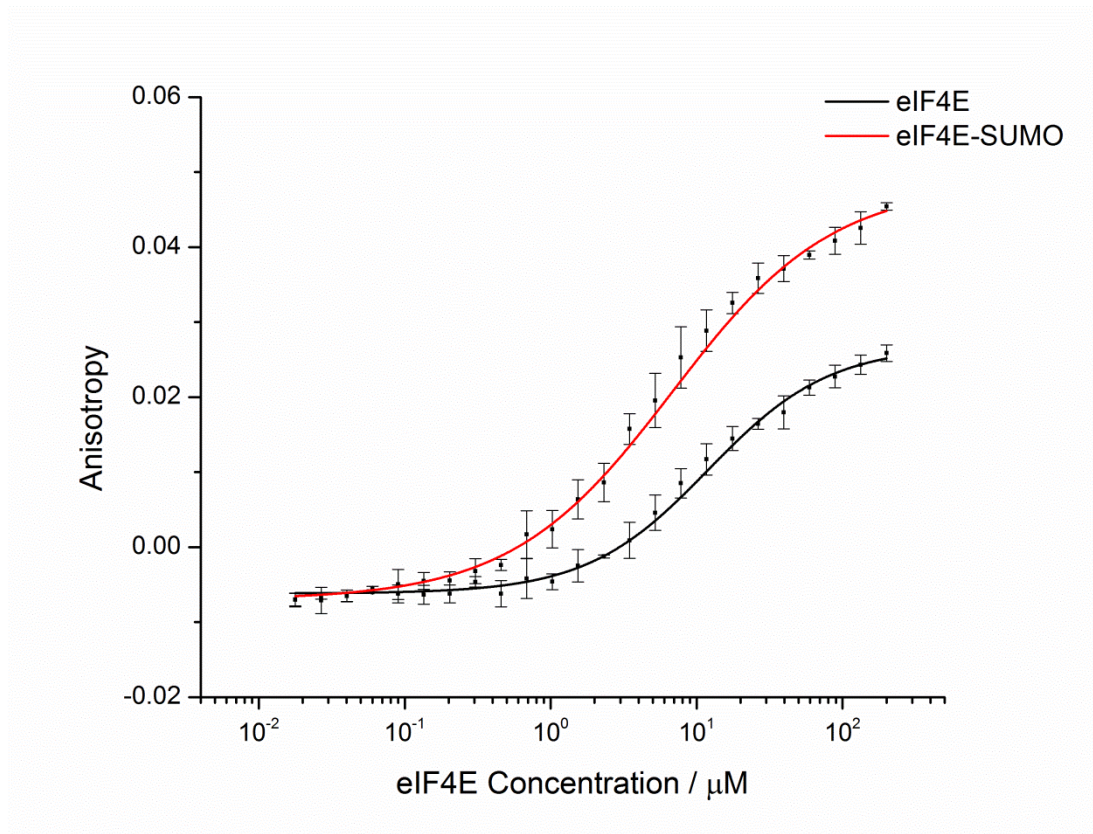
Figure 37 - Comparison of labelled eIF4G binding to eIF4E and eIF4E-SUMO by anisotropy (80 nM FITC-eIF4G, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

2.3.2 Time and Temperature Dependence of eIF4E/eIF4G Interaction

Incubation time was assessed as shown in Figure 38 and indicated that 40 minutes was an adequate time to reach equilibrium as very little change was observed after 5 hours. It also appears that the protein is unstable as the data begins to deviate and become less consistent at 18 hours of incubation.
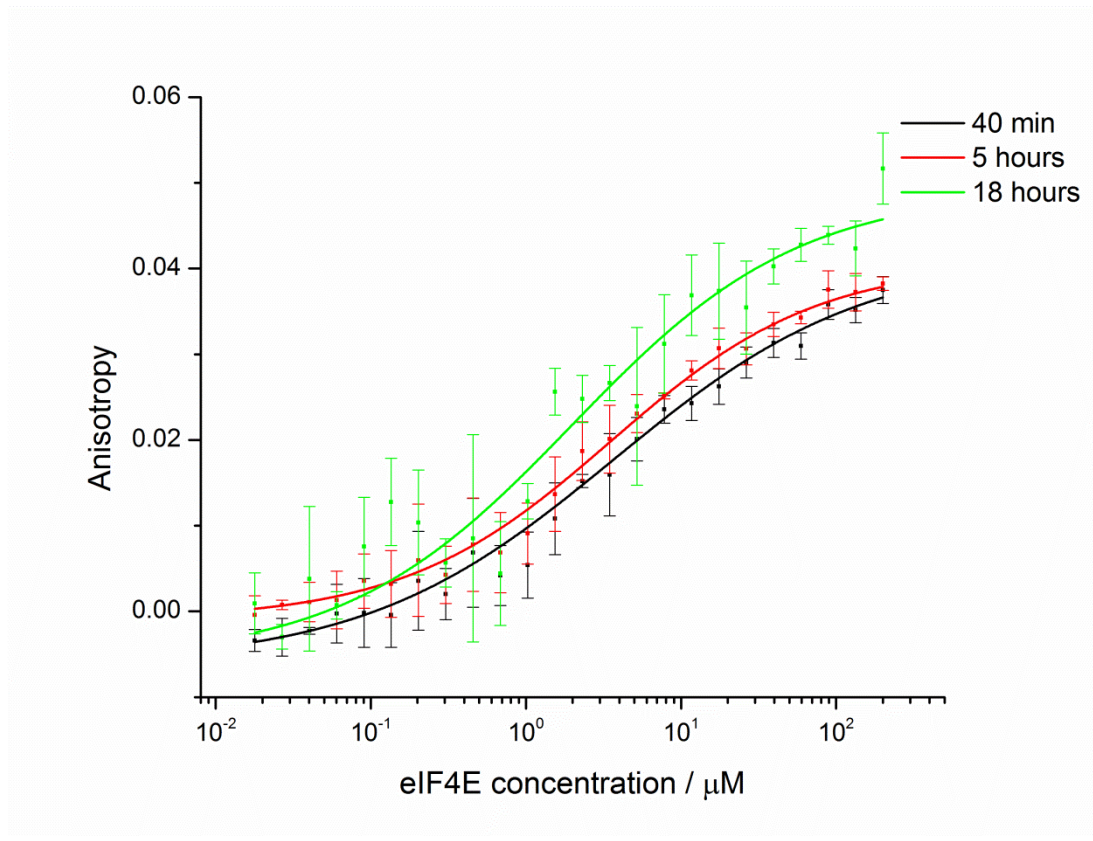
Figure 38 - Incubation time dependence on fluorescence anisotropy assays between eIF4G and eIF4E (80 nM FITC-eIF4G, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats

Similarly, temperature dependence studies were performed. Again the increased temperature increased the general rate of tumbling in solution causing a decrease in the apparent maximum anisotropy which can be seen in Figure 39. The binding constants increased slightly as the temperature increased and performing a Van't Hoff analysis (Figure 40) showed a slightly negative slope – suggesting a slightly endothermic interaction. The magnitude of the gradient suggests the interaction is mostly driven by entropic contribution. This effect may potentially stem from the displacement of ordered water both in the binding site and surrounding the peptide in solution.[107]
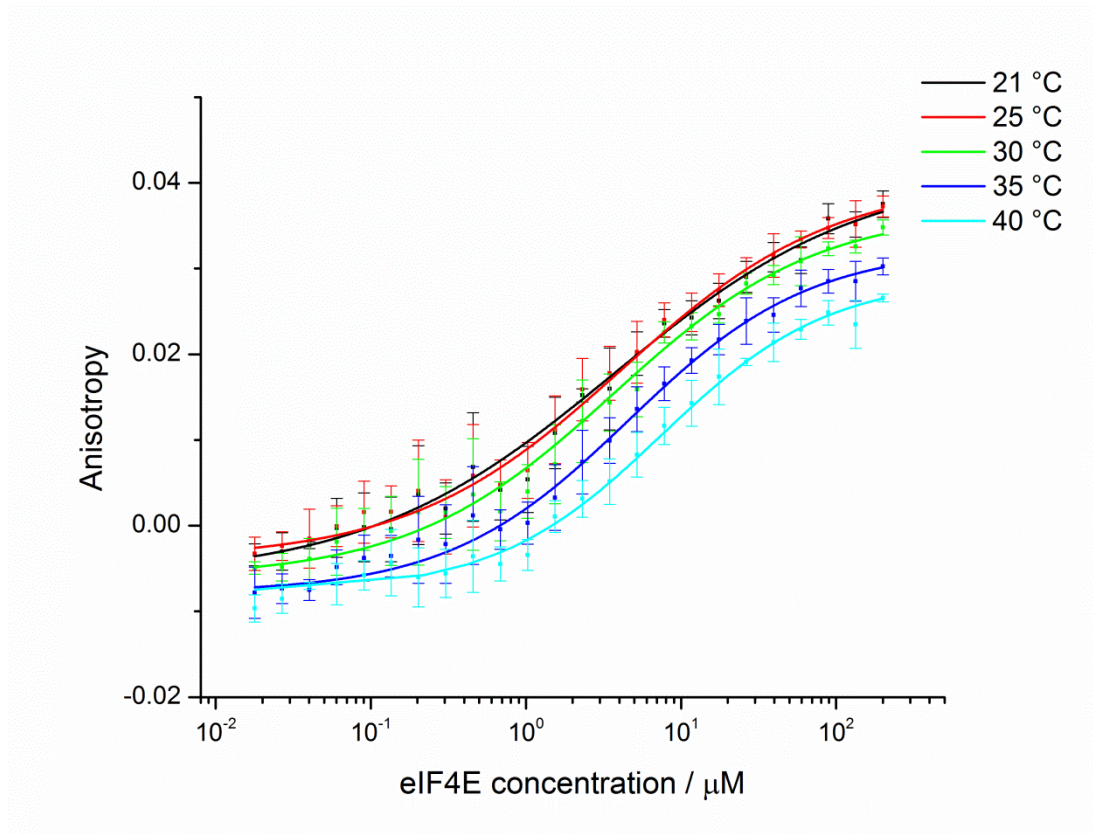
Figure 39 - Temperature dependence of eIF4G/eIF4G interaction (80 nM FITC-eIF4G, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats
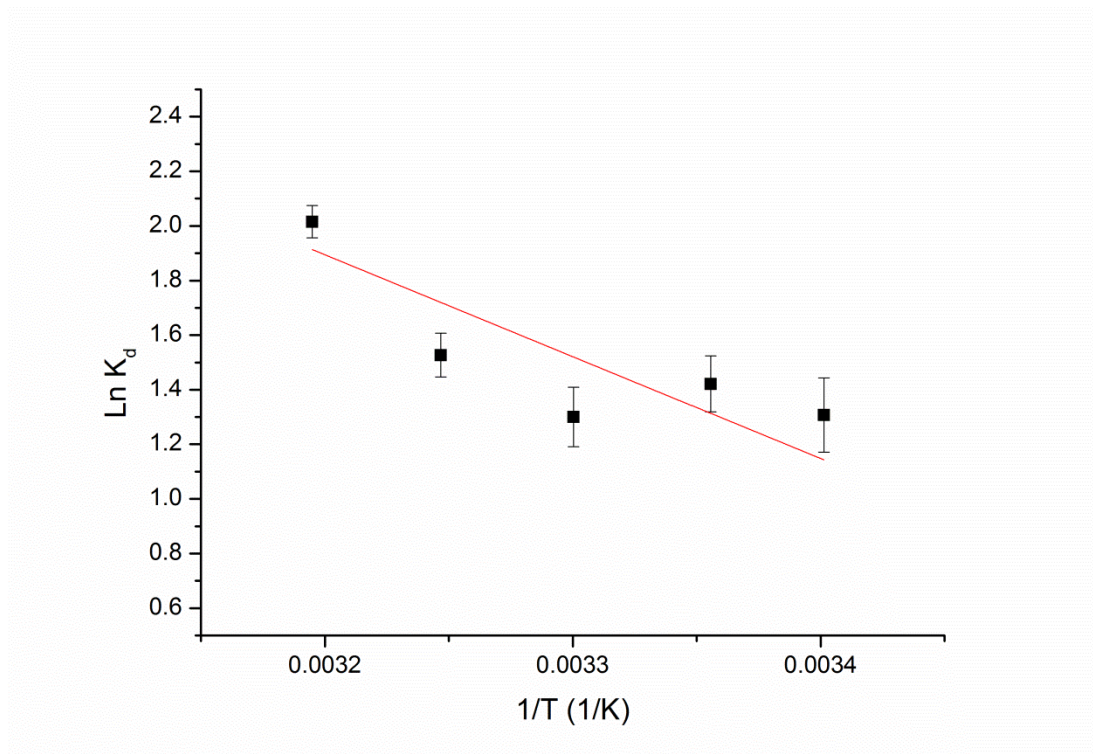
Figure 40 - Van't Hoff Analysis for the eIF4E/eIF4G interaction

2.3.3 Cap dependence of eIF4E/eIF4G interaction

Since the eIF4E/eIF4G interaction is key to the regulation of 5′ cap mRNA translation (discussed in Chapter 1) it was also assessed if the presence of cap or a cap analogue, structures shown in Figure 41, would affect the interaction.
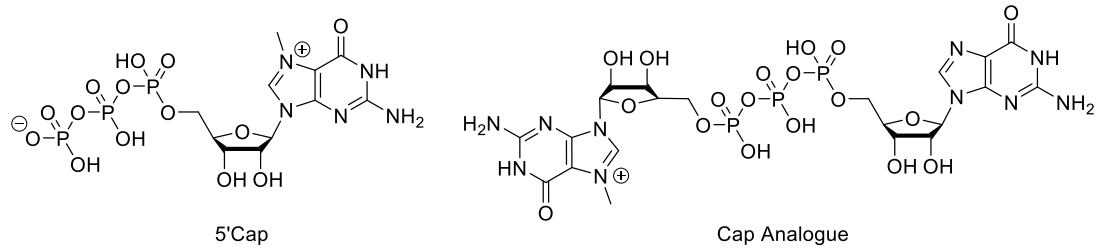
Figure 41 - Structures of 5' Cap and Cap Analogue used in binding investigation

The assay was performed in the presence of 1 mM cap or its analogue but no significant change in affinity was observed as shown in Figure 42. Given this result no cap or analogue was included in subsequent screening experiments. This result was subsequently corroborated by others.[71]
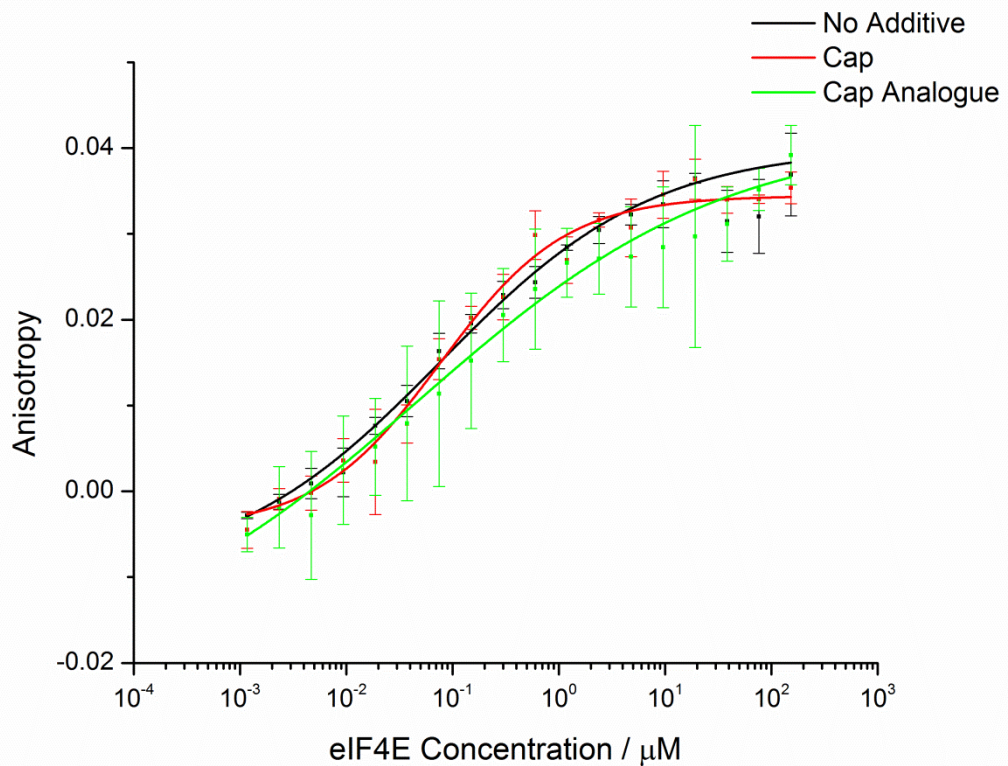
Figure 42 - Effect of cap or cap analogue on eIF4E/eIF4G binding interaction (80 nM FITC-eIF4G, 40 mM sodium phosphate, 200 mM NaCl, 1mM Cap or Analogue, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats

2.3.5 Competition assay for inhibitors of the eIF4E/eIF4G interaction

Subsequently, a competition assay was developed whereby the unlabelled eIF4G peptide was titrated into a mixture of the labelled peptide and the eIF4E protein. It was shown that the unlabelled peptide could disrupt the interaction with an IC_{50} of 36 ± 4.5 µM (Figure 43) which when converted to an inhibition constant (K_i) gives 2.36 ± 0.63 µM.[93]
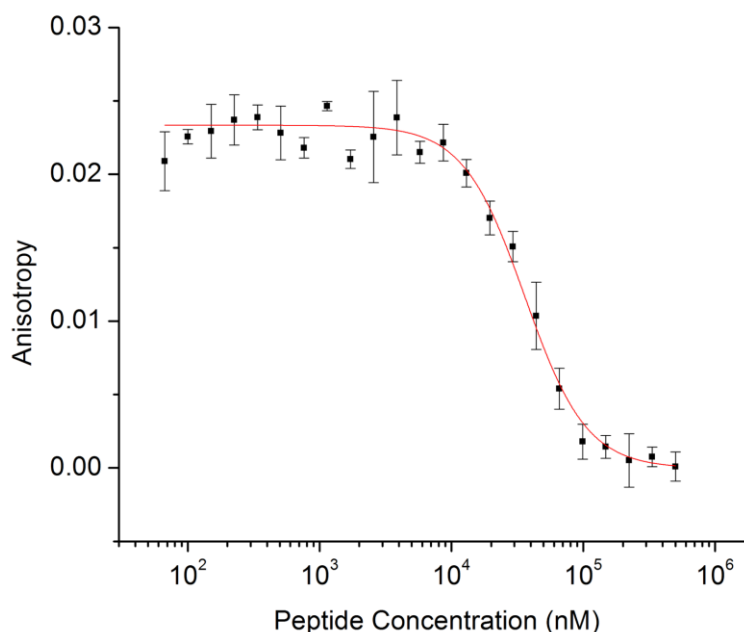
Figure 43 - Competition titration of eIF4G vs FITC-eIF4G with eIF4E followed by fluorescence anisotropy (80 nM FITC-eIF4G, 3 µM eIF4E, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

This positive control experiment validated this fluorescent anisotropy assay for the identification of inhibitors of the eIF4E/eIF4G protein-protein interaction and a further control experiment showed that the assay was uneffected by DMSO up to 20% v/v in buffer as a mock inhibitor.

4EGI-1

4E1R-Cat
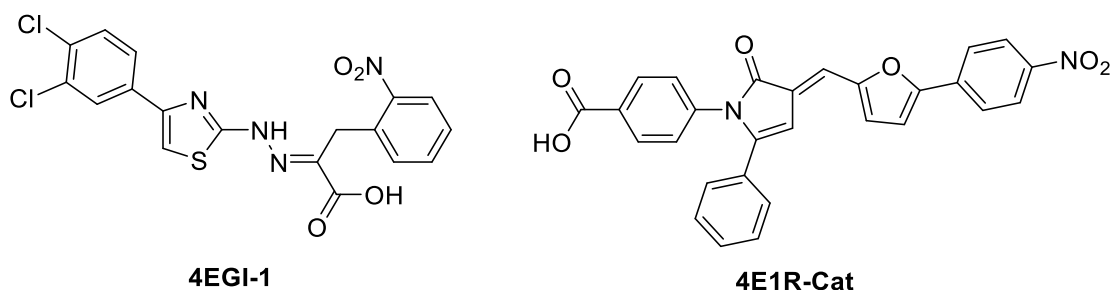
Figure 44 - Structure of known eIF4E/eIF4G inhibitors

Finally, as several compounds have been reported in the literature as inhibitors of eIF4E/eIF4G, the assay was validated by titration with known inhibitors (Figure 44).

4EGI-1[80] was found to have intrinsic fluorescence which interefered with the assay but 41R-Cat inhibitied the interaction with a IC_{50} value of 7.7 ± 0.4 µM (Figure 45) which is in good agreement with the reported value of ~4 µM.[82]
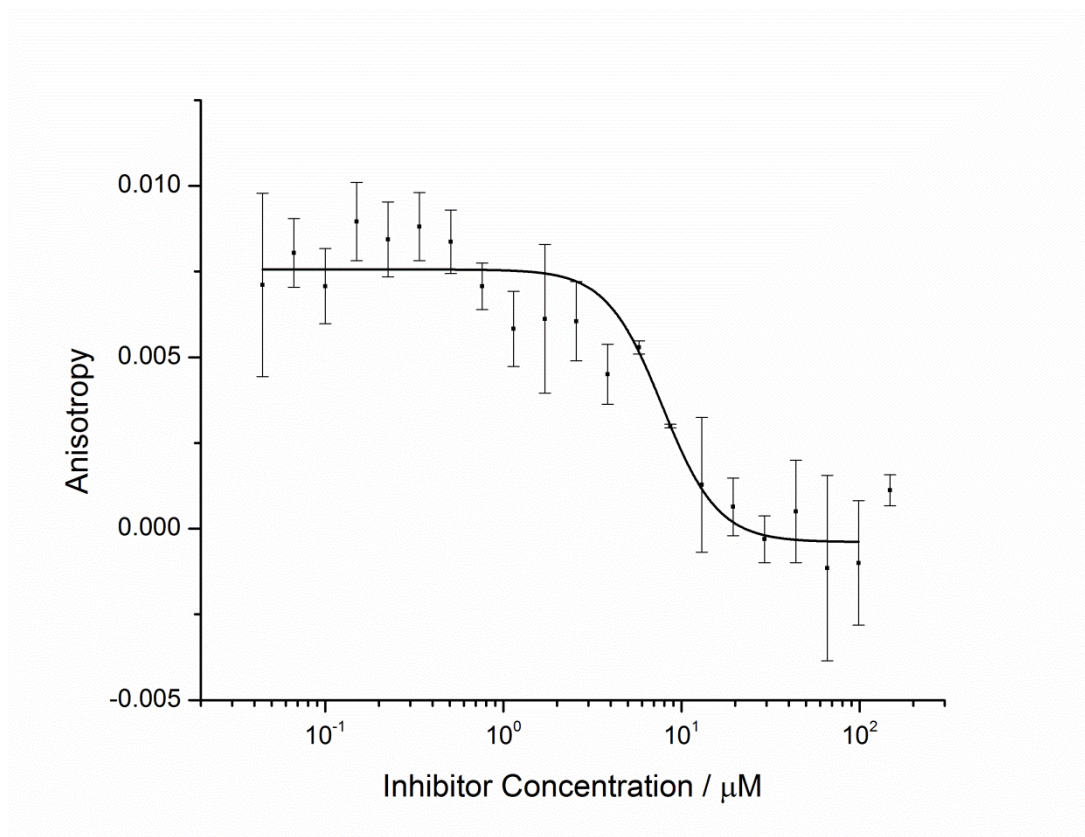
Figure 45 - Competition titration of 41R-Cat vs FITC-eIF4G with eIF4E followed by fluorescence anisotropy (80 nM FITC-eIF4G, 3 µM eIF4E, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats**.**

2.4 Summary and Conclusions

In summary, the interaction between the HIF-1α CTAD and the CH1 domain of p300 was assessed by both fluorescence anisotropy and ITC. These experiments gave a K_d of approximately 40 nM which is in good agreement across both techniques and the relevant literature.[29, 33, 108] The effects of time, temperature and DMSO co-solvent were assessed. Furthermore, an investigation into the requirements for HIF-1α/p300 binding revealed the importance of the Helix 3 region and the role of chelation.

The interaction between a peptide derived from eIF4G and eIF4E protein was assessed by fluorescence anisotropy. The K_d was determined to be approximately 1.7 µM which is in good agreement the relevant literature.[80, 109] The effects of time, temperature, protein construct, assay additives and DMSO co-solvent were assessed. A known ligand was also used to validate the assay.

Subsequently this information was utilised to develop robust fluorescence anisotropy assays for both interactions which could allow the identification of inhibitors of these PPIs (see Chapters 3 and 4).

Chapter 3 – Computational Approaches to Inhibitors

The candidate confirms that the work submitted is his 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.

Portions of the work described in this chapter were submitted as part of a research paper describing the synthesis of 1,3-amino alcohols (I. Colomer, O. Adeniji, G.M. Burslem, P. Craven, M.O. Rasmussen, A. Willaume, T. Kalliokoski, R. Foster, S. Marsden, A. Nelson, Bioorganic and Medicinal Chemistry, 2014). G.M.B. and P.C. proposed the synthetic approach, I.C., O.A. and G.M.B. performed the initial synthesis validation, M.O.R and A.W. performed the library synthesis, T.K. performed the chemoinformatics, R.F., S.M. and A.N. supervised the project. I.C. prepared a draft manuscript which was then edited into its present from (attached) by A.N. with comments from all authors. Selected examples prepared by GMB are included in this chapter.

The *de novo* compound library was enumerated by Adam Nelson (University of Leeds). Paul Faulder (AstraZeneca) performed the initial docking of the *de novo* library against p300 which was repeated by G.M.B. The X-ray crystal structure was solved by Chris Pask (University of Leeds).

3.1 Computational Chemistry in Lead Discovery

Computational drug design has become an important tool in recent years[110] and can provide valuable information for medicinal chemistry efforts. Whilst the main advantage of computational modelling has previously been ligand optimisation and structure-based drug design using X-ray structures, it has increasingly been used to identify lead compounds as a complementary method to high-throughput screening.[111]

Molecular docking of proposed ligand structures into known protein structures[112] is the most common form of computational lead discovery; either docking of commercially-available compounds or by *de novo* design of putative inhibitors.[113] Both approaches were employed in this project as discussed below.

3.2 Identification of Commercially Available Inhibitors

To allow the identification of inhibitors of the HIF-1α/p300 and eIF4E/eIF4G PPIs, a virtual library of commercially available compounds was provided by AstraZeneca. The virtual library was derived from the eMolecules collection of commercially available compounds (www.emolecules.com), from which duplicates and compounds with undesirable functionalities were removed by AstraZeneca staff. Two approaches were employed to computationally identify compounds to be purchased and tested for inhibitory activity in the assays described in Chapter 2. One approach was to employ docking in a helix accepter/protein-centric approach whilst the other was to employ a helix donor/ligand-centric approach using the known pharmacophore from the peptide/protein complex.

3.3 Docking Virtual Libraries

Prior to docking, the protein receptor must be prepared and putative binding sites identified. The proteins were prepared using the protein preparation wizard in Maestro 9.3 (Schrodinger) by removing the helix donor binding partner, defining the binding site and setting any relevant constraints, optimising the hydrogen bonding within the remaining protein, removing any non-hydrogen bonding residual water molecules and performing a restrained minimisation.[114]

3.3.1 HIF-1α/p300 Docking

For the HIF-1α/p300 system, two docking sites were prepared from the complex NMR structure (PDB ID: 1L8C)[30] using Glide (Schrodinger)[114, 115] – one comprising the helix 2 binding region, the other the helix 3 binding region.
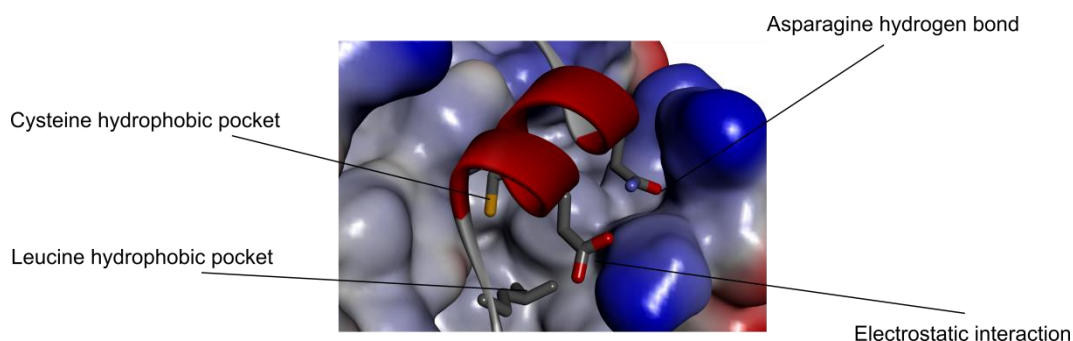
Figure 46 - Helix 2 binding region

For the helix 2 binding cleft, shown in Figure 46, the grid comprised the region in which the helical portion binds and also specifies two key interactions made between HIF-1α and p300. Ligands must interact favourably with protein, fit within the defined space and make at least one of two key interactions; a hydrogen bond made by the key asparagine residue of HIF-1α and/or an interaction with the surface lysine.

Another receptor grid was defined for the helix 3 binding region. The helix 3 binding region makes less of an obvious pocket and is more "featureless", as shown in Figure 47 and was therefore considered less tractable for docking.
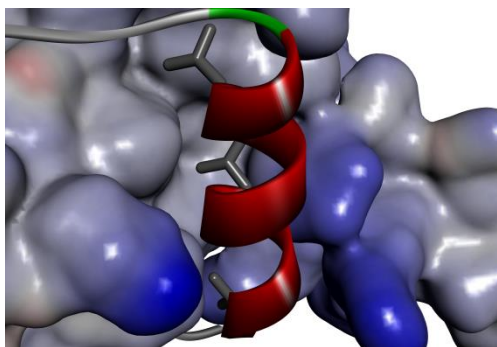
Figure 47 - Helix 3 binding region

The virtual library for docking was prepared by generating up to 200 low energy 3D conformations for each compound using Omega (OpenEye)[116] and physiologically relevant protonation states using LigPrep (Schrodinger).[117] Low energy conformers were generated in advance of the docking to allow rigid docking to be performed. This approach can give more realistic docked poses and reduces the computational time required to dock large libraries.[118]

Initial runs were performed using virtual high-throughput screening settings and the top compounds were then subjected to higher precision docking to give greater confidence in the results. Compounds were inspected visually and selected for screening based on the docking score, recapitulation of important peptide features and consistency of pose across multiple iterations. For example, if the highest scoring compound was only observed once in the top 50 hits, but another compound

was observed 5 times in similar poses in the top 20, then the latter compound was prioritised for screening (Table 3).

3.3.2 eIF4E/eIF4G Docking
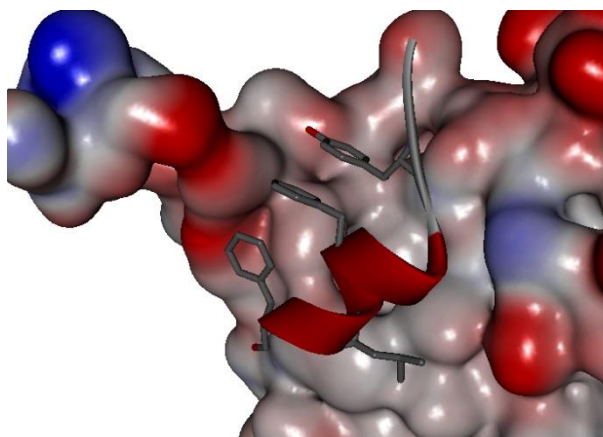
Figure 48 – eIF4G binding region

A similar approach was employed to dock into the helix binding region of eIF4E – a receptor grid was generated from the co-crystal structure (PDB ID:2W97)[68] using Glide (Schrodinger)[114, 115] A docking grid was defined based upon the location of binding of the eIF4G peptide in the crystal structure, shown in Figure 48. As there is a known inhibitor **4EGI-1**,[80] discussed in Chapter 1 and shown in Figure 49, for the eIF4E/eIF4G interaction this compound was used to validate the docking model – at the time of receptor preparation the ligand-protein complex was not known.
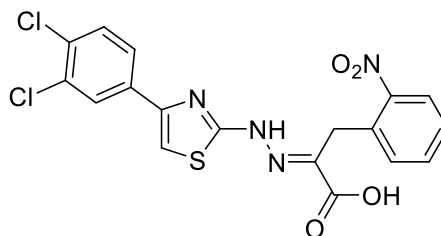
4EGI-1

Figure 49 - Structure of known eIF4E/eIF4G inhibitor, **4EGI-1** used in docking validation

To validate docking results from this system, the known ligand was docked into the system and, by incorporating a hydrophobic receptor constraint, the known ligand docked in a consistent pose which matched with the HSQC shifts reported in the literature for this compound binding to eIF4E.[80] The hydrophobic constraint ensures a non-polar region of the molecule binds into a small depression in the surface of eIF4E and is consistent with the HSQC shifts. Figure 50 shows the overlay of 10 different docked poses with relatively little difference between them and the nitro-aryl ring bound into the hydrophobic depression. This receptor grid was then used

for all subsequent docking. Unfortunately it was subsequently shown by X-ray crystallography that this inhibitor has an allosteric mode of action.[71]
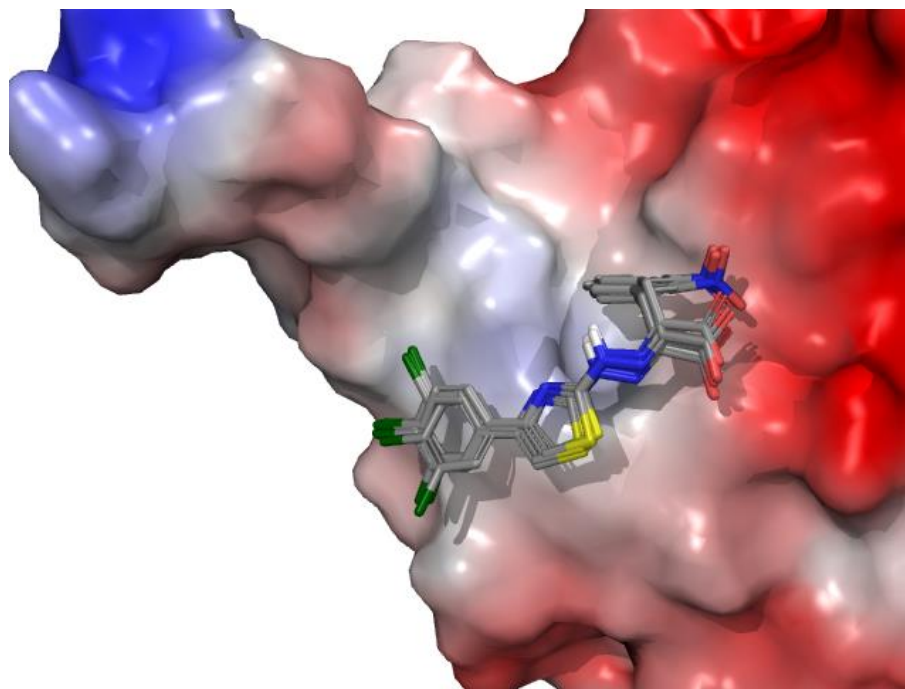
Figure 50 - Ten overlaid structures of known eIF4E/eIF4G inhibitor, **4EGI-1**

3.4 Shape-Based Screening

Another approach utilised in medicinal chemistry is pharmacophore modelling whereby a known inhibitor/binder is used to define a model of what is required to bind to the protein of interest.[119] This approach is often used in "scaffold-hopping" to switch to a different chemical structure which is still capable of making the same interactions.[120] It was envisioned that this approach could be used to identify PPI inhibitors based on the pharmacophore of the helix donor protein/peptide in the case of HIF-1α/p300 or eIF4E/eIF4G. To do this, the helical donor regions were extracted from the complex structures and used as queries for Rapid Overlay of Chemical Shape (ROCS) analysis.[121] Similar approaches have previously been applied to PPIs by the groups of Arora[8] and Burgess[122, 123].

After optimisation, it was found that the best results were achieved by excising the donor region of interest, removing side chains from non-bonding residues and converting the peptide backbone to an all carbon chain whilst maintaining the 3D structure: an example of this process for the helix 3 region of HIF-1α is shown in Figure 51. This approach allows only the desired sidechains to be included in the pharmacophore model and prevented overemphasis of the amide functionality in the backbone.
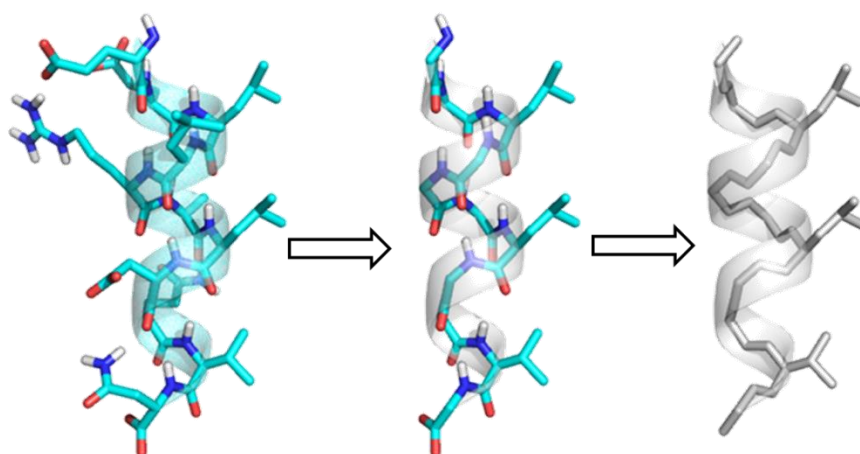
Figure 51 - ROCS query preparation. From left to right: helical region of peptide extracted from 1L8C, helical region with non-bonding sidechains removed, helical region with backbone converted to carbon chain.

This approach to query preparation was employed for helix 2 and 3 of HIF-1α and the eIF4G helix. Important contacts were also highlighted as desired features providing the structures shown in Figure 52. For helix 2 of HIF-1α, two hydrophobic residues were highlighted along with a hydrogen bond donor and acceptor analogous to the docking constraints defined above. The helix 3 query highlights the *i*, *i*+4 and *i*+7 residues thought to be crucial as key hydrophobic regions. For the eIF4G query the key hydrophobic rings identified in the literature were defined as well as a potential hydrogen bond donor/acceptor on tyrosine.
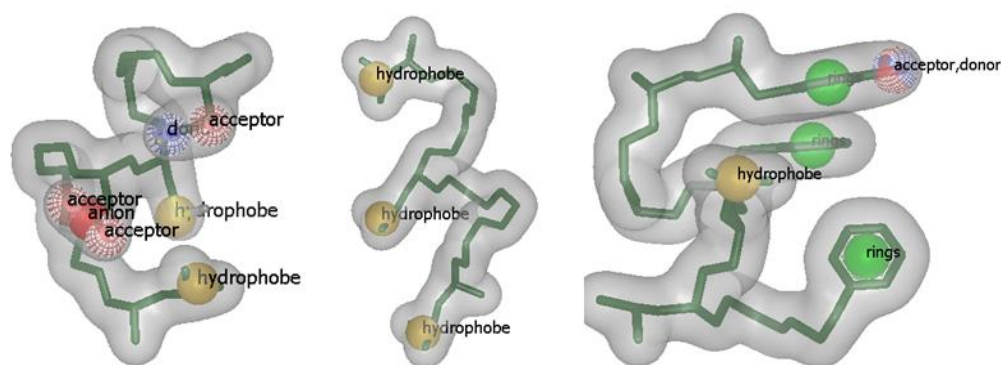
Figure 52 - ROCS queries employed in pharmacophore searching. From left to right: HIF-1α helix 2, HIF-1α helix 3, eIF4G
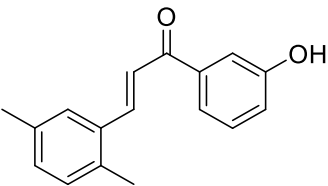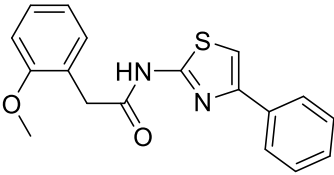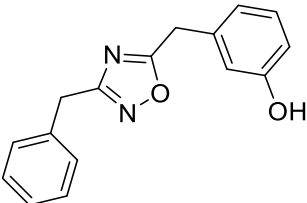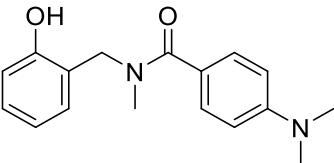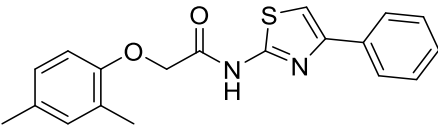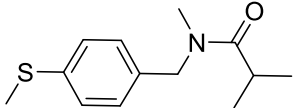
These structures were then employed as ROCS queries and screened against the commercially available eMolecules library, discussed above, to identify potential inhibitors for screening. The top 100 scoring matches for each query were visually inspected and compounds of interest selected based on their ability to match the defined pharmacophores as well as the flexibility of the ligand. Compounds with
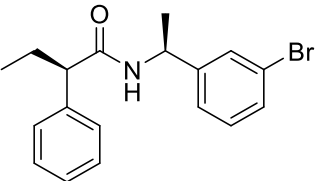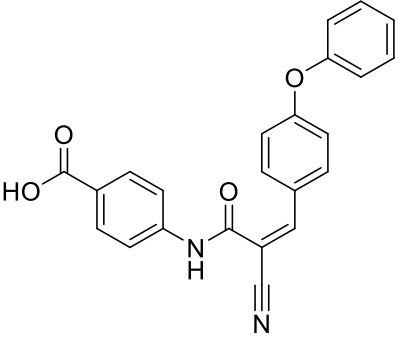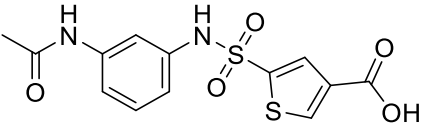
more rigid structures that matched the query well were prioritised over flexible compounds capable of matching the query (Table 3)

3.5 Compound Selection

From both of the above approaches, docking and pharmacophore searching, large sets of compounds were suggested as potential inhibitors. A small set of compounds which met the defined criteria were selected and purchased. The compound structures selected for screening along with the method and target used to select them are shown below in Table 3.

Table 3 – Compounds selected for screening from commercially available libraries

Compound	Structure	Method	Target
3		Docking	eIF4E
4		Docking	eIF4E
5		Docking	eIF4E
6		ROCS	eIF4E
7		ROCS	eIF4E
8		Docking	p300 Helix 3

9		Docking	p300 Helix 3
10		ROCS	p300 Helix 3
11		Docking	p300 Helix 2

3.6 Compound Screening

The compounds detailed in Table 3 were screened in the competition assays discussed in Chapter 2. All of the compounds were tested in both competition assays (HIF-1α/p300 and eIF4E/eIF4G) despite being selected separately to allow an assessment of selectivity.

3.6.1 HIF-1α/p300 Screening

Screening of the compounds in the HIF-1α/p300 assay identified only one compound that showed a decrease in anisotropy consistent with inhibition: compound **10**. This compound showed an apparent IC_{50} value of 67 ± 13 μM from the inhibition curve shown in Figure 53.
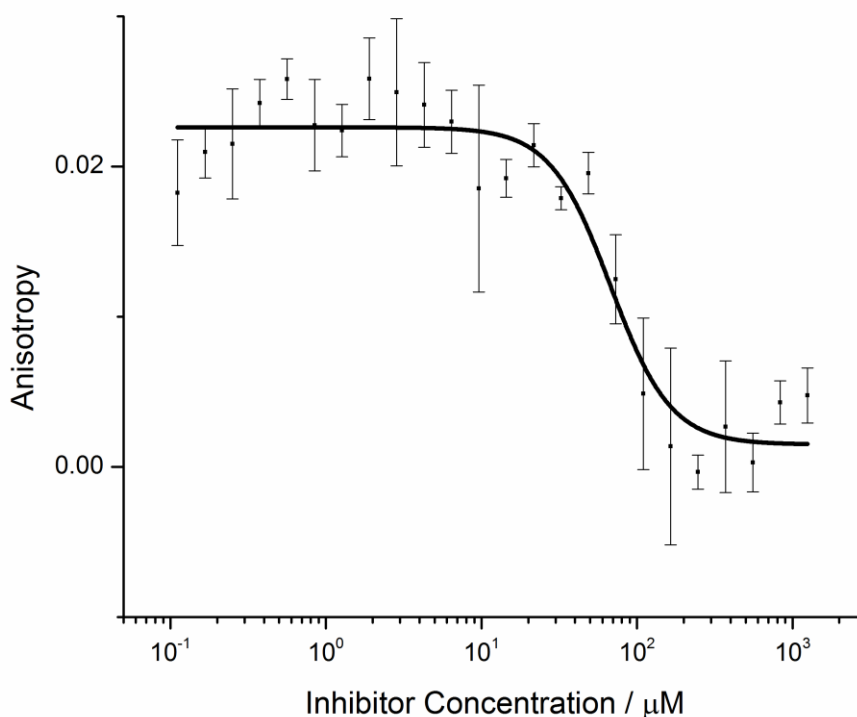
Figure 53 - Competition Assay of Compound **10** against HIF-1α/p300 (80 nM FITC-HIF-1α CTAD, 0.1 μM p300, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

Whilst initially exciting, comparison with the data for the eIF4E/eIF4G interaction for this compound showed a similar IC_{50} value. Inspection of the structure of compound **10** suggests it might possess the ability to act as a covalent inhibitor and the fact it inhibits both interactions assayed suggest it is a non-selective covalent inhibitor. To confirm this, the protein was subjected to mass spectrometry after incubation with the compound. The analysis showed a time dependent increase in molecular weight as shown in Figure 54. Prior to treatment with compound **10**, mass spectrometry of the protein gives a series of peaks representing different charge states. After 30 minutes of treatment, with compound **10**, another set of peaks appeared indicating covalent labelling of the protein whilst the initial unlabelled protein peaks are still observed. This most likely occurs through the reaction of a nucleophilic residue on p300 reacting with the Michael accepter within compound **10**.
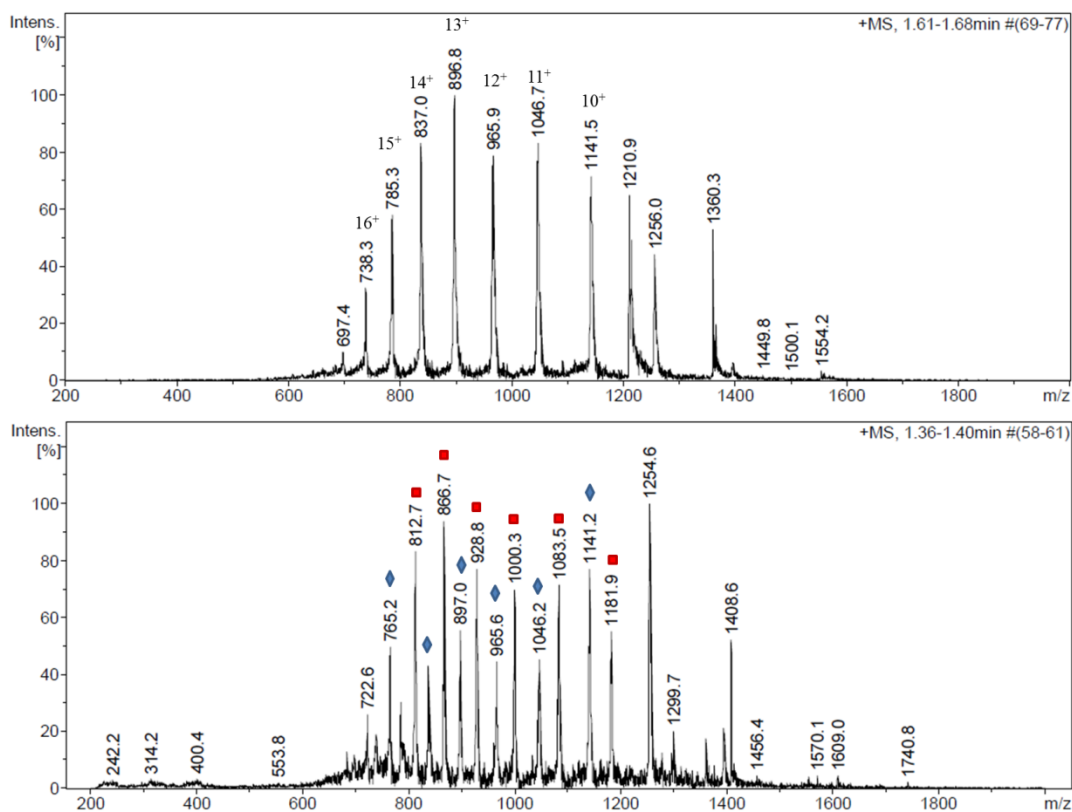
Figure 54 – Mass spectrometric analysis of p300 (100 µM) and compound **10**. Top panel – protein mass spectrum with no compound, Bottom panel – after treatment for 30 minutes with 100 uM compound **10** for 30 minutes. Blue diamonds denote protein peaks, red squares denote protein-compound complex

Having shown the only compound eliciting an effect in the HIF-1α/p300 competition assay was both non-selective and a competitive inhibitor, compound **10** was not progressed further. The approaches discussed above did not provide any other inhibitors for further investigation.

3.6.2 eIF4E/eIF4G Screening

Screening of the compounds described in Table 3 in the fluorescence anisotropy competition assay previously described, identified compound **10** (which was discarded as discussed above) and compound **4**. The assay was repeated and compound **4** was shown to have an IC_{50} of 612 ± 59 µM (Figure 55).
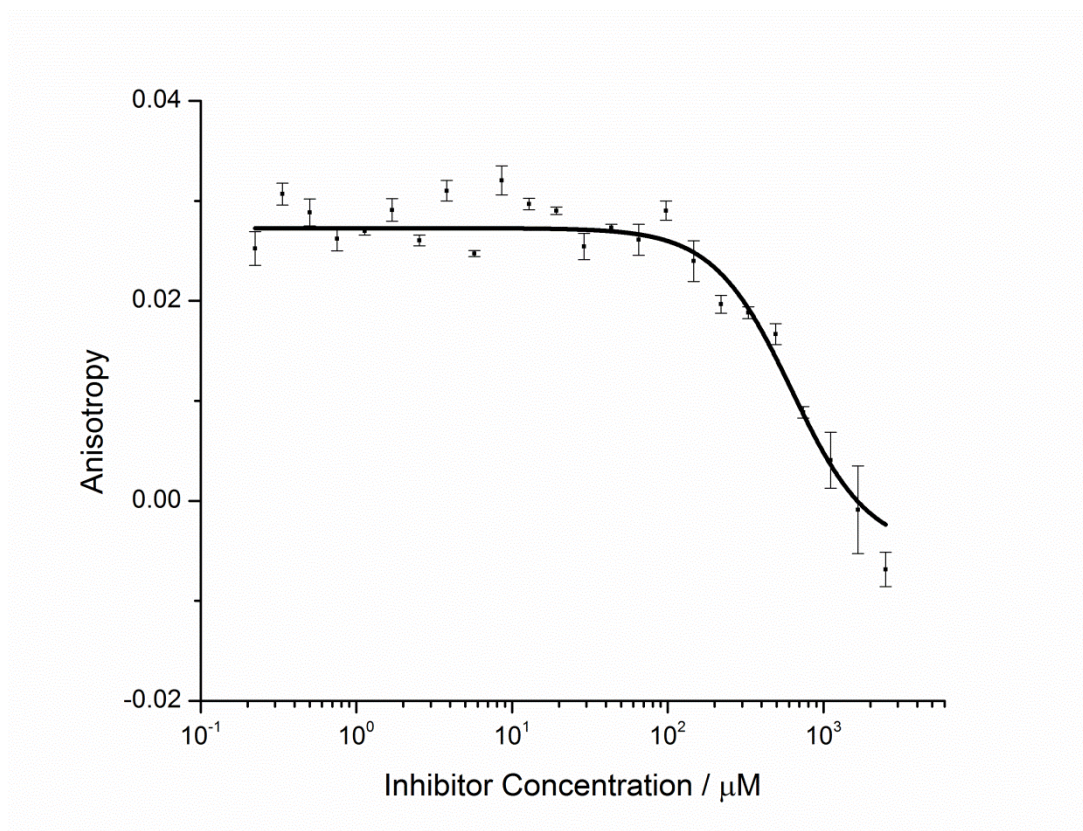
Figure 55 - Competition curve for compound **4** in the eIF4E/eIF4G competition assay. IC_{50} - 612 ± 59 μM. 80 nM FITC-eIF4G, 3 μM eIF4G, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

Whilst not a high affinity compound, its relatively low molecular weight (324 Da), and potential for development, made it a promising starting point for further elaboration and optimisation. Inspection of the docked pose for this compound, shown in Figure 56, suggested several possible modifications of compound **4** to explore the structure activity relationship.
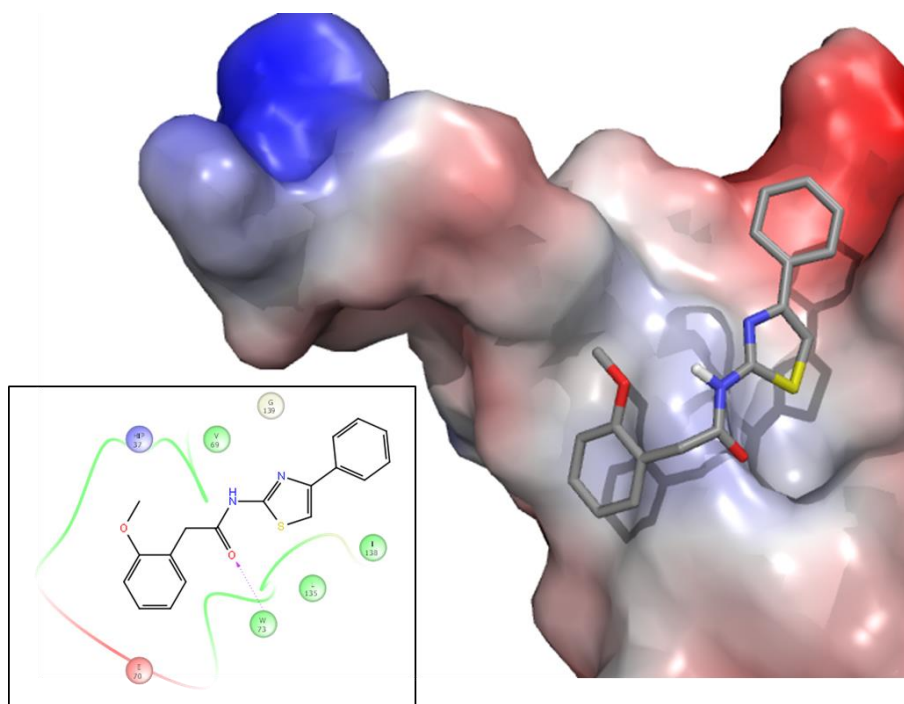
Figure 56 - Docked pose for compound **4** in the eIF4G binding domain of eIF4E with ligand interaction diagram (inset). Docking performed using Glide

Compound **4** is structurally similar to the known inhibitor **4EGI-1** (see Figure 49) and binds in a similar way during docking. This may be an artefact of the docking procedure/validation or may suggest a common pharmacophore/chemotype for eIF4E binders. Both compounds contain a central aminothiazole ring flanked by two aromatic rings. This lead to the design of a small analogue library which is summarised in Figure 57.
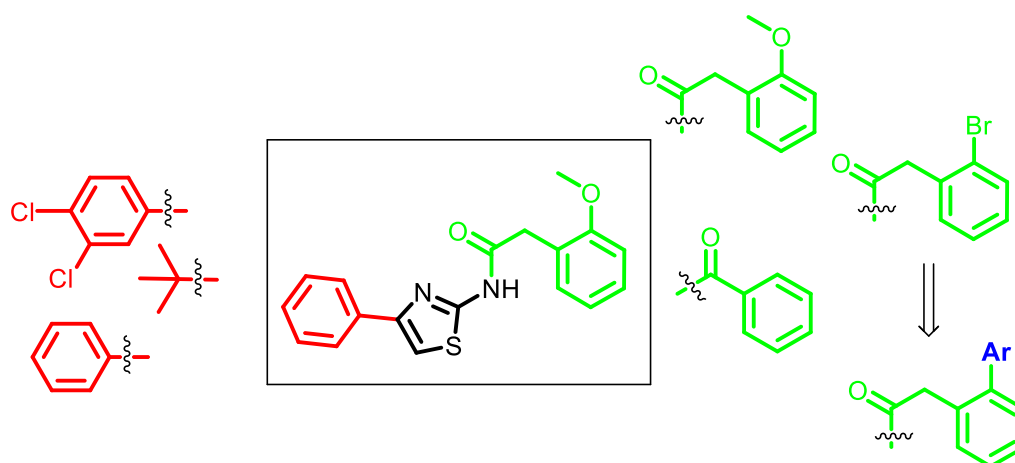
Figure 57 - Analogues of compound **4**

Initial optimisation focussed on maintaining the aminothiazole core while altering the pendant aromatic rings based on the docking and similarity to the known inhibitor. The 3,4-dichlorophenyl group was incorporated as its presence in the previously reported compounds provided an increase in activity.[124] The t-butyl

group was selected to include a more 3D hydrophobic moiety. On the other side of the compound (shown in green) the length of the linker to the aromatic ring was varied and the methoxy group replaced with a bromide. The bromide was included both as a potential inhibitor and to allow the introduction of additional aromatic functionality (blue) via cross-coupling reactions.

3.6.3 Analogue synthesis

A small series of compounds based on the design outlined in Figure 57 was prepared using the route shown in Scheme 1. The aminothiazoles **14-16** were prepared through the reaction of α-haloketones **11-13** and thiourea either under thermal conditions (for aryl ketones[125]) or *via* copper catalysis (for the t-butyl variant, **16**).[126] The resulting aminothiazoles then underwent amide bond forming reactions, in this case carbodiimide mediated coupling or reaction with an acid chloride. Finally, the compound **18** containing an aryl bromide were further derivatized by Suzuki-Miyaura cross coupling[127] reactions to introduce an additional aromatic ring. The library is summarized in Table 4.

11 R^1 - Ph
12 R^1 - 3,4-dichlorophenyl
13 R^1 - t-butyl

14 R^1 - Ph 76%
15 R^1 - 3,4-dichlorophenyl 76%
16 R^1 - t-butyl 66%

See Table 4
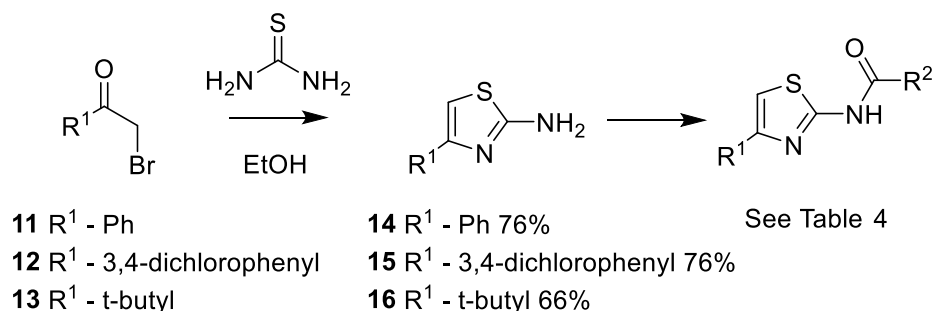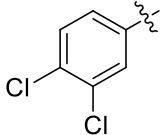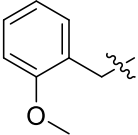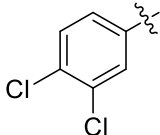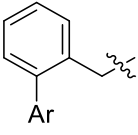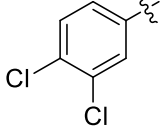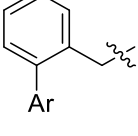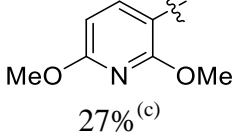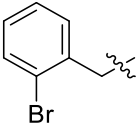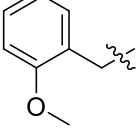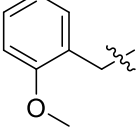
Scheme 1 - Preparation of aminothiazoles

Table 4 – Compound 4 analogue library prepared
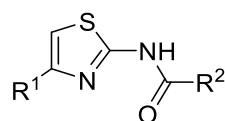
Compound	R^1	R^2 yield	Ar yield
17		52 %$^{(a)}$	n/a
18			n/a

		60% [b]	
19	Cl, Cl substituted phenyl	ortho-methoxy benzyl, 36% [b]	n/a
20	Cl, Cl substituted phenyl	Ar-substituted benzyl, 60% [b]	phenyl, 41% [c]
21	Cl, Cl substituted phenyl	Ar-substituted benzyl, 60% [b]	MeO—N—OMe pyridyl, 27% [c]
22	phenyl	ortho-Br benzyl, 41% [b]	n/a
23	phenyl	ortho-methoxy benzyl, 18% [b]	n/a
24	tert-butyl	ortho-methoxy benzyl, 86% [b]	n/a

(a) prepared from amine and benzoyl chloride (1.1 eq) (b) prepared from amine and carboxylic acid (1.1 eq.), N-(3-Dimethylaminopropyl)-N′-ethylcarbodiimide (2 eq.), DMAP (10 mol%) and diisopropylethylamine (5 eq.) (c) prepared from bromide and Ar-B(OH)$_2$ (1 eq.), Pd(PPh$_3$)$_4$ (5 mol%) and caesium carbonate (3 eq.)

Compounds **17-24** were then analysed for their ability to inhibit the eIF4E/eIF4G compared to the initial hit compound. Unfortunately all of the analogues prepared had IC$_{50}$ values greater than 1 mM against eIF4E/eIF4G. Given this lack of improvement no further work was performed on this compound series. It should also

be noted that 2-aminothiazoles have been reported as potentially problematic compounds in screening collections – 2-aminothiazoles are included on lists of pan assay interference compounds (PAINS)[83] and have been found to be promiscuous in biophysical assays.[128]

3.7 Design of Bespoke Compound Libraries

In parallel to screening libraries of commercially available compounds to provide a starting point for inhibitor optimisation, a *de novo* approach was employed.[113] This approach was only employed for the helix 2 binding site on p300 due to the amount of computational time required. The helix 2 binding site was selected as it appeared to be the most amenable to docking.

Initially, a virtual library of compounds, previously enumerated within the Nelson group, containing 80,000 diverse compounds derived from reliable chemistry developed or validated within the group and commercially available starting materials was used. This library was docked against the helix 2 binding site (performed by Paul Faulder, AstraZeneca). Examples of the best docked compounds are shown in Figure 58. From these virtual high-throughput hits a series of compounds were identified for further analysis.
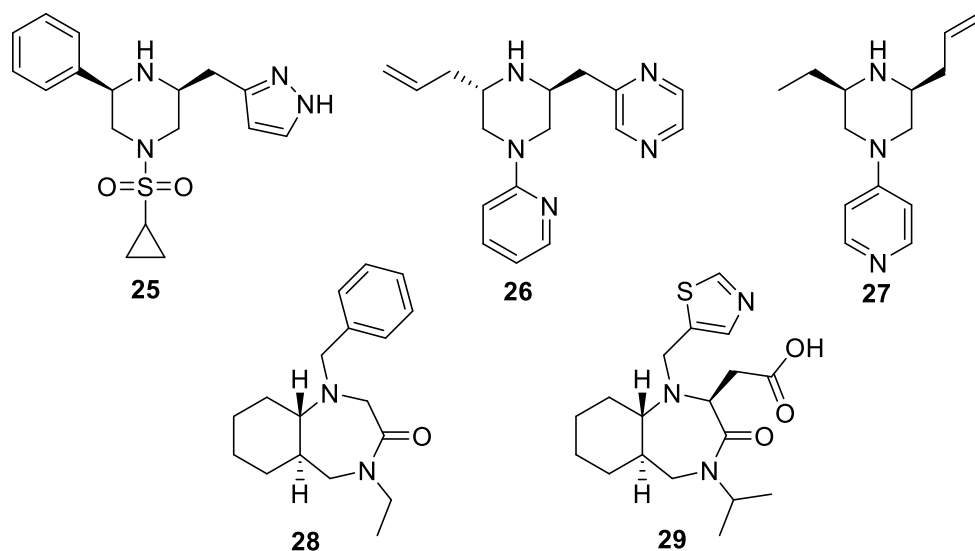
Figure 58 - Initial hits from docking of *de novo* library against helix 2 binding region

The bicyclic system, **28** and **29**, was selected for further study as it was predicted to make the same contacts as the peptide in the desired region as shown in Figure 60. To enable further computational studies, a focussed virtual library based on the scaffold shown in Figure 59 was enumerated. The virtual library was enumerated using Pipeline Pilot[129] by programming virtual reactions and providing readily

available building blocks to yield compounds of the type shown in Figure 59 (full protocol and building block libraries are included in the experimental).
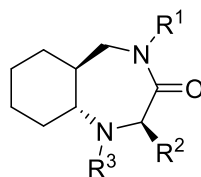
Figure 59 – Core structure of enumerated bicyclic compound library.

The enumerated libraries were then docked against the helix 2 binding region and the top 100 structures inspected. Examples of some of the best docked structures are shown Figure 60.
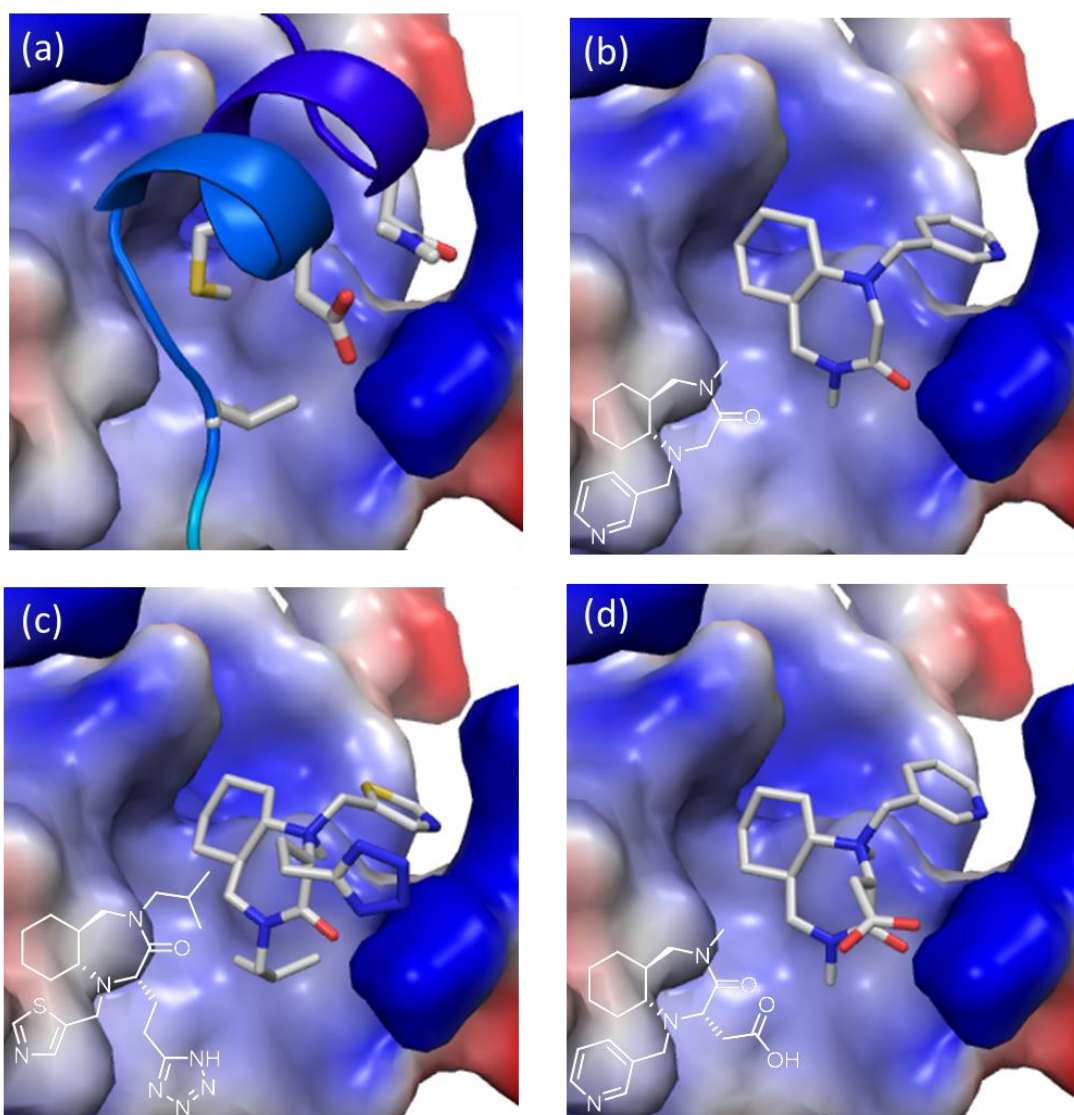
Figure 60 - Example docked structures of *de novo* designed compounds in the helix 2 binding site of p300. (a) helix 2 in the binding (b-d) docked compound structures

As the bicyclic 6,7 cis-fused ring system consistently docked with a conserved pose, it was decided to prepare selected analogues. An examination of the literature also revealed that this fused ring system containing both an amine and an amide had not been previously reported. Given the similarity in structure to benzodiazepines[130] and the frequency at which nitrogen heterocycles appear in pharmaceuticals,[131] this was surprising and provided further encouragement to develop synthetic routes to these compounds.

3.7.1 Synthesis of selected scaffold

After several routes were explored with limited success (shown in Appendix II), the route shown in Scheme 2 successfully provided the scaffold. This route exploited the aminomethyl-hydroxylation reaction initially reported by Katritsky[132] between a cyclohexene **30** and *N*-hydroxymethylbenzamide **31** to yield the cis-oxazine **32**. Reductive ring opening of the oxazine **32** with lithium aluminum hydride provided the 1,3-aminoalcohol **33**. [133]
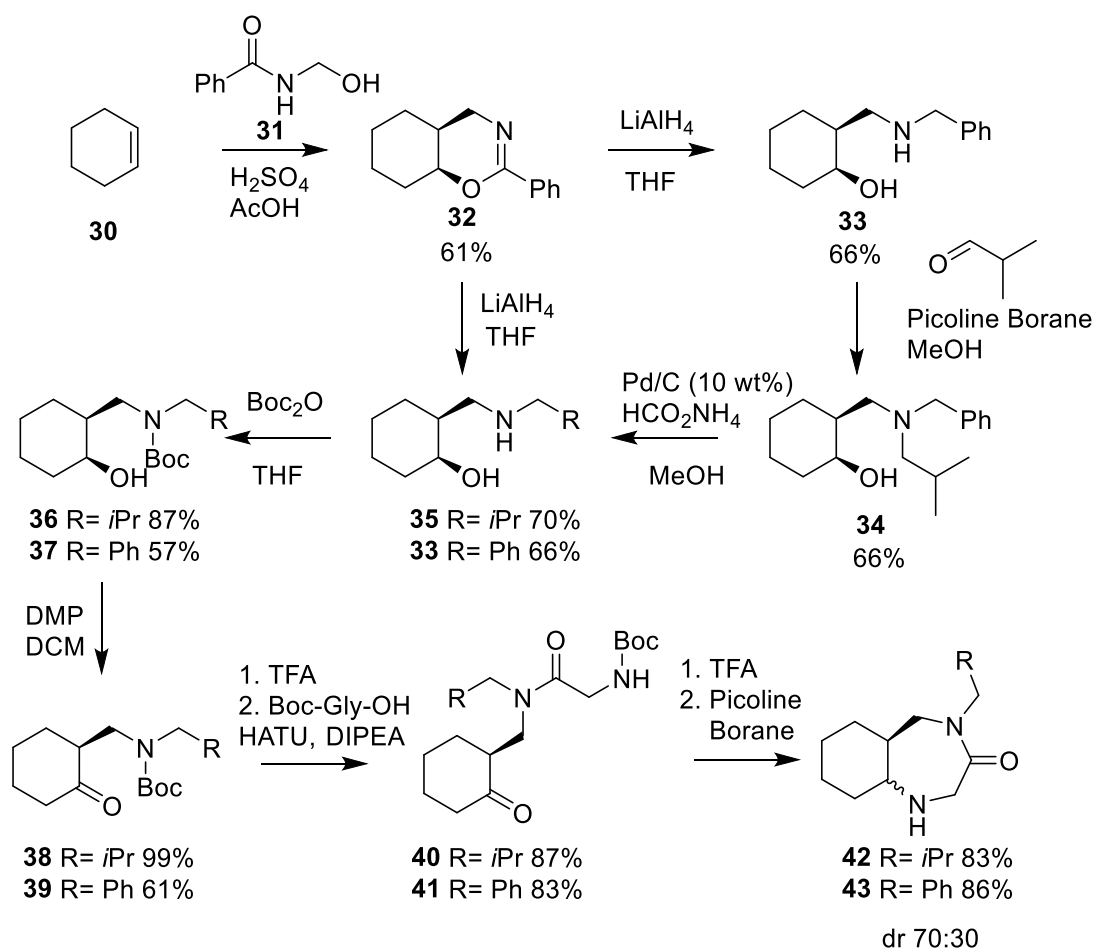
Scheme 2 – Synthetic route to bicyclic target compound

The *N*-benzyl 1,3-aminoalcohol **33** was then carried on to final compounds containing benzyl functionality or the alkyl group exchanged via reductive amination to **34** followed by debenzylation to give **35**. Efforts were also made at this stage to convert the alcohol into a masked amine via Mitsunobu reactions with various nucleophiles – converting the compound to the desired trans compound. Unfortunately under Mitsunobu conditions the alcohol was prone to eliminate to the cyclohexene rather than undergo the desired substitution.[134]

The desired alkyl-amino alcohol **33/35** was protected as the t-butyl carbamates (Boc) **36/37** before oxidation of the alcohol with Dess-Martin periodinane[135] to give cyclohexanones **38/39**. The protection of the amine was not necessary for the oxidation to proceed but it was found to improve yields and ease purification. The Boc group was removed with TFA and Boc-glycine coupled onto the resulting secondary amines in one operation to provide **40/41**. Finally the glycine moiety was deprotected and underwent reductive cyclisation to give the desired scaffolds **42/43** as a mixture of separable diastereomers. A crystal structure confirmed the identity of the *cis* isomer (Figure 61 and Appendix III). An investigation of reducing agents was also performed to assess control of diastereoselectivity (Appendix IV). Finally the separated diastereomers were functionalised on the secondary amine via reductive amination with various aldehydes mediated by picoline borane **44-50** or by reaction with sulfonyl chlorides **51** as shown in Table 5.
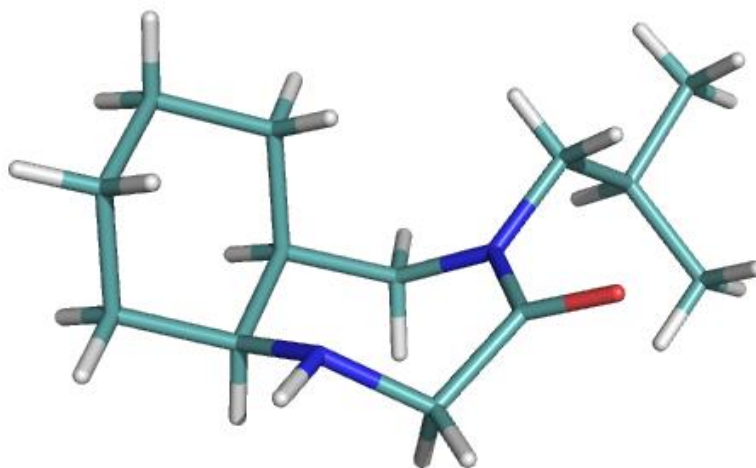
Figure 61 - X-ray crystal structure of compound *cis*-**42**

3.7.2 Screening of Designed Compounds

The library of designed compounds was screened in the HIF-1α/p300 fluorescence anisotropy competition assay and their IC_{50} values determined as shown in Table 5.

Unfortunately, none of these compounds were particularly active with all but one having IC_{50} values greater than 1 mM.

Table 5 - Bicyclic compounds prepared from computationally designed library
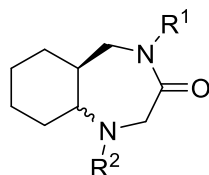
Compound	R^1	R^2	Diastereomer	$IC_{50}/\mu M$
44			cis	>1000
45			trans	>1000
46			cis	689 ± 125
47			trans	>1000
48			cis	>1000
49			cis	>1000
50			cis	>1000
51			cis	>1000

Compound **46** shows a decrease in anisotropy consistent with the HIF-1α CTAD, albeit at much higher concentrations, suggesting it is able to inhibit the HIF-1α/p300 interaction with an IC_{50} value of 689 ± 125 (Figure 62). Whilst this potency is modest, compound **46** provides an potentially very exciting starting point for inhibitor design. The low molecular weight (349 Da) and relatively high degree of

aqueous solubility make it very attractive. Indeed, this compound falls within the ideal "lead-like" parameters (cLogP between -1 and 3, number of heavy atoms 14-26).[136] A considerable amount of optimisation will be required to increase the potency of compound **46** and elucidate its mode of action.
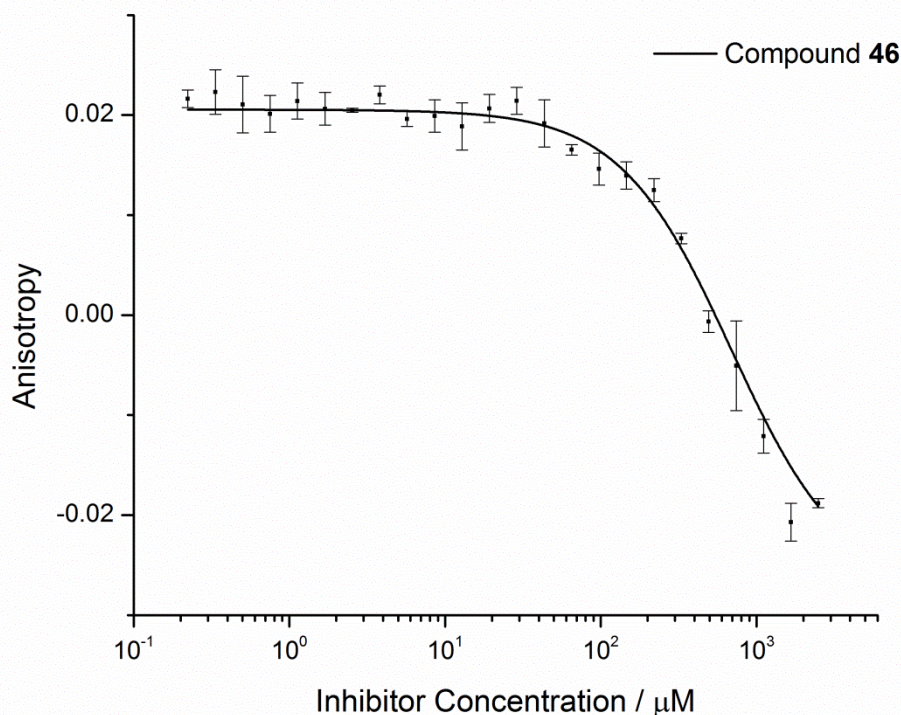
Figure 62 - Competition Assay of Compound **46** against HIF-1α/p300 (80 nM FITC-HIF-1α CTAD, 0.1 µM p300, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

3.8 Summary and Conclusions

The approaches to inhibitor identification discussed in this Chapter failed to identify high affinity inhibitors of the interactions of interest. This may have been a result of the relatively small numbers of compounds selected for screening. Similar approaches have previously been employed successfully to identify inhibitors of PPIs[8, 137] but only weak binders or covalent inhibitors were identified in this work. Attempts to optimise compound **4** were unsuccessful. It is also unfortunate that the small molecule inhibitor **4EGI-1** used in docking grid construction for eIF4E was later shown to act allosterically.[69]

However, the bespoke compound library provided a potential starting point for inhibitor development by identifying a novel scaffold with weak inhibitory activity against HIF-1α/p300. Further work will focus on the preparation of larger libraries

based on compound **46** to optimise affinity and use of orthogonal assays to confirm the activity of compound **46**.

Chapter 4
Development of Oligobenzamide α-Helix Mimetics as Inhibitors of the HIF-1α/p300 protein-protein interaction

The candidate confirms that the work submitted is his 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.

Elements of the work presented in this chapter formed the basis for a research article that was accepted for publication in February 2014 (G.M. Burslem, H.F. Kyle, A.L. Breeze, T. A. Edwards, A. Nelson, S.L. Warriner and A.J. Wilson, ChemBioChem, 2014). The contributions of the authors were as follows: GMB (the candidate) performed all synthetic chemistry, peptide synthesis and modelling, HFK performed all protein expression, GMB and HFK performed all assays jointly. ALB, TAE, AN, SLW and AJW supervised the project. GMB prepared a draft of the manuscript which was then edited into its present form (attached) by all authors. The draft originally prepared by GMB was then rewritten as portions of this chapter.

A review based on the synthesis of oligobenzamides as α-helix mimetics was also published by the candidate in January 2014 (G.M. Burslem and A.J. Wilson, *Synlett*, 2014, **25**, 324). GMB prepared a draft of the manuscript which was then edited into its present form (attached) by AJW. Portions of the original draft have been rewritten and included in this chapter.

4.1 The Concept of Helix Mimicry

The most common secondary structure in proteins is the α-helix and as such α-helices are often found at the interface between proteins.[6] Therefore, there has been significant interest in designing compounds which mimic α-helices as potential inhibitors of Protein-Protein Interactions (PPIs). These compounds, often referred to as proteomimetics, are attractive because truncated versions of the native peptides may or may not be active due to propensity of short peptides to form random coils rather than more defined secondary structures due to the entropic penalties involved. Even if they prove to be active, peptides may not be absorbed into cells and they can be prone to enzymatic degradation.
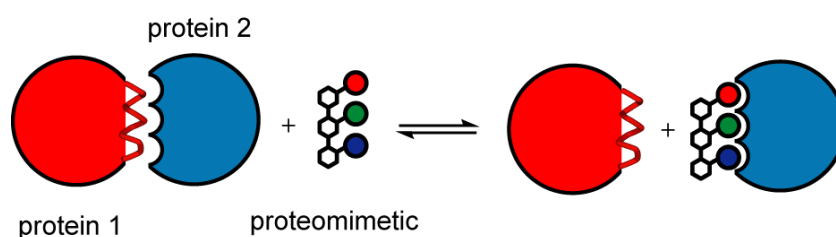
Figure 63 - Schematic representation of the proteomimetic concept

Therefore, rigid small molecules capable of recapitulating the vectoral representation of side chains of an α-helical peptide have been sought. Several different scaffolds have been developed as α-helix mimetics which present the side chain residues of one face of the helix – the i, $i+4$ and $i+7$ residues. These scaffolds are sometimes referred to as foldamers – oligomers that have well defined conformations.[138] The first scaffold was the Hamilton group's terphenyl structure which has been shown to be efficacious as a PPI inhibitor.[139] Since then, there have been several other scaffolds exemplified in the literature including oligo-benzamides[87], terephthalamides,[140, 141] enaminones[142] and benzoyl ureas[143] the structures of which are shown in Figure 64. The control of the conformation generally arises from intramolecular hydrogen bonding. The R groups depicted can be altered to present side chains which are crucial for binding depending on the targeted interaction.
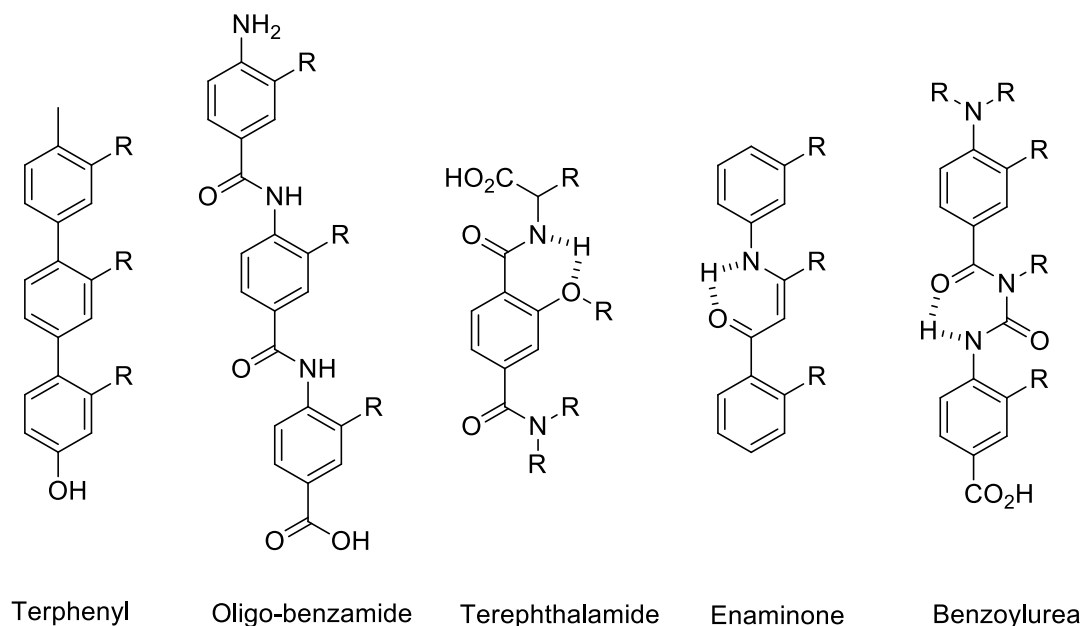
Figure 64 - Structures of foldamer helix mimetics

Two main classes of oligo-benzamide helix mimetics have been developed: *O*-alkylated[144] analogues in which the side chain mimic is tethered to the scaffold via a pendant oxygen and *N*-alkylated[145] analogues tethered via the amide nitrogen (Figure 65).The *O*-alkylated oligo-benzamides have 2 possible regioisomers depending on the building blocks used to prepare them.[146] They also employ intramolecular hydrogen bonds to promote the desired conformation (Figure 65).[147] *N*-Alkylated oligobenzamides contain no such intramolecular hydrogen bonding capability; and as such the lowest energy conformation is the *cis*-amide but they have been shown to mimic helices, presumably by adopting the *trans*-amide confirmation when bound to a protein. Recently, N-alkylated oligobenzamides have been shown to be potent cellular inhibitors of PPIs.[148] *S*-alkylated oligobenzamide analogues have been prepared and used as antibacterial compounds[149] and C-alkylated oligo-benzamides have been prepared and shown to act as inhibitors of p53/*h*DM2[150] and KSHV protease dimerization.[151] Oligo-benzamides are also advantageous as the synthesis is modular and can be performed on solid phase.[152, 153] Variations of the oligo-benzamide scaffold exist such as the oligo-pyridylamide scaffold which exhibits further hydrogen bonding with the ring nitrogen.[97]
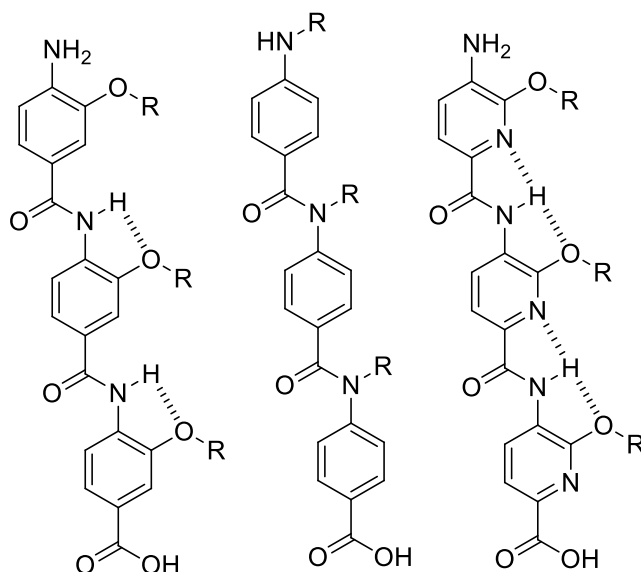
Figure 65 - Structures of 3-*O*-alkylated oligo-benzamides, *N*-alkylated oligo-benzamides and the oligopyridylamides

Crystal structures of the oligo-pyridylamides show that the bifurcated hydrogen bond between the ring nitrogen and the amide N-H imparts a curve to the backbone – a deviation from the rod like structure of oligo-benzamides.[146, 147] However the conformational stability of the oligo-pyridylamides is greater than that of the oligo-benzamides and as such there is proposed to be a lower entropic penalty for binding. Due to this entropic advantage, the oligo-pyridylamides are more effective inhibitors of amyloid fibril formation than the corresponding analogous oligo-benzamides.[154]

Terephthalamides were developed as an easily-accessible analogue of the terphenyls which can be synthetically challenging.[140] They also have an advantage over the terphenyls because of the conformational influence of hydrogen bonding. Terephthalamides have been used to inhibit the interaction between Bak and Bcl-x_L, a target implicated in cancer.[141]

Enaminones were also developed to replace terphenyls; again they have intramolecular hydrogen bonds to control conformation and they are also more polar, aiding the solubility, due to the replacement of the central phenyl ring.[142] Several studies on the synthesis and conformation of enaminones as helix mimetics have been published but no biological data has been published so far.[155, 156]

Benzoyl ureas were developed as a synthetically more accessible and water soluble alternative to terphenyls.[143] They have a slightly shorter distance between the side chains and also potentially allow the mimicking of more than the just *i*, *i*+4 and *i*+7 residues.[6] Benzoyl urea foldamers have also been used to inhibit the Bak/Bcl-x_L interaction[98] - displaying higher affinity than the binding region of the native peptide.

Another family of important helix mimetics are β-peptides;[157] peptides comprising β amino acids – analogues of α-amino acids with an extra methylene between the acid and amine groups. β-peptides adopt a similar structure to a helix based on regular α amino acids but are more stable to proteolytic digestion.[158] These extended peptides are expected to have similar problems to α-peptides with absorption.[159] However, there are examples of cell permeable β-peptides that inhibit the p53/hDM2 interaction.[160]

Although not strictly peptidomimetics, constrained peptides should also be considered in the context of helix mimicry. Constrained peptides are usually based on the native amino acid sequence but are held in a helical conformation thus negating the entropic penalty for binding by pre-organising the molecule.[161] Constraints can also impart resistance to proteolysis by preventing unfolding.[162] Peptides can also be constrained in other conformations.[163]

These molecules are generally constrained by introducing a covalent link between the residue aligned vertically – the i to $i+4$ or the i to $i+7$ residues.[6] The previously-discussed (Chapter 1) hydrogen bond surrogates (HBS) introduce a covalent link to replace the hydrogen bond between the i carbonyl and the $i+4$ amide N-H as shown in Figure 66.[164] The HBS approach has been successfully employed for HIF-1α/p300[54, 55] and Bak/Bcl-x_L interactions.[162]

α−Helix   Hydrogen Bond Surrogate   Stapled peptide
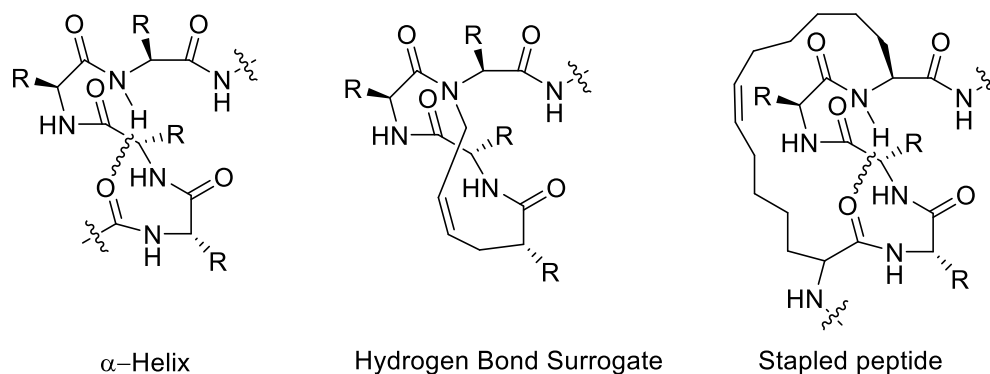
Figure 66 - Structures of constrained peptides

Stapled peptides maintain the hydrogen bond but introduce conformational stability by linking the side chains of two unnatural amino acid residues as shown in Figure 66.[165, 166] These links are usually introduced by ring closing metathesis of two appropriately spaced alkenyl amino acid residues. There are several examples and variants[88] of stapled peptides in the literature including inhibitors of the Bcl-2/BID interaction which promote the apoptosis of leukaemia cells in mice.[167]

4.2 Application to the HIF-1α/p300 Protein-Protein Interaction

HIF-1α forms a key PPI with the protein p300 which leads to the hypoxic response cascade, as discussed in Chapter 1.[168] Upon inspection of the NMR structure of the HIF-1α/p300 complex,[29] it was noted that it contains helical binding regions with the reported key residues [33] displayed on one face of the α-helices, making them an obvious target for inhibition using appropriately functionalised α-helix mimetics.

Having developed an assay to allow the screening of inhibitors of this PPI and identified a potential binding site for inhibitors, putative inhibitors were designed based on the amino acid sequence and secondary structure of the Helix 3 region of HIF-1α.

For this work, the 3-*O*-alkylated oligo-benzamide scaffold was chosen due to:

(i) The robust modular synthesis which is amenable to rapid combinatorial library generation either in solution[169, 170] or on solid support[152, 153]

(ii) The ability to adopt the desired conformation in solution[89, 144]

(iii) Proven ability to inhibit α-helix mediated PPIs.[87, 89, 171]

4.2.1 Helix 3 Mimicry

Initially two compounds, shown in Figure 67 were designed to mimic the *C*-terminal helix of HIF-1α. The compounds were based on the previously reported modular oligo-benzamide PPI inhibitors.[144]

Figure 67 – Rational design of helix 3 mimics. From left to right: Helix mimetic 3D structure, Helix 3 excised from the NMR structure of p300 in complex with the C-terminal transactivation domain of HIF-1α (PDB ID: 1L8C), superimposed 3D structures and chemical structures of designed oligobenzamide mimics

Compounds **1** and **52,** shown in Figure 67, recapitulate the Leu, Leu, Val side chains in both the matched and mismatched *N* to *C* direction with respect to the helix dipole. Modelling studies have shown that both the parallel and anti-parallel mimics are potentially effective mimics.[172]

4.2.1.1 Monomer Preparation

Before the oligo-benzamides could be assembled in a modular fashion, the monomers shown in Figure 68 had to be prepared.
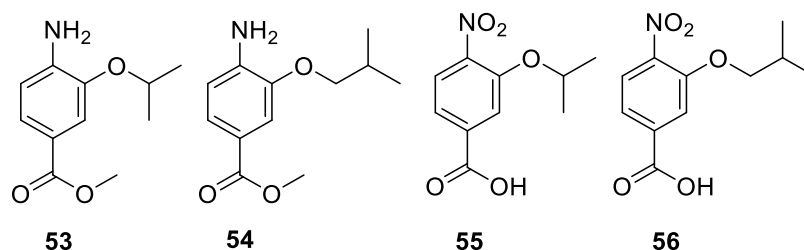
Figure 68 - Structures of monomers

Initially, commercially available 3-hydroxy-4-nitrobenzoic acid **57** was protected as the methyl ester **58** by treatment with *p*-toluenesulfonic acid in methanol, and then the appropriate side chain was introduced by alkylation. Alkyl side chains were introduced using the commercially available bromides as shown in Scheme 3.
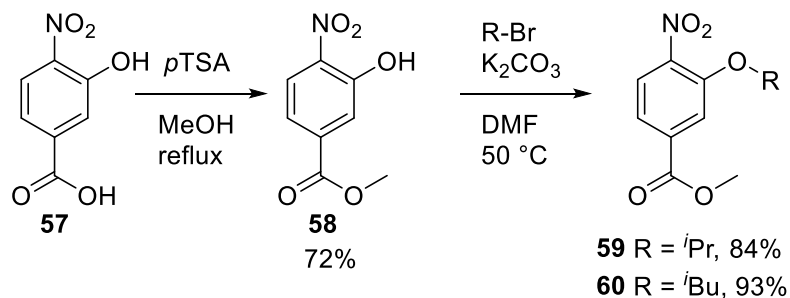
Scheme 3 – Preparation and alkylation of methyl 3-hydroxy-4-nitrobenzoate, **58**

As both the amine coupling partners and the acid coupling partners were required for synthesis of the oligo-amides the material resulting from each alkylation step was split into two portions (Scheme 4). For a portion of the material, the nitro group was reduced to the amine whilst for the remainder the methyl ester was hydrolysed to the acid. The saponification was simply performed using aqueous sodium hydroxide in methanol and the reduction was performed with stannous chloride.
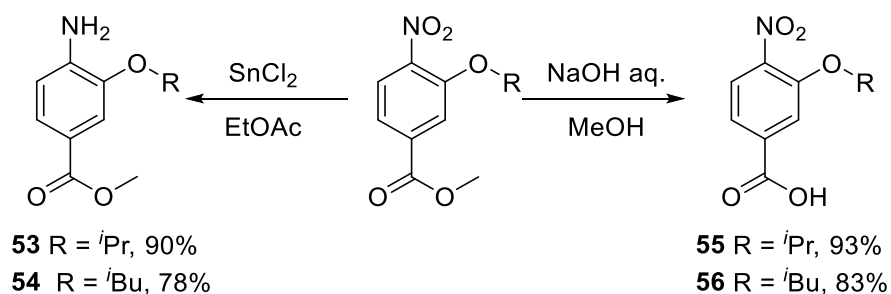
Scheme 4 - Monomer differentiation via either reduction or ester hydrolysis.

4.2.2.2 Modular Assembly of Oligo-Benzamides

The desired oligo-benzamides were constructed using iterative cycles of amide bond formation and reduction. Once the desired length is obtained, a final reduction and methyl ester hydrolysis affords the desired compound as shown below in Scheme 5.

The couplings between the derivatized benzoic acids and anilines are more difficult to achieve than most amide bond formations due to the low nucleophilicity of the aniline. Standard amide bond forming reagents, such as those employed in peptide synthesis, often return only starting materials. However formation of acyl chlorides, either by pre-activation or *in situ* formation followed by subsequent attack of the aniline is generally more successful.

Initially, couplings were attempted, with moderate success, using dichlorotriphenylphosphorane, which forms the acid chloride *in situ*.[173] In light of the modest yields, other conditions were investigated but the initial phosphorane coupling proved to be the best. The use of Mukaiyama's reagent[174] consumed the starting materials but none of the desired product was observed by crude NMR. Preforming an acyl chloride using thionyl chloride followed by reaction with an aniline[154] gave none of the desired product by crude NMR. The dehydration reaction using PCl$_3$ under high temperature microwave reaction conditions[175] gave ~5% conversion by crude NMR. Therefore the desired oligobenzamides **1** and **52** were successfully prepared using the dichlorotriphenylphosphorane coupling conditions.
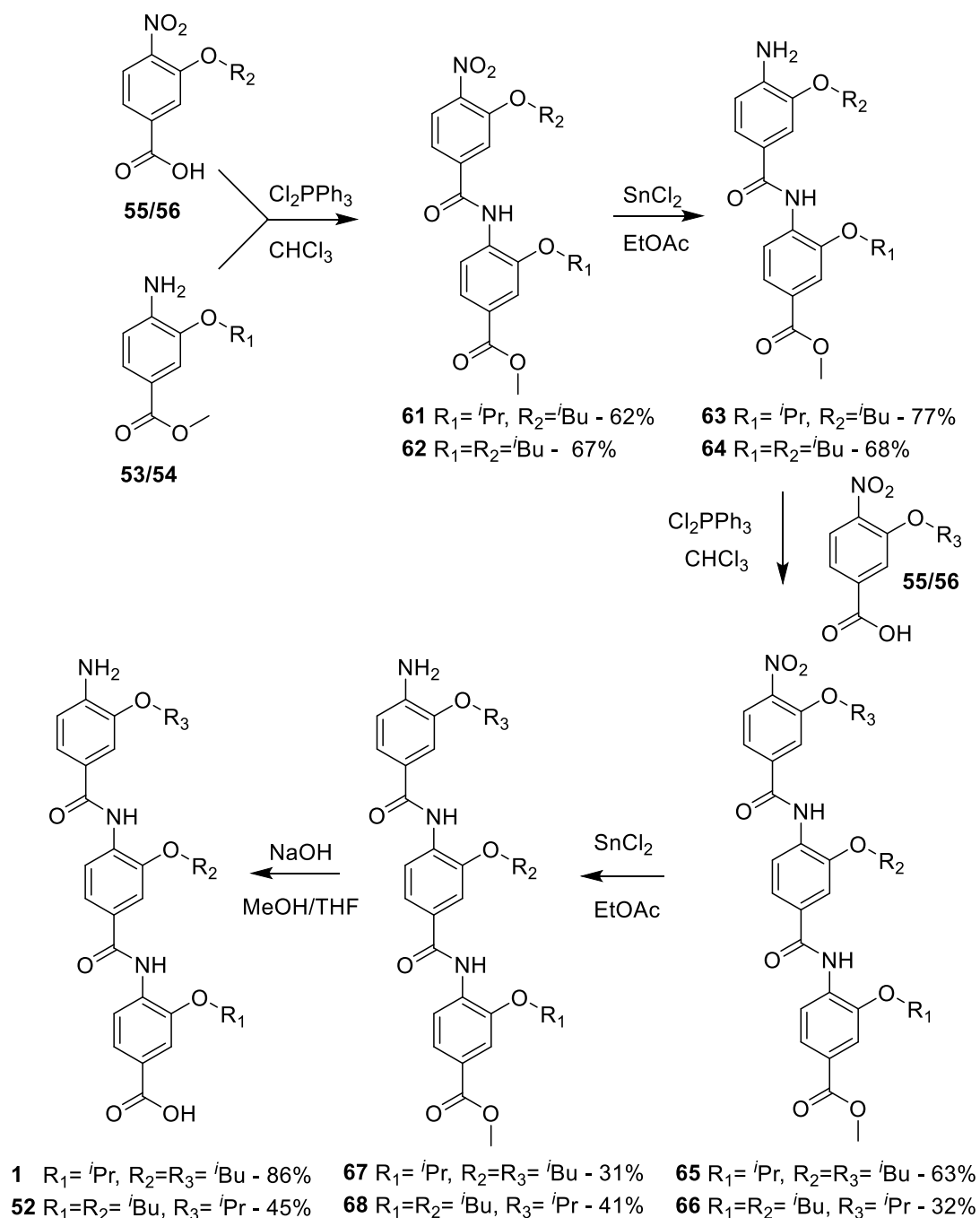
Scheme 5 - Assembly of oligobenzamide helix mimetics *via* iterative coupling and reduction cycles ending with a final hydrolysis

4.2.2.3 Biological Assessment

Once the oligobenzamides were synthesized, they were tested in the previously discussed fluorescence anisotropy competition assay (See Chapter 2). Upon titrating the putative inhibitors, a decrease in anisotropy associated with the disruption of the interaction was observed (Figure 69), in a similar manner to the decrease observed with the unlabelled peptide. In our assay the IC_{50} value of the unlabelled peptide was 0.28 ± 0.02 μM and compounds **1** and **52** gave IC_{50} values of 9.1 ± 0.8 and 24 μM ± 1 respectively.
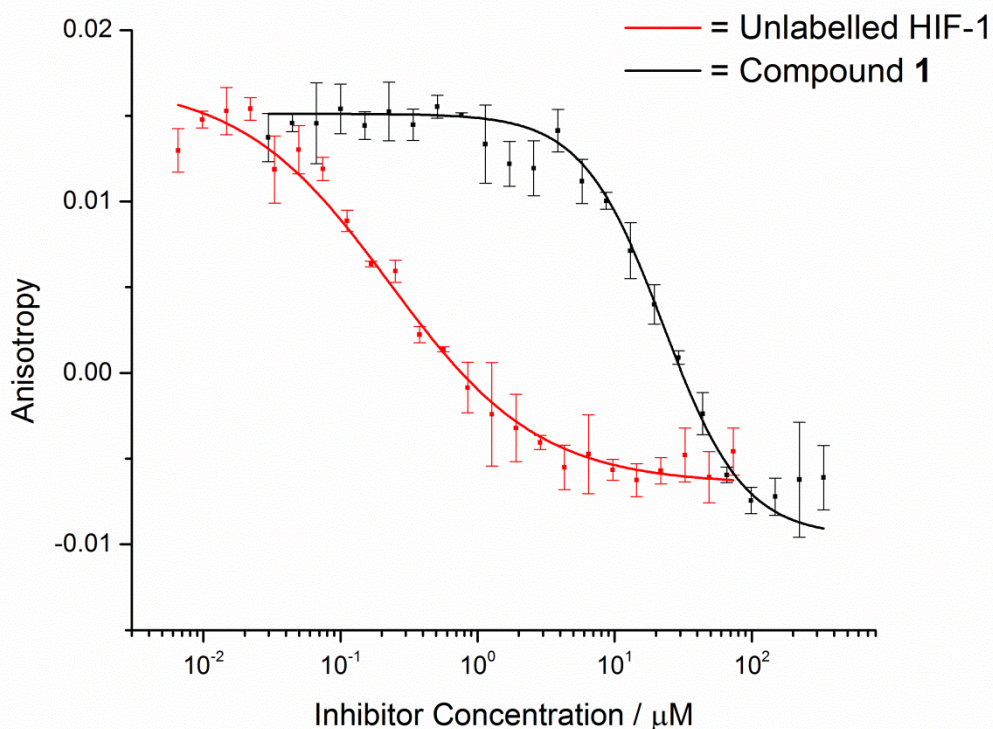
Figure 69 - Representative fluorescence anisotropy titration data for the unlabelled HIF-1α (red) and compound **1** (black). (40 mM sodium phosphate buffer pH 7.5, 80 nM FITC-HIF-1α, 100 nM p300)

Encouraged by these preliminary results, a small library of compounds with alternative side chains was prepared to probe the relative importance of side chain size and polarity. The additional compounds were prepared via the same methods described above: alkylation of the phenol with the corresponding halide, reduction/hydrolysis and coupling with dichlorotriphenylphosphorane. The compounds and their IC_{50} values are summarised in Table **6**. Crucially, a peptide derived from the helix 3 region of the HIF-1α CTAD was incapable of inhibiting this interaction ($IC_{50} > 500\mu M$)

Table 6 - Structures and IC$_{50}$ values for Oligobenzamide library
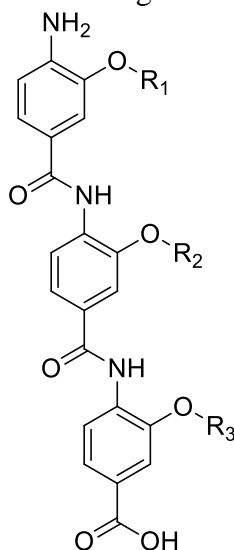
Compound	R$_1$	R$_2$	R$_3$	IC$_{50}$ (µM)
1	^{i}Bu	^{i}Bu	^{i}Pr	9.1 ± 0.8
52	^{i}Pr	^{i}Bu	^{i}Bu	24 ± 1
69	Me	^{i}Pr	^{i}Bu	216 ± 16
70	^{i}Bu	^{i}Bu	^{i}Bu	9.8 ± 1.3
71	^{i}Bu	^{i}Pr	^{i}Bu	13.4 ±1.5
72	Benzyl	Benzyl	Benzyl	56 ± 6
73	^{i}Pr	^{i}Pr	^{i}Pr	39 ± 4
74	^{i}Bu	^{i}Pr	^{i}Pr	17 ± 1
75	Benzyl	^{i}Pr	^{i}Pr	20 ± 1
76	2-hydroxyethyl	^{i}Pr	^{i}Pr	416 ± 64

The best compound identified from this library is the exact mimic of the helix side chains with the same *C* to *N* directionality as the peptide, **1**. The next best compound, **70**, has the same top and bottom residues but with a very small difference (^{i}Bu to ^{i}Pr) in the central position. Incorporation of larger aromatic side chains, **72** in all positions seems to have a significant negative effect on the binding affinity. It should also be noted that the nitro-ester precursor **65** (shown in Scheme 5) to **1** has an IC$_{50}$ greater than 1 mM suggesting that the amine, the acid or both are imparting significant affinity along with the improved solubility.

In an attempt to rationalise the observed SAR, the best compound was docked computationally into the binding site of the corresponding peptidic region. This simulation and the structure activity relationship studies suggest that the R_1 side chain is important and effectively fills a hydrophobic pocket within the binding cleft as shown in Figure 70. This is backed up by the reduced binding of **69** as a methyl group is insufficient to fill the hydrophobic pocket and **76** as the introduction of polar functionality into a hydrophobic binding site is disfavoured. Interestingly, as well as suggesting our designed binding mode could be successful, computational docking proposed an alternative binding mode (Figure 70 B) in which the oligoamide inhibitors sit transverse to the native peptide binding with the *N*-terminal region bound into the Leu 141 pocket but the *C*-terminal portion sitting in hydrophobic pocket made up by Trp 403, Cys 406 and Leu 417. Both binding poses are consistent with the observations from experimentation.
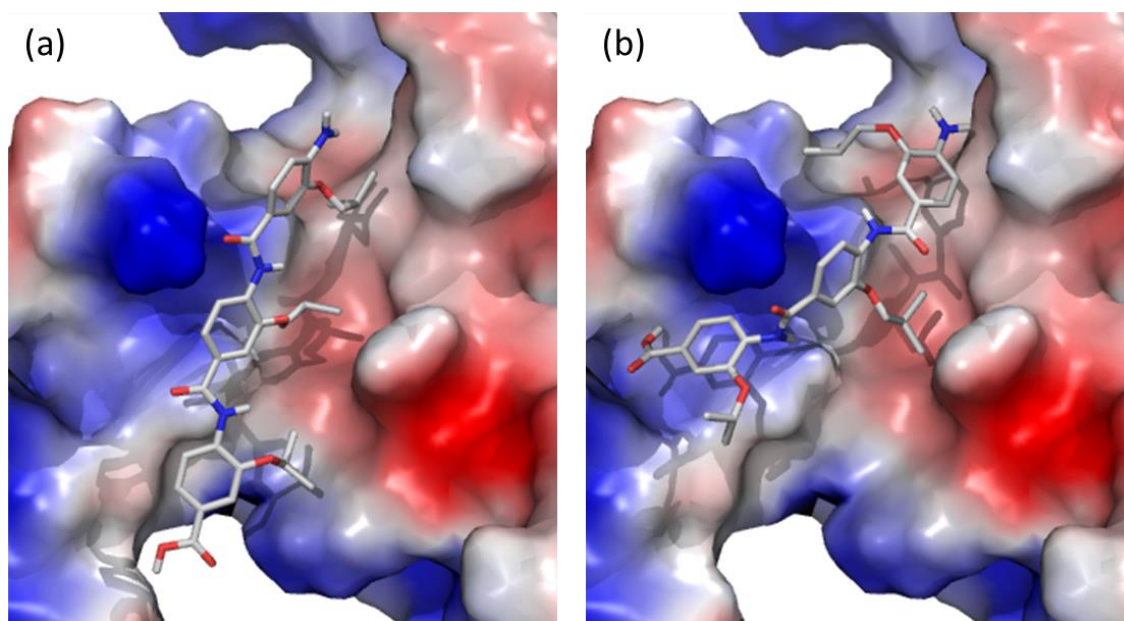
Figure 70 - Proposed binding modes of compound **1** in the Helix 3 binding cleft

Having developed inhibitors of this interaction we sought to compare our compounds with those reported in the literature. To this end the sulfonamide inhibitor, **KCN-1**, discussed in Chapter 1, was prepared via a modified route (Scheme 6).
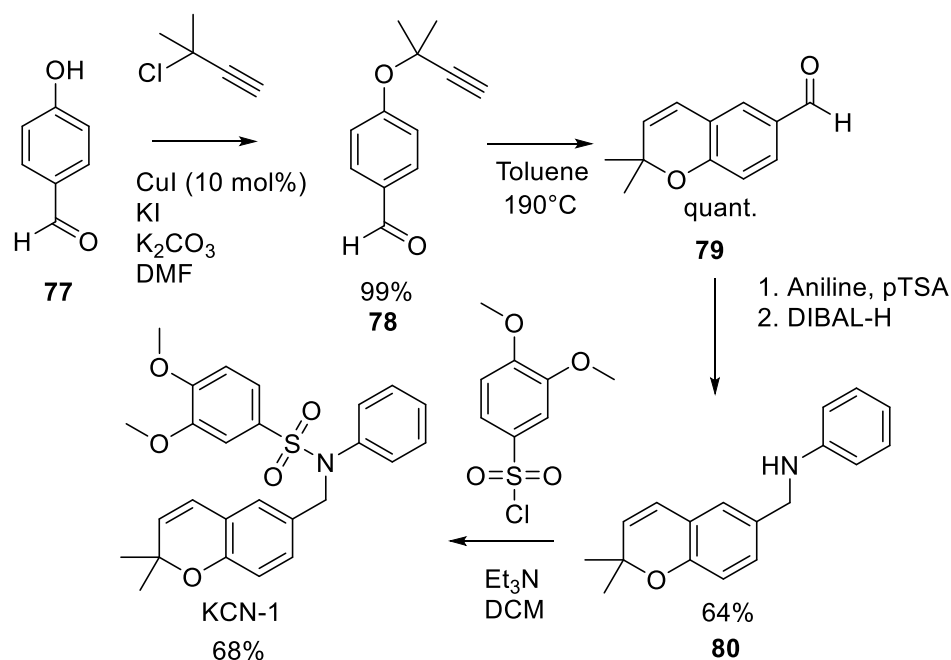
Scheme 6 – Synthesis of sulfonamide, KCN-1

Alkylation of 4-hydroxybenzaldehye **77** via an Ullman type coupling[176] followed by a Claisen cyclisation under thermal conditions provided the chromene **79** in excellent yield. Subsequent reductive amination with aniline afforded the secondary amine **80** which when treated with 3,4-dimethoxybenzenesulfonyl chloride gave the desired sulphonamide compound **KCN-1** in significantly improved yields compared to the original preparation (cf. 15% over the first 2 steps).[177]

KCN-1 was subsequently tested in the fluorescence anisotropy assay to give a comparison to the best oligoamide inhibitor **1**. Due to poor solubility, the sulfonamide could only be tested up to a maximum concentration of 25 μM. At this concentration, no inhibition of the HIF-1α/p300 PPI was observed whilst at similar concentrations the best oligoamide compounds showed significant inhibition. Given the reported cellular IC_{50} values for this compound is between 0.6-1.6 μM this suggests that the effects of KCN-1 observed in cellular assays may result, at least partially, from off-target effects.

The selectivity of the oligobenzamide **1** was then assessed by testing the inhibitory activity on another therapeutically relevant and α-helix mediated PPI, eukaryotic initiation factor (eIF) 4E/4G.[80] As Figure 71 illustrates, the oligobenzamide compound **1** does inhibit this alternative protein-protein interaction but only weakly with an IC_{50} of approximately 1 mM, giving over 100-fold selectivity for the HIF-1α/p300 interaction.
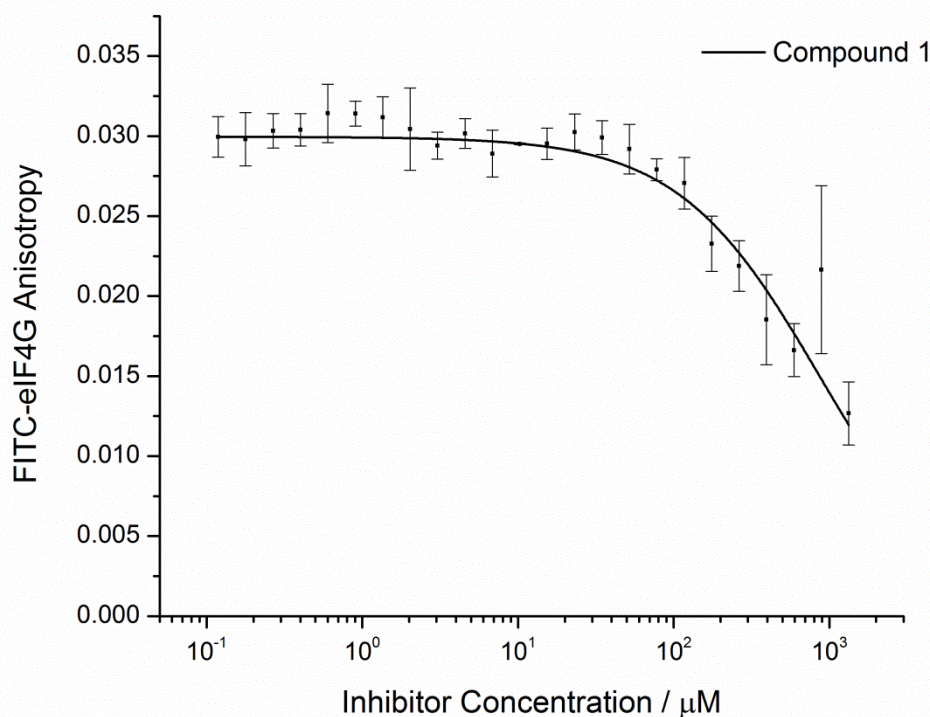
Figure 71 - Competition assay for compound **1** against the eIF4E/eIF4G Protein-Protein Interaction (80 nM FITC-eIF4G, 3 μM eIF4E, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats

Given the success of the *O*-alkylated oligo-benzamides as inhibitors of the HIF-1α/p300 interaction, a compound with the requisite side chains displayed on an alternative proteomimetic scaffold was investigated. An *N*-alkylated oligo-benzamide **81** (prepared by K. Long[178]) should also be capable of recapitulating the helix 3 sidechains (Figure 72) and therefore inhibit the interaction. However, compound **81** was shown to be unable to disrupt the interaction between HIF-1α and p300 ($IC_{50} > 250$ μM) suggesting that both the sidechains and the scaffold are crucial for inhibition.
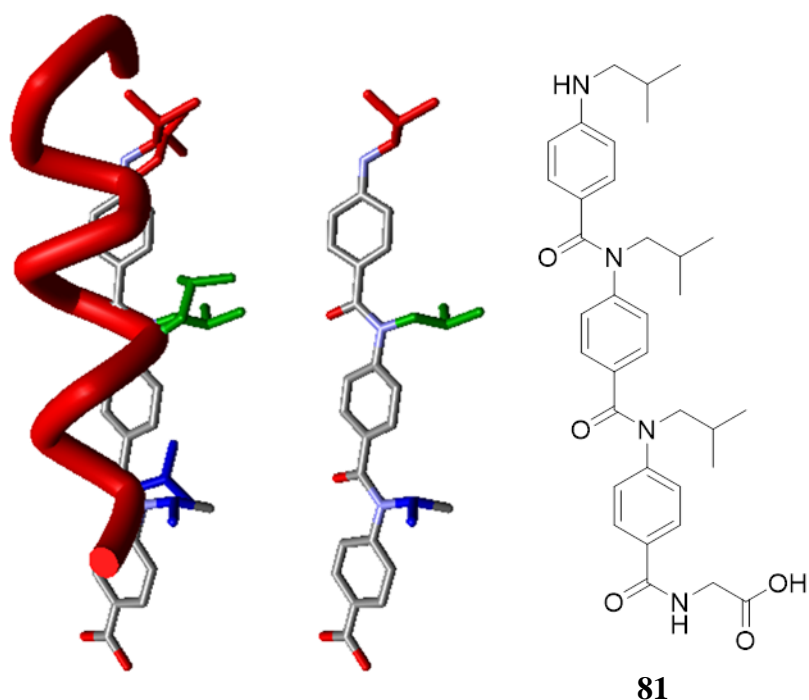
Figure 72 - *N*-alkylated analogues. From left to right, superimposed structures of N-alkylated proteomimetic and the HIF-1 Helix 3 excised from the NMR structure of p300 in complex with the C-terminal transactivation domain of HIF-1α (PDB ID: 1L8C), 3d model of *N*-alkylated oligo-benzamide in the extended conformation, structure of *N*-alkylated oligo-benzamide **81**.

4.2.2 Helix 2 Mimicry

To further validate the theory that the *C*-terminal helix (helix 3) of the HIF-1α CTAD is crucial for the interaction (see Chapter 2), and that the compounds described above act as helix 3 mimetic inhibitors, a series of compounds was designed to mimic helix 2. The helix 2 regions contains two key residues, identified by others[108], on one face of the helix – namely Cys_{800} and Asn_{804}.
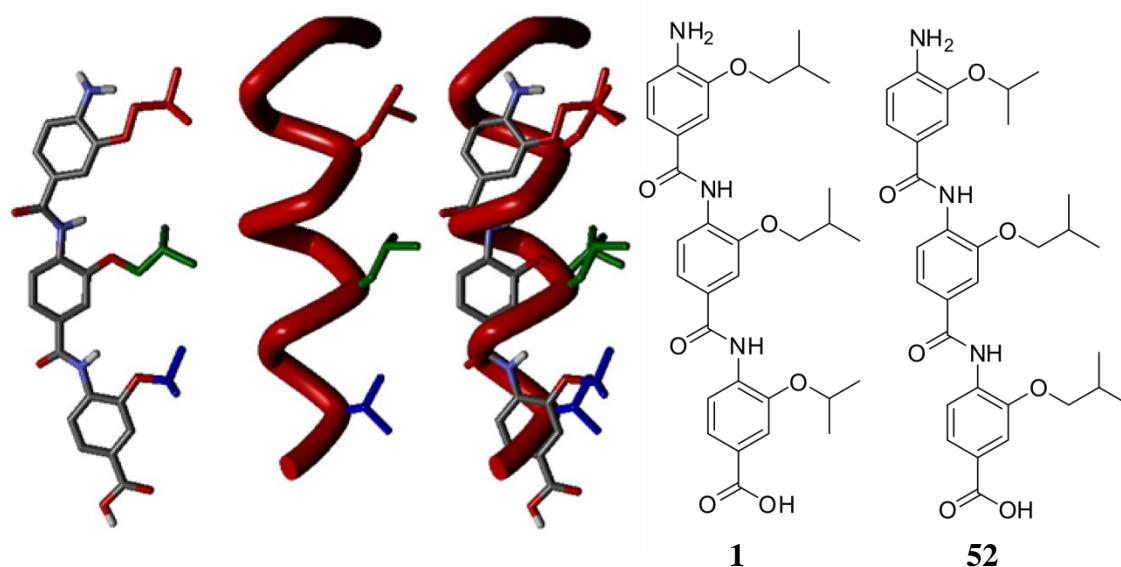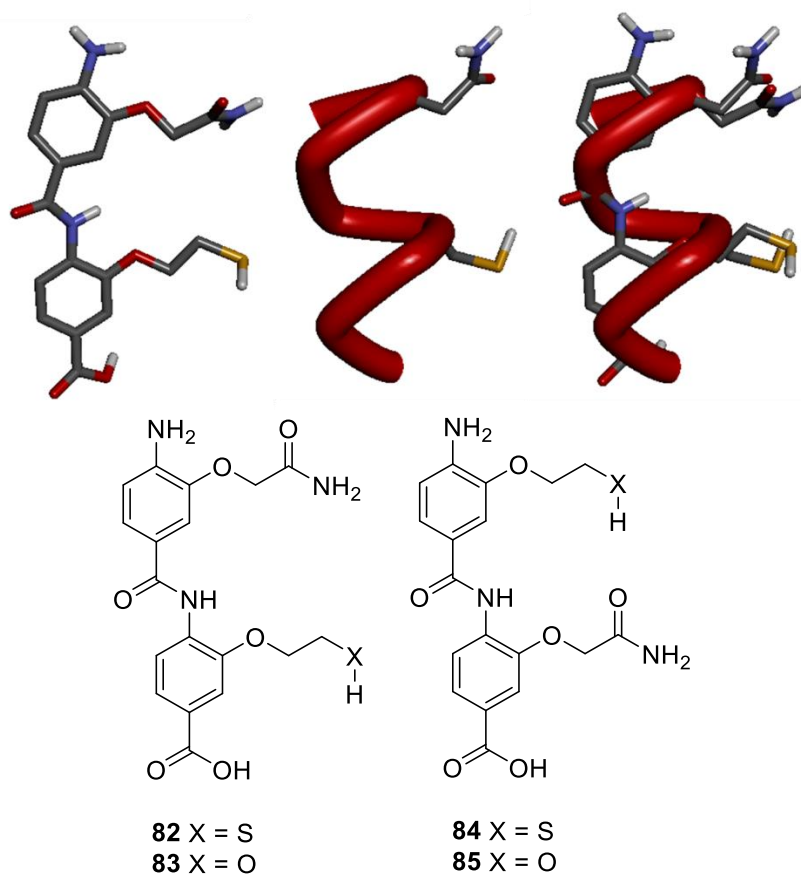
Figure 73 – Rational design of helix 2 mimetics. Top from left to right: Helix mimetic 3D structure, Helix 2 excised from the NMR structure of p300 in complex with the C-terminal transactivation domain of HIF-1α (PDB ID: 1L8C), superimposed 3D structures. Bottom chemical structures of designed oligobenzamide mimics

To mimic two turns of an alpha helix, only a dimeric benzamide is required. The compounds were designed to mimic the relevant side chains with both the matched **82/83** and mismatched directionality **84/85** with respect to the dipole of the native helix (Figure 73). Also the oxy-analogue (cf serine) was incorporated **83/85** , in place of the cysteine mimic, to prevent any possible issues with disulfide bond formation.

4.2.2.1 Monomer Preparation

The preparation of these compounds required the synthesis of novel monomers with protecting groups to prevent the nucleophilic side chains from interfering with the coupling reactions. They were prepared in an analogous manner those described above. The asparagine mimic building blocks **95/98** was prepared by the reaction of trityl amine **86** and bromoacetyl bromide to give the protected amide moiety **87**. This could then be used to alkylate methyl 3-hydroxy-4-nitrobenzoate **58** giving the nitro ester compound **92** (Scheme 7).
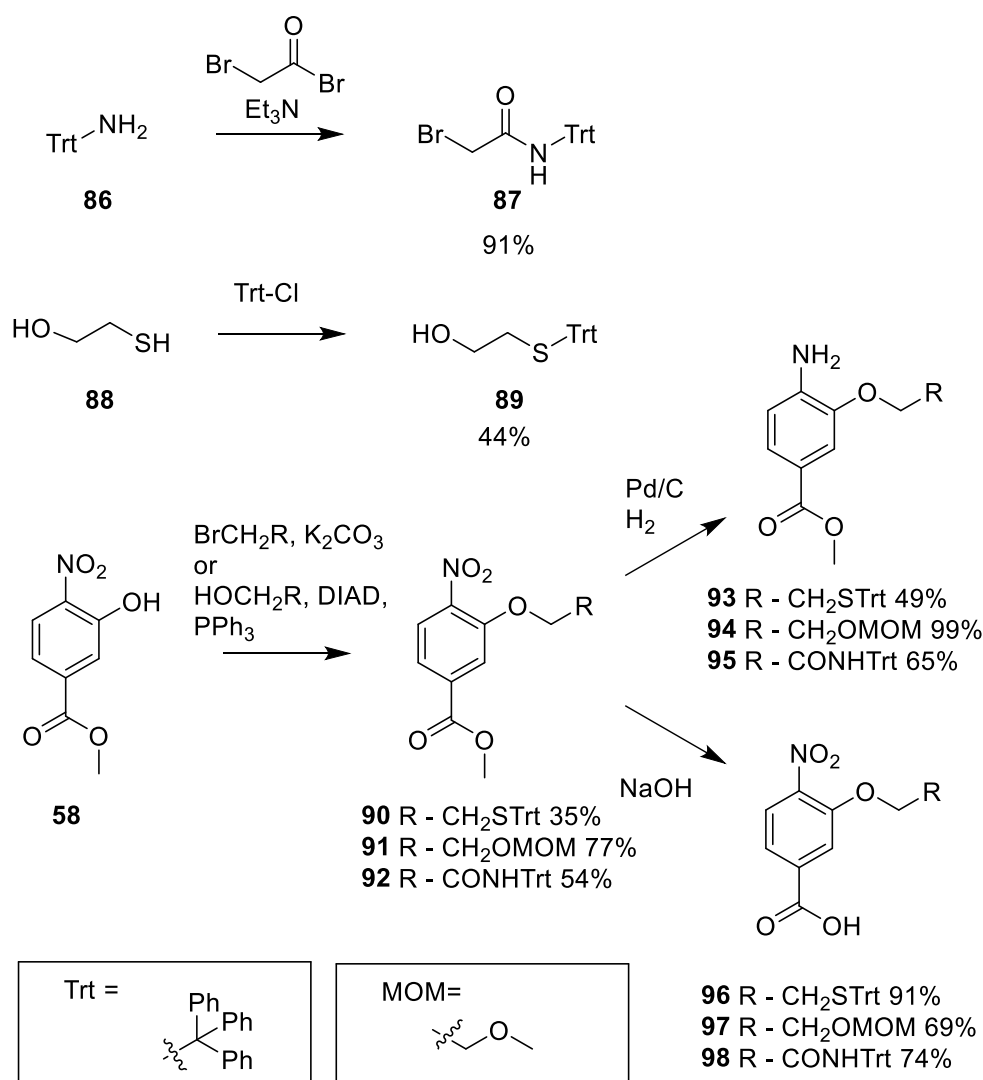
Scheme 7 - Synthesis of protected helix 2 monomers

The cysteine mimic building block **88** was prepared via the selective trityl protection of mercaptoethanol[179] using trityl chloride followed by a Mitsunobu reaction to introduce the side chain **89** into nitro ester **90**. The serine mimic was prepared via the alkylation of **58** using commercially available MOM-protected bromo-ethanol. The nitro esters **90-92** were subsequently either reduced to the aniline building blocks **93-95** or hydrolysed to give the acid blocks **96-98.**

Significantly reduced yields were observed in the alkylation steps, in comparison to the simple alkyl chains described previously. Two-dimensional TLC showed the products to be unstable on silica. Switching from silica to alumina for chromatography led to an increase in isolated yield in later steps.

4.2.2.3 Dimer Assembly

With the requisite building blocks prepared, attempts were made to assemble them into the designed compounds. Due to the acid lability of the side chains several coupling conditions were trialled as described in Table 7.

Table 7 – Results from coupling experiments towards helix 2 mimetics
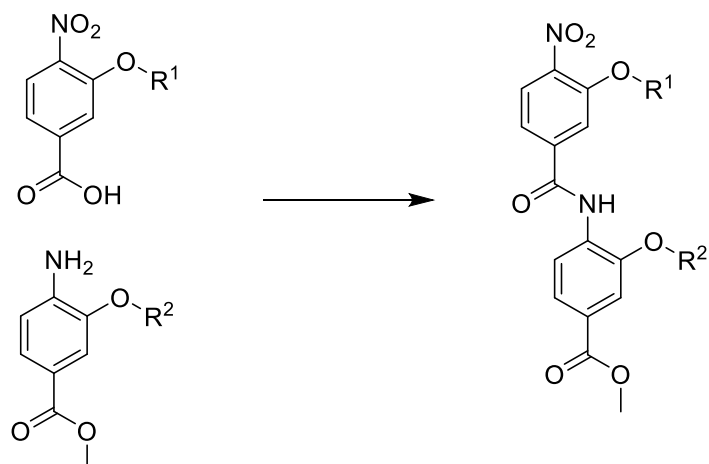
Entry	R^1	R^2	Conditions	Conversion by crude NMR
1	(O, NH-Trt amide side chain)	(OMOM side chain)	HATU, DIPEA	No conversion
2	(O, NH-Trt amide side chain)	(OMOM side chain)	SOCl$_2$	No conversion
3	(O, NH-Trt amide side chain)	(OMOM side chain)	Ghosez's Reagent	~5%
4	(O, NH-Trt amide side chain)	(OMOM side chain)	Cl$_2$PPh$_3$	No conversion
5	(OMOM side chain)	(O, NH-Trt amide side chain)	Ghosez's Reagent	~5%
6	(OMOM side chain)	(isobutyl side chain)	Cl$_2$PPh$_3$	Complete conversion
7	(isopropyl side chain)	(O, NH-Trt amide side chain)	Cl$_2$PPh$_3$	Complete conversion

None of the coupling conditions between two exemplar helix 2 mimetic building blocks gave satisfactory coupling (entries 1-4). The building block side chains were then reversed to rule out an incompatibility with the examples chosen for the

screening but coupling was still insufficient (entry 5). Finally it was shown that the building blocks could be coupled with other building blocks (entries 6-7) but not each other.

In order to prepare the dimers for screening an alternative synthetic method was required. To reduce the steric effects on the amide bond forming reaction an un-functionalised monomer unit (methyl 4-amino-3-hydroxybenzoate) **99** was coupled to the monomer **97** using PyBOP.[153, 180] This gave a 9:1 mixture of the ester **101** to the desired amide **100**. When treated with caesium carbonate, the ester undergoes a *O*- to *N*-acyl transfer giving the desired amide **100** in moderate yield suggesting that *O*-acylation is kinetically favoured but that the ester rearranges to the thermodynamically more stable amide. The desired side chain can then be introduced via the alkylation of the phenol to give **101**. Subsequent catalytic hydrogenation, ester hydrolysis and side chain deprotection afforded the desired dimer **102** for screening.
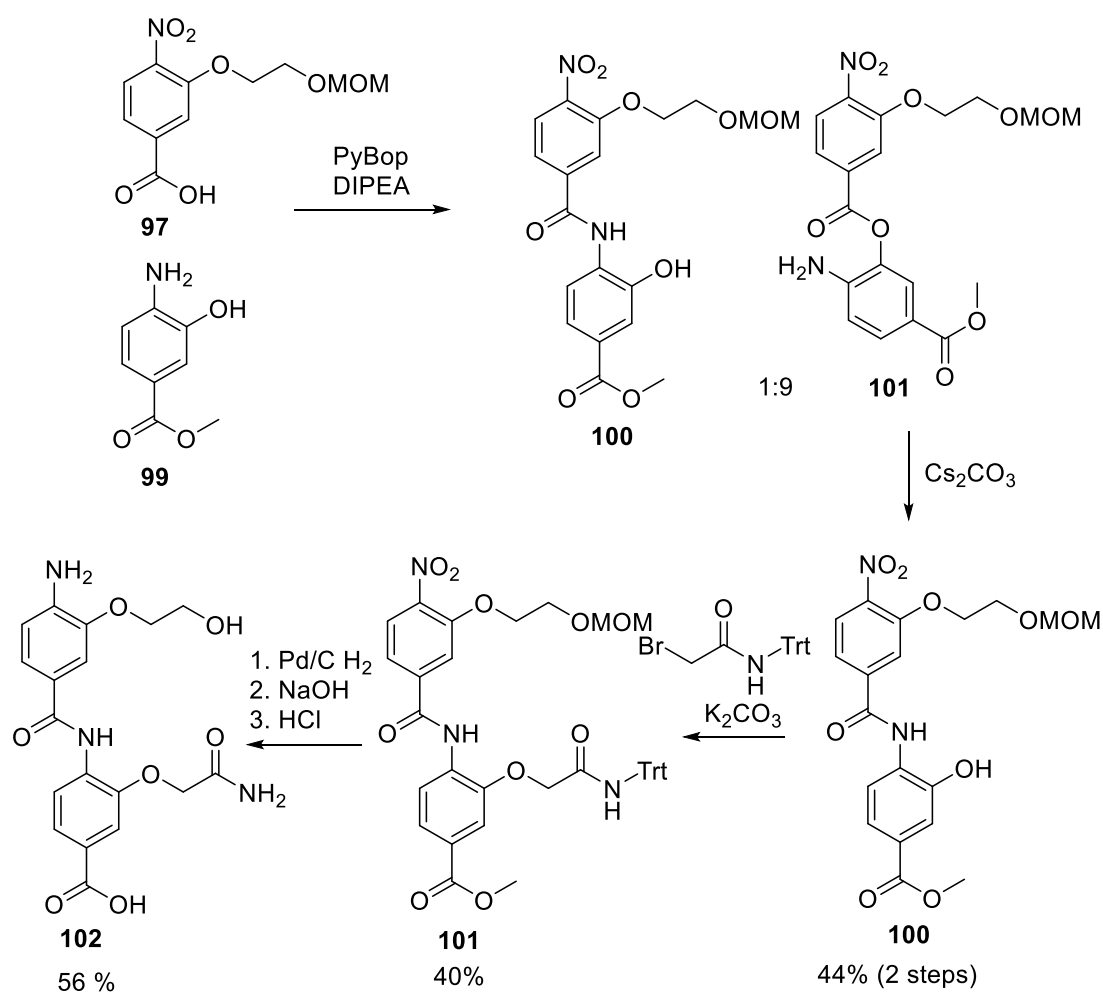
Scheme 8 - Synthesis of dimers *via* acyl transfer

4.2.2.3 Assessment of Inhibitory Activity

The prepared helix 2 mimetic compound **102** was screened in HIF-1α/p300 competition assay and no inhibitory activity observed up to 1 mM (higher concentrations were not screened) – given the complete lack of activity, and time constraints, no further analogues were prepared. This lack of potency provides additional evidence for the theory that helix 3 drives the binding of the HIF-1α CTAD to p300. Although a body of previous work shows compounds of this type are good mimics of α-helices there is no evidence that the compounds used here are capably of mimicking this particular helix. The lack of activity of these compounds, featuring the same backbone scaffold as the active compounds described in section 4.2, also provides additional evidence that the side chains are crucial for the recognition of oligobenzamide **1** by p300.

4.3 Late Stage Functionalization

Late stage functionalization is important in the exploration of structure activity relationships (SAR) – allowing rapid access to many analogues from a single precursor which can be prepared on a reasonable scale. The development of synthetic methodology in the area of late stage functionalization has recently expanded[181] particularly through the use of C-H functionalization.[182]

In able to allow access to a larger set of oligo-benzamide compounds with more diverse side chains to explore the SAR of their HIF-1α/p300 inhibition, it was desirable to develop a late stage functionalization approach. This abrogates the requirement to build large sets of monomers and assemble them; whilst this approach is useful for library production[152] and initial screening,[148, 170] it disfavours making small changes to compound structures.

Employing the best inhibitor identified from the initial library of HIF-1α/p300 inhibitors **1**, as a starting point for SAR studies, and with a secondary aim of improving the solubility, a series of compounds with cleavable protecting groups was prepared. Phenolic allyl ethers are readily cleaved *via* the formation of a π-allyl complex with palladium (0) but are relatively inert to other conditions. This made then ideal for use in the late stage functionalization of oligobenzamides.
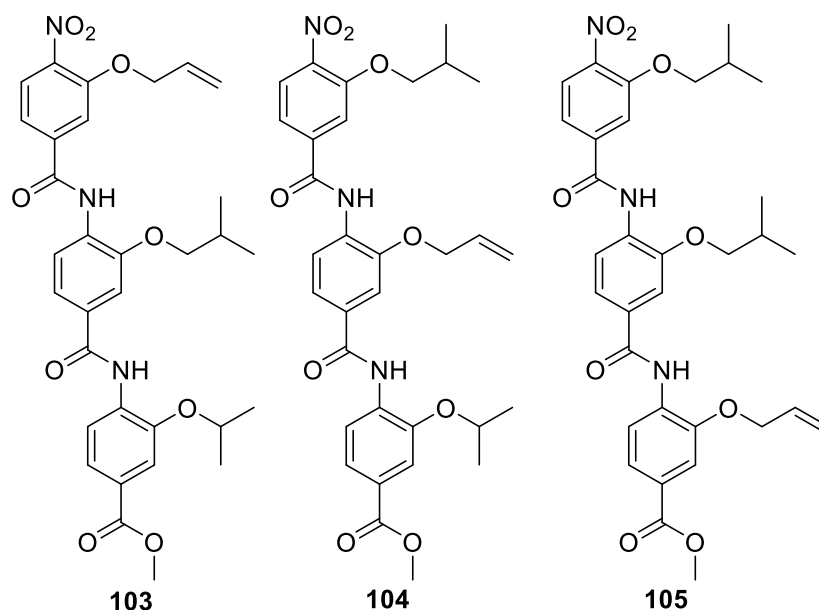
Figure 74 - Allyl precursors for late stage functionalization

The allyl side chain monomer was produced via the methods previously described for monomer synthesis.[144, 152] This allyl monomer can then be carried through iterative trimer assembly to give the allyl protected trimers **103-105** shown in Figure 74 which have each of the side chain positions sequentially substituted from the trimer **1** − the best inhibitor of HIF-1α/p300.

The allyl group can then be cleaved by treatment with palladium tetrakis triphenylphosphine in the presence of a sulfinic acid scavenger as shown in Scheme 9.[183] This reveals the phenol which can then be functionalized via Mitsunobu[184] or S_N2 alkylation reactions. As well as incorporating additional functionality at the phenolic position, it can be left functionalised to assess the individual contributions of sidechains after reduction and hydrolysis. An analogous sidechain deletion approach was employed to elucidate the mechanism of amphotericin B activity.[185]
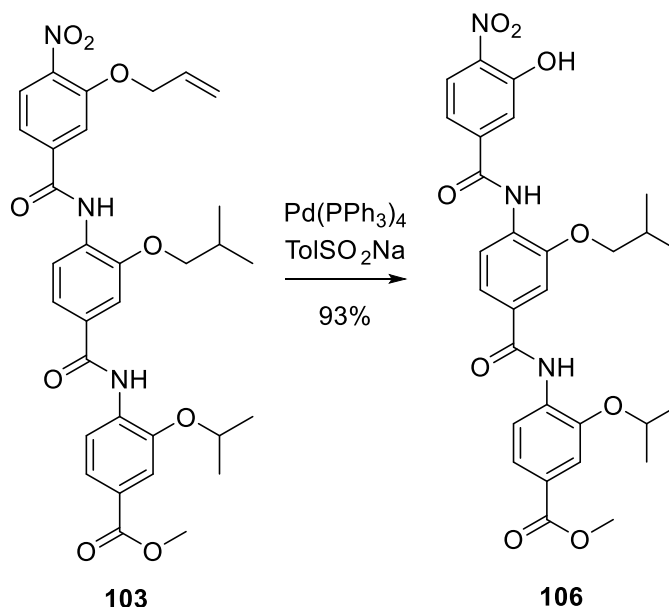
Scheme 9 - Example de-allylation reaction of trimer **103**

In an attempt to increase the aqueous solubility of the oligobenzamides, a variety of solubilising side chains were incorporated into the scaffolds. The compounds were then reduced and hydrolysed to provide another library of compounds with more diverse functionality without having to produce all the monomers separately and construct the compounds individually. An example of a synthetic route is shown in Scheme 10.
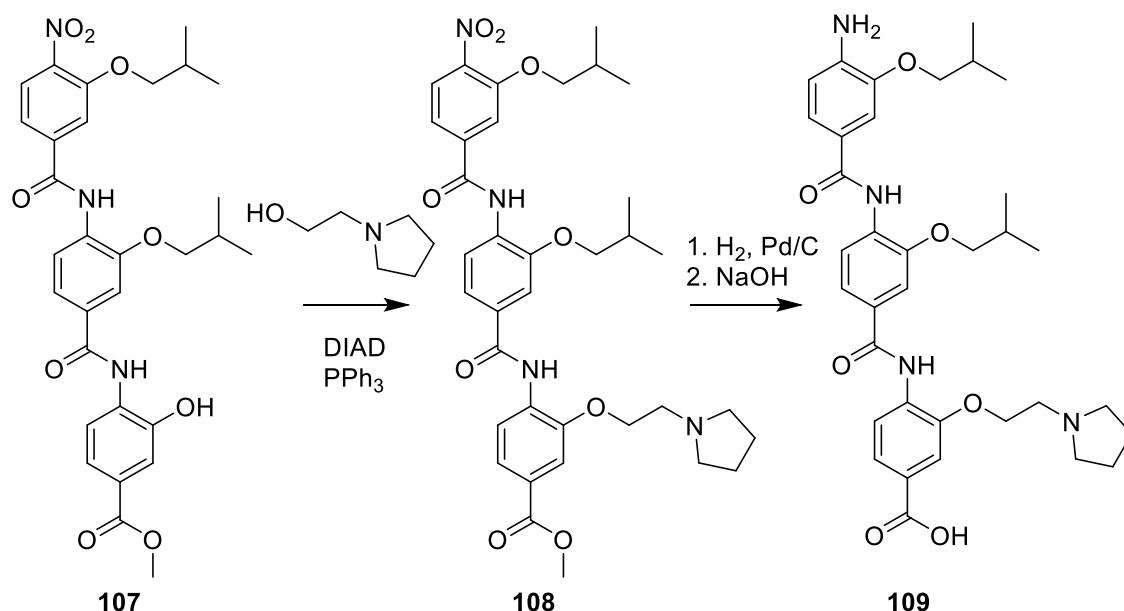
Scheme 10 - Example late stage functionalization

This approach reduces the number of synthetic steps required to make a library of 5 trimeric compounds with variation in one position from 42 to 27 steps from 3-

hydroxy-4-nitrobenzoic acid **58**. The compounds prepared were then screened in the HIF-1α/p300 competition assay and the results are summarised in the table below.
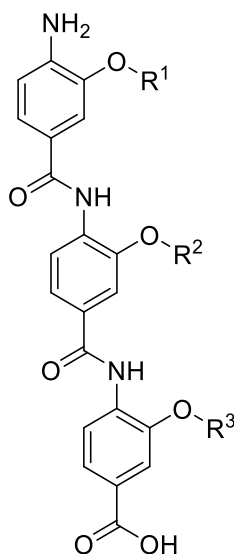
Compound	R^1	R^2	R^3	IC$_{50}$ / μM
110	H	^{i}Bu	^{i}Pr	>500
111		^{i}Bu	^{i}Pr	521 ± 16
112	^{i}Bu	H	^{i}Pr	341 ± 15
113	^{i}Bu		^{i}Pr	>500
114	^{i}Bu		^{i}Pr	168 ± 22
115	^{i}Bu	^{i}Bu	H	137 ± 12
116	^{i}Bu	^{i}Bu		>500
109	^{i}Bu	^{i}Bu		322 ±15
117	^{i}Bu	^{i}Bu		163 ± 10

Unfortunately none of the compounds exhibited activities anywhere near those of the previously described compounds. This provides further evidence that the hydrophobic nature of the sidechains is crucial for inhibition of the HIF-1α/p300 interaction. The amine functionalized compounds were observed to be significantly more soluble than the previous compounds. Some of the compound appeared to interfere with the assay so accurate IC$_{50}$ values could not be determined.

4.4 Photo-crosslinking Compounds

Another advantage of the late stage functionalization approach is that sensitive functional groups can be incorporated later in the synthesis and therefore carried through fewer steps. This allowed the preparation of a trifluoromethyl diazarine photocrosslinking group[186] functionalized trimer in an attempt to confirm active helix mimetic binding sides via covalent crosslinking and proteomics as shown in Figure 75.
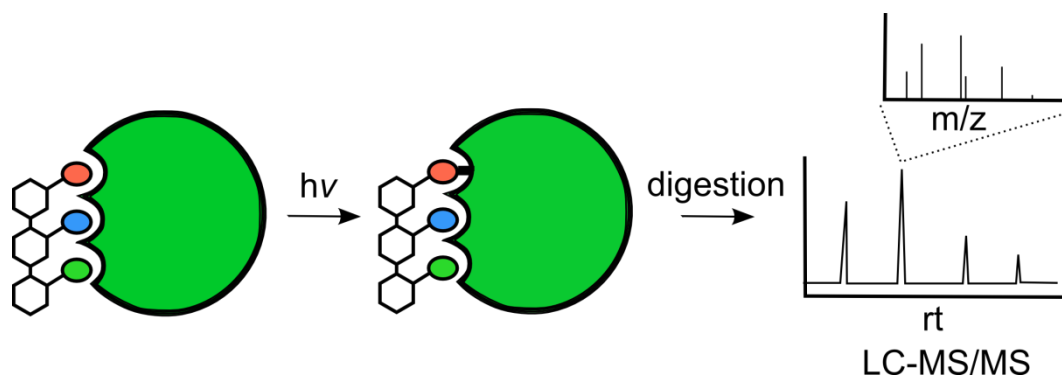
Figure 75 - Schematic representation of photo-crosslinking approach

The trifluorodiazarine functional group extrudes nitrogen when irradiated with UV light and generates a singlet carbene. This carbene is highly reactive and promiscuously forms covalent bonds with proximal groups. It has been utilised to explore biological interactions, including amyloid formation[187] and prolyl hydroxylase enzymes,[188] via crosslinking then mass spectrometry. It was envisioned that by functionalising compound **1** with trifluoromethyl diazarines in various positions it would be possible to identify the binding site of the compounds by inducing crosslinks, digesting the protein into peptide fragments and identifying the crosslinked region by LC-MS/MS.

The trifluoromethyl diazarine moiety was prepared according to a literature procedure[189] as shown in Scheme 11. Briefly, 4-bromobenzyl alcohol **118** was silyl protected **119** then lithiated and trapped with diethyltrifluoroacetamide to introduce the trifluoromethyl ketone functionality **120**. The ketone was sequentially converted to the oxime **121**, tosyl-oxime **122** and finally the diaziridine **123**. The diaziridine **123** was oxidised to the diazarine **124** with iodine, the alcohol deprotected to **125** and converted to the iodide **126**.
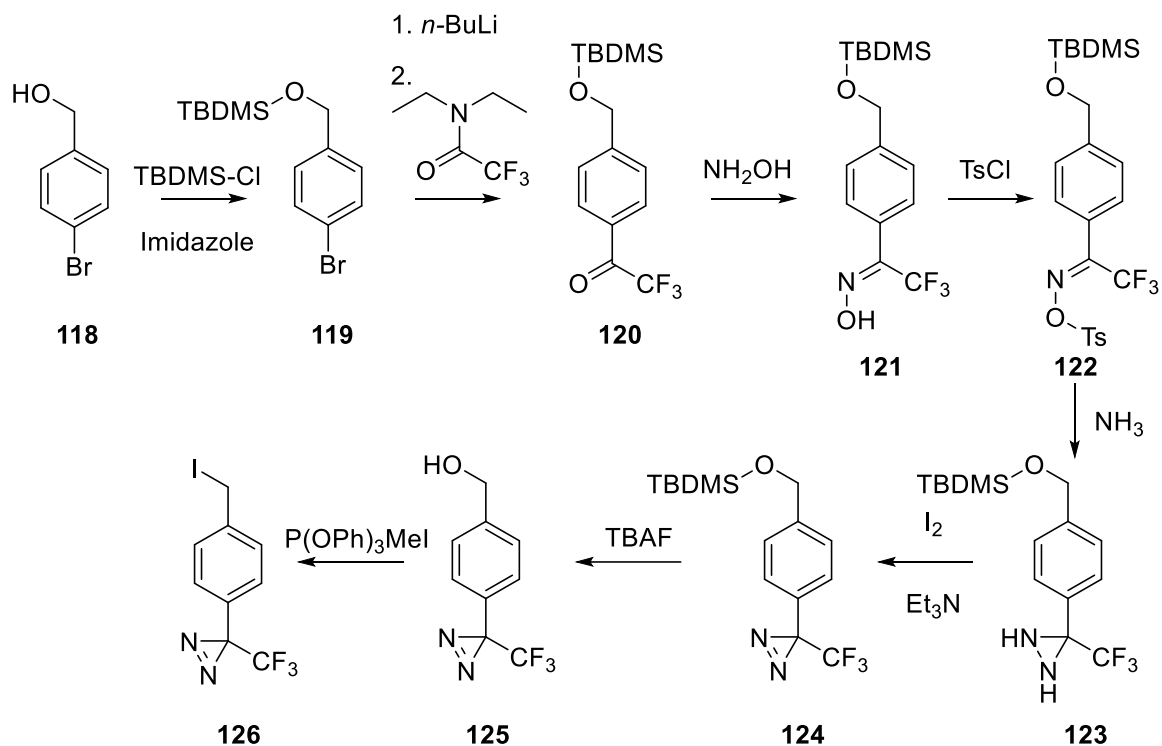
Scheme 11 - Synthesis of trifluoromethyldiazarine building block

The fragile diazarine would likely not survive the harsh coupling conditions required to prepare the oligobenzamide backbone and would certainly not be stable to the reductive conditions required to reduce the nitro groups. It would also significantly lengthen the number of linear steps required to produce the compounds. However using the late stage methodology described above, the aryl trifluoromethyl diazarine can be incorporated after all the required reductions as shown in Scheme 12.
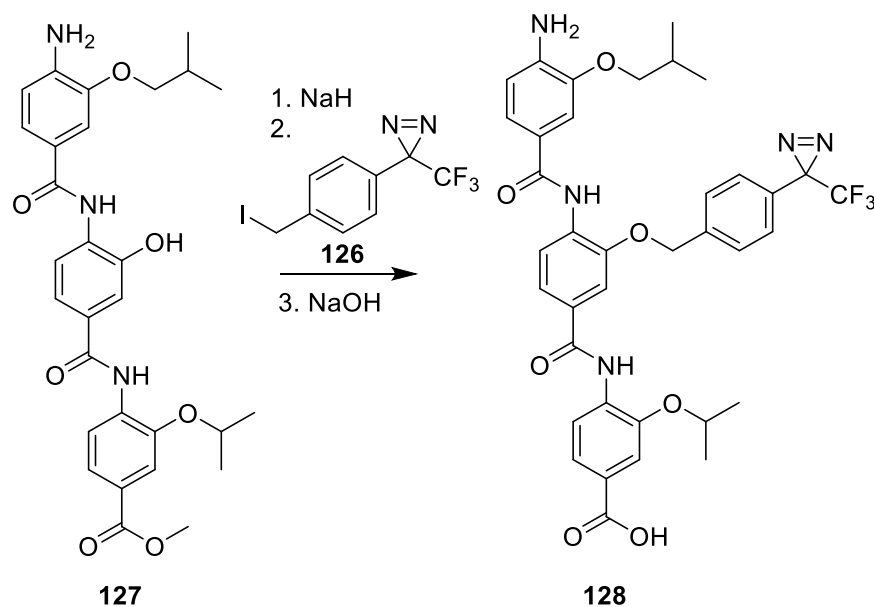
Scheme 12 - Late stage one-pot synthesis of photocrosslinking helix mimetic

By selectively, and irreversibly, deprotonating the phenol with sodium hydride then treating with the iodide, the desired functionality could be introduced. Final ester hydrolysis was performed *in situ* by treating the reaction mixture with sodium hydroxide. This afforded the compounds shown in Figure 76 with the aryl group in the middle **128** and bottom positions **129** but failed to produce the compound functionalized in the top position – possibly due to the proximity of the aniline to the phenol.
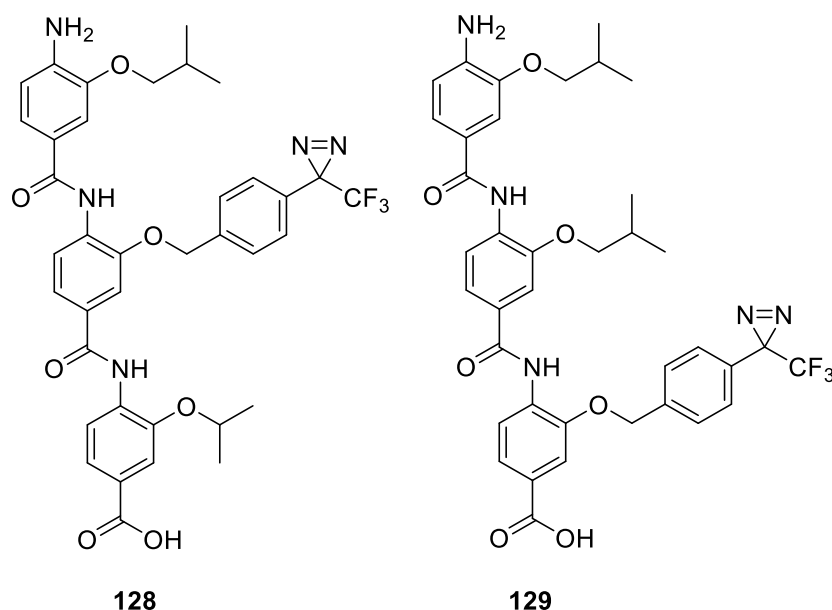
Figure 76 - Structure of cross linking compounds prepared

The photo-crosslinking compounds were screened for activity in the HIF-1α/p300 competition assay and unfortunately, although not surprisingly, the presence of the cross-linking functionality impaired the ability to inhibit the interaction. The degree of impairment was position dependant and consistent with the models provided in Figure 70. Incorporation in the middle position, compound **128**, was particularly poorly tolerated with an estimated IC_{50} of 985 ± 102 µM. Functionalization at the *C*-terminal position **129** was tolerated considerably better with an IC_{50} of 136 ± 8.7 µM. Compound **129** was therefore selected for preliminary crosslinking studies.

4.4.1 Protein Digestion Analysis

To enable the identification of protein-compound cross links an unlabelled protein digestion and analysis was required as background. The p300 protein was dialysed into ammonium bicarbonate buffer, denatured thermally and disulfides reduced prior to the cysteines being alkylated with iodoacetamide. The protein was then treated with trypsin for 4 hours, quenched with TFA and analysed by LC-MS/MS. The LC trace and example MS/MS data are shown below in Figure 77. This analysis provided approximately 70% coverage of the p300 primary sequence without further

optimisation suggesting that it was likely the binding sites could be identified using this method.
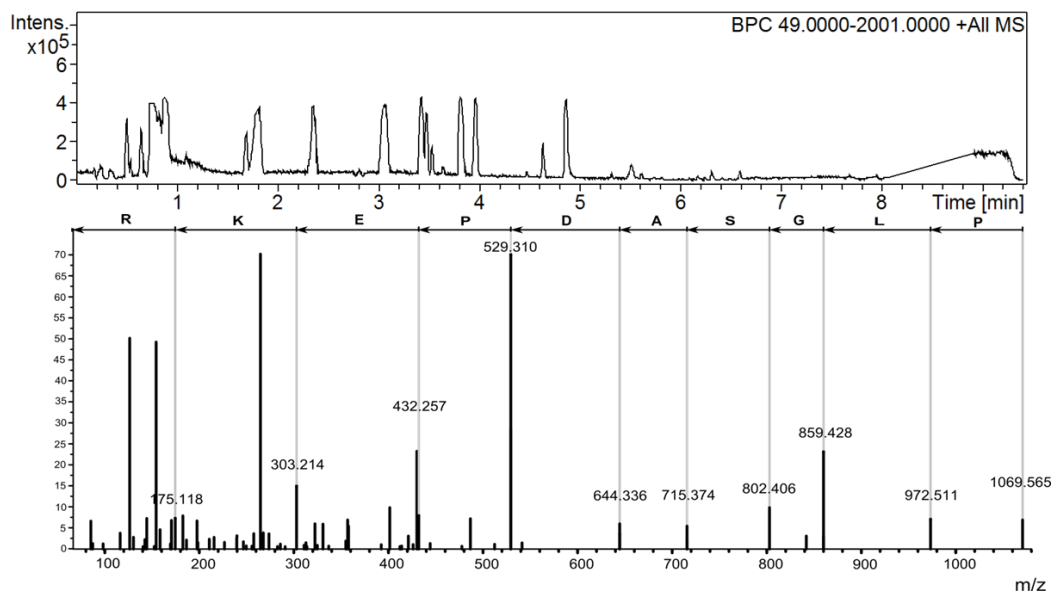
Figure 77 - LC trace and example MS/MS for p300 digestion

Compound **129** was incubated with p300 protein (100 μM) at 150 μM, a concentration above its IC_{50} value for HIF-1α/p300 inhibition, and irradiated with 360 nM light for 1 hour. Crosslinking was initially monitored by intact protein electrospray mass spectrometry and LC-MS. Prior to UV irradiation, samples containing protein only, compound **129** only and the protein/compound mixture were analysed by LC-MS. These samples were then irradiated with 365 nm UV light for 1 hour whilst cooled in an ice bath and reanalysed. LC-MS analysis showed that that compound 129 had been almost completely consumed in both the compound only sample and the protein/compound sample however mass spectrometry showed no protein adducts (Figure 78).
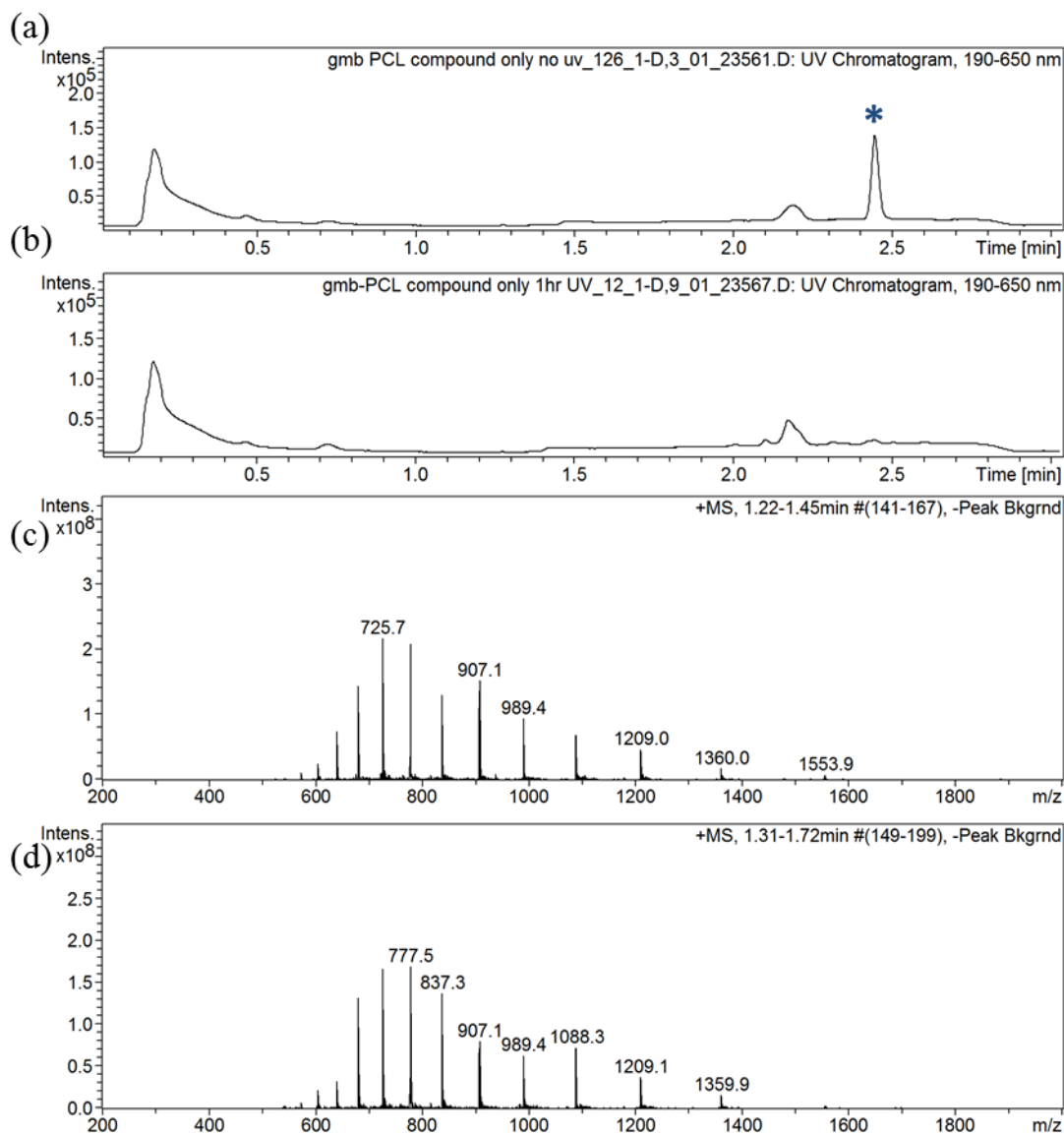
Figure 78 - Analytical data for photo-crosslinking experiments. (a) Chromatogram of 150 μM compound **129** in buffer prior to UV irradiation. The peak corresponding to compound **129** is denoted by the asterisk (b) Chromatogram of 150 μM compound **129** in buffer post UV irradiation (365 nm, 1 hour) (c) mass spectrum of protein before UV irradiation (365 nm, 1 hour) (d) mass spectrum of protein after UV activation (365 nm, 1 hour)

Upon further inspection, several additional peaks were evident in the LC-MS trace corresponding to various degradation products arising from carbene formation and reaction with water. The water quenched carbene products and the lack of crosslinking observed suggest that the trifluoromethyl diazarine is exposed to solvent when compound **129** is bound to p300 rather than buried in the protein surface. This could explain why incorporation at the bottom position is tolerated better than in the middle position. Unfortunately, this approach did not provide any further information about the binding mode of these oligoamides as hoped.

4.5 Protein Prosthesis

A long term goal in chemical biology and synthetic biology is the incorporation of non-natural functionality into biomolecules such that the natural function of the original molecule is enhanced (protein prosthesis).[190] There are many ways in which this has been attempted; chemical mutagenesis of enzyme active sites,[191, 192] the synthesis of all mirror image proteins[193] and the replacement of canonical amino acids with β-peptides.[194] Nature, for the most part, uses the 20 canonical amino acids but protein prosthesis allows the incorporation of any number of unnatural functionalities.

Often in this field, biomolecules are engineered to either improve stability[195] or to change substrate specificity.[196] It is rare for the activity to be equal to the wild type but there are examples where activity is retained[197] or only mildly depleted[198] when unnatural functionality is incorporated. A common approach to protein prosthesis is the replacement of individual residues[199] or small sections[198] of the protein primary structure with unnatural amino acids. β-turns have also been replaced by small molecules with locked conformations.[200] There are examples of β-peptide sequences replacing helical regions of proteins without complete loss of activity[201] but, so far, no examples of helix mimetic small molecule incorporation have been reported.

4.5.1 Oligobenzamide-Peptide Hybrids

Having observed that the oligobenzamide proteomimetic compounds, discussed in Chapter 5, provided better inhibitors that the short (helix 3) peptide that they mimic, it was envisaged that incorporation of the helix mimetic into a longer peptide sequence could provide an inhibitor with enhanced affinity that retained the specificity imparted by a long peptide sequence. To this end, synthetic methodology to develop oligobenzamide helix mimetic/peptide hybrids was developed based on the previously developed solid phase oligobenzamide synthesis shown in Scheme 13.[152] Fluorenylmethoxy carbamate (Fmoc)-protected monomers were activated as the acid chloride then added to glycine loaded Wang resin. The Fmoc can then be subsequently removed with piperidine and additional monomers added to the growing chain in an method analogous to standard Fmoc-SPPS.
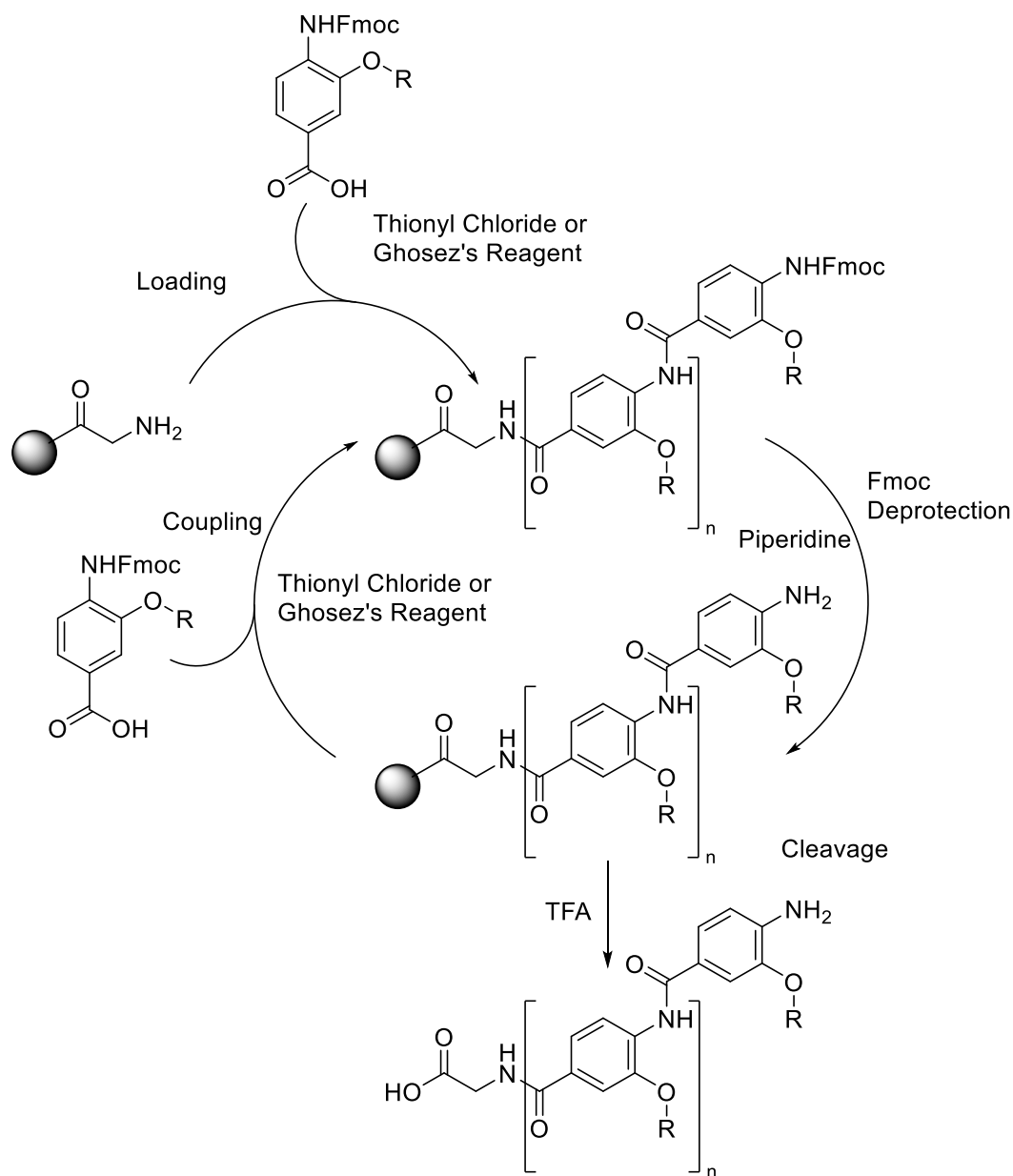
Scheme 13 - Previously developed iterative solid phase synthesis of oligobenzamides[152]

4.5.2 Method Development

Previously developed solid-phase synthesis methods for the preparation of oligobenzamides allows the rapid preparation of helix mimetic trimers (Scheme 13)[152] but the conditions were not suitable for the preparation of peptide-conjugates. The poor nucleophilicity of the anilines requires harsh acid chloride mediated conditions likely to racemize and/or deprotect conventional amino acid building blocks. Attempts were made to couple Fmoc-glycine to the top of a trimer **130** using either HATU or *via* Fmoc-glycine acid chloride, but no coupling product **131** was observed (Scheme 14).
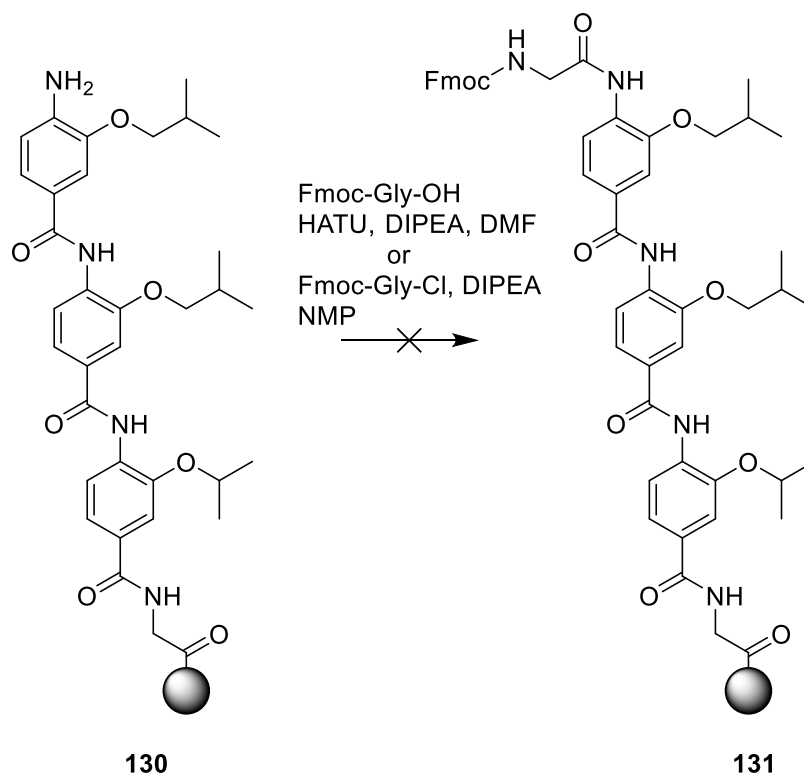
Scheme 14 - Attempted coupling of Fmoc-Glycine to the *N*-terminus of an oligobenzamide trimer

To circumvent the poor nucleophilicity of trimer anilines, Fmoc-glycine was coupled to an isobutyl monomer **134** using dichlorotriphenylphosphorane (Scheme 15). The use of an allyl ester facilitated the ester cleavage in the presence of the base labile Fmoc protecting group as shown in Scheme 15.
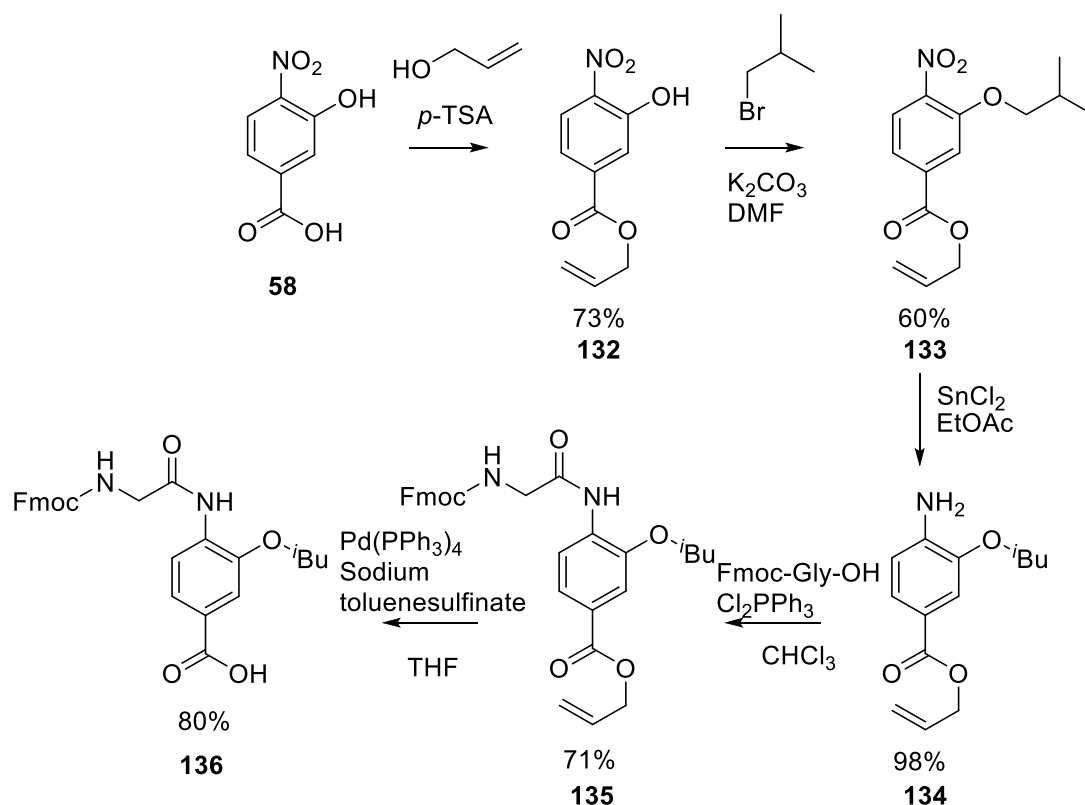
Scheme 15 - Synthesis of Fmoc-Gly iBu monomer, **136**

Synthesis of key monomer **136**, began with Fischer esterification of nitro acid **58** with allyl alcohol to give ester **132** before alkylation with isobutyl bromide to give alkoxy nitro ester **133**. The aryl nitro group was reduced to the aniline **134** using stannous chloride. The aniline **134** was then acylated with Fmoc-glycine *via in situ* acid chloride formation with dichlorotriphenylphosphorane. Finally the ester **135** was revealed with palladium (0) tetrakis triphenylphosphine and sodium toluenesulfinate scavenger[183] to give the final building block **136** in good yield.
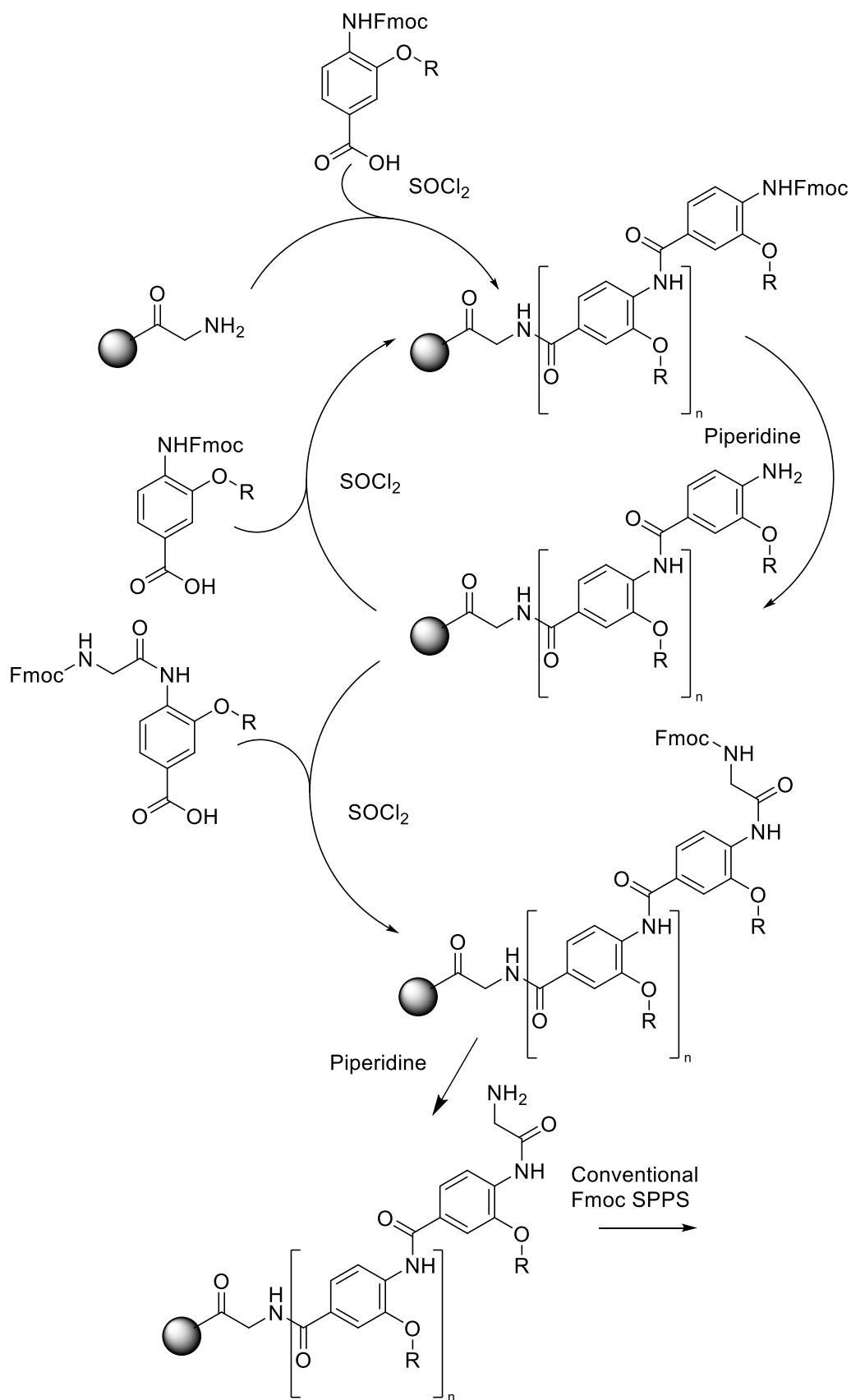
Scheme 16 - Solid phase synthesis of oligobenzamide peptide hybrids

4.5.3 Application to HIF-1α/p300

The Fmoc-Gly-^{i}Bu monomer **136** was then used in the modified oligomer synthesis depicted in Scheme 16. Initially the method is the same as the one shown in Scheme 13, however when the terminal benzamide residue was reached, the glycine conjugated monomer was included. Deprotection of the glycine amine with piperidine provided a handle for standard solid phase peptide synthesis. The pendant peptide was then extended using coupling reagent HCTU in a conventional manner.

This methodology was applied to the helix 2 and 3 region of the *C*-terminal transactivation domain (CTAD) of HIF-1α. The helix 3 region was replaced with the previously identified helix mimetic (see Chapter 5) and the rest of the amino acids kept unchanged as shown in Figure 79. The preparation of compound **137** was performed in a single run on an automated peptide synthesiser (CEM Liberty).

137
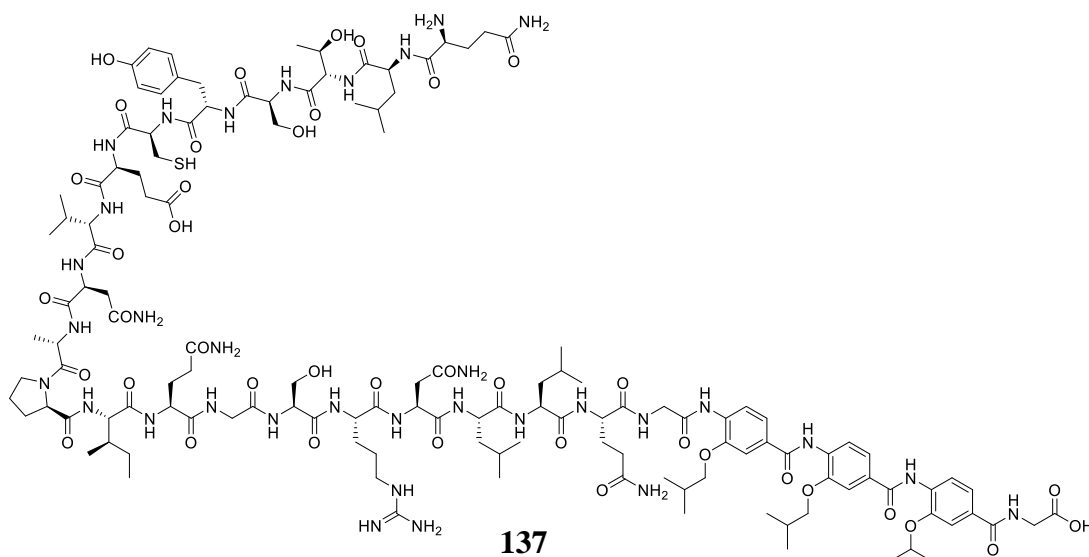
Figure 79 - Structure of oligobenzamide peptide hybrid **137**

The hybrid compound **137** was prepared via the methods discussed above and purified by preparative-HPLC three times, however each time it was unstable and could not be stored for any amount of time thus precluding the testing of its inhibitory activity. Due to the instability of hybrid **137**, several structural modifications were proposed to increase stability. The *N*-terminal amine was acetylated to curtail reactivity and the cysteine residue was replaced with a serine residue to give hybrid **138** (Figure 80). Hybrid **138** was prepared via the same methodology, purified by HPLC and exhibited significantly improved stability.

138

Figure 80 - Structure of oligobenzamide peptide hybrid **138**

4.5.4 Assessment of Inhibitory Activity

The increased stability of hybrid **138** allowed it to be tested in the HIF-1α/p300 fluorescence anisotropy competition assays (described in Chapter 1). The hybrid was able to inhibit the interaction to almost exactly the same extent as the helix 2-3 peptide (Chapter 2) as shown in Figure 81. The similar IC_{50} values (Hybrid: 83 ± 1.8 μM; peptide: 89 ± 2.8 μM) were encouraging, despite the fact that they were not as potent as the helix mimetics alone (cf. 10 μM), because no potency is lost when 11 residues were replaced with a small molecule. Crucially, the peptide does not bear the cysteine to serine modification. As the mutation of this cysteine residue to a serine has been shown to reduce activity by approximately 10-fold,[35] it is conceivable that the hybrid has improved potency provided by the unnatural portion but reduced potency from the serine residue. Further compounds in this series could further elucidate this effect and a direct comparison with the cysteine to serine modified peptide should be performed. It should also be noted that this affinity does not come from the peptidic region alone as the "Helix 2 plus linker" peptide comprises the same residues and shows no inhibitory activity (Chapter 4).

Figure 81 - Fluorescence anisotropy inhibitor curves for hybrid **138** and the helix 2-3 peptide

The decreased affinity, compared to the helix mimetics portion alone, implies that the peptide chain may be forcing the compound to bind in a suboptimal position or in a different position all together. Optimisation of the spacing/linker between the small molecule and the peptide sequence could be performed, especially if crystallographic data could be acquired. Molecular modelling of the complex (Figure 82), based on the previously docked oligoamide compound, suggests that further truncating the peptide in the region between helix 2 and 3 may provide a better fit. By beginning the peptide region at the indicated leucine residue rather than glycine, a higher affinity binder could conceivably be achieved.
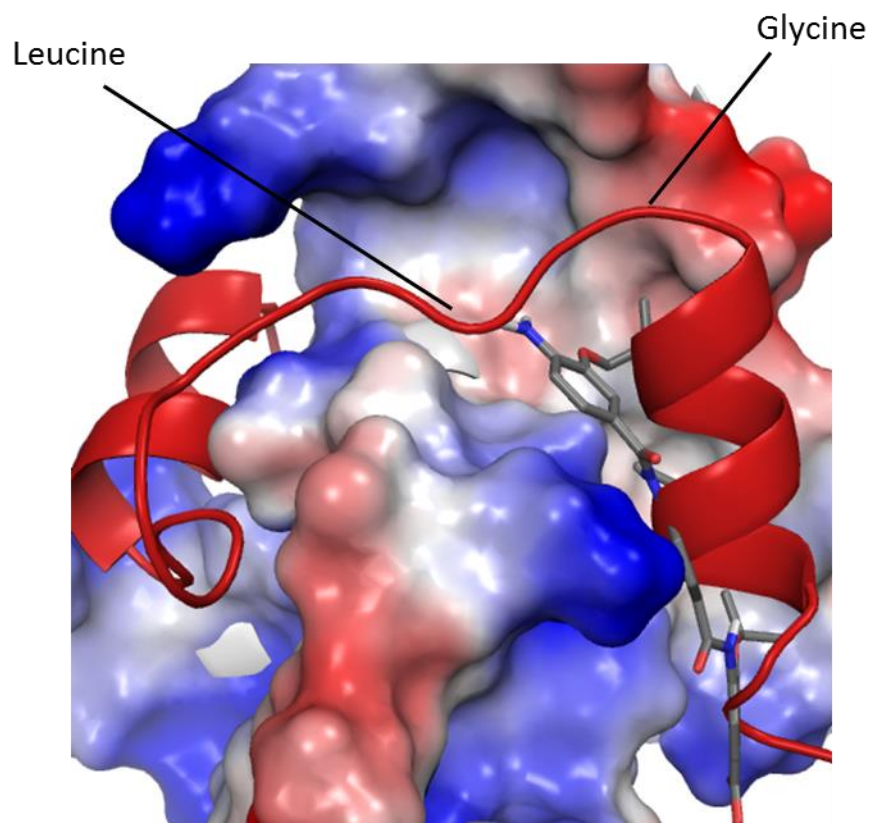
Figure 82 - Overlaid NMR structure and docked pose showing mismatch at the *N*-terminus of helix 3

4.6 Summary and Conclusions

This chapter introduced the concept of helix mimicry and then goes on to describe the design, synthesis and screening of small molecule helix mimetic inhibitors of the HIF-1α/p300 with low micromolar activity. The effect of scaffold is considered followed by a novel synthetic procedure for the preparation of oligobenzamides featuring designed against a different region of the interaction. Additionally a late stage functionalization procedure capable of incorporating diverse functionality is described and applied to developing analogues with increase solubility and photocrosslinking potential. Finally, synthetic methodology to provide novel peptide-oligobenzamide hybrid molecules has been developed. This methodology was then applied to the CTAD of HIF-1α to provide a compound with equivalent potency to the native peptide sequence it replaces. This provides further evidence that oligobenzamides can successfully mimic helices.

The work presented in this chapter demonstrated the applicability of helix mimetic strategies to a new target and provided the first biophysically characterised inhibitors of a therapeutically relevant PPI. Whilst the extended SAR studies and photo-crosslinking analogues failed to further elucidate the binding mode, the

methodology reported herein will hopefully prove useful for other targets/projects. Further work could focus on further increasing affinity and increasing solubility in a fashion which does not reduce activity.

Thesis Summary and Future Work

This thesis began with a description of two biological pathways featuring α-helix mediated protein-protein interactions, both of which provide potential points for therapeutic intervention during cancer progression. The HIF-1α/p300 interaction is crucial for oxygen homeostasis and provides solid tumours a pathway through which hypoxia can be lifted. The eIF4E/eIF4G interaction provides control over expression of proteins by limiting the recruitment of mRNA to the ribosome. Overexpression of eIF4E is a common feature of cancers and allows the overexpression of anti-apoptotic proteins, thus preventing programmed cell death. The currently available inhibitors of these interactions were reviewed and discussed.

Seeking to identify and develop new inhibitors of the HIF-1α/p300 and eIF4E/eIF4G interactions, binding assays were developed and the requirements for binding investigated. It was found that the complete CTAD of HIF-1α was required for high affinity binding to p300 but that the *C*-terminal Helix 3 provided a significant amount of this affinity and could be a potential target for inhibitor design. An investigation into cooperativity between the Helix 2 and Helix 3 binding regions was carried out and the effect of the chelate effect investigated. Some of the results in Chapter 2 hint at the importance of the previously disregarded Helix 1 region which requires further investigation. A full alanine scan of the HIF-1α CTAD followed by biophysical analysis of the binding to p300 has been proposed.

The requirements for eIF4E/eIF4G binding are better known in the literature but the studies reported here suggest that the binding is independent of m7G Cap binding at a distal position. Further work is warranted to fully understand the reported binding of both 4E-BP and **4EGI-1** at an allosteric site and the potential of this allosteric site for inhibitor development.

Subsequently, competition assays capable of identifying inhibitors were developed for both PPIs of interest. These screening assays should allow the discovery of potent inhibitors and future work may focus on the high-throughput screening of compound libraries. These assays were used to screen compounds identified from virtual libraries of commercially available and bespoke diverse compound libraries. Both docking and pharmacophore modelling were used to identify compounds for purchase or further virtual enumeration. Purchased compounds were screened and a non-selective covalent inhibitor **10** identified as well as a weak inhibitor of the eIF4E/eIF4G interaction, **4**. A small library of compounds based on compound **4** was prepared but no clear SAR could be determined. Further work could focus on the further exploration of this compound, particularly moving away from the

aminothiazole core. Additional docking and screening could also provide more starting points for inhibitor development.

A novel, drug like, small molecule scaffold emerged from the docking of bespoke libraries against HIF-1α/p300 and underwent optimisation by iterative virtual library enumeration and docking before compounds were selected for synthesis. A small library of novel compounds was prepared and screened against HIF-1α/p300 resulting in the identification of a potential starting point for a medicinal chemistry campaign. Further work will focus on optimising the potency, screening for selectivity and elucidating the binding mode.

A proteomimetic approach was also employed for HIF-1α/p300. Oligobenzamides capable of recapitulating the Helix 3 sidechains were prepared and screened, thus identifying small molecule inhibitors of the HIF-1α/p300 interaction. A library of analogues was prepared to explore the SAR and a binding model proposed based on docking and the results from the library screening. A literature compound, **KCN-1**, was also prepared *via* an improved route and found not to inhibit the HIF-1α/p300 interaction. Further work would focus on improving the potency and physical properties of the oligobenzamide inhibitor **1**. This would be aided hugely by structural studies of the compounds in complex with p300, either by NMR or X-ray crystallography.

New methodology was developed to allowed the mimicry of the Helix 2 region of the HIF-1α CTAD, however this compound proved to be inactive. A late stage functionalization protocol for oligobenzamides was developed to allow easier access to large libraries for SAR studies and the incorporation of delicate functionalities. The late stage functionalization protocol was employed to provide another set of oligobenzamide analogues with amine sidechains which proved to be weaker inhibitors, albeit with increased aqueous solubility. The late stage functionalization procedure also allowed the incorporation of a trifluoromethyl diazarine moiety for photo-crosslinking. Unfortunately, no photo-crosslinking could be observed to p300. Despite the fact that the late stage functionalization protocol failed to produce improved inhibitors or crosslinking compounds, the methodology should provide useful for the preparation of larger libraries for SAR studies for this and other targets in the future. Incorporation of the photocrosslinking capabilities in other positions should be carried out to attempt to validate the binding mode of these compounds.

Finally, a new hybrid compound containing both a helix mimetic compound and a peptide was proposed and prepared. New solid phase methodology was employed to prepare two examples of this type. Screening of the hybrid compounds revealed them to be as potent as the peptide they were designed to mimic but not as potent as the small molecule alone. This provides further evidence that oligobenzamides are

capable of mimicking helices. However, further work is clearly required to improve the potency to the level of the small molecules and potentially beyond.

In summary, binding requirements for two PPIs were investigated allowing the development of screening assays. These assays were used, alongside computational approaches, to identify several small molecule inhibitors which may provide useful starting points for inhibitor development. An secondary structure mimicry approach also provided inhibitors of the HIF-1α/p300 interaction. New methodology was developed to further explore these proteomimetic inhibitors.

Chapter 5 – Experimental Section

The candidate confirms that the work submitted is his 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.

5.1 General Experimental

All commercial solvents and reagents were used without further purification unless stated otherwise. All non-aqueous reactions were performed under an atmosphere of nitrogen and using anhydrous solvents. Water-sensitive reactions were performed in oven-dried glassware, cooled under nitrogen before use, or flame dried and cooled, under vacuum if stated. Solvents were removed under reduced pressure using a Büchi rotary evaporator. Ether refers to diethyl ether and petrol refers to petroleum spirit (b.p. 40-60 °C). Flash column chromatography was carried out using silica (35-70 μm particles) or alumina (neutral, Brockman activity 1), with crude reaction mixtures loaded in the initial solvent system or its least polar constituent. Thin layer chromatography was carried out on commercially available silica pre-coated aluminium plates (Kieselgel 60 F254, Merck) or commercially available alumina pre-coated glass plates (neutral, Brockman activity 1). Strong cation exchange columns were carried out using SCX, 5.0 g pre-packed cartridge, Supelco.

Proton and carbon NMR spectra were recorded on a Bruker Avance 500, Avance DPX300 or DRX500 spectrophotometer with an internal deuterium lock. Carbon NMR spectra were recorded with composite pulse decoupling using the waltz 16 pulse sequence. Chemical shifts are quoted in parts per million downfield of tetramethylsilane, and coupling constants (J) are given in Hz. NMR spectra were recorded at 300 K unless otherwise stated. Infra-red spectra were recorded using a Perkin–Elmer Spectrum One FT-IR spectrophotometer. Melting points were determined using a Griffin and George melting point apparatus and are uncorrected. Nominal mass spectrometry was routinely performed on a Bruker HCT Ultra spectrometer using electrospray (+) ionization. Nominal and accurate mass spectrometry using electrospray ionisation was carried out by staff or the candidate in the School of Chemistry using a Micromass LCT-KA111, Bruker MicroTOF or Bruker MaXis Impact TOF mass spectrometer. Mass-directed HPLC purifications were run on an Agilent 1260 Infinity Preparative HPLC system equipped with a Waters XBridge™ Prep C18 19 × 100 mm column, 5 μm particle size, on an acetonitrile or methanol/water gradient (5-95% acetonitrile or methanol over 8 minutes) and an Agilent 6120 Quadrupole system equipped with a quadrupole MS detector, using electrospray ionisation (ESI).

5.1.1 Oligobenzamide Nomenclature
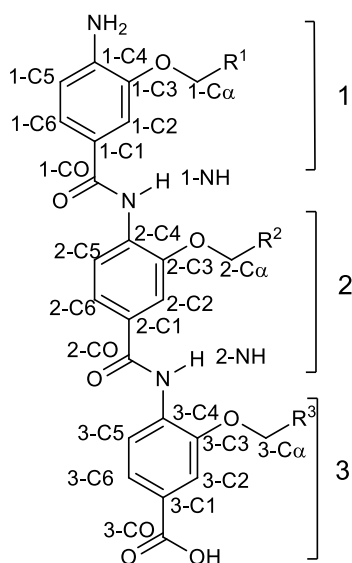
To simplify the numbering and NMR assignment of oligobenzamides, we have devised a sequential nomenclature, where each of the monomer building blocks is considered separately. The monomers are numbered from 1 to 3 starting from the N-terminal. Within each monomer, the numbering is the same: the carbons from the aminobenzoic acid are numbered using the standard system (the aromatic carbon bearing the carboxylic acid is C1, the one bearing the amine is C4). Then, the lateral chain is numbered: the carbon attached to the oxygen is the Cα, and the numbering of the aliphatic part of the side chain continues with Cβ, etc. In the case of aromatic side chains, the aromatic carbons are numbered CAr1, CAr2, etc. The numbering of the protons is based on the carbon numbering. To differentiate each individual carbon/proton, the monomer number is added as a prefix to the carbon/proton number representative examples are given above.

5.2 Experimental for Chapter 2

5.2.1 HIF-1α CTAD Peptides

The HIF-1α *C*-terminal transactivation domain (CTAD) peptides were purchased from Proteogenix, France. The sequence is shown below.

Ac-SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN-NH$_2$

The labelled analogue has the same sequence with an additional aminohexanoic acid (Ahx) residue as the *N*-terminus was functionalised with fluorescein isothiocyanate (FITC).

FITC-Ahx-SMDESGLPQLTSYDCEVNAPIQGSRNLLQGEELLRALDQVN-NH$_2$

5.2.2 Determination of the Binding of Labelled HIF-1α Peptide to p300 Protein

Protein was serially diluted into a solution of labelled peptide (80 nM) and the plates incubated for 30 minutes at room temperature. Each experiment was run in triplicate and the fluorescence anisotropy measured using a EnVision 2103 MultiLabel plate reader (Perkin Elmer) with excitation at 480 nm and emission at 535 nm (5 nm bandwidths). The Intensity was calculated for each point using Eq. 1 and used to calculated anisotropy using Eq. 2. From a plot of anisotropy against protein concentration, the minimum and maximum anisotropies were obtained using a logistic sigmoidal fit in OriginPro 8.6.
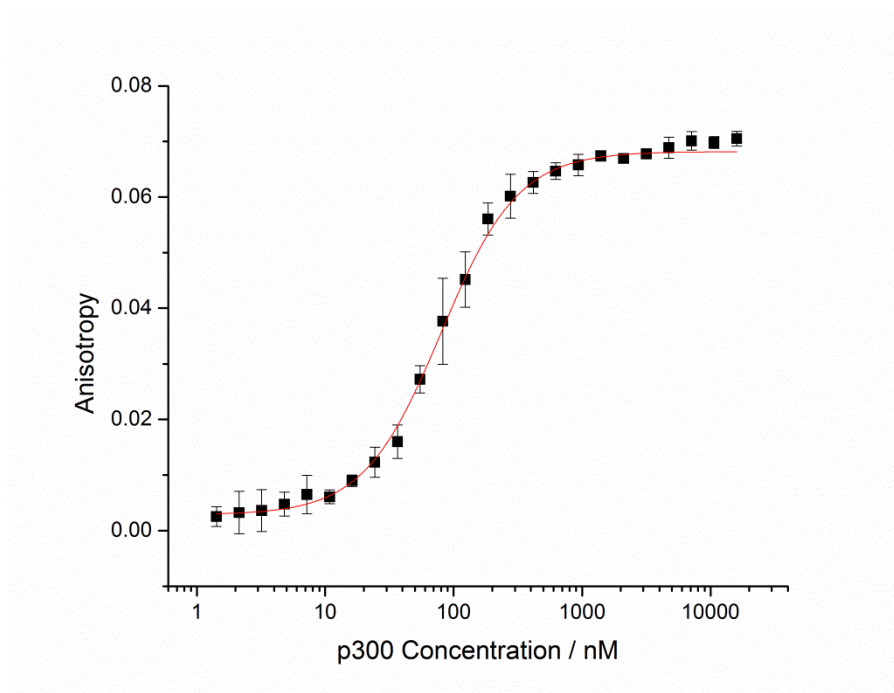
Determination of minimum and maximum anisotropy (80 nM FITC-HIF-1α CTAD, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol). Error bars represent the standard deviation of 3 repeats.

These values were used to convert anisotropy into amount of bound labelled peptide using Eq. 3. This data was then fitted using Eq. 4 in OriginPro 8.6 to determine the dissociation constant, K_D.

$$I = (2PG) + S \qquad \text{(Eq.1)}$$

$$r = \frac{S - PG}{1} \qquad \text{(Eq. 2)}$$

$$L_b = \frac{r - r_{min}}{\lambda(r_{max} - r) + r - r_{min}} \qquad \text{(Eq. 3)}$$

$$y = \frac{(K_D + x + [FL]) - \sqrt{[\{(K_D + x + [FL])\}^2 - 4x[FL]\}]}}{2} \qquad \text{(Eq. 4)}$$

R= anisotropy, I= total intensity, P=perpendicular intensity, S= parallel intensity, G = an instrument factor set to 1, L_b = fraction ligand bound, $\lambda = I_{bound}/I_{unbound} = 1$, [FL] = concentration of fluorescent peptide, K_D = dissociation constant, y = L_b multiplied by [FL], x = protein concentration
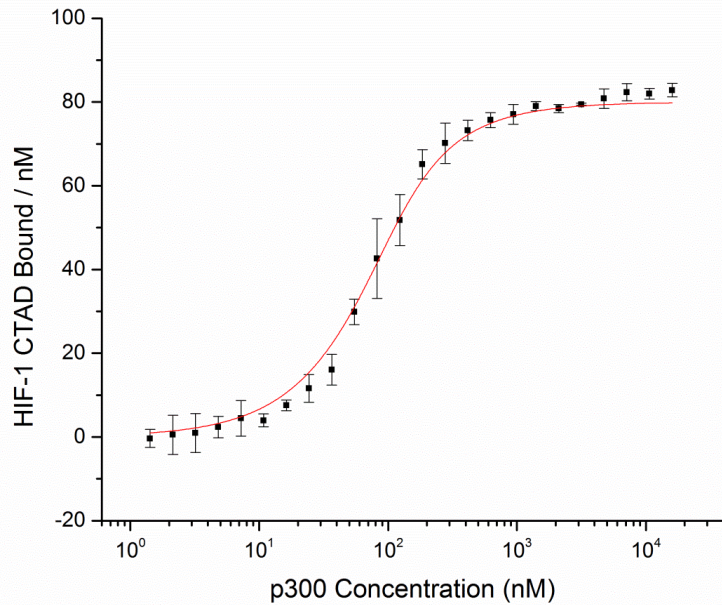
Direct titration of p300 into FITC-HIF-1α C-TAD followed by fluorescence anisotropy $K_D = 37.4 \pm 4.4$ nM (80 nM FITC-HIF-1α CTAD, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol). Error bars represent the standard deviation of 3 repeats.

Isothermal Titration Calorimetry

The HIF-1α peptide (100 μM) was titrated into p300 protein (10 μM) in matched ITC buffer (40 mM sodium phosphate, 100 mM sodium chloride and 5% glycerol). An initial injection of 10 μL peptide was made followed by 10 μL injections, over 20 seconds, made every 300 seconds for 30 injections. Peptide was also titrated in to buffer to give a baseline integration. This baseline was then subtracted to give a baseline-subtracted thermogram. The thermogram was then modeled by one-site interaction ITC data analysis, performed with Origin 5.0 software (MicroCal Inc.) The titration was conducted at 25 °C using a VP-ITC titration calorimeter from MicroCal Inc. All solutions were degassed prior to use.

5.2.3 Competition Assays

Compounds to be tested were serially diluted across a 384 well plate in buffer containing the minimum DMSO (final concentration <5%) before labelled peptide and protein were added sequentially. In parallel, a control experiment was performed in which no labelled peptide was added and the volume made up with additional buffer. Each experiment was performed in triplicate using a EnVision 2103 MultiLabel plate reader (Perkin Elmer) with excitation at 480 nm and emission at 535 nm (5 nm bandwidths) at 25 °C. Intensity and anisotropy were calculated as above using Eq. 1 and Eq. 2 respectively. Plots of anisotropy against compound concentration were fitted to a logistic sigmoidal dose response model to determine IC_{50} values. This procedure was used for all screening described below.

A positive control was performed with the unlabelled version of the labelled tracer peptide which is shown below.
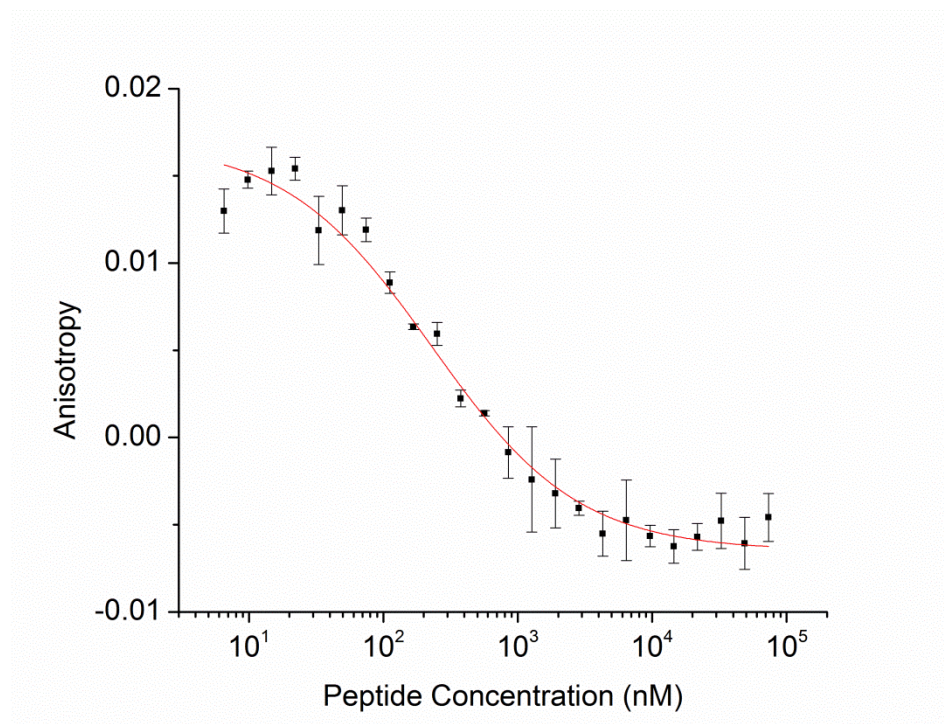
Competition titration of p300/HIF-1α with HIF-1α CTAD followed by fluorescence anisotropy $IC_{50} = 228 \pm 25$ nM (80 nM FITC-HIF-1α CTAD, 0.1 μM p300, 40 mM sodium phosphate, 100 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

5.2.4 Calculation of occluded surface area

Using PyMol, the 1L8C protein structure (downloaded from the PDB) was separated into component parts. The solvent accessible surface area for each component was measured using the command:

get_area, object

The settings for this command were as follows:

dot_solvent, 1 (calculates solvent accessible surface area rather than molecular surface area)

dot_density, 4 (the highest level of density)

solvent_radius, 1.4 (approximates water with a radius of 1.4 Å)

The difference between the sum of the surface areas of the individual components and the surface area of the complex gives an approximation of the occluded surface area. The results are summarised in the table below.

Component	Solvent accessible surface area / Å^2
p300 CH1 domain	7929
HIF-1α CTAD	5585
Complex	9304
Occluded surface area	4210

5.2.5 Peptide Synthesis

Peptides were synthesized by manual or automated Fmoc solid phase peptide synthesis methods on Rink Amide MBHA resin. For each coupling step 5 equivalents of the Fmoc-Amino acid in DMF was added to the resin along with 5 equivalents of HCTU and 5 equivalents of DIPEA in DMF and mixed for 2 hours followed by washing the resin 3 times for 2 minutes with DMF. For the Fmoc deprotection the resin was treated with an excess of 20% Piperidine in DMF 5 times for 2 minutes each followed by washing 5 times with DMF for 2 minutes each.

Acetylation

The resin bound peptide was treated with 10 equivalents of acetic anhydride and 10 equivalents of DIPEA and mixed overnight.

Fluorescein Labelling

Fmoc-amino hexanoic acid was coupled to the resin bound peptide and deprotected as before and it was then treated with fluorescein isothiocyanate (1.2 equivalents) overnight in the minimum volume of 12:7:5 pyridine/DMF/CH_2Cl_2. The pyridine was distilled over CaH_2 immediately before use.

Deprotection and Cleavage

The resin was washed with DMF 3 times and DCM 3 times prior to cleavage and deprotection. The peptides were globally deprotected and cleaved from the resin with TFA/thioanisole/water/phenol/EDT (83:5:5:5:2) [Reagent K], concentrated *in vacuo*, precipitated with ice cold ether, washed 3 times with ice cold ether and dried *in vacuo*. Peptides were purified by reverse phase HPLC as required.

Helix 2 Peptides

Ac-GTQLTSYDCEVNAAAG-NH$_2$

LC-MS *m/z* (ES) 1640.6 [M+H]$^+$, 821.6 [M+2H]$^{2+}$. HRMS Found: 1662.7046; $C_{67}H_{105}N_{19}NaO_{27}S$ requires [M+ Na]$^+$ 1662.7046.
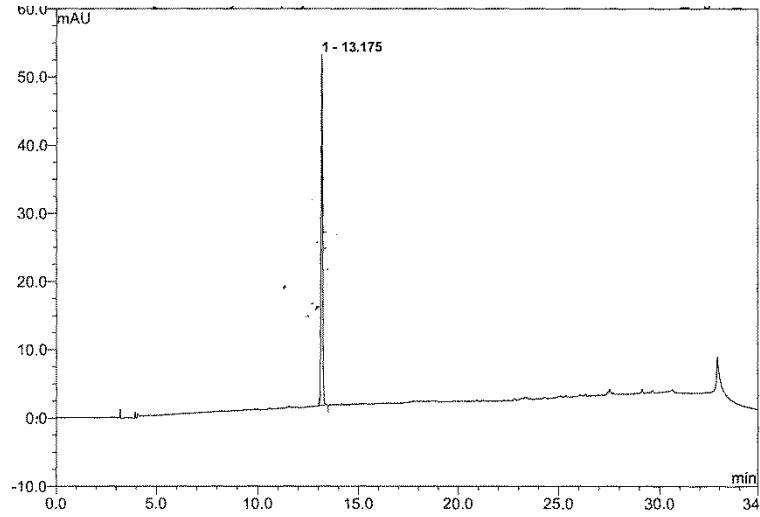
FITC-Ahx-GTQLTSYDCEVNAAAG-NH$_2$

LC-MS *m/z* (ES) 1051.3 [M+2H]$^{2+}$; MS Found: 2100.79: $C_{92}H_{125}N_{21}O_{32}S_2$ require [M+H]$^+$ 2100.82 (outside calibration range for HRMS)
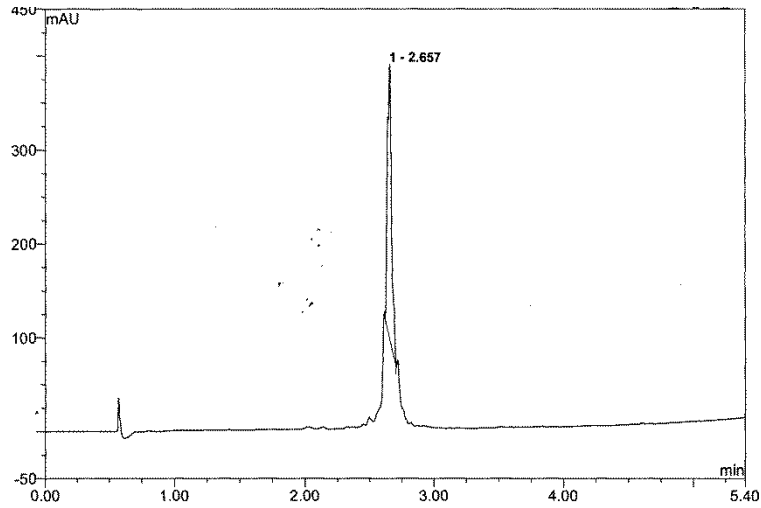
Helix 3 Peptides

Ac-GTEELLRALDQVNAAG-NH$_2$

LC-MS *m/z* (ES) 1697.7 [M+H]$^+$, 849.8 [M+2H]^2 HRMS Found: 849.4462, 1697.8843; C$_{70}$H$_{120}$N$_{22}$O$_{26}$ requires [M+2H]$^{2+}$ 849.4444, [M+H]$^+$ 1697.8851
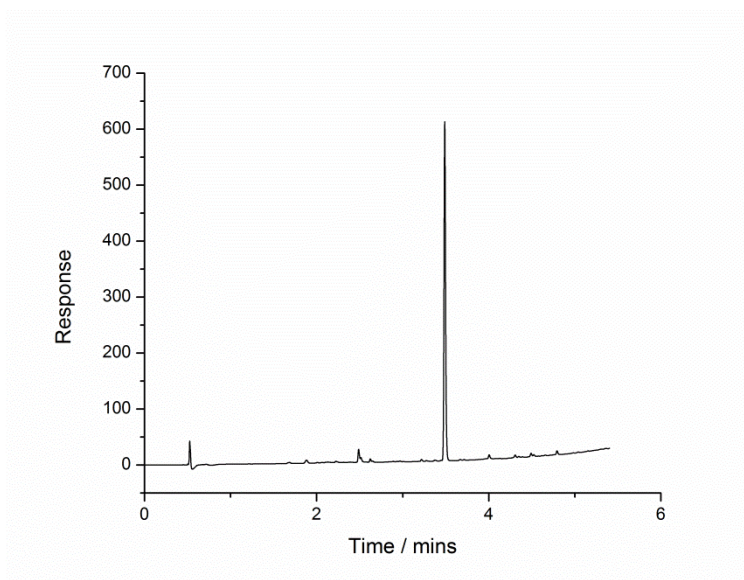
FITC-Ahx-GTEELLRALDQVNAAG-NH$_2$

LC-MS *m/z* (ES) 1079.8 [M+2H]$^{2+}$, 720.2 [M+3H]$^{3+}$ HRMS Found: 1080.0015 C$_{96}$H$_{140}$N$_{24}$O$_{31}$S requires [M+2H]$^{2+}$ 1080.0008
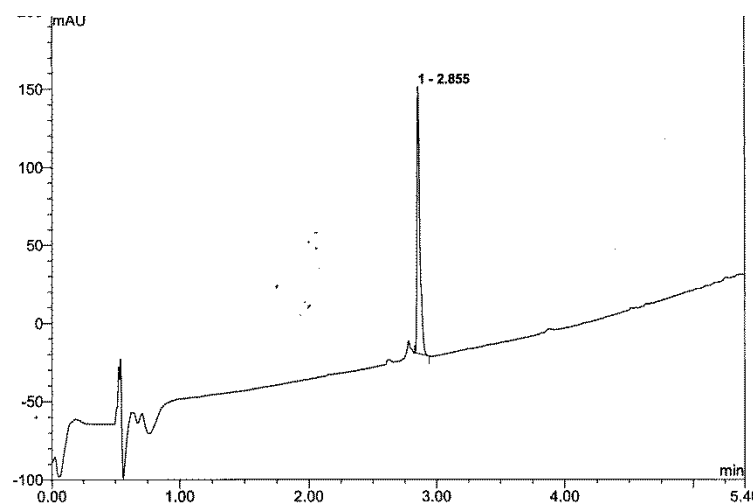
5.2.6 Peptide-PEG Conjugates

Fmoc-Amino-PEG$_{8/12/24}$-COOH

Amino-PEG$_{8/12/24}$-COOH (100 mg) was dissolved in 1:1 water/dioxane (6 ml) and sodium carbonate (3 eq.) was added followed by Fmoc-Cl (1.2 eq) and the reaction stirred overnight. The reaction mixture was then acidified with conc. HCl and extracted with ethyl acetate (3 × 5 ml). The combined organic layers were dried over

MgSO$_4$ and concentrated *in vacuo*. The product was checked by LC-MS and used without further purification.

Fmoc-Amino-PEG$_8$-COOH
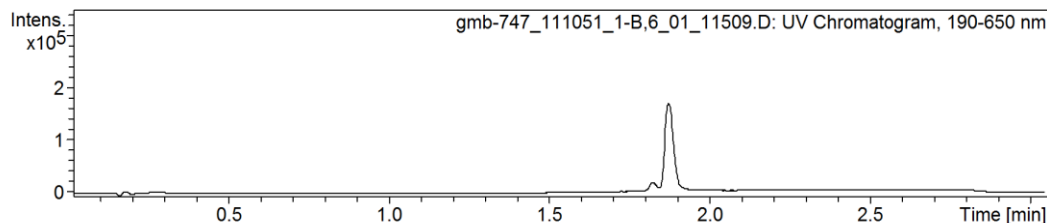
LC-MS *m/z* (ES) 664.5 [M+H]$^+$

Fmoc-Amino-PEG$_{12}$-COOH
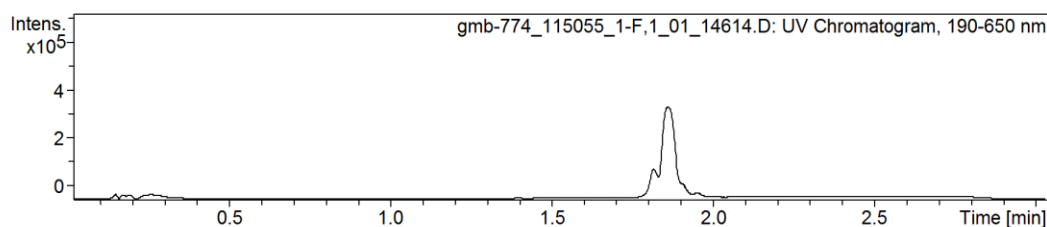
LC-MS *m/z* (ES) 840.7 [M+H]$^+$

Fmoc-Amino-PEG$_{24}$-COOH
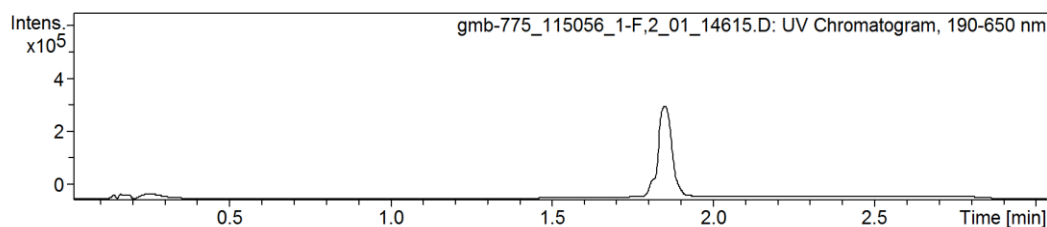
LC-MS *m/z* (ES) 1391.1 [M+Na]$^+$

Peptide-PEG Conjugates

The C-terminal portion of the peptide was synthesised using standard Fmoc-SPPS as above on 0.3 mmol scale and checked by cleaving a small amount and performing LC-MS analysis. The resin was then split into 3 equal portions.

To each portion was added the Fmoc-Amino-PEG-COOH monomer (PEG$_8$ − 2 eq., PEG$_{12}$ − 1.6 eq. and PEG$_{24}$ − 1.1 eq.), HCTU (5 eq.) and DIPEA (5 eq.) in DMF (2 ml). The reaction mixture was agitated for 16 hours and drained. The resin was washed 5 × 2 minutes with DMF and the remainder of the peptide synthesised using standard Fmoc-SPPS.

FITC-Ahx-QLTSYDCEVAN-PEG$_8$-EELLRALDQVN

LC-MS *m/z* (ES) 1150.7 [M+3H]$^{3+}$, 872 [M+4H]$^{4+}$ HRMS Found: 1724.2922 C$_{152}$H$_{231}$N$_{33}$O$_{54}$S$_2$ requires [M+2H]$^{2+}$ 1724.2922
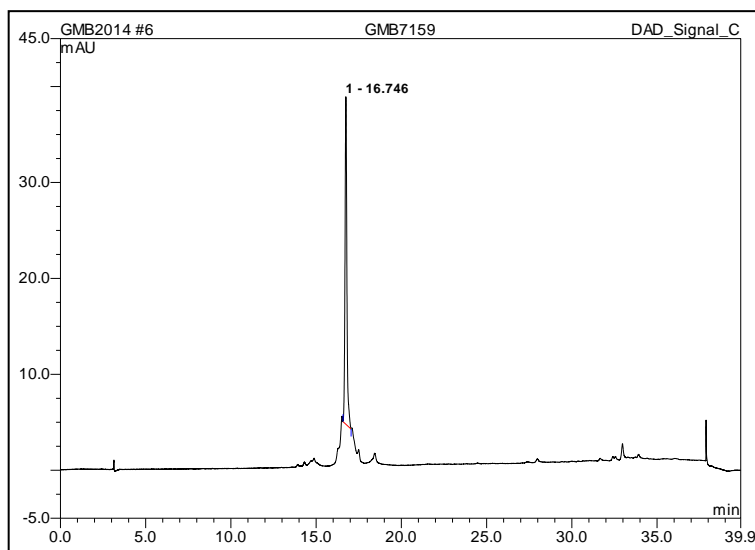
FITC-Ahx-QLTSYDCEVAN-PEG$_{12}$-EELLRALDQVN

LC-MS *m/z* (ES) 1813.1 [M+2H]$^{2+}$, 1209 [M+3H]$^{3+}$ HRMS Found: 1813.3517 C$_{160}$H$_{247}$N$_{33}$O$_{58}$S$_2$ requires [M+2H]$^{2+}$ 1813.3411
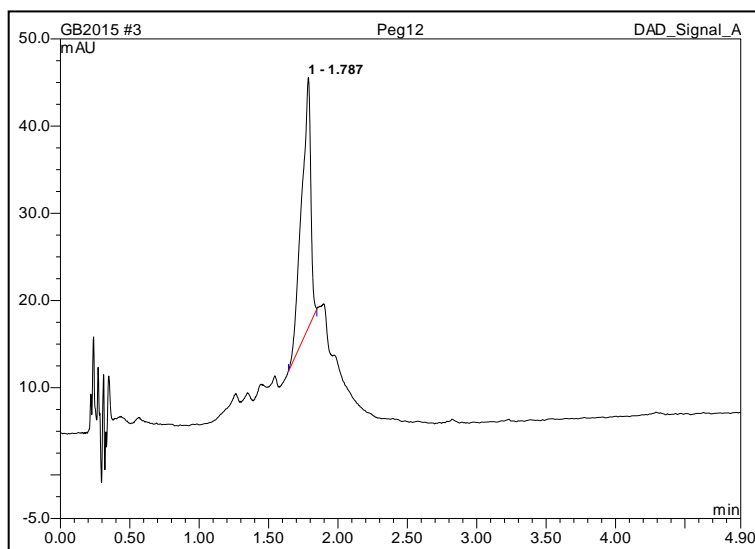
FITC-Ahx-QLTSYDCEVAN-PEG$_{24}$-EELLRALDQVN

LC-MS *m/z* (ES) 1385.5 [M+3H]$^{3+}$, 1039.5 [M+4H]$^{4+}$ HRMS Found: 2077.4836 C$_{184}$H$_{295}$N$_{33}$O$_{70}$S$_2$ requires [M+2H]$^{2+}$ 2077.4984
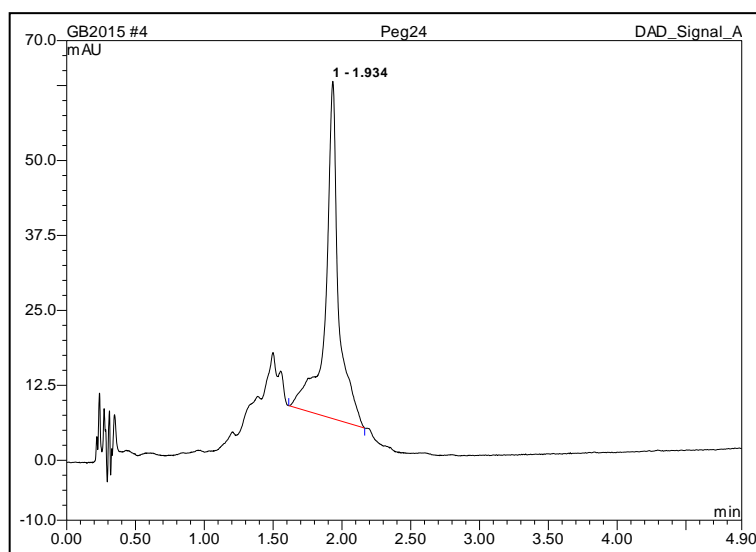
5.2.7 eIF4G Peptide Synthesis

Ac-RYDREFLLGFQ-NH$_2$

LC-MS *m/z* (ES) 1484.7 [M+H]$^+$, 743.0 [M+2H]$^{2+}$; HPLC r.t. 2.49 min (5-95% H$_2$O-MeCN + 0.1% TFA, 5.40 min, 0.5ml/min, Ascentis Peptide). HRMS Found: 742.8860; C$_{69}$H$_{101}$N$_{19}$O$_{18}$ requires [M+2H]$^{2+}$ 742.8786.
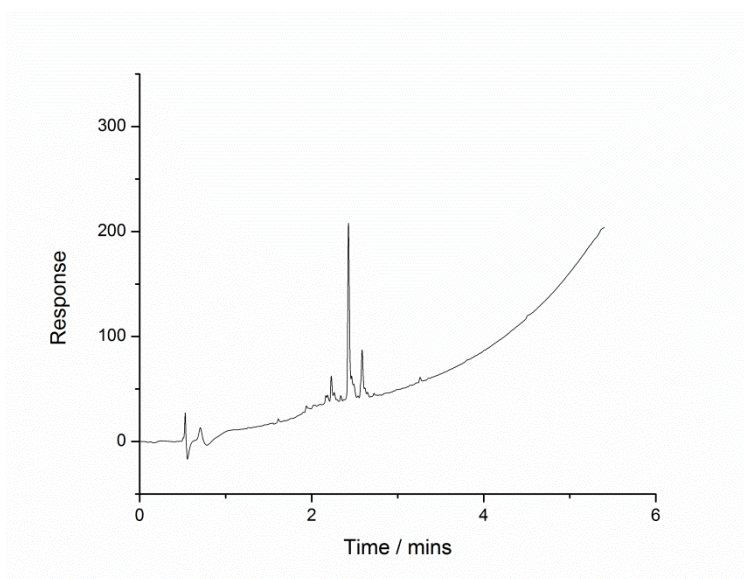
Fluorescein-Ahx-RYDREFLLGFQ-NH$_2$

LC-MS m/z (ES) 1945.6 $[M+H]^+$, 973.4 $[M+2H]^{2+}$, 649.7 $[M+3H]^{3+}$; HRMS Found: 1966.8350; $C_{94}H_{121}N_{21}NaO_{23}S$ requires $[M+Na]^+$ 1966.8557.
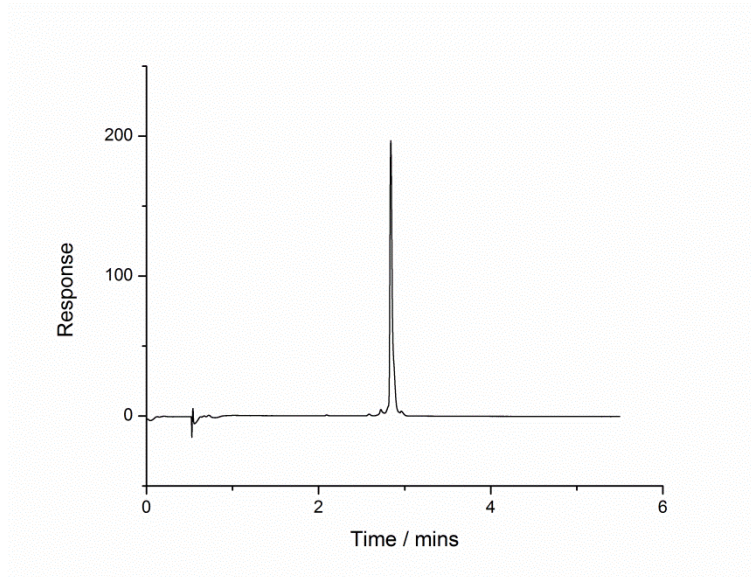
5.2.8 Determination of the Binding of Labelled eIF4G Peptide to eIF4E Protein

Development of a fluorescence anisotropy binding assay for eIF4E/eIF4G was performed as for the HIF-1a/p300 interaction described above.
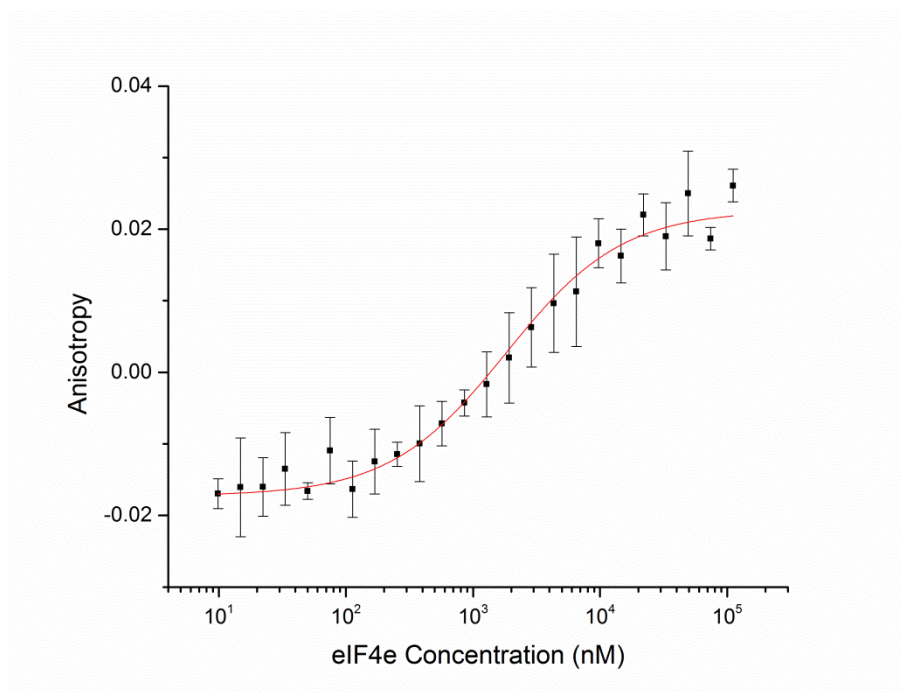
Determination of minimum and maximum anisotropy values by direct titration of eIF4E into FITC-eIF4G followed by fluorescence anisotropy $K_D = 1.76 \pm 0.17$ µM

(80 nM FITC-eIF4G, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.
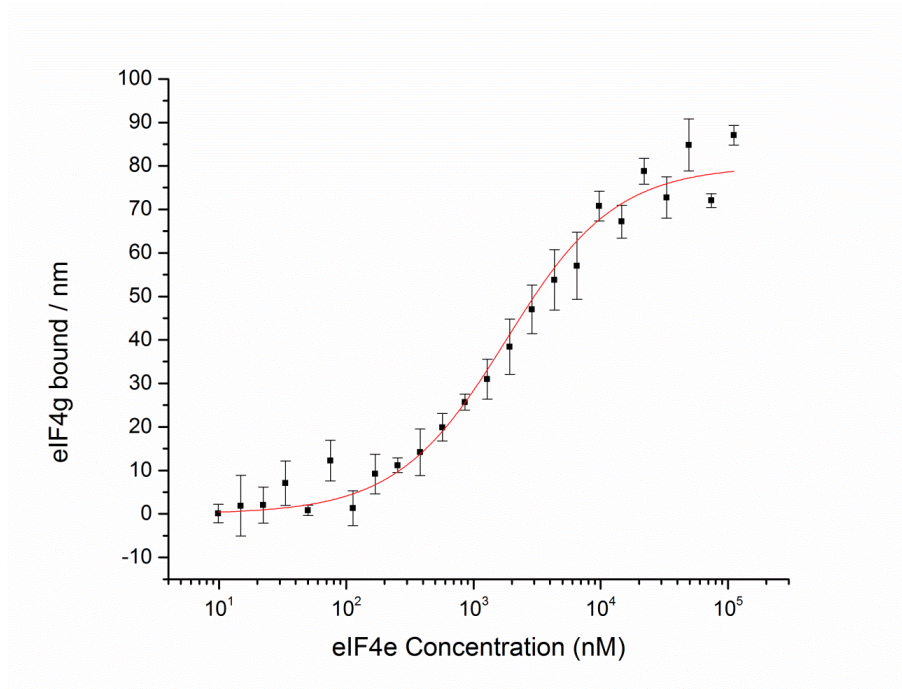
Direct titration of eIF4E into FITC-eIF4G followed by fluorescence anisotropy $K_D = 1.76 \pm 0.17$ µM (80 nM FITC-eIF4G, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

5.2.9 Competition Assays

The assay was validated for identification of competitive inhibitors through competition titration with unlabelled peptide as described for the HIF-1α/p300 interaction above.
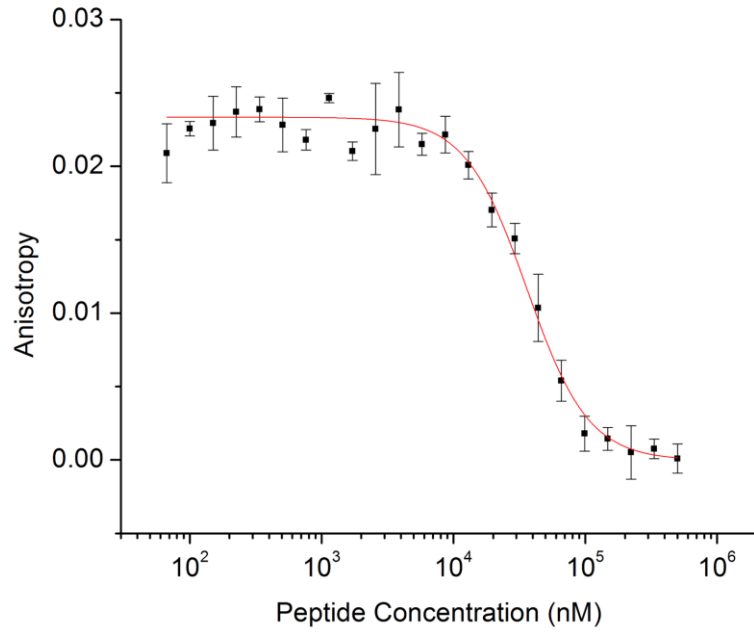
Competition titration of eiF4E/FITC-eIF4G with eIF4G followed by fluorescence anisotropy $IC_{50} = 36 \pm 4.5$ μM (80 nM FITC-eIF4G, 3 μM eIF4E, 40 mM sodium phosphate, 200 mM NaCl, 1 mM DTT, 5% glycerol, 0.1% triton). Error bars represent the standard deviation of 3 repeats.

5.3 Experimental for Chapter 3

5.3.1 Docking

Compound libraries generated in Pipeline Pilot 8.5 (Accelrys) were converted to 3D conformer libraries of up to 200 conformers per compound using Omega (OpenEye). The proteins were prepared using the protein preparation wizard in Maestro 9.3 (Schrodinger) by removing the helix donor binding partner, defining the binding site and setting any relevant constraints, optimising the hydrogen bonding within the remaining protein, removing any non-hydrogen bonding residual water molecules and performing a restrained minimisation.

Conformer libraries were then prepared using Ligprep (Schrodinger) and docked rigidly into binding sites of interest using Glide (Schrodinger). Compounds were selected based on docking score, recapitulation of important residues and consistency of pose.

eIF4E preparation – PDB ID: 2W97

Two hydrophobic constraints were defined based on peptide binding and known molecule NMR ^{15}N-HSQC shifts.
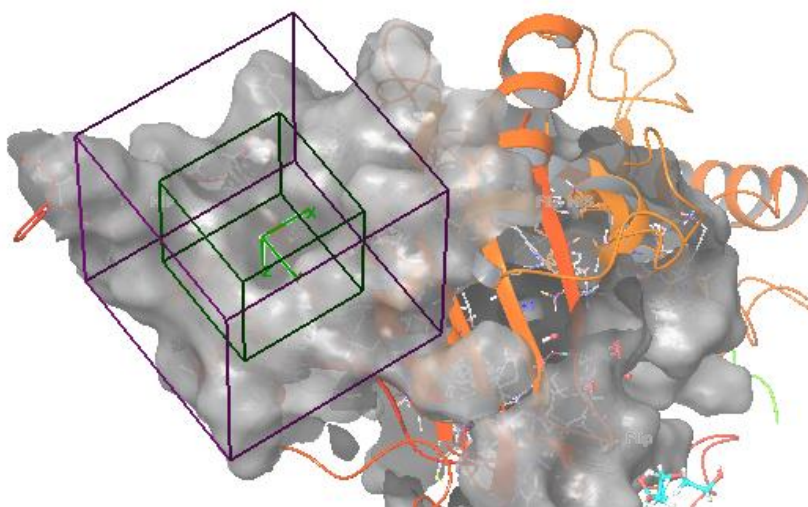
p300 preparation – PDB ID: 1L8C

Helix 2 binding site – two hydrogen bonding constraints were defined based on peptide protein interactions observed in the NMR structure and residues thought to be important from the literature.

Asp – hydrogen bond acceptor – polar hydrogen

Lys – hydrogen bond acceptor – polar hydrogen
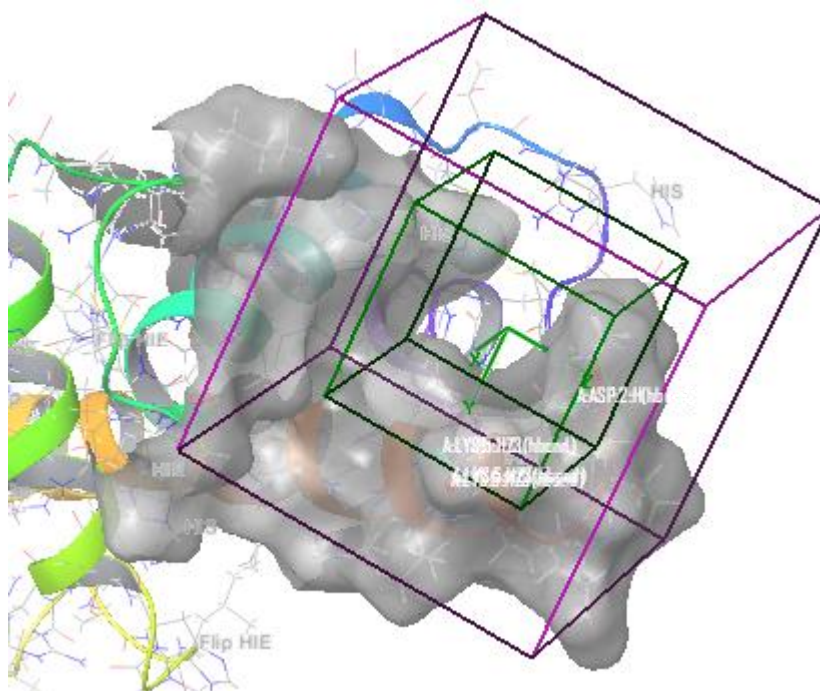
Helix 3 binding site – no constraints applied
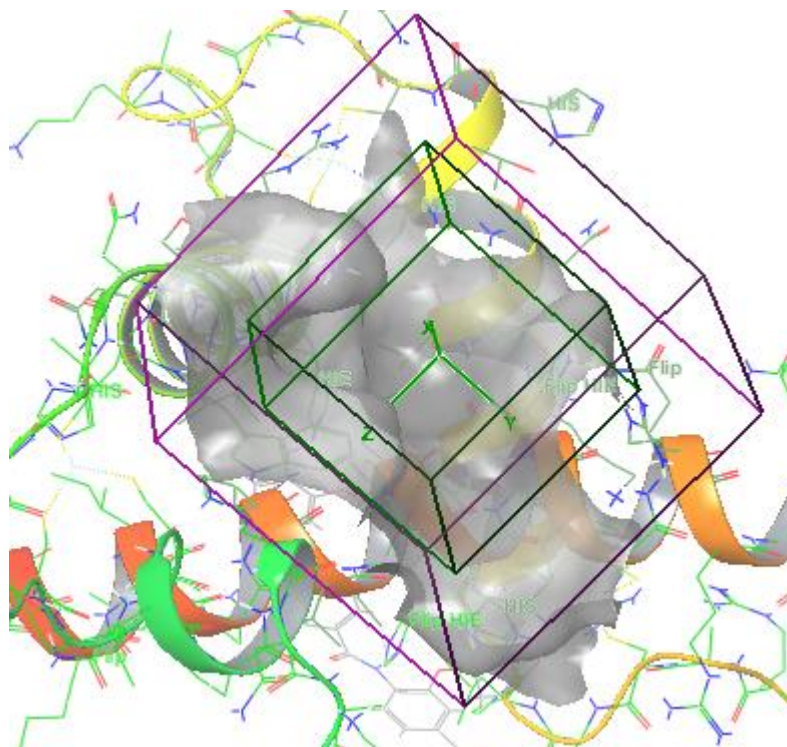
5.3.2 Pharmacophore modelling

Helical regions of the peptide donor binding partner were excised from the NMR or crystal structure using Maestro 9.3 (Schrodinger), the side chains on the non-binding faces of the helix were removed and the backbone converted to carbon whilst retaining the structural integrity of the helix. These structures were then imported into vROCs (Openeye) and used to generate queries based on matching the relevant important side chains which are shown below.
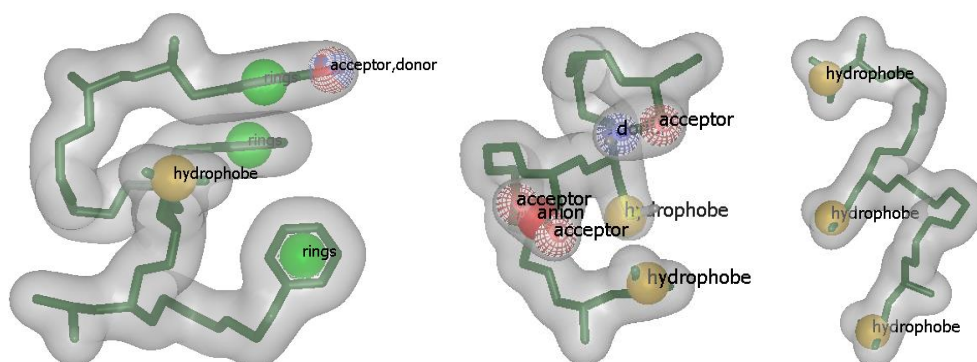
Multi-conformer libraries generated using Omega (Openeye) from the eMolecules collection of commercially available compounds were used as the database for ROCS analysis using the manually generated queries to search for molecules from a database which could match the desired structural elements. The top 100 scoring matches for each query were visually inspected and compounds of interest selected based on their ability to match the defined pharmacophores as well as the flexibility of the ligand. Compounds with more rigid structures that matched the query well were prioritised over flexible compounds capable of matching the query.

5.3.3 Library Enumeration

Virtual libraries were enumerated and manipulated using Accelrys Pipeline Pilot version 8.5 (Pipeline Pilot v8.5.0.200, Accelrys© Software Inc., 2011). A virtual library was enumerated based on the route shown in Scheme 17. The protocol (Figure 83) and building blocks used are shown below (Figures 2-5).
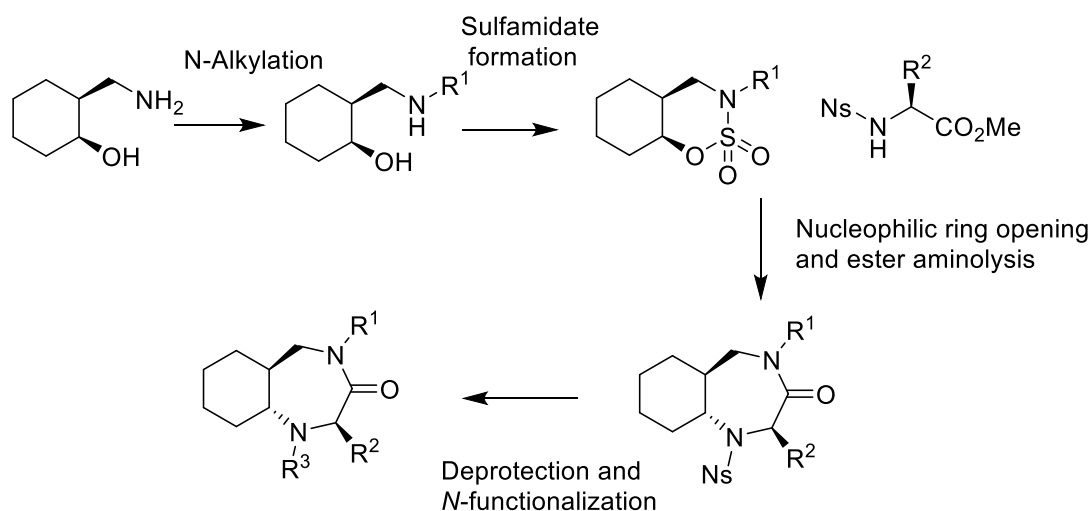
Scheme 17 - Route for virtual enumeration
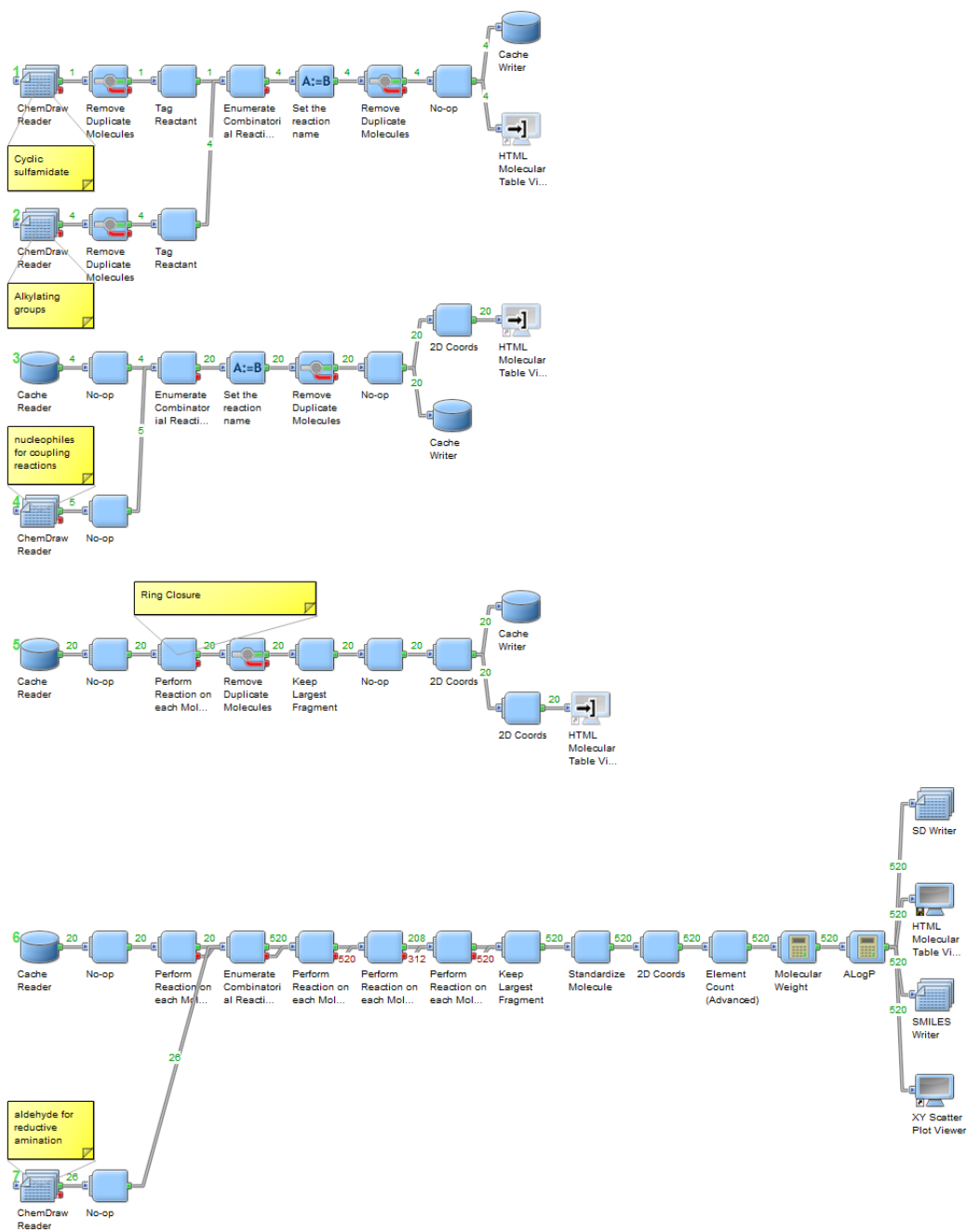
Figure 83 - Library enumeration protocol
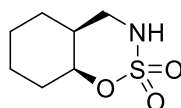
Figure 84 - Starting cyclic sulfamidate for bicyclic core enumeration

Figure 85 - Alkylating groups for bicyclic core enumeration
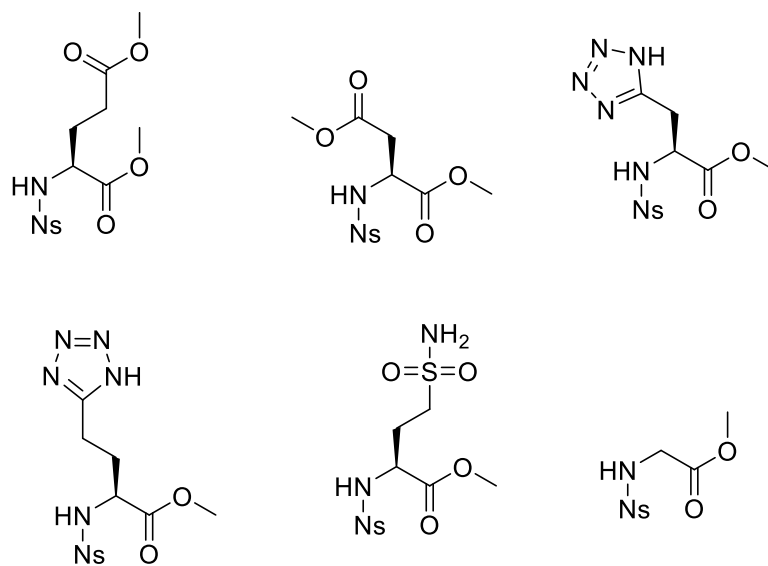
Figure 86 - Protected amino esters for bicyclic core enumeration
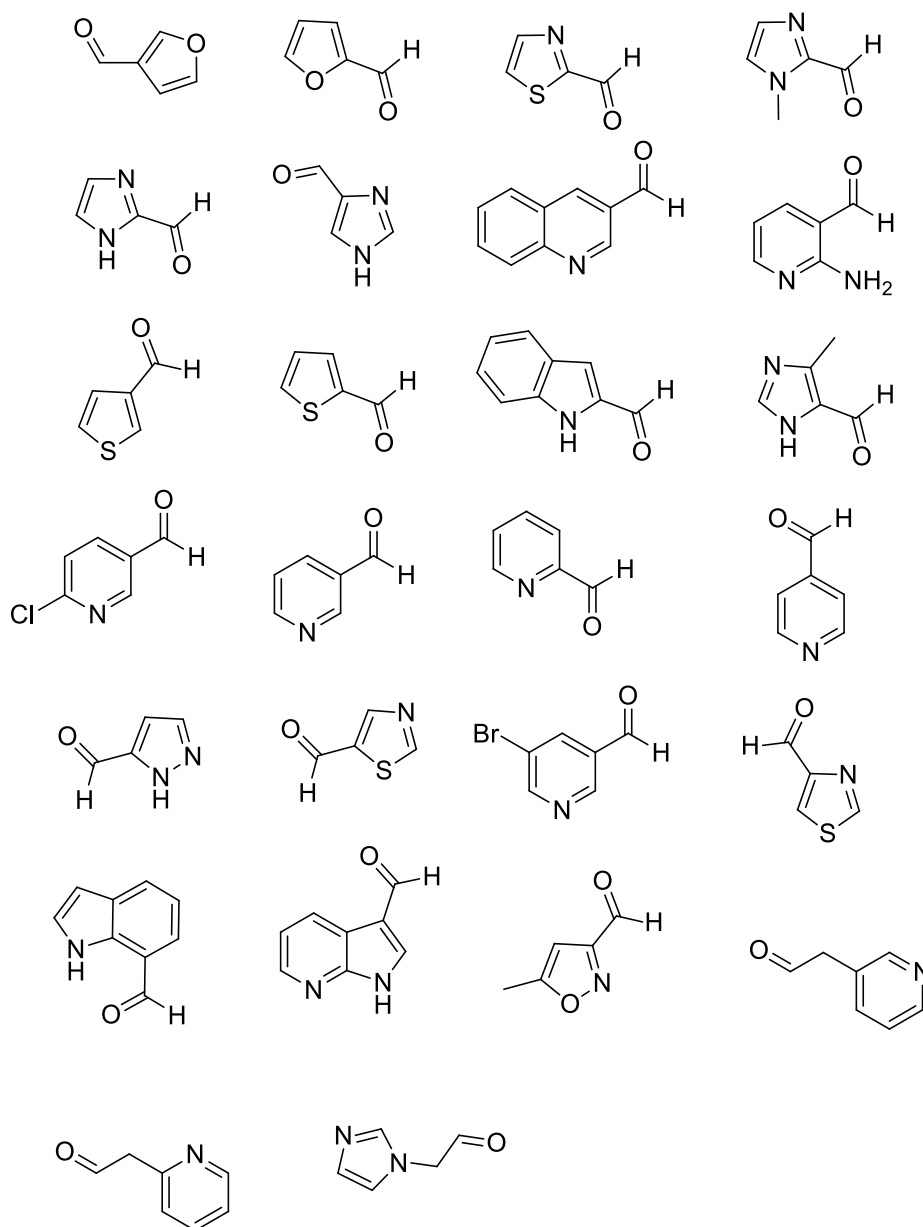
Figure 87 - Capping groups for bicyclic core enumeration

5.3.4 Synthesis of Compound 4 analogues

4-phenyl-1,3-thiazol-2-amine, 11
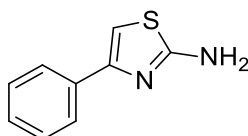
Following a literature procedure, [125] bromoacetophenone (1 g, 5 mmol) and thiourea (425 mg, 5.58 mmol) were heated to reflux in ethanol (50 ml) for 1 hour. The reaction mixture was then allowed to cool to r.t. and concentrated *in vacuo*, suspended in saturated aqueous sodium bicarbonate (30 ml) and extracted with DCM (3 × 20 ml). The combined organic layers were washed with brine, dried with MgSO$_4$ and concentrated *in vacuo*. The residue was stirred in petrol for 30 minutes

and the solid collected by filtration and dried *in vacuo* to give the desired product as an off-white solid (672 mg, 76%).^{1}H NMR (500 MHz, d^6-DMSO) δ 7.78 (m, 2H, Ar-H), 7.38 – 7.30 (m, 2H, Ar-H), 7.26 – 7.22 (m, 1H, Ar-H), 7.02 (br s, 2H, NH$_2$), 6.98 (s, 1H, 5-CH); ^{13}C NMR (125 MHz, d^6-DMSO) δ 168.14, 149.84, 134.91, 128.41, 127.12, 125.50, 101.44; LC-MS: 177.0 [M+H]

4-(3,4-dichlorophenyl)-1,3-thiazol-2-amine, 12[125]
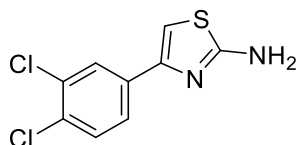
Following a literature procedure, [125] 2-bromo-1-(3,4-dichlorophenyl)ethan-1-one (600 mg, 2.25 mmol) and thiourea (188 mg, 2.47 mmol) were heated to reflux in ethanol (50 ml) for 1 hour. The reaction mixture was then allowed to cool to r.t. and concentrated *in vacuo*, suspended in saturated aqueous sodium bicarbonate (30 ml) and extracted with DCM (3 × 20 ml). The combined organic layers were washed with brine, dried with MgSO$_4$ and concentrated *in vacuo*. The residue was stirred in petrol for 30 minutes and the solid collected by filtration and dried *in vacuo* to give the desired product as an off-white solid (420 mg, 76%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.82 (d, J = 1.8 Hz, 1H, 2'C-H), 7.52 (dd, J = 8.4, 1.8 Hz, 1H, 6'C-H), 7.36 (d, J = 8.4 Hz, 1H, 5'C-H), 6.68 (s, 1H, 5-CH), 4.97 (br s, 2H, NH$_2$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 167.28, 149.02, 134.61, 132.76, 131.44, 130.50, 127.92, 125.10, 104.23; LC-MS: 244.9 [M+H]

4-tert-butyl-1,3-thiazol-2-amine, 13
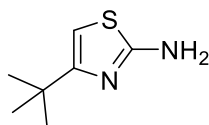
Following a literature procedure,[126] copper (II) acetate (120 mg, 0.65 mmol) was added to a solution of 1-bromopinacolone (0.88 ml, 6.57 mmol) and thiourea (500 mg, 6.57 mmol) in ethanol (50 ml) and stirred at r.t. for 30 minutes. The reaction was then diluted with water (30 ml) and extracted with ethyl acetate (3 × 20 ml). The combined organic layers were washed with brine, dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 10% ethyl acetate/petrol to give the desired product as an off-white solid (680 mg, 66%). Data matches literature.[202] R$_f$ 0.2 (10% ethyl acetate/petrol); ^{1}H NMR (500 MHz, CDCl$_3$) δ 6.11 (s, 1H, 5-CH), 5.16 (br s, 2H, NH$_2$), 1.28 (s, 9H, t-butyl); ^{13}C NMR (125 MHz, CDCl$_3$) δ 166.95, 162.60, 99.86, 34.53, 29.70; LC-MS: 157.0 [M+H]

N-(4-phenyl-1,3-thiazol-2-yl)benzamide, 17
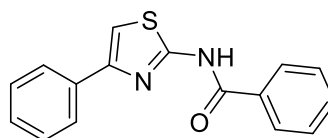
Benzoyl chloride (78 μl, 0.68 mmol) was added dropwise to a cooled solution (0 °C) of 4-phenyl-1,3-thiazol-2-amine (100 mg, 0.57 mmol) and diisopropylethylamine (0.3 ml, 1.7 mmol) in THF (5 ml) and allowed to warm to r.t. overnight with stirring. Reaction mixture concentrated *in vacuo* and purified by mass-directed chromatography to give the title compound as an off-white amorphous solid (83.4 mg, 52%). ν_{max}/cm^{-1} (solid state) 3126, 2780, 1667; ^{1}H NMR (500 MHz, CDCl$_3$) δ 10.86 (s, 1H, Amide NH), 7.84 (d, J = 7.7 Hz, 2H, Ar-H), 7.77 (d, J = 7.7 Hz, 2H, Ar-H), 7.51 (t, J = 7.4 Hz, 1H, Ar-H), 7.42 – 7.32 (m, 4H, Ar-H), 7.31 – 7.25 (m, 1H, Ar-H), 7.22 (s, 1H, 5C-H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 165.17, 158.94, 150.21, 134.14, 132.75, 131.89, 128.79, 128.70, 128.07, 127.43, 126.06, 108.09; HRMS Found: 281.0745; C$_{16}$H$_{12}$N$_2$OS [M+H]$^+$ requires 281.0743.

2-(2'-bromophenyl)-*N*-[4-(3'',4''-dichlorophenyl)-1,3-thiazol-2-yl]acetamide, 18
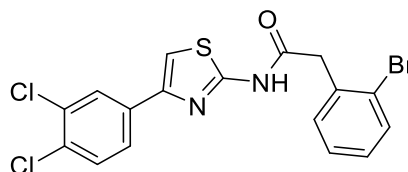
N-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide (481 mg, 2.52 mmol), DMAP (15 mg, 0.12 mmol) and diisopropylethylamine (1 ml, 6.3 mmol) were added sequentially to a solution of 2-bromophenylacetic acid (300 mg, 1.39 mmol) in DCM (20 ml). The reaction was stirred for 30 minutes at r.t. followed by the addition of 4-(3,4-dichlorophenyl)-1,3-thiazol-2-amine (310 mg, 1.26 mmol) after which the reaction was stirred overnight. The reaction mixture was washed with 2M HCl (aq.) (2 × 20 ml) and saturated aqueous sodium bicarbonate (2 × 20 ml), dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 10% ethyl acetate/petrol to give the desired compound as a colourless amorphous solid (337 mg, 60%). R$_f$ 0.3 (10% ethyl acetate/petrol); ν_{max}/cm^{-1} (solid state) 3172, 3077, 2997, 1697; ^{1}H NMR (500 MHz, d^6-DMSO) δ 12.62 (s, 1H, Amide NH), 8.15 (d, J = 1.9 Hz, 1H, 2''C-H), 7.89 (dd, J = 8.4, 1.9 Hz, 1H, 6''C-H), 7.83 (s, 1H, 5C-H), 7.70 (d, J = 8.4 Hz, 1H, 5''C-H), 7.63 (d, J = 7.9 Hz, 1H, 3'C-H), 7.45 (d, J = 7.5 Hz, 1H, 6'C-H), 7.38 (t, J = 7.4 Hz, 1H, 5'C-H), 7.26 (t, J = 7.7 Hz, 1H, 4'C-H), 4.03 (s, 2H, acetamide CH$_2$); ^{13}C NMR (126 MHz, d^6-DMSO) δ 168.53, 158.21, 146.22, 134.81, 134.69, 132.42, 132.32, 131.55, 131.01, 130.02, 129.11, 127.70, 127.32, 125.68, 124.52, 110.14, 41.90; HRMS found 442.9205; C$_{17}$H$_{12}$BrCl$_2$N$_2$OS [M+H]$^+$ requires 442.9205

N-[4-(3'',4''-dichlorophenyl)-1,3-thiazol-2-yl]-2-(2'-methoxyphenyl)acetamide, 19
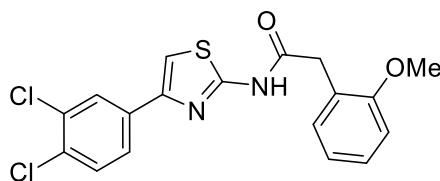
N-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide (156 mg, 0.82 mmol), DMAP (5 mg, 0.04 mmol) and diisopropylethylamine (0.35 ml, 2.05 mmol) were added sequentially to a solution of 2-methoxyphenylacetic acid (136 mg, 0.82 mmol) in DCM (20 ml). The reaction was stirred for 30 minutes at r.t. followed by the addition of 4-(3,4-dichlorophenyl)-1,3-thiazol-2-amine (100 mg, 0.41 mmol) after which the reaction was stirred overnight. The reaction mixture was concentrated *in vacuo,* re-dissolved in ethyl acetate, washed with 2M HCl (aq.) (2 × 20 ml) and saturated aqueous sodium bicarbonate (2 × 20 ml), dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 10% ethyl acetate/petrol and triturated with hexane to give the desired compound as a colourless amorphous solid (58 mg, 36%). R$_f$ 0.25 (10% ethyl acetate/petrol); v_{max}/cm^{-1} (solid state) 3187, 1655, 1549; ^{1}H NMR (500 MHz, CDCl$_3$) δ 9.25 (s, 1H, Amide NH), 7.91 (d, *J* = 2.0 Hz, 1H, 2''C-H), 7.61 (dd, *J* = 8.4, 2.0 Hz, 1H, 6''C-H), 7.45 (d, *J* = 8.4 Hz, 1H, 5''C-H), 7.36 (td, *J* = 8.0, 1.7 Hz, 1H, 6'C-H), 7.30 (dd, *J* = 7.4, 1.7 Hz, 1H, 4'C-H), 7.12 (s, 1H, 5C-H), 7.01 (m, 2H, 3'C-H and 5'C-H), 3.97 (s, 3H, OMe), 3.84 (s, 2H, acetamide CH$_2$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 169.06, 157.98, 157.04, 147.27, 134.20, 132.91, 131.80, 131.51, 130.62, 129.68, 127.93, 125.11, 121.58, 111.21, 108.93, 55.81, 38.84; HRMS Found: 393.0236; C$_{18}$H$_{14}$Cl$_2$N$_2$O2$_S$ [M+H]$^+$ requires 393.0225

N-[4-(3'',4''-dichlorophenyl)-1,3-thiazol-2-yl]-2-(2'-phenylphenyl)acetamide, 20
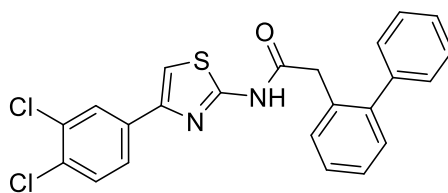
A microwave reaction vessel was charged with 2-(2-bromophenyl)-*N*-[4-(3,4-dichlorophenyl)-1,3-thiazol-2-yl]acetamide (20 mg, 0.045 mmol), phenylboronic acid (5.4 mg, 0.045 mmol), tetrakis(triphenylphosphine) palladium (2.6 mg, 5 mol%), caesium carbonate (44 mg, 0.135 mmol) and dioxane (3 ml). The reaction was then sealed and heated to 120 °C for 1 hour in the microwave, filtered through Celite, concentrated *in vacuo* and purified by mass directed chromatography (8 mg, 41%). v_{max}/cm^{-1} (solid state) 3425, 2926, 2853, 1750; ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.52 (s, 1H, Amide NH), 7.92 (d, *J* = 1.9 Hz, 1H, Ar-H), 7.62 (dd, *J* = 8.4, 2.0 Hz,

1H, Ar-H), 7.50 – 7.38 (m, 8H, Ar-H), 7.34 – 7.30 (m, 2H, Ar-H), 7.15 (s, 1H, 5'C-H), 3.84 (s, 2H, acetamide CH$_2$); HRMS Found 439.0440, C$_{23}$H$_{16}$Cl$_2$N$_2$OS [M+H]$^+$ requires 439.0433.

N-[4-(3'',4''-dichlorophenyl)-1,3-thiazol-2-yl]-2-[2'-(2''',6'''-dimethoxypyridin-3-yl)phenyl]acetamide, 21
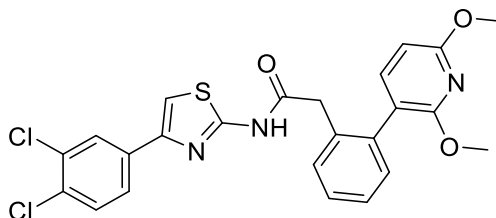
A microwave reaction vessel was charged with 2-(2-bromophenyl)-*N*-[4-(3,4-dichlorophenyl)-1,3-thiazol-2-yl]acetamide (20 mg, 0.045 mmol), 2,6-Dimethoxy-3-pyridineboronic acid (10 mg, 0.054 mmol), tetrakis(triphenylphosphine) palladium (2.6 mg, 5 mol%), caesium carbonate (44 mg, 0.135 mmol) and dioxane (3 ml). The reaction was then sealed and heated to 120 °C for 1 hour in the microwave, filtered through Celite, concentrated *in vacuo* and purified by mass directed chromatography to give the title compound as a yellow solid (6 mg, 27 %). v_{max}/cm^{-1} (solid state) 3303, 2906, 1678, 1532; ^{1}H NMR (500 MHz, CDCl$_3$) δ 9.07 (s, 1H, Amide N-H), 7.90 (d, *J* = 2.0 Hz, 1H, Ar-H), 7.60 (dd, *J* = 8.4, 2.1 Hz, 1H, Ar-H), 7.44 (d, *J* = 8.4 Hz, 1H, Ar-H), 7.43 – 7.40 (m, 3H, Ar-H), 7.39 (d, *J* = 8.0 Hz, 1H, Ar-H), 7.13 (s, 1H, Thiazole-CH), 6.42 (d, *J* = 8.0 Hz, 1H, Ar-H), 3.97 (s, 3H, OMe), 3.94 (s, 3H, OMe), 3.69 (s, 2H, acetamide CH$_2$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 169.21, 163.04, 158.95, 157.81, 147.53, 142.43, 137.53, 134.39, 132.89, 132.38, 131.68, 130.61, 130.05, 128.73, 128.29, 127.88, 125.03, 113.93, 108.82, 101.80, 53.76, 53.67, 41.29; HRMS Found: 500.06062; C$_{24}$H$_{19}$Cl$_2$N$_3$O$_3$S [M+H]$^+$ requires 500.0597

2-(2-bromophenyl)-*N*-(4-phenyl-1,3-thiazol-2-yl)acetamide, 22
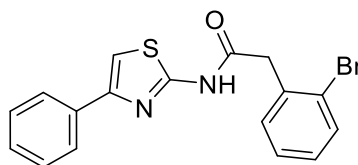
N-(3-Dimethylaminopropyl)-*N*′-ethylcarbodiimide (1.3 g, 6.8 mmol), DMAP (20 mg, 0.17 mmol) and diisopropylethylamine (1.5 ml, 8.5 mmol) were added sequentially to a solution of 2-bromophenylacetic acid (400 mg, 1.87 mmol) in DCM (30 ml). The reaction was stirred for 30 minutes at r.t. followed by the addition of 4-phenyl-1,3-thiazol-2-amine (300 mg, 1.7 mmol) after which the reaction was stirred overnight. The reaction mixture was washed with 2M HCl (aq) (2 × 20 ml) and saturated aqueous sodium bicarbonate (2 × 20 ml), dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by column

chromatography eluting with 10% ethyl acetate/petrol to give the desired compound as a colourless amorphous solid (261 mg, 41%). R_f 0.3 (10% ethyl acetate/petrol); v_{max}/cm^{-1} (solid state) 3290, 1635, 1539; ^{1}H NMR (500 MHz, CDCl$_3$) δ 10.21 (s, 1H, Amide NH), 7.84 (d, J = 7.9 Hz, 2H, Ar-H), 7.60 (d, J = 7.9 Hz, 1H, Ar-H), 7.44 (t, J = 7.6 Hz, 2H, Ar-H), 7.39 – 7.24 (m, 3H, Ar-H), 7.24 – 7.15 (m, 2H, Ar-H), 3.75 (s, 2H, acetamide CH$_2$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 167.91, 158.78, 149.81, 134.28, 133.31, 133.10, 131.78, 129.46, 128.95, 128.28, 127.92, 126.24, 125.03, 108.16, 43.12; HRMS Found: 373.0011; C$_{17}$H$_{13}$BrN$_2$OS [M+H]$^+$ requires 373.0004

2-(2'-methoxyphenyl)-*N*-(4-phenyl-1,3-thiazol-2-yl)acetamide, 23
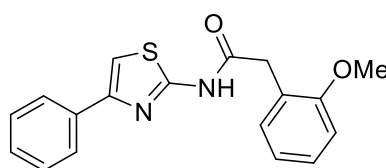
N-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide (435 mg, 2.28 mmol), DMAP (7 mg, 0.057 mmol) and diisopropylethylamine (0.5 ml, 2.85 mmol) were added sequentially to a solution of 2-(2-methoxyphenyl)acetic acid (104 mg, 0.625 mmol) in DCM (20 ml). The reaction was stirred for 30 minutes at r.t. followed by the addition of 4-phenyl-1,3-thiazol-2-amine (100 mg, 0.57 mmol) after which the reaction was stirred overnight. The reaction mixture was washed with 2M HCl (aq) (2 × 20 ml) and saturated aqueous sodium bicarbonate (2 × 20 ml), dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 10% ethyl acetate/petrol to give the desired compound as a colourless amorphous solid (33 mg, 18%). R_f 0.2 (10% ethyl acetate/petrol) v_{max}/cm^{-1} (solid state) 3168, 3069, 2970, 1650; ^{1}H NMR (500 MHz, CDCl$_3$) δ 9.39 (s, 1H, Amide NH), 7.82 (d, J = 7.4 Hz, 2H, Ar-H), 7.42 (t, J = 7.6 Hz, 2H, Ar-H), 7.38 – 7.30 (m, 3H, Ar-H), 7.13 (s, 1H,5-CH), 7.01 (dd, J = 15.5, 7.9 Hz, 2H, Ar-H), 3.95 (s, 3H, OMe), 3.82 (s, 2H, Acetamide CH$_2$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 169.03, 157.72, 157.11, 149.85, 134.42, 131.50, 129.56, 128.72, 127.98, 126.06, 121.74, 121.44, 111.09, 107.76, 55.71, 38.72; HRMS Found: 325.1007; C$_{18}$H$_{16}$N$_2$O$_2$S [M+H]$^+$ requires 325.1005

### *N*-(4-tert-butyl-1,3-thiazol-2-yl)-2-(2-methoxyphenyl)acetamide, 24
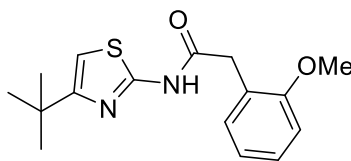
N-(3-Dimethylaminopropyl)-*N'*-ethylcarbodiimide (1.0 g, 5.12 mmol), DMAP (16 mg, 0.128 mmol) and diisopropylethylamine (1.1 ml, 6.4 mmol) were added sequentially to a solution of 2-(2-methoxyphenyl)acetic acid (637 mg, 3.84 mmol) in DCM (30 ml). The reaction was stirred for 30 minutes at r.t. followed by the addition of 4-*t*-butyl-1,3-thiazol-2-amine (200 mg, 1.28 mmol) after which the reaction was stirred overnight. The reaction mixture was washed with 2M HCl (aq) (2 × 20 ml) and saturated aqueous sodium bicarbonate (2 × 20 ml), dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 10% ethyl acetate/petrol to give the desired compound as a colourless amorphous solid (335 mg, 86%). R$_f$ 0.3 (10% ethyl acetate/petrol); ν_{max}/cm^{-1} (solid state) 3183, 2962, 1683; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.42 – 7.23 (m, 2H, Ar-H), 7.09 – 6.92 (m, 2H, Ar-H), 6.52 (s, 1H, 5-CH), 3.92 (s, 3H, OMe), 3.83 (s, 2H, CH$_2$), 1.30 (s, 9H, t-butyl); ^{13}C NMR (125 MHz, CDCl$_3$) δ 168.93, 157.22, 131.46, 129.38, 121.88, 121.25, 110.95, 104.98, 55.60, 38.55, 34.29, 29.77; HRMS Found: 305.1327; C$_{16}$H$_{20}$N$_2$O$_2$S [M+H]$^+$ requires 305.1318.

5.3.5 Synthesis of designed compounds

cis-2-Phenyl-4a,5,6,7,8,8a-hexahydro-4H-1,3-benzoxazine, 32
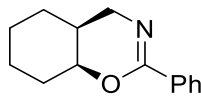
Following a literature procedure,[132] cyclohexene (1.0 ml, 10 mmol) and *N*-hydroxymethyl benzamide (1.5 g, 10 mmol) were stirred in acetic acid (10 ml) at r.t. for 10 minutes then cooled to 10 °C and a sulphuric acid (1.5 ml) in acetic acid (12.5 ml) was added dropwise maintaining a temperature between 10-15 °C. After the addition was complete, the reaction was heated to 70 °C for 1 hour, allowed to cool then poured into ice water (300 ml). The resulting solution was washed with ether (2 × 50 ml) and the organics discarded. The remaining aqueous layer was adjusted to pH 7.5 by the addition of 25 % aqueous sodium carbonate with vigorous stirring then extracted with ether (3 × 30 ml). The combined organics were washed with saturated aqueous sodium bicarbonate (3 × 20 ml), dried over $MgSO_4$ and concentrated in vacuo to a pale yellow oil (1.3 g, 61%). Analytical data matches literature. IR (v_{max} /cm^{-1}) 3300, 2931, 1655; ^{1}H NMR (500 MHz; CDCl$_3$) δ 7.86 (2 H, d, *J* = 7.3, Ar-H), 7.30 (3 H, m, Ar-H), 4.31 (1H, m, 8a-H), 3.57 (1 H, dd, *J* = 17, 5.5, 4-H$_a$), 3.30 (1 H, dd, *J* = 17, 2.3, 4-H$_b$), 1.98 (1 H, m, 8-H), 1.79 (1 H, m, 4a-H), 1.61 (1 H, m, Cy-CH), 1.45 (5 H, m, Cy-CH), 1.25 (1 H, m, Cy-CH); ^{13}C NMR (125 MHz; CDCl$_3$) δ 155.0, 134.1, 130.1, 127.9, 126.9, 72.7, 48.9, 32.1, 30.5, 25.4, 24.4, 20.3; HRMS found: 216.1392; C$_{14}$H$_{18}$NO [M+H]$^+$ requires 216.1383.

Cis-2-(benzylamino)methyl cyclohexan-1-ol, 33
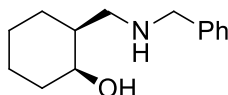
Following a literature procedure[133], *cis*-2-Phenyl-4a,5,6,7,8,8a-hexahydro-4H-1,3-benzoxazine (5 g, 23 mmol) was added dropwise as a solution in tetrahydrofuran (10 ml) to a stirred suspension of lithium aluminium hydride (1.8 g, 46 mmol) in anhydrous tetrahydrofuran (200 ml) under N$_2$. The reaction was heated to reflux overnight then cooled to 0°C and quenched by the dropwise addition of water (2 ml). The resulting precipitate was removed by filtration, the filtrate dried over MgSO$_4$ and concentrated *in vacuo* to give the *title compound* as a yellow oil (3.33 g, 66%).

^{1}H NMR (500 MHz; CDCl$_3$) δ 7.28 (4 H, d, *J* = 4.6, Ar-H), 7.22 (1 H, m, Ar-H), 4.1 (1 H, br s, NH), 3.97 (1 H, m, 1-H), 3.73 (2 H, s, 1'-H$_2$), 2.92 (1 H, dd, *J* = 12.0, 6.4, 7-H$_a$), 2.66 (1 H, dd, *J* = 12.0, 3.4, 7-H$_b$), 1.72 – 1.44 (6 H, m, Cy-CH$_2$), 1.36 – 1.21 (3 H, m, CH$_2$ and 2-H); ^{13}C NMR $_C$ (125 MHz; CDCl$_3$) δ 138.9, 128.4, 126.2,

127.2, 70.7, 54.1, 52.4, 39.8, 32.4, 25.5, 24.6, 21.1; IR (ν_{max} /cm^{-1}): 3299, 2931, 2856, 1655; HRMS found: 220.1698; $C_{14}H_{22}NO$ [M+H]$^+$ requires 220.1696.

Cis-2-[benzyl(2-methylpropyl)amino]methyl cyclohexan-1-ol, 34
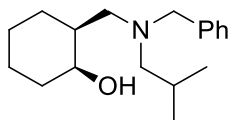
Isobutyraldehyde (0.6 ml, 6.4 mmol) and picoline borane complex (670 mg, 6.4 mmol) were added sequentially to a solution of *cis*-2-(benzylamino)methyl cyclohexan-1-ol (700 mg, 3.2 mmol) in methanol (30 ml) and stirred at r.t. overnight. Upon completion the reaction was quenched with water (20 ml) and acidified to pH 2 with 2M aqueous HCl, washed with ethyl acetate (3 × 10 ml), neutralised with saturated aqueous sodium bicarbonate and extracted with ethyl acetate (3 × 15 ml). Final extractions were combined, washed with brine, dried over MgSO$_4$ and concentrated *in vacuo* to yield the *title compound* as a colourless oil (585 mg, 66%). IR (ν_{max} /cm^{-1}): 2930, 2865, 2801, 1452; ^{1}H NMR (500 MHz; CDCl$_3$) δ 7.35 (4 H, m, Ar-H), 7.31 – 7.27 (1 H, m, Ar-H), 3.94 – 3.88 (2 H, m, 1'-CH$_2$), 3.22 (1H, d, *J* = 11.9, 7-H$_a$), 2.89 (1H, t, *J* = 11.5, 7-H$_b$), 2.39 – 2.32 (1H, m, 1-H), 2.19 – 2.10 (3H, m, 2"CH$_2$ and Cy-CH), 1.93 – 1.91 (1H, m, 2"CH), 1.63 – 1.56 (2H, m, Cy-CH), 1.48 – 1.43 (3H, m, Cy-CH), 1.37 – 1.26 (3H, m, Cy-CH), 0.93 (6H, dd, *J* = 19.2, 6.4, 2 × 3"-H$_3$); ^{13}C NMR (125 MHz; CDCl$_3$): δ 129.5, 126.4, 127.3, 70.79, 63.6, 59.4, 56.5, 36.5, 31.6, 27.6, 26.0, 22.9, 21.2, 20.7; HRMS Found: 276.2335 $C_{18}H_{30}NO$ [M+H]$^+$ requires 276.2322.

Cis-2-[[(2-methylpropyl)amino]methyl]cyclohexan-1-ol, 35
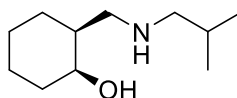
Cis-2-[benzyl(2-methylpropyl)amino]methyl cyclohexan-1-ol (1 g, 3.6 mmol) was dissolved in methanol (50 ml) under an N$_2$ atmosphere. Pd/C (100 mg, 10 wt%) was added followed by ammonium formate (1.15 g, 18.1 mmol) and the reaction heated to reflux for 2 hours. The reaction was allowed to cool to RT and filtered through Celite and concentrated *in vacuo*. The residue was triturated with acetonitrile and purified by strong cation exchange (SCX) column chromatography eluting with saturated ammonia in methanol to give the title compound as a colourless oil (461 mg, 70%). ν_{max}/cm^{-1} (solid state) 3318, 2927, 1589; ^{1}H NMR (500 MHz, CDCl$_3$) δ 4.07 – 3.93 (m, 1H, 1-CH), 2.97 (dd, *J* = 12.0, 6.0 Hz, 1H, 7-CH$_a$), 2.67 (dd, *J* = 12.0, 2.9 Hz, 1H, 7-CH$_b$), 2.39 (app. qd, *J* = 11.5, 6.7 Hz, 2H, 1'-CH$_2$, rot.), 1.79 – 1.61 (m, 4H, Cy-Ch and 2'-CH), 1.60 – 1.46 (m, 2H, Cy-CH), 1.42 – 1.22 (m, 3H, Cy-CH), 0.89 (d, *J* = 4.3 Hz, 6H, rot. 2 × 3'-CH$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ

71.74, 58.32, 53.68, 39.65, 32.38, 28.15, 25.62, 24.57, 21.30, 20.57, 20.52; HRMS Found 186.1860 $C_{11}H_{23}NO$ $[M+H]^+$ requires 186.1852

tert-butyl N-([cis-2-hydroxycyclohexyl]methyl)-N-(2-methylpropyl)carbamate, 36
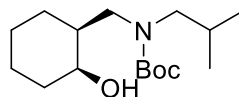
2-{[(2-methylpropyl)amino]methyl}cyclohexan-1-ol (1.12 g, 6.1 mmol) was dissolved in THF (50 ml). (Boc)$_2$O (1.46 g, 6.7 mmol) and diisopropylethylamine (2.6 ml, 15.25 mmol) were added and the reaction stirred overnight and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 10% ethyl acetate in petrol to give the title compound as a colourless oil (1.42 g, 87%). R$_f$ 0.25 (10% ethyl acetate/petrol); v_{max}/cm^{-1} (solid state) 3446, 3186, 2932, 1656; 1H NMR (500 MHz, CDCl$_3$) δ 4.22 (s, 1H, 1-CH), 3.66 (m, 2H, 2-CH and 7-CH$_a$), 3.17 (dd, J = 13.9, 7.6 Hz, 1H, 1'-CH$_a$), 2.70 (dd, J = 14.0, 7.3 Hz, 1H, 1'-CH$_b$), 2.55 (dd, J = 14.5, 4.1 Hz, 1H, 7-CH$_b$), 1.90 (m, 2H, Cy-CH and 2'-CH), 1.73 – 1.46 (m, 3H, Cy-CH), 1.43 (s, 6H, t-butyl), 1.40 – 1.08 (m, 4H, Cy-CH), 0.88 (d, J = 6.7 Hz, 3H, 3'-CH$_3$), 0.85 (d, J = 6.9 Hz, 3H, 3'-CH$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 157.53, 79.90, 64.44, 54.55, 49.22, 40.17, 32.16, 28.33, 27.57, 25.83, 24.46, 20.21, 20.09, 20.06; HRMS Found 308.2205 $C_{16}H_{31}NO_3$ [M+Na] requires 308.2196

tert-butyl *N*-(2-methylpropyl)-*N*-{[-2-oxocyclohexyl]methyl}carbamate, 38
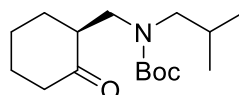
tert-butyl *N*-{[(1S,2S)-2-hydroxycyclohexyl]methyl}-*N*-(2-methylpropyl)carbamate (1 g, 3.51 mmol) was dissolved in DCM (30 ml) and Dess-Martin Periodinane (1.6 g, 3.86 mmol) was added. The reaction was stirred at r.t. for 3 hours then quenched with saturated sodium thiosulfate solution (10 ml) and saturated sodium hydrogen carbonate solution (10 ml). The organic layer was separated, and the aqueous layer extracted with DCM (3 × 10 ml). The organic layers were combined, dried over MgSO$_4$ and concentrated *in vacuo* to a pale yellow oil (990 mg, quant). v_{max}/cm^{-1} (solid state) 2960, 2868 1693, ^{1}H NMR (500 MHz, CDCl$_3$) δ 3.51 (m, 1H, 7-CH$_a$), 3.30 – 3.17 (m, 1H, 7-CH$_b$), 3.13 – 2.93 (m, 2H, 1'-CH$_2$), 2.88 – 2.57 (m, 1H, 2-CH), 2.41 (d, J = 13.1 Hz, 1H, Cy-CH), 2.32 (m, 1H, Cy-CH), 2.10 (dd, J = 44.3, 5.7 Hz, 2H, Cy-CH), 2.01 – 1.80 (m, 2H, 2'-CH and Cy-CH), 1.66 (tt, J = 11.9, 5.9 Hz, 2H, Cy-CH), 1.45 (s, 9H, t-butyl), 1.43 – 1.29 (m, 1H, Cy-CH), 0.87 (app. t, J = 7.2 Hz, 6H, 2 × 3'-CH$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 212.79, 156.04, 79.21, 55.67, 54.70, 50.37, 49.72, 47.57, 46.89, 42.11, 32.04, 28.46, 27.96, 27.11, 24.68, 20.10; HRMS found 284.2217, $C_{16}H_{29}NO_3$ requires $[M+H]^+$ 284.2217

tert-butyl **N-{[(2-methylpropyl)({[(1S)-2-oxocyclohexyl] methyl})carbamoyl] methyl}carbamate, 40**
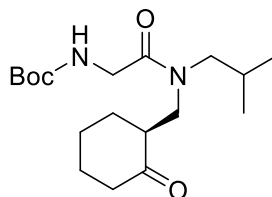
tert-butyl N-(2-methylpropyl)-N-{[-2-oxocyclohexyl]methyl}carbamate (990 mg, 3.8 mmol) was dissolved in 10% TFA in DCM (40 ml) and stirred overnight and concentrated *in vacuo*. The crude TFA salt was then re-dissolved in DCM (30 ml) to which was added Boc-glycine (798 mg, 4.56 mmol), HATU (1.7 g, 4.56 mmol) and diisopropylethylamine (1.9 ml, 11.4 mmol) and the reaction stirred overnight. The reaction mixture was then washed with H_2O (20 ml) and 1M HCl (20 ml), dried over $MgSO_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 20% ethyl acetate in petrol to give the title compound as a colourless oil (1.13 g, 87%). R_f 0.3 (20% ethyl acetate/petrol); ν_{max}/cm^{-1} (solid state) 3228, 1699, 1626; ^{1}H NMR (500 MHz, $C_2Cl_4D_2$) δ 3.95 (d, J = 3.3 Hz, 1H, 2"-CH_2), 3.55 (dd, J = 13.7, 6.9 Hz, 1H, 7-CH_a), 3.35 (dd, J = 13.8, 6.0 Hz, 1H, 7-CH_b), 3.25 (dd, J = 14.7, 7.4 Hz, 1H, 2'-CH_a), 2.97 (dd, J = 14.7, 7.9 Hz, 1H, 2'-CH_b), 2.84 (m, 1H, 2-CH), 2.37 (m, 1H, Cy-CH), 2.16 – 2.03 (m, 2H, 2'-CH and Cy-CH), 1.94 (m, 2H, Cy-CH), 1.67 (m, 2H, Cy-CH), 1.46 (s, 9H, t-butyl), 1.41 – 1.32 (m, 1H, Cy-CH), 0.95 (d, J = 6.6 Hz, 3H, 3'-CH_3) 0.92 (d, J = 6.6 Hz, 3H, 3'-CH_3); ^{13}C NMR (125 MHz, $C_2Cl_4D_2$) δ 212.09, 169.02, 155.63, 79.60, 55.22, 49.31, 46.95, 42.64, 42.13, 41.13, 32.38, 28.54, 27.76, 24.85, 23.99, 21.01, 20.14, 17.60; HRMS Found 341.24429 $C_{18}H_{32}N_2O_4$ [M+H]$^+$ requires 341.2434.

4-(2-methylpropyl)-decahydro-1H-1,4-benzodiazepin-3-one, 42
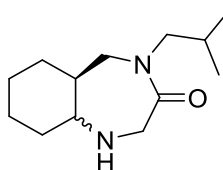
tert-butyl   N-{[(2-methylpropyl)   ({[-2-oxocyclohexyl]   methyl})carbamoyl] methyl}carbamate (200 mg, 0.58 mmol) was dissolved in 20% TFA in DCM (20 ml) and stirred at r.t. for 1 hour. The reaction mixture was then concentrated *in vacuo* and the residue re-suspended in methanol (20 ml) and picoline borane complex (75 mg, 0.70 mmol) added and reaction stirred for 3 hours, concentrated *in vacuo* and residue purified by column chromatography eluting with DCM then 10% methanol in DCM to give the title compound as a mixture of diastereomers (125 mg, 83%).

The diastereomers were separated by column chromatography eluting with 1.5% saturated NH$_3$ in methanol in DCM.

HRMS found 225.1968, C$_{13}$H$_{24}$N$_2$O [M+H]$^+$ requires 225.1961

v$_{max}$/cm^{-1} (solid state) 2923, 2853, 1639

cis-4-(2-methylpropyl)-decahydro-1H-1,4-benzodiazepin-3-one, cis-42
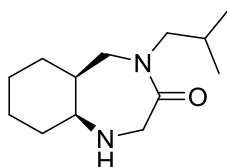
^{1}H NMR (500 MHz, CDCl$_3$) δ 3.82 – 3.71 (m, 1H, 1'-CH$_a$), 3.72 – 3.60 (m, 2H, 2-CH$_a$ and 5-CH$_a$), 3.54 (d, J = 14.8 Hz, 1H, 2-CH$_b$), 3.26 (d, J = 11.1 Hz, 1H, 5-CH$_b$), 3.02 (q, J = 3.4 Hz, 1H, 7-CH), 2.62 (br, s, 1H, 1'-CH$_b$), 2.07 – 1.87 (m, 1H, 2'-CH), 1.84 – 1.67 (m, 3H, Cy-CH), 1.69 – 1.54 (m, 2H, Cy-CH), 1.44 (ddd, J = 16.8, 9.9, 3.4 Hz, 3H, Cy-CH), 1.33 (tdd, J = 16.3, 12.5, 4.0 Hz, 1H, Cy-CH), 0.93 (d, J = 6.7 Hz, 3H, 3'-CH$_3$), 0.88 (d, J = 6.7 Hz, 3H, 4'-CH$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 175.98, 60.10, 57.89, 54.67, 54.04, 40.63, 33.18, 28.31, 26.41, 24.74, 21.61, 20.55, 20.24.

trans-4-(2-methylpropyl)-decahydro-1H-1,4-benzodiazepin-3-one, trans-42
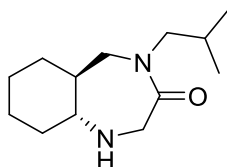
^{1}H NMR (500 MHz, d^4-MeOD) δ 3.58 (d, J = 14.6 Hz, 1H, 2-CH$_a$), 3.49 (dd, J = 15.1, 10.1 Hz, 1H, 5-CH$_a$), 3.34 (d, J = 14.6 Hz, 1H, 2-CH$_b$), 3.21 (m, 1H, 2'-CH$_a$), 2.99 (dd, J = 13.2, 7.6 Hz, 1H, 2'CH$_b$), 2.87 (dd, J = 15.1, 1.4 Hz, 1H, 5-CH$_b$), 2.24 (ddd, J = 11.3, 9.6, 3.4 Hz, 1H, 7-CH), 1.81 (dp, J = 13.9, 6.9 Hz, 1H, 2'-CH), 1.73 – 1.60 (m, 3H, Cy-CH), 1.57 – 1.40 (m, 1H, Cy-CH), 1.32 – 1.16 (m, 3H, Cy-CH), 1.16 – 1.06 (m, 1H, Cy-CH), 0.94 (qd, J = 12.8, 3.7 Hz, 1H, Cy-CH), 0.84 (d, J = 6.7 Hz, 3H, 3'CH$_3$), 0.79 (d, J = 6.7 Hz, 3H, 4'CH$_3$); ^{13}C NMR (125 MHz, d^4-MeOD) δ 176.10, 67.05, 56.92, 56.69, 54.66, 45.60, 34.26, 31.24, 28.34, 26.72, 26.24, 20.45, 20.37.

tert-butyl N-benzyl-N-{[cis-2-hydroxycyclohexyl]methyl}carbamate, 37
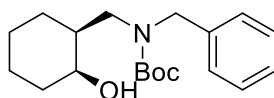
Cis-2-(benzylamino)methyl cyclohexan-1-ol (1.0 g, 4.56 mmol) was dissolved in THF (30 ml). (Boc)$_2$O (1.1 g, 5.02 mmol) and diisopropylethylamine (1.98 ml, 11.4

mmol) were added and the reaction stirred overnight and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 10% ethyl acetate in petrol to give the title compound as a colourless oil (832 mg, 57%). R*f* 0.3 (10% ethyl acetate/petrol); ν_{max}/cm^{-1} (solid state) 3446, 2930, 2857, 1657; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.35 (m, 2H, Ar-H), 7.32 − 7.17 (m, 3H, Ar-H), 4.69 (d, *J* = 15.6 Hz, 1H, 1'-CH$_a$), 4.14 (m, 2H, 1'-CH$_b$ and 1-CH), 3.78 (s, 1H, OH), 3.73 − 3.65 (m, 1H, 7-CH$_a$), 2.61 (dd, *J* = 14.5, 4.0 Hz, 1H, 7-CH$_b$), 1.96 (d, *J* = 13.3 Hz, 1H, 2-CH), 1.77-1.69 (m, 1H, Cy-CH), 1.67 − 1.51 (m, 3H, Cy-CH), 1.47 (s, 9H), 1.41 − 1.30 (m, 2H), 1.28 − 1.17 (m, 1H), 0.97 − 0.84 (m, 1H); ^{13}C NMR (125 MHz, CDCl$_3$) δ 157.30, 138.13, 128.52, 127.28, 127.20, 80.60, 64.57, 50.79, 48.65, 39.93, 32.17, 28.41, 28.32, 25.78, 24.40, 20.06; HRMS found 320.2221, C$_{19}$H$_{29}$NO$_3$ requires [M+H]$^+$ 320.2220.

tert-butyl *N*-benzyl-*N*-{[2-oxocyclohexyl]methyl}carbamate, 39
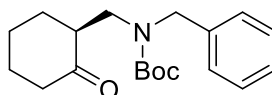
tert-butyl *N*-benzyl-*N*-{[*cis*-2-hydroxycyclohexyl]methyl}carbamate (832 mg, 2.6 mmol) was dissolved in DCM (40 ml) and Dess-Martin Periodinane (1.2g, 2.86 mmol) was added. The reaction was stirred at r.t. for 3 hours then quenched with saturated sodium thiosulfate solution (10 ml) and saturated sodium hydrogen carbonate solution (10 ml). The organic layer was separated, and the aqueous layer extracted with DCM (3 × 10 ml). The organic layers were combined, dried over MgSO$_4$ and concentrated *in vacuo* to a pale yellow oil (506 mg, 61%). ν_{max}/cm^{-1} (solid state) 2934, 2863, 1696; ^{1}H NMR (500 MHz, *d*^{4}-MeOD) δ 7.19 (m, 2H, Ar-H), 7.13 (m, 3H, Ar-H), 4.42 − 4.19 (m, 2H, 1-CH$_2$), 3.39 (dd, *J* = 14.3, 5.9 Hz, 1H, 7-CH$_a$), 3.23 − 3.02 (m, 1H, 7-CH$_b$), 2.60 (s, 1H, 2-CH), 2.31 − 2.10 (m, 2H, Cy-CH), 2.00 − 1.90 (m, 1H, Cy-CH), 1.92 − 1.84 (m, 1H, Cy-CH), 1.72 (dd, *J* = 9.0, 4.5 Hz, 1H, Cy-CH), 1.61 − 1.45 (m, 2H, Cy-CH), 1.34 (s, 9H, t-butyl), 1.23 (m, 1H, Cy-CH); ^{13}C NMR (125 MHz, *d*^{4}-MeOD) δ 214.06, 157.73, 139.59, 129.59, 127.75, 81.25, 52.55, 51.33, 50.80, 48.18, 47.66, 47.51, 42.84, 33.35, 28.87, 25.77; HRMS found 340.1877, C$_{19}$H$_{27}$NO$_3$ requires [M+Na]$^+$ 318.2063

tert-butyl *N*-{[benzyl({[2-oxocyclohexyl]methyl})carbamoyl]methyl}carbamate, 41
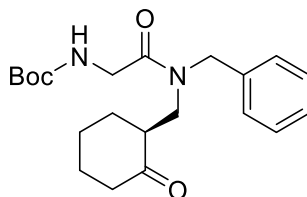
tert-butyl *N*-benzyl-*N*-{[2-oxocyclohexyl]methyl}carbamate (500 mg, 1.6 mmol) was dissolved in 20% TFA in DCM (20 ml) and stirred overnight then concentrated *in vacuo*. The crude TFA salt was then re-dissolved in DCM (30 ml) to which was added Boc-glycine (332 mg, 1.89 mmol), HATU (722 mg, 1.89 mmol) and diisopropylethylamine (0.82 ml, 4.73 mmol) and the reaction stirred overnight. The reaction mixture was then washed with H_2O (20 ml) and 1M HCl (20 ml), dried over $MgSO_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 40% ethyl acetate in petrol to give the title compound as a colourless oil (490 mg, 83%). R_f 0.2 (40% ethyl acetate/petrol); ν_{max}/cm^{-1} (solid state) 3418, 2934, 2863, 1704, 1647; ^{1}H NMR (500 MHz, d^4-MeOD) δ 7.27 (m, 1H, Ar-H), 7.23 − 7.08 (m, 4H, Ar-H), 4.69 (s, 2H, 2''-CH_2), 4.61 (dd, *J* = 15.9, 6.6 Hz, 1H, 1'-CH_a), 4.47 (dd, *J* = 33.3, 16.1 Hz, rot. 1'-CH_b) 4.00 (dd, *J* = 63.2, 16.9 Hz, 1H, rot. 7-CH_a), 3.81 (dd, *J* = 39.7, 16.8 Hz, 1H, rot, 7-CH_a), 3.62 (dd, *J* = 15.2, 6.6 Hz, 1H, 2-CH), 3.44 − 3.30 (m, 1H, Cy-CH), 3.26 − 2.98 (m, 1H, Cy-CH), 2.83 − 2.60 (m, 1H, Cy-CH), 2.31 − 2.08 (m, 2H, Cy-CH), 2.07 − 1.86 (m, 2H, Cy-CH), 1.74 (d, *J* = 8.5 Hz, 1H, Cy-CH), 1.66 − 1.47 (m, 2H, Cy-CH), 1.35 (d, *J* = 16.9 Hz, 9H, t-butyl), 1.27 − 1.10 (m, 1H, Cy-CH); ^{13}C NMR (125 MHz, d^4-MeOD) δ 214.45, 214.10, 171.93, 158.11, 138.11, 129.95, 129.54, 128.82, 128.63, 127.68, 80.64, 52.43, 50.50, 47.32, 43.48, 42.72, 33.25, 28.75, 25.62; HRMS found 375.2275, $C_{21}H_{30}N_2O_4$ requires [M+H]$^+$ 375.2278.

4-benzyl-decahydro-1H-1,4-benzodiazepin-3-one, 43
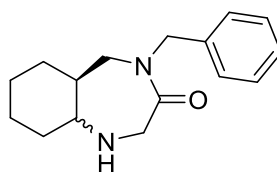
(400 mg, 1.1 mmol) was dissolved in 20% TFA in DCM (20 ml) and stirred at r.t. for 1 hour. The reaction mixture was then concentrated *in vacuo* and the residue re-suspended in DCM and stirred with $MgSO_4$ overnight. The mixture was then filtered, concentrated *in vacuo* and re-dissolved in methanol (30 ml). Sodium borohydride (58 mg, 1.57 mmol) was added and the reaction stirred for 3 hours, concentrated *in vacuo* and residue purified by SCX column chromatography eluting

with saturated ammonia in methanol to give the *title compound* as a 50:50 mixture of diastereomers (246 mg, 86%).

The diastereomers were separated by column chromatography eluting with 3% saturated NH$_3$ in methanol in DCM.

HRMS Found 259.1803; C$_{16}$H$_{22}$N$_2$O [M+H]$^+$ requires 259.1804

v_{max}/cm^{-1} (solid state) 2935, 2848, 1637

Cis-4-benzyl-decahydro-1H-1,4-benzodiazepin-3-one, *cis*-43
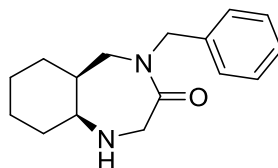
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.37 – 7.30 (m, 2H, Ar-H), 7.30 – 7.23 (m, 3H, Ar-H), 5.16 (d, *J* = 14.6 Hz, 1H, benzyl CH$_a$), 4.01 (d, *J* = 14.6 Hz, 1H benzyl CH$_b$), 3.75 (d, *J* = 15.0 Hz, 1H, 2-CH$_a$), 3.58 (d, *J* = 15.0 Hz, 1H, 2-CH$_b$), 3.51 (d, *J* = 15.1 Hz, 1H, 5-CH$_a$), 3.22 (dt, *J* = 16.8, 8.5 Hz, 1H, 5-CH$_b$), 2.99 (dd, *J* = 6.8, 3.4 Hz, 1H, 7-CH), 1.78 – 1.51 (m, 5H, Cy-CH), 1.50 – 1.39 (m, 2H, Cy-CH), 1.35 – 1.18 (m, 2H, Cy-CH); ^{13}C NMR (126 MHz, CDCl$_3$) δ 174.28, 137.57, 137.56, 128.54, 128.22, 127.36, 59.10, 53.96, 52.34, 52.08, 39.47, 32.85, 25.45, 23.77, 20.72

Trans-4-benzyl-decahydro-1H-1,4-benzodiazepin-3-one, *trans*-43
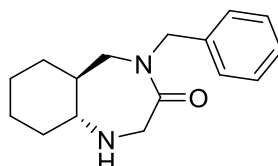
^{1}H NMR (500 MHz, CDCl$_3$) δ 7.38 – 7.32 (m, 2H, Ar-H), 7.32 – 7.24 (m, 3H, Ar-H), 4.80 (d, *J* = 14.7 Hz, 1H, benzyl CH$_a$), 4.41 (d, *J* = 14.7 Hz, 1H benzyl CH$_b$), 3.74 (d, *J* = 14.8 Hz, 1H, 2-CH$_a$), 3.68 (d, *J* = 14.8 Hz, 1H, 2-CH$_a$), 3.40 (dd, *J* = 15.2, 10.0 Hz, 1H, 5-CH$_a$), 2.88 (dd, *J* = 15.2, 1.0 Hz, 1H, 5-CH$_a$), 2.29 (ddd, *J* = 11.2, 9.6, 3.5 Hz, 1H, 7-CH), 1.86 – 1.72 (m, 2H, Cy-CH), 1.67 (d, *J* = 14.5 Hz, 1H, Cy-CH), 1.47 – 1.39 (m, 1H, Cy-CH), 1.37 – 1.06 (m, 4H, Cy-CH), 0.91 (qd, *J* = 12.8, 3.7 Hz, 1H, Cy-CH); ^{13}C NMR (125 MHz, CDCl$_3$) δ 174.44, 137.52, 128.58, 128.17, 127.41, 66.12, 54.52, 54.36, 53.44, 51.36, 45.34, 34.03, 30.27, 25.64, 25.39

Standard Procedure A – Reductive amination

Amine (1 eq.) was dissolved in MeOH (10 ml per mmol) and the relevant aldehyde (2 eq.) added followed by picoline borane complex (2 eq.) and the reaction stirred at r.t. until LC-MS confirmed completion. The reaction mixture was then concentrated *in vacuo* and the residue purified by mass-directed HPLC.

Standard Procedure B – Sulfonamide formation

Triethylamine (2 eq.) and the relevant sulfonyl chloride (1.2 eq.) were added to a solution of the amine (1 eq.) in DCM (10 ml per mmol) and the reaction stirred at r.t. until LC-MS confirmed completion. The reaction mixture was then concentrated *in vacuo* and the residue purified by mass-directed HPLC.

Cis-**4-(2'-methylpropyl)-1-(pyridin-3-ylmethyl)-decahydro-1H-1,4-benzodiazepin-3-one, 44**
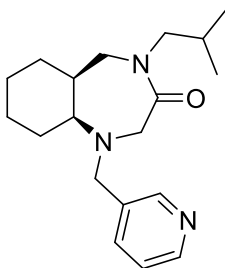
Prepared using general procedure A (16.2 mg 53%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.58 (s, 1H, 2-pyridyl CH), 8.52 (d, J = 4.2 Hz, 1H, 6-pyridyl CH), 7.77 (d, J = 7.8 Hz, 1H, 4-pyridyl CH), 7.33 (dd, J = 7.7, 4.9 Hz, 1H, 5-pyridyl CH), 4.01 (dd, J = 14.7, 10.6 Hz, 1H, 2-CH$_a$), 3.90 (dd, J = 14.6, 10.5 Hz, 2H, 5-CH$_a$), 3.83 (d, J = 14.3 Hz, 1H,7-CH), 3.33 (dd, J = 13.2, 6.9 Hz, 1H, 1'CH$_b$), 3.29 – 3.11 (m, 2H, 1'CH$_a$ and 5-CH$_b$), 2.85 (d, J = 15.2 Hz, 1H,2-CH$_b$), 2.82 – 2.66 (m, 1H, Cy-CH), 2.28 (d, J = 5.8 Hz, 1H, Cy-CH), 1.97 – 1.81 (m, 3H, 2'C-H and Cy-CH), 1.75 (d, J = 13.3 Hz, 1H, Cy-CH), 1.61 (d, J = 11.1 Hz, 1H, Cy-CH), 1.52 – 1.41 (m, 2H, Cy-CH), 1.35 – 1.24 (m, 1H, Cy-CH), 1.16 (t, J = 13.6 Hz, 1H, Cy-CH), 1.00 (d, J = 6.6 Hz, 3H, 3'-CH$_3$), 0.94 (d, J = 6.7 Hz, 3H, 4'-CH$_3$); HRMS Found: 316.2393; C$_{19}$H$_{29}$N$_3$O [M+H]$^+$ requires 316.2383.

Trans-**4-(2'-methylpropyl)-1-(pyridin-3-ylmethyl)-decahydro-1H-1,4-benzodiazepin-3-one, 45**
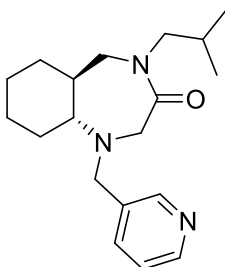
Prepared using general procedure A (9.2 mg, 58%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.63 (s, 1H, 2-pyridyl CH), 8.53 (d, J = 4.9 Hz, 1H, 6-pyridyl CH), 7.83 (d, J = 7.7 Hz, 1H, 4-pyridyl CH), 7.36 (dd, J = 7.7, 4.9 Hz, 1H, 5-pyridyl CH), 3.98 (d, J = 14.4 Hz, 1H, benzylic CH$_a$), 3.67 (d, J = 15.5 Hz, 1H, 2-CH$_a$), 3.53 – 3.38 (m, 3H, benzylic C-H$_b$, 2-CH$_b$ and 5-CH$_a$), 3.34 – 3.23 (m, 2H, 1'CH$_2$), 2.93 (d, J = 14.8 Hz,

1H, 5-CH$_b$), 2.60 (td, J = 11.6, 3.0 Hz, 1H, 7-CH), 2.00 − 1.81 (m, 3H, 2'CH and Cy-CH), 1.81 − 1.62 (m, 3H, Cy-CH), 1.60 − 1.42 (m, 1H, Cy-CH), 1.29 (tt, J = 25.5, 12.7 Hz, 2H, Cy-CH), 1.00 (d, J = 6.6 Hz, 3H, 3'-CH$_3$), 0.95 (d, J = 6.6 Hz, 3H, 4'-CH$_3$); HRMS found 316.2390; C$_{19}$H$_{29}$N$_3$O [M+H]$^+$ requires 316.2383.

Cis-4-benzyl-1-(pyridin-3-ylmethyl)-decahydro-1H-1,4-benzodiazepin-3-one, 46
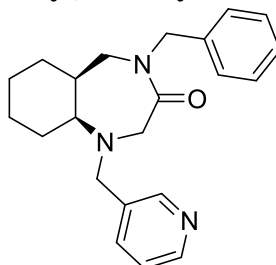
Prepared using general procedure A (13.5 mg, 50%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.57 (s, 1H, Ar-H), 8.54 (d, J = 4.1 Hz, 1H, Ar-H), 7.77 (d, J = 7.8 Hz, 1H, Ar-H), 7.42 − 7.31 (m, 6H, Ar-H), 4.65 (s, 2H, pyridyl CH$_2$), 4.00 − 3.85 (m, 3H, 2-CH$_a$, benzyl-CH$_a$ and 5-CH$_a$), 3.80 (d, J = 14.3 Hz, 1H, benzyl-CH$_b$), 3.33 (d, J = 15.4 Hz, 1H, 2-CH$_b$), 2.87 (d, J = 15.1 Hz, 1H, 5-CH$_b$), 2.77 − 2.65 (m, 1H, 7-CH), 2.04 (d, J = 6.5 Hz, 1H, Cy-CH), 1.92 − 1.72 (m, 2H, Cy-CH), 1.57 (app. t, J = 11.7 Hz, 2H, Cy-CH), 1.46 − 1.18 (m, 3H, Cy-CH), 1.05 (q, J = 13.3 Hz, 1H, Cy-CH); HRMS Found 350.2232; C$_{22}$H$_{27}$N$_3$O [M+H]$^+$ requires 350.2226.

Trans-4-benzyl-1-(pyridin-3-ylmethyl)-decahydro-1H-1,4-benzodiazepin-3-one, 47
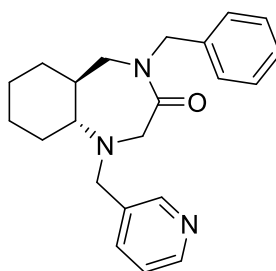
Prepared using general procedure A (16.5 mg, 61%. ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.61 (s, 1H, 2-pyridyl CH), 8.53 (d, J = 4.9 Hz, 1H,6-pyridyl CH), 7.80 (d, J = 7.8 Hz, 1H, 4-pyridyl CH), 7.41 − 7.31 (m, 6H, Ar-H), 4.67 (d, J = 6.5 Hz, 2H, benzylic CH$_2$), 3.94 (d, J = 14.3 Hz, 1H, pyridylmethyl CH$_a$), 3.70 (d, J = 15.8 Hz, 1H, 2-CH$_a$), 3.60 − 3.49 (m, 2H, pyridylmethyl CH$_b$ and 2-CH$_b$), 3.39 (dd, J = 14.8, 10.1 Hz, 1H, 5-CH$_a$), 2.96 (d, J = 14.8 Hz, 1H, 5-CH$_b$), 2.60 − 2.51 (m, 1H, 7-CH), 1.85 (t, J = 12.0 Hz, 2H, Cy-CH), 1.69 (d, J = 13.5 Hz, 1H, Cy-CH), 1.60 − 1.38 (m, 3H, Cy-CH), 1.33 − 1.22 (m, 1H, Cy-CH), 1.21 − 1.10 (m, 1H, Cy-CH), 0.94 (qd, J = 12.6, 3.8 Hz, 1H, Cy-CH); HRMS Found 350.2235; C$_{22}$H$_{27}$N$_3$O [M+H]$^+$ requires 350.2226.

cis-4-benzyl-1-(1H-imidazol-4-ylmethyl)-decahydro-1H-1,4-benzodiazepin-3-one, 48
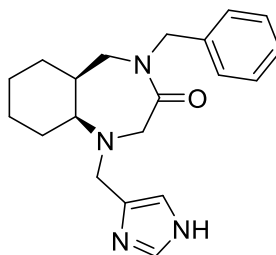
Prepared using general procedure A, with 1-Tritylimidazole-4-carboxyaldehyde, followed by treatment with 20% TFA in DCM, concentrated *in vacuo* and the residue purification by mass directed HPLC (13.5 mg, 52%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.64 (s, 1H, Ar-H), 7.50 – 7.34 (m, 3H, Ar-H), 7.30 (d, $J = 7.5$ Hz, 3H, Ar-H), 4.66 (s, 2H, imidazol-4-ylmethyl CH$_2$), 4.39 (dd, $J = 31.4, 14.4$ Hz, 2H, benzylic CH$_2$), 4.27 (d, $J = 14.6$ Hz, 1H, 2-CH$_a$), 4.02 (dd, $J = 14.8, 10.9$ Hz, 1H, 5-CH$_a$), 3.67 (d, $J = 14.7$ Hz, 1H, 2-CH$_b$), 3.35 (d, $J = 9.6$ Hz, 1H, 7-CH), 3.04 (d, $J = 15.4$ Hz, 1H, 5-CH$_b$), 2.28 (d, $J = 8.0$ Hz, 1H, 6-CH), 2.07 – 1.80 (m, 3H, Cy-CH), 1.65 (d, $J = 12.1$ Hz, 1H, Cy-CH), 1.58 – 1.39 (m, 2H, Cy-CH), 1.40 – 1.25 (m, 1H, Cy-CH), 1.24 – 0.98 (m, 1H, Cy-CH); HRMS Found: 339.2185; C$_{20}$H$_{26}$N$_4$O [M+H]$^+$ requires 339.2179.

cis-1-(furan-2-ylmethyl)-4-(2-methylpropyl)-decahydro-1H-1,4-benzodiazepin-3-one, 48
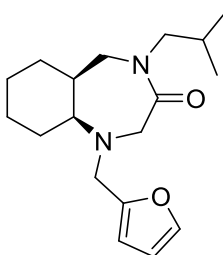
Prepared using general procedure A (9.9 mg, 36%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.42 – 7.39 (m, 1H, 5-furanyl CH), 6.35 (dd, $J = 3.0, 1.9$ Hz, 1H, 4-furanyl CH), 6.31 (d, $J = 3.1$ Hz, 1H, 3-furanyl CH), 4.01 (d, $J = 14.7$ Hz, 1H, furanyl methyl CH$_a$), 3.87 (t, $J = 17.7$ Hz, 3H, furanyl methyl CH$_b$, 2-CH$_a$ and 5-CH$_a$), 3.43 (d, $J = 15.4$ Hz, 1H, 2-CH$_b$), 3.25 (s, 2H, 1'CH$_2$), 2.90 (d $J = 11.3$, 1H, 5-CH$_b$), 2.83 (d, $J = 10.0$ Hz, 1H, 7-CH), 2.22 (d, $J = 18.2$ Hz, 1H, 6-CH), 2.01 – 1.73 (m, 4H, 2'-CH and Cy-CH), 1.68 (d, $J = 11.9$ Hz, 1H, Cy-CH), 1.47 (dd, $J = 17.9, 8.4$ Hz, 2H, Cy-CH), 1.42 – 1.10 (m, 2H, Cy-CH), 0.97 (d, $J = 6.7$ Hz, 3H, 3'-CH$_3$), 0.93 (d, $J = 6.7$ Hz, 3H, 4'-CH$_3$); HRMS Found 305.2230; C$_{18}$H$_{28}$N$_2$O$_2$ [M+H]$^+$ requires 305.2223.

Cis-1-[(4-fluorophenyl)methyl]-4-(2-methylpropyl)-decahydro-1H-1,4-benzodiazepin-3-one, 50
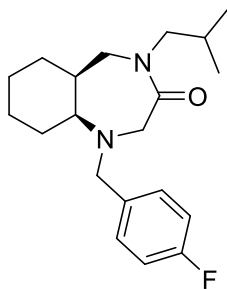
Prepared using general procedure A (26 mg, 87%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.35 − 7.27 (m, 2H, Ar-H), 7.07 − 6.96 (m, 2H, Ar-H), 4.00 (dd, J = 14.6, 10.8 Hz, 1H, 5-CH$_a$), 3.94 − 3.82 (m, 2H, benzylic CH$_a$ and 2-CH$_a$), 3.76 (d, J = 13.8 Hz, 1H, benzylic CH$_b$), 3.38 − 3.17 (m, 3H, 2-CH$_b$ and 2'CH$_2$), 2.84 (d, J = 15.0 Hz, 1H, 5-CH$_b$), 2.78 (dt, J = 11.6, 3.8 Hz, 1H, 7-CH), 2.27 (d, J = 5.4 Hz, 1H, 6-CH), 1.99 − 1.79 (m, 3H, 2'CH and Cy-CH), 1.74 (d, J = 12.8 Hz, 1H, Cy-CH), 1.65 − 1.53 (m, 1H, Cy-CH), 1.52 − 1.39 (m, 2H, Cy-CH), 1.36 − 1.22 (m, 1H, Cy-CH), 1.16 (t, J = 13.5 Hz, 1H, Cy-CH), 1.00 (d, J = 6.6 Hz, 3H, 3'-CH$_3$), 0.94 (d, J = 6.7 Hz, 3H, 4'-CH$_3$); HRMS Found: 333.2347; C$_{20}$H$_{29}$N$_2$FO [M+H]$^+$ requires 333.2336.

cis-1-[(1-methyl-1H-imidazol-4-yl)sulfonyl]-4-(2-methylpropyl)-decahydro-1H-1,4-benzodiazepin-3-one, 51
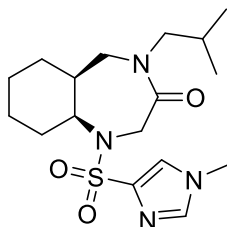
Prepared using general procedure B (23 mg, 69%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.49 (d, J = 1.2 Hz, 1H, Ar-H), 7.45 (s, 1H, Ar-H), 4.49 (dt, J = 9.2, 4.2 Hz, 1H, 7-CH), 4.24 (d, J = 16.3 Hz, 1H, 2-CH$_a$), 4.08 − 3.90 (m, 2H, 2-CH$_b$ and 5-CH$_a$), 3.73 (s, 3H, NMe), 2.98 (ddd, J = 20.3, 13.3, 7.4 Hz, 2H, diastereomeric 1'CH$_2$), 2.80 (d, J = 15.3 Hz, 1H, 5-CH$_b$), 2.29 (d, J = 10.7 Hz, 1H, 6-CH), 1.98 − 1.81 (m, 2H), 1.80 − 1.61 (m, 4H, 2'CH and Cy-CH), 1.46 (dd, J = 30.0, 13.0 Hz, 2H, Cy-CH), 1.18 − 0.99 (m, 1H, Cy-CH), 0.84 (d, J = 6.7 Hz, 3H, 3'-CH$_3$), 0.78 (d, J = 6.6 Hz, 3H, 4'-CH$_3$). HRMS Found: 369.1961; C$_{17}$H$_{28}$N$_4$O$_3$S [M+H]$^+$ requires 369.1954

5.4 Experimental for Chapter 4

Methyl 3-hydroxy-4-nitrobenzoate, 58
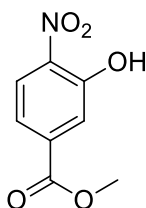
Following a literature procedure, *Para*-toluene sulphonic acid (*p*TSA) (415 mg, 2.18 mmol, 20 mol%) was added to a suspension of 3-hydroxy-4-nitrobenzoic acid (2 g, 10.9 mmol) in methanol (40 ml) and the reaction heated at reflux for 18 hours. After which time the reaction was concentrated *in vacuo* to afford a yellow solid which was then partitioned between H_2O (20 ml) and ethyl acetate (40 ml), the organic layer was washed with H_2O (2 × 20 ml) and saturated sodium bicarbonate solution (20 ml), dried over $MgSO_4$ and concentrated *in vacuo* to yield the *title compound* as yellow needles (1.8 g, 87%); R_f 0.3 (20% EtOAc in Petrol); δ_H (500 MHz; $CDCl_3$) 10.50 (1H, s (broad), OH), 8.16 (1H, d, $J = 8.7$ Hz, Ar-H), 7.83 (1H, d, $J = 1.6$ Hz, Ar-H), 7.62 (1H, dd, $J = 8.7, 1.6$ Hz, Ar-H), 3.95 (3H, s, OCH_3), δ_C (125 MHz; $CDCl_3$) 164.8, 154.6, 137.9, 135.8, 125.3, 121.6, 120.5, 52.9; *m/z* (ES) $[M+H]^+$ 197.

Methyl 3-isopropyloxy-4-nitrobenzoate 59
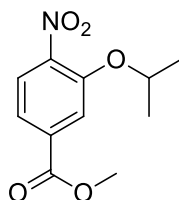
Potassium carbonate (2.1 g, 15.2 mmol) was added to a solution of methy-3-hydroxy-4-nitrobenzoate (600 mg, 3.04 mmol) in DMF (5 ml), followed by 2-bromopropane (0.40 ml, 4.26 mmol) and the reaction was stirred at 50 °C. After 20 hours the reaction was diluted with ethyl acetate (50 ml), washed with H_2O (2 × 50 ml) and brine (50 ml), dried over $MgSO_4$ and concentrated *in vacuo* to a pale yellow oil. The resulting residue was purified by flash column chromatography eluting with 1:1 Petrol–DCM to give the *title compound* as a yellow oil which crystallized on standing – yellow needles (396 mg, 54%), m.p. 54-56 °C (DCM); R_f 0.3 (1:1 Petrol–DCM); v_{max}/cm^{-1} (solid state) 3437, 3119, 2990, 2516, 2159; δ_H (500 MHz; $CDCl_3$) 7.70 (2H, m, ArH), 7.59 (1H, dd, $J = 8.4, 1.4$ Hz, ArH), 4.74 (1H, spt, $J = 6.2$ Hz, $C_\alpha H$), 3.91 (3H, s, OCH_3), 1.36 (6H, d, $J = 6.2$ Hz, $C_\beta H_3$ and $C_\gamma H_3$); δ_C (125 MHz; $CDCl_3$) 165.2, 150.6, 143.6, 134.4, 124.9, 121.0, 116.8, 72.9, 52.6, 21.6; HRMS *m/z* (ESI) Found 262.0686, $C_{12}H_{15}NO_5$ $[M+Na]^+$ requires 262.0691.

Methyl 3-(2-methyl-propyloxy)-4-nitrobenzoate, 60
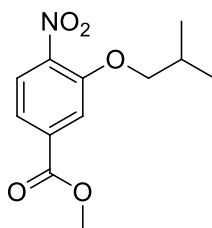
Potassium carbonate (4.20 g, 30.50 mmol) was added to a solution of methy-3-hydroxy-4-nitrobenzoate **2** (1.20 g, 6.09 mmol) in DMF (10 ml), followed by 1-bromo-2-methylpropane (0.93 ml, 8.52 mmol) and the reaction was stirred at 50 °C. After 20 hours a further aliquot of 1-bromo-2-methylpropane (0.93 ml, 8.52 mmol) was added and the reaction stirred at 50 °C for a further 24 hours. After this time the reaction was diluted with ethyl acetate (100 ml), washed with H_2O (2 × 100ml) and brine (100ml), dried over magnesium sulphate and concentrated *in vacuo* to an orange oil which crystallised on standing. The resulting mixture was purified by flash column chromatography eluting with 1:1 Petrol–DCM to give the *title compound* (964 mg, 63%) as a yellow crystalline solid; m.p. 69-71 °C (from DCM–Petrol); R_f 0.3 (1:1 Petrol–DCM); v_{max}/cm^{-1} (solid state) 3101, 2469, 2159, 1727; δ_H (500 MHz; CDCl$_3$) 7.78 (1H, d, J = 8.2 Hz, Ar-H), 7.69 (1H, s, Ar-H), 7.62 (1H, d, J = 8.2 Hz, Ar-H), 3.94 (3H, s, CO$_2$CH$_3$), 3.90 (2H, d, J = 6.3 Hz, C$_\alpha$H$_2$), 2.13 (1H, sept, J = 6.7 Hz, C$_\beta$H), 1.03 (6H, d, J = 6.7 Hz, C$_\gamma$H$_3$ and C$_\delta$H$_3$); δ_C (125 MHz; CDCl$_3$) 165.1, 151.9, 142.3, 134.6, 125.7, 125.0, 120.9, 115.3, 52.6, 28.1, 18.8 (2C); m/z (ESI) 276.1 (100%, [M+Na]$^+$); HRMS Found: 276.0835; C$_{12}$H$_{15}$NO$_5$[M+Na]$^+$ requires 276.0842.

5.4.1 Synthesis of Oligobenzamides

Standard Procedure C - Ester Hydrolysis

An aqueous sodium hydroxide solution (2 M, 1 ml per 100 mg of ester) was added to a solution of the ester in methanol (~5 ml per 100 mg of ester) and stirred at room temperature until the starting material had been consumed, as observed by TLC. The reaction mixture was concentrated by half *in vacuo* then adjusted to pH 3 by the addition of 1M HCl (aq.); the resulting precipitate was isolated by filtration, dried *in vacuo* and used without further purification in subsequent steps.

Standard Procedure D – Tin mediated nitro reduction

Tin (II) chloride dihydrate (5 equivalents) was added in one portion to a solution of the nitro compound in ethyl acetate (5 ml per 100 mg) and the reaction stirred at 50 °C under a calcium chloride drying tube for 24 hours. The reaction was then allowed to cool to room temperature and poured into 2 M sodium hydroxide solution (5 ml

per 100 mg of starting material). The organic layer was separated, washed with 2 M sodium hydroxide solution (2 × 5 ml per 100 mg of starting material) and brine (5 ml per 100 mg of starting material), dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by flash column chromatography to give the desired compound.

Standard Procedure E – Nitro reduction by hydrogenation

Palladium on charcoal (10 %) was added against a flow of nitrogen to a solution of the nitro compound in methanol (10 ml per 100 mg) under a nitrogen atmosphere, the atmosphere was then replaced with hydrogen and the reaction stirred vigorously until complete by TLC (typically 2 hours). The hydrogen atmosphere was vented and the reaction mixture filtered through a pad of Celite with methanol, concentrated *in vacuo* and purified by flash column chromatography.

Standard Procedure F - Coupling

Dichlorotriphenylphosphorane (4.5 equivalents) was added to a solution of nitro-acid compound (1.2 equivalents) in chloroform (5 ml per 100 mg of amine) and the reaction heated to reflux with stirring under nitrogen. After 2 hours at reflux, the amine ester compound (1 equivalent) was added as solution in chloroform (1 ml) and the reaction was heated to reflux for a further 24 hours. The reaction mixture was then concentrated *in vacuo* and partitioned between ethyl acetate (5 ml per 100 mg of amine) and H$_2$O (5 ml per 100 mg of amine). The organic layer was separated and washed with saturated aqueous sodium bicarbonate solution (5 ml per 100 mg of amine), dried over magnesium sulphate and concentrated *in vacuo*. The resulting residue was purified by flash column chromatography to give the desired compound.

3-Isopropyloxy-4-nitrobenzoic acid, 55
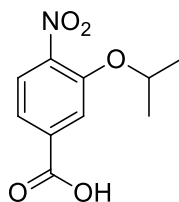
Prepared using procedure C from 3-isopropyloxy-4-nitrobenzoate (100mg, 0.42 mmol) to give pale yellow needles (41.5 mg, 44%); R_f 0.2 (DCM); m.p. 166-170 °C (H$_2$O); ν_{max}/cm^{-1} (solid state) 2970, 2534, 2159, 1737; δ_H (500 MHz; CDCl$_3$) 7.80 (1H, m, Ar-H), 7.78 (1H, s, Ar-H), 7.74 (1H, m, Ar-H), 4.79 (1H, spt, J = 6.0 Hz, C$_\alpha$H), 1.43 (6H, d, J = 6.4 Hz, C$_\beta$H$_3$ and C$_\gamma$H$_3$); δC (125 MHz; CDCl$_3$) 170.1, 150.8, 144.4, 133.2, 125.1, 121.8, 117.3, 73.1, 21.7; HRMS Found: 225.0622; C$_{10}$H$_{11}$NO$_5$ [M+H]$^+$ requires 225.0637

3-(2-Methyl-propyloxy)-4-nitrobenzoic acid, 56
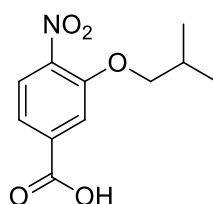
Prepared using procedure C from methyl 3-(2-methyl-propyloxy)-4-nitrobenzoate (400mg, 1.58 mmol) to give pale yellow needles (345 mg, 73%); R_f 0.2 (DCM); m.p. 188-191 °C (MeOH); ν_{max}/cm^{-1} (solid state) 2962 (broad), 1715, 1610; δ_H (500 MHz; CDCl$_3$) 7.87 (1H, d, J = 8.3 Hz, ArH), 7.81 (1H, d, J = 1.4 Hz, ArH), 7.78 (1H, dd, J = 8.3, 1.4 Hz, ArH), 3.97 (2H, d, J = 6.4 Hz, C$_\alpha$H$_2$), 2.20 (1H, sept, J= 6.4 Hz, C$_\beta$H), 1.10 (6H, d, J = 6.7 Hz, C$_\gamma$H$_3$ and C$_\delta$H$_3$); δ_C (125 MHz; CDCl$_3$) 169.6, 152.1, 143.2, 133.6, 125.3, 121.8, 115.8, 91.9, 28.2, 19.0; HRMS Found: 239.0805; C$_{12}$H$_{15}$NO$_5$ [M+Na]$^+$ requires 239.0794.

Methyl 3-isopropyloxy-4-aminobenzoate, 53
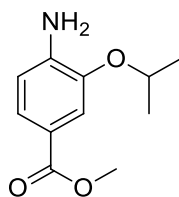
Prepared using procedure D from methyl 3-isopropyloxy-4-nitrobenzoate (100mg, 0.42 mmol) and purified by flash column chromatography eluting with DCM to give a pale orange oil (72 mg, 77%); ν_{max}/cm^{-1} (solid state) 3488, 3371, 2978, 1703; δ_H (500 MHz, CDCl$_3$) 7.55 (dd, J = 8.15, 1.8 Hz, 1 H, Ar-H), 7.49 (d, J = 1.6 Hz, 1 H, Ar-H), 6.69 (d, J = 7.9 Hz, 1 H, Ar-H), 4.66 (spt, J = 6.0 Hz, 1 H, C$_\alpha$H), 1.39 (d, J = 5.9 Hz, 6 H, C$_\beta$H$_3$ and C$_\gamma$H$_3$); δ_C (125 MHz, CDCl$_3$) 167.6, 144.1, 142.2, 123.8, 119.4, 114.0, 113.4, 70.8, 51.6, 22.1; HRMS Found: 210.1122; C$_{11}$H$_{16}$NO$_3$ [M+H]$^+$ requires 210.1125.

Methyl 3-(2-methyl-propyloxy)-4-amino-benzoate, 54
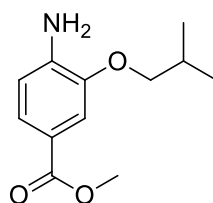
Prepared using procedure D from methyl 3-(2-methyl-propyloxy)-4-nitrobenzoate (200mg, 0.79 mmol) to give a pale yellow oil which crystallized on standing (125 mg, 71%). R_f 0.5 (DCM); m.p. 61-62 °C (from DCM); ν_{max}/cm^{-1} (solid state) 3462, 3341, 3202, 2951; δ_H (500 MHZ; CDCl$_3$) 7.45 (1H, dd, J = 8.2, 1.8 Hz, ArH) 7.35 (1H, d, J = 1.7 Hz, ArH), 6.58 (1H, d, J = 8.2 Hz, ArH), 4.18 (2H, s, broad, NH$_2$),

3.77 (3H, s, OCH$_3$), 3.73 (2H, d, J = 6.5 Hz, C$_\alpha$H$_2$), 2.04 (1H, sept, J = 6.7 Hz, C$_\beta$H), 0.96 (6H, d, J = 6.7 Hz, C$_\gamma$H$_3$ and C$_\delta$H$_3$); δ_C (125 MHz, CDCl$_3$) 167.4, 145.6, 141.4, 123.9, 119.3, 113.1, 74.7, 51.6, 28.3, 19.3; HRMS Found: 223.1212; C$_{11}$H$_{16}$NO$_3$ [M+H]$^+$ requires 223.1208.

O$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Pr (3-HABA)]-COOMe, 61
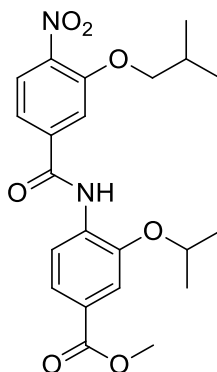
Prepared using procedure F on a 0.34 mmol scale to give a pale yellow oil (94 mg, 46%); R_f 0.25 (DCM); ν_{max}/cm^{-1} (film) 3421, 2979, 2504, 2159, 2030, 1976, 1717; δ_H (500 MHz; CDCl$_3$) 8.80 (1H, s, Amide NH), 8.59 (1H, d, J = 8.2 Hz, Ar-H), 7.93 (1H, d, J = 8.2 Hz, Ar-H), 7.74 (1H, dd, J = 8.5, 1.6 Hz, Ar-H), 7.68 (1H, d, J = 1.4 Hz, Ar-H), 7.61 (1H, d, J = 1.8 Hz, Ar-H), 7.37 (1H, dd, J = 8.2, 1.4 Hz, Ar-H), 4.78 (1H, spt, J = 6.1 Hz, 2-C$_\alpha$H), 3.97 (2H, d, J = 6.4 Hz, 1-C$_\alpha$H$_2$), 3.32 (3H, s, OMe), 2.19 (1H, spt, J = 6.6 Hz, 2-C$_\alpha$H), 1.44 (6H, d, J = 6.0 Hz, 2-C$_\beta$H$_3$ and 1-C$_\gamma$H$_3$), 1.07 (6H, d, J = 6.4 Hz, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$); δ_C (125 MHz; CDCl$_3$) 166.6, 163.0, 152.7, 145.6, 141.6, 139.7, 132.1, 125.8, 123.2, 118.8, 117.2, 114.9, 114.0, 113.1, 76.1, 71.8, 52.1, 28.1, 21.8, 18.9; HRMS Found 431.1816, C$_{22}$H$_{26}$N$_2$O$_7$ [M+H]$^+$ requires 431.1813.

O$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Bu (3-HABA)]-COOMe, 62
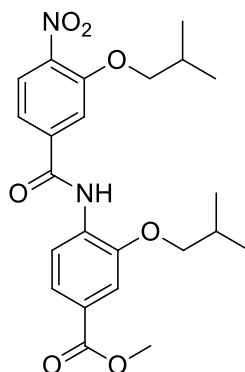
Prepared using procedure F on a 0.42 mmol scale to give pale yellow needles (92.3 mg, 49%); R_f 0.2 (DCM); m.p. 146-149 °C (DCM); ν_{max}/cm^{-1} (solid state) 3431, 3085, 2971, 1722; δ_H (500 MHz; CDCl$_3$) 8.70 (1H, s, Amide N-H), 8.50 (1H, d, J = 8.7 Hz, Ar-H), 7.83 (1H, d, J = 7.9 Hz, Ar-H), 7.66 (1H, d, J = 8.3 Hz, Ar-H), 7.56

(1H, s, Ar-H), 7.50 (1H, s, Ar-H), 7.31 (1H, d, J = 7.9 Hz, Ar-H), 3.87 (4H, m, 1-$C_\alpha H_2$ and 2- $C\alpha H_2$), 3.84 (3H, s, OMe), 2.12 (2H, m, 1-$C_\beta H$ and 2-$C_\beta H$), 1.03 (6H, d, J = 6.8 Hz, 1-$C_\gamma H_3$ and 1-$C_\delta H_3$), 0.99 (6H, d, J = 6.8 Hz, 2-$C_\gamma H_3$ and 2-$C_\delta H_3$); δ_C (125 MHz; CDCl$_3$) 166.5, 163.0, 152.8, 147.1, 141.6, 139.7, 131.4, 125.9, 125.8, 123.3, 118.6, 117.3, 113.7, 111.6, 76.1, 75.1, 52.1, 28.3, 28.2, 19.3, 19.0; HRMS Found 467.1789, $C_{23}H_{28}N_2O_7$ [M+Na]$^+$ requires 467.1789.

H$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Pr(3-HABA)]-COOMe, 63
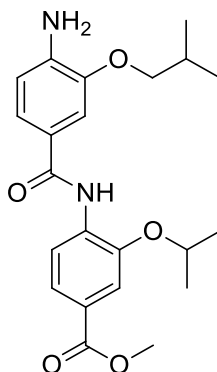
Prepared using procedure D from O$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Pr (3-HABA)]-COOMe (94 mg, 0.22 mmol) to give a colourless oil (47 mg, 53%); R_f 0.3 (2% Et$_2$O in DCM); ν_{max}/cm^{-1} (film) 3490, 3430, 3370, 2959, 1713; δ_H (500 MHz; CDCl$_3$) 8.79 (1H, s, Amide NH), 8.65 (1H, d, J = 8.2 Hz, Ar-H), 7.74 (1H, dd, J = 8.7 Hz, 1.4, Ar-H), 7.61 (1H, s, Ar-H), 7.45 (1H, d, J = 1.8 Hz, Ar-H), 7.31 (1H, d, J = 1.4 Hz, Ar-H), 6.76 (1H, d, J = 8.2 Hz, Ar-H), 4.78 (1H, spt, J = 6.0 Hz, 2-$C_\alpha H$), 4.27 (2H, s, NH$_2$), 3.93 (3H, s, OMe), 3.89 (2H, d, J = 6.4 Hz, 1-$C_\alpha H_2$), 2.19 (1H, spt, J = 6.6 Hz, 1-$C_\beta H$), 1.46 (6H, d, J = 6.0 Hz, 2-$C_\beta H_3$ and 2-$C_\gamma H_3$),1.09 (6H, d, J = 6.9 Hz, 1-$C_\gamma H_3$ and 1-$C_\delta H_3$); δ_C (125 MHz; CDCl$_3$) 166.9, 165.0, 146.2, 145.5, 140.5, 133.5, 124.4, 124.1, 123.4, 119.8, 118.4, 113.3, 113.1, 111.5, 74.7, 71.7, 52.0, 28.2, 19.3, 19.2; HRMS Found 423.1897, $C_{22}H_{28}N_2O_5$ [M+Na]$^+$ requires 423.1890.

H$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Bu (3-HABA)]-COOMe, 64
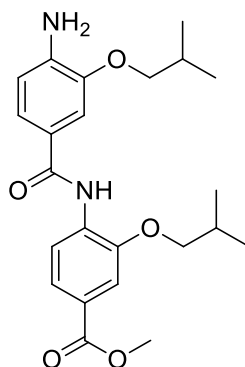
Prepared using procedure D from O$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Bu (3-HABA)]-COOMe (63 mg, 0.14 mmol) to give off-white needles (41.4 mg, 71%); R_f 0.3 (10%

MeOH in DCM); m.p. 147-148 °C (DCM); ν_{max}/cm^{-1} (solid state) 3461, 3435, 3335, 2961, 2529, 2159, 2028; δ_H (500 MHz; CDCl$_3$) 8.73 (1H, s, Amide NH), 8.63 (1H, d, J 8.7, Ar-H), 7.72 (1H, dd, J = 8.8, 1.6 Hz, Ar-H), 7.55 (1H, d, J = 1.8 Hz, Ar-H), 7.40 (1H, d, J = 1.8 Hz, Ar-H), 7.30 (1H, dd, J = 8.2, 1.8 Hz, Ar-H), 6.73 (1H, d, J = 7.8 Hz, Ar-H), 4.23 (2H, s (broad), NH$_2$), 3.92 (2H, d, J = 6.4 Hz, 2-C$_\alpha$H$_2$), 3.91 (3H, s, OMe), 3.85 (2H, d, J = 6.4 Hz, 2-C$_\alpha$H$_2$), 2.18 (2H, m, 1-C$_\beta$H and 2-C$_\beta$H), 1.11 (6H, d, J = 6.9 Hz, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.06 (6H, d, J = 6.9 Hz, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); δ_C (125 MHz; CDCl$_3$) 166.9, 165.1, 146.85 146.2, 140.5, 132.8, 124.4, 124.1, 123.5, 120.1, 118.2, 113.4, 111.4, 110.3, 75.0, 74.7, 52.0, 28.3, 28.3, 19.4, 19.3; HRMS Found: 437.2042; C$_{23}$H$_{30}$N$_2$O$_5$ [M+Na]$^+$ requires 437.2047.

O$_2$N-[*O-^{i}*Bu(3-HABA)]-[*O-^{i}*Bu (3-HABA)]-[*O-^{i}*Pr(3-HABA)]-COOMe, 65
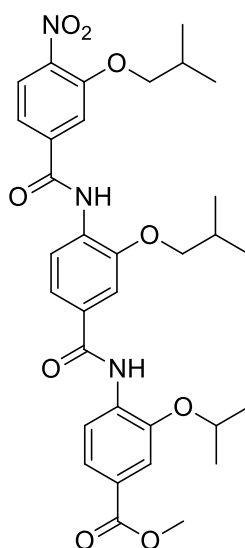
Prepared using procedure F on a 0.11 mmol scale to give a yellow solid (21.3 mg, 30%); R_f 0.25 (2% Et$_2$O in DCM); m.p. 189-191 °C (DCM); ν_{max}/cm^{-1} (solid state) 2966, 2514, 2159, 2026, 1706, 1680; δ_H (500 MHz; CDCl$_3$) 8.90 (1H, s, Amide NH), 8.81 (1H, s, Amide NH), 8.69 (1H, d, J = 8.2 Hz, Ar-H), 8.65 (1H, d, J = 8.7 Hz, Ar-H), 7.96 (1H, d, J = 8.2 Hz, Ar-H), 7.76 (1H, dd, J = 8.7, 1.4 Hz, Ar-H), 7.69 (1H, s, Ar-H), 7.64 (2H, d, J = 6.9 Hz, Ar-H), 7.46 (2H, dd, J = 12.1, 8.5 Hz, Ar-H), 4.81 (1H, spt, J = 6.0 Hz, 3-C$_\alpha$H), 4.01 (4H, dd, J = 10.5, 6.4 Hz, 1-C$_\alpha$H$_2$ and 2-C$_\alpha$H$_2$), 3.94 (3H, s, OMe), 2.24 (2H, m, 1-C$_\beta$H and 2-C$_\beta$H), 1.49 (6H, d, J = 6.4 Hz, 3-C$_\gamma$H$_3$ and 3-C$_\delta$H$_3$), 1.15 (6H, d, J = Hz 6.9, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.11 (6H, d, J = 6.9 Hz, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); δ_C (125 MHz; CDCl$_3$) 166.7, 164.2, 163.1, 152.8, 147.8, 145.8, 141.7, 139.6, 132.9, 130.7, 130.6, 125.9, 125.1, 123.3, 118.9, 118.9, 118.6, 117.4, 113.7, 113.1, 110.6, 76.1, 75.2, 71.8, 52.1, 28.2, 28.2, 22.2, 19.3,19.0; HRMS Found: 644.2596; C$_{33}$H$_{39}$N$_3$O$_9$ [M+Na]$^+$ requires 644.2579.

O$_2$N-[O-^{i}Pr(3-HABA)]-[O-^{i}Bu (3-HABA)]-[O-^{i}Bu(3-HABA)]-COOMe, 66
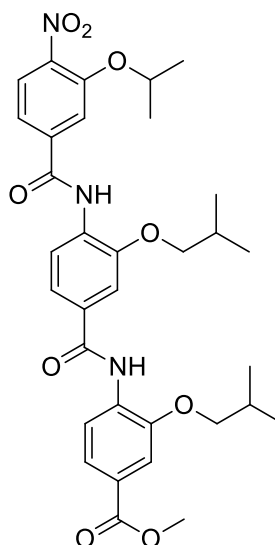
Prepared using procedure F on a 0.1 mmol scale and further purified by trituration with hexane to give pale yellow needles (5.5 mg, 9%); R_f 0.3 (99:1 DCM–Et$_2$O); m.p. 195-196 °C (Hexane); v_{max}/cm^{-1} (solid state) 3437, 2961, 2874, 1709, 1680; δ_H (500 MHz; CDCl$_3$) 8.88 (1H, s, Amide NH), 8.78 (1H, s, Amide NH), 8.66 (1H, d, J = 8.2 Hz, Ar-H), 8.63 (1H, d, J = 8.2 Hz, Ar-H), 7.88 (1H, s, Ar-H), 7.87 (1H, s, Ar-H), 7.75 (1H, dd, J = 8.5, 1.1 Hz, Ar-H), 7.70 (1H, s, Ar-H), 7.58 (1H, s, Ar-H), 7.46 (1H, d, J = 7.8 Hz, Ar-H), 7.38 (1H, d, J = 8.2 Hz, Ar-H), 4.82 (1H, spt, J = 6.0 Hz, 1-C$_\alpha$H), 3.98 (2H, d, J = 6.9 Hz, 1-C$_\alpha$H$_2$), 3.95 (2H, d, J = 6.4 Hz, 2-C$_\alpha$H$_2$), 3.92 (3H, s, OMe), 2.23 (2H, m, 1-C$_\beta$H and 2-C$_\beta$H), 1.45 (6H, d, J =6.0 Hz, 3-C$_\beta$H$_3$ and 3-C$_\gamma$H$_3$), 1.13 (6H, d, J = 6.9 Hz, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.11 (6H, d, J = 6.9 Hz, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); δ_C (125 MHz; CDCl$_3$) 166.7, 164.3, 163.2, 151.6, 147.8, 147.0, 143.2, 143.0, 139.3, 132.1, 130.8, 130.6, 125.8, 125.1, 123.4, 118.9, 118.4, 117.1, 115.4, 111.5, 110.5, 75.2, 75.1, 73.2, 52.1, 28.3, 28.2, 21.8, 19.38, 19.3; HRMS Found: 644.2567; C$_{33}$H$_{39}$N$_3$O$_9$ [M+Na]$^+$ requires 644.2579.

H₂N-[*O-ⁱBu*(3-HABA)]-[*O-ⁱBu* (3-HABA)]-[*O-ⁱPr*(3-HABA)]-COOMe, 67
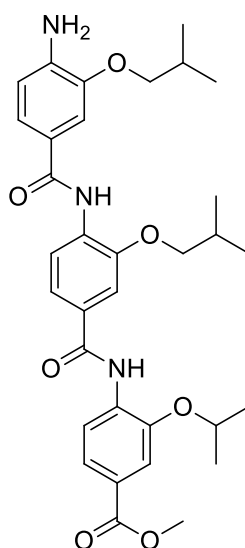
Prepared using procedure D from O₂N-[O-ⁱBu(3-HABA)]-[O-ⁱBu (3-HABA)]-[O-ⁱPr(3-HABA)]-COOMe (15 mg, 24 μmol) to give a colourless solid (4.4 mg, 31%). m.p. 197-198 °C (CH₂Cl₂); R_f 0.1 (2:98 Et₂O–CH₂Cl₂); ν_{max}/cm⁻¹ (solid state) 3491, 3436, 3350, 2960, 1714; δ_H (500 MHz; CDCl₃) 8.88 (1H, s, Amide NH), 8.73 (1H, s, Amide NH), 8.70 (1H, d, J = 8.5 Hz, Ar-H), 8.63 (1H, d, J = 8.5 Hz, Ar-H), 7.73 (1H, dd, J = 8.5 Hz, 1.7, Ar-H), 7.60 (2H, dd, J = 4.7, 1.7 Hz, Ar-H), 7.42 (2H, m, Ar-H), 7.33 (1H, dd, J = 8.1, 1.7 Hz, Ar-H), 6.75 (1H, d, J = 8.1 Hz, Ar-H), 4.77 (1H, spt, J = 6.1 Hz, 3-C$_α$H₂), 4.23 (2H, br. s, NH₂), 3.98 (2H, d, J = 6.4 Hz, 1-C$_α$H₂), 3.92 (3H, s, OCH₃), 3.87 (2H, d, J = 6.8 Hz, 2-C$_α$H₂), 2.23 (1H, spt, J =6.4 Hz, 1-C$_β$H), 2.18 (1H, spt, J = 6.8 Hz, 2-C$_β$H), 1.46 (6H, d, J = 6.0 Hz, 3-C$_β$H₃ and 3-C$_γ$H₃), 1.13 (6H, d, J = 6.8 Hz, 1-C$_γ$H₃ and 1-C$_δ$H₃), 1.08 (6H, d, J = 6.4 Hz, 2-C$_γ$H₃ and 2-C$_δ$H₃); δ_C (125 MHz; CDCl₃) 166.8, 165.1, 164.6, 147.6, 146.2, 145.7, 140.6, 133.2, 132.0, 129.1, 124.9, 124.0, 1?23.3, 120.1, 119.0, 118.5, 118.4, 113.4, 113.1, 110.4, 110.3, 75.1, 74.8, 71.9, 52.0, 28.4, 28.3, 22.2, 19.4, 19.3; HRMS Found: 614.2825; C₃₃H₄₁N₃O₇ [M+Na]⁺ requires 614.2837.

H$_2$N-[O-^{i}Pr(3-HABA)]-[O-^{i}Bu (3-HABA)]-[O-^{i}Bu(3-HABA)]-COOMe, 68
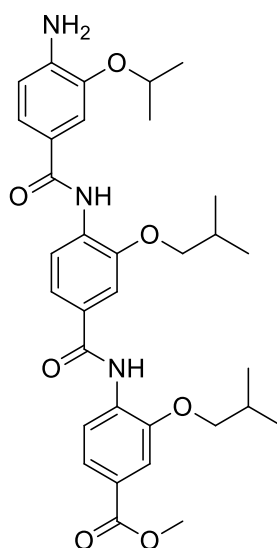
Prepared using procedure D from O$_2$N-[O-^{i}Pr(3-HABA)]-[O-^{i}Bu (3-HABA)]-[O-^{i}Bu(3-HABA)]-COOMe (89 mg, 0.14 mmol) to give a colourless solid (34 mg, 41%). m.p. 210-212 °C (CH$_2$Cl$_2$); R_f 0.25 (1:99 MeOH–CH$_2$Cl$_2$); ν_{max}/cm^{-1} (solid state) 3496, 3443, 3376, 2962, 2893, 1697, 1673; δ$_H$ (500 MHz; CDCl$_3$) 8.90 (1H, s, Amide NH), 8.77 (1H, s, Amide NH), 8.73 (1H, d, J = 8.7 Hz, Ar-H), 8.67 (1H, d, J = 8.7 Hz, Ar-H), 7.77 (1H, d, J = 8.7 Hz, Ar-H), 7.61 (2H, dd, J = 4.6, 1.4 Hz, Ar-H), 7.49 (1H, d, J =1.4 Hz, Ar-H), 7.46 (1H, dd, J = 8.5, 1.4 Hz, Ar-H), 7.32 (2H, dd, J = 8.5, 1.4 Hz, Ar-H), 6.80 (1H, d, J = 8.2 Hz, Ar-H), 4.72 (1H, spt, J =6 Hz, 1-C$_\alpha$H), 3.99 (2H, d, J =6.4 Hz, 2-C$_\alpha$H$_2$), 3.97 (2H, d, J = 6.4 Hz, 3-C$_\alpha$H$_2$), 3.95 (3H, s, CO$_2$Me), 2.25 (2H, m, 2-C$_\beta$H and 3-C$_\beta$H), 1.43 (6H, d, J =6.0 Hz, 1-C$_\beta$H$_3$ and 1-C$_\gamma$H$_3$), 1.15 (12H, overlapping doublets, J = 6.0 Hz, 2-C$_\gamma$H$_3$, 2-C$_\delta$H$_3$, 3-C$_\gamma$H$_3$ and 3C$_\delta$H$_3$); δ$_C$ (125 MHz; CDCl$_3$) 166.8, 165.1, 164.6, 147.6, 147.0, 145.0, 132.4, 132.0, 129.1, 124.9, 123.4, 119.9, 119.0, 118.5, 118.3, 113.9, 112.3, 111.5, 110.2, 75.1, 70.9, 52.0, 28.4, 28.3, 22.2, 19.4, 19.3; HRMS Found: 614.2834; C$_{33}$H$_{41}$N$_3$O$_7$ [M+Na]$^+$ requires 614.2837.

H$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Bu (3-HABA)]-[O-^{i}Pr(3-HABA)]-COOH, 1
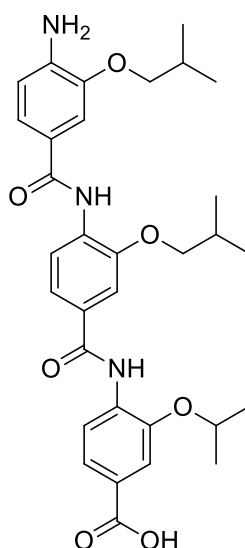
Prepared using procedure A from H$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Bu (3-HABA)]-[O-^{i}Pr(3-HABA)]-COOMe (43 mg, 0.07 mmol) to give the product as a pale yellow solid (18.4 mg, 45%); m.p. 211-213 °C (CH$_2$Cl$_2$–MeOH); v_{max}/cm^{-1} (solid state) 3365, 3480, 2958, 2874, 2552, 1682; δ_H (500 MHz; d_6-DMSO) 9.33 (1H, s, NH), 9.10 (1H, s, NH), 8.25 (1H, d, J = 8.5 Hz, Ar-H), 8.19 (1H, d, J = 8.1 Hz, Ar-H), 7.61 (1H, dd, J = 8.1, 1.7 Hz, Ar-H), 7.58 (3H, m, Ar-H), 7.41 (1H, dd, J = 8.1, 2.1, Ar-H), 7.38 (1H, d, J = 1.7 Hz, Ar-H), 6.86 (1H, d, J = 8.1 Hz, Ar-H), 4.74 (1H, spt, J = 6.1 Hz, 3-C$_\alpha$H), 3.97 (2H, d, J = 6.4 Hz, 1-C$_\alpha$H$_2$), 3.83 (2H, d, J = 6.8 Hz, 2-C$_\alpha$H$_2$), 2.17 (1H, spt, J = 6.4 Hz, 1-C$_\beta$H), 2.11 (1H, spt, J = 6.8 Hz, 2-C$_\beta$H), 1.37 (6H, d, J = 6.1 Hz, 3-C$_\beta$H$_3$ and 3-C$_\gamma$H$_3$), 1.06 (6H, d, J = 6.4 Hz, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.04 (6H, d, J = 6.8 Hz, 2-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$); δ_C (125 MHz; d_6-DMSO) 166.8, 164.5, 164.4, 148.5, 148.4, 147.4, 144.8, 132.5, 131.3, 129.2, 126.6, 126.5, 122.2, 121.3, 120.4, 120.0, 119.9, 113.9, 112.5, 110.5, 109.8, 74.5, 74.0, 71.4, 27.8, 28.7, 21.6, 19.1, 19.0; HRMS Found: 600.2673; C$_{32}$H$_{39}$N$_3$O$_7$ [M+Na]$^+$ requires 600.2680.

H$_2$N-[O-^{i}Pr(3-HABA)]-[O-^{i}Bu (3-HABA)]-[O-^{i}Bu(3-HABA)]-COOH, 52
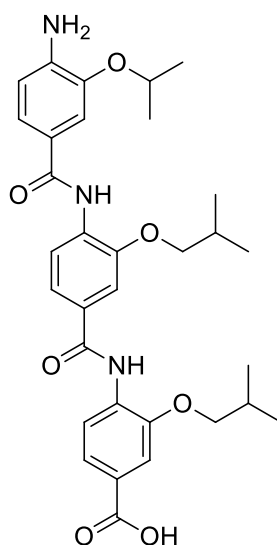
Prepared using procedure A from H$_2$N-[O-^{i}Pr(3-HABA)]-[O-^{i}Bu (3-HABA)]-[O-^{i}Bu(3-HABA)]-COOMe (24 mg, 0.04 mmol) to give the product as a pale yellow solid (20 mg, 86%). m.p. 209-211 °C (CH$_2$Cl$_2$); ν_{max}/cm^{-1} (solid state) 3379, 2959, 2927, 2549, 1715; δ_H (500 MHz; d_6-DMSO) 9.42 (1H, s, NH), 9.16 (1H, s, NH), 8.15 (1H, d, J = 8.1 Hz, Ar-H), 8.08 (1H, d, J = 8.5 Hz, Ar-H), 7.60 (2H, m, Ar-H), 7.57 (1H, dd, J = 8.5 Hz, 1.4, Ar-H), 7.55 (1H, d, J = 1.7 Hz, Ar-H), 7.44 (1H, d, J = 1.4 Hz, Ar-H), 7.39 (1H, dd, J = 8.3, 1.9 Hz, Ar-H), 6.92 (1H, d, J = 7.7 Hz, Ar-H), 4.64 (1H, spt, J = 6.1 Hz, 1-C$_\alpha$H), 3.93 (2H, d, J = 6.4 Hz, 2-C$_\alpha$H$_2$), 3.90 (2H, d, J = 6.4 Hz, 3-C$_\alpha$H$_2$), 2.13 (1H, spt, J = 6.4 Hz, 2-C$_\beta$H$_2$), 2.11 (1H, spt, J = 6.4 Hz, 3-C$_\beta$H$_2$), 1.31 (6H, d, J = 6.1 Hz, 1-C$_\beta$H$_3$ and 1-C$_\gamma$H$_3$), 1.02 (6H, d, J = 6.4 Hz, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$), 1.01 (6H, d, J = 6.4 Hz, 3-C$_\gamma$H$_3$ and 3-C$_\delta$H$_3$); δ_C (125 MHz; d_6-DMSO) 166.8, 164.4, 164.2, 149.4, 149.0, 147.2, 143.9, 132.2, 131.5, 129.1, 127.0, 126.8, 122.1, 121.8, 121.3, 121.1, 121.0, 120.0, 112.7, 112.2, 110.6, 110.5, 74.5, 74.4, 70.6, 27.8, 27.7, 21.8, 19.1, 19.0; HRMS Found: 600.2675; C$_{32}$H$_{39}$N$_3$O$_7$ [M+Na]$^+$ requires 600.2680.

5.4.2 Library Synthesis

The compound library was prepared in parallel using the methods and building blocks described above, checking at pertinent times during the synthesis by crude NMR and LC-MS, to afford the below compounds. Compounds were either pure following final precipitation or purified by preparative HPLC. Where the amount of material produced allowed ^{13}C-NMR spectra were collected otherwise ^{1}H-NMR and high resolution mass spectrometry were used to characterise the compound library.

H₂N-[*O*-Me(3-HABA)]-[*O*-*ᶦ*Pr (3-HABA)]-[*O*-*ᶦ*Bu(3-HABA)]-COOH, 69
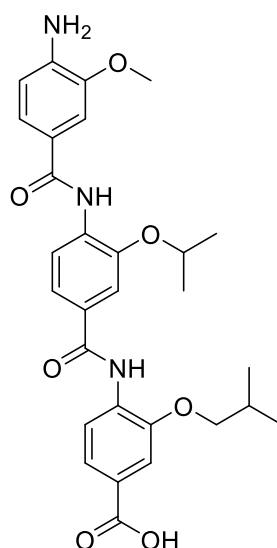
δ_H (500 MHz; d_6-DMSO) 9.35 (1H, s, Amide NH), 8.97 (1H, s, Amide NH), 8.28 (1H, d, J = 8.5 Hz, Ar-H), 8.07 (1H, d, J = 8.5 Hz, Ar-H), 7.59 (4H, m, Ar-H), 7.35 (2H, m, Ar-H), 6.70 (1H, d, J = 8.5 Hz, Ar-H), 4.78 (1H, spt, J = 6.0 Hz 2-C$_\alpha$H), 3.90 (2H, d, J = 6.4 Hz, 3-C$_\beta$H$_2$), 3.85 (3H, s, 1-C$_\alpha$H$_3$), 2.10 (1H, spt, J = 6.6 Hz, 3-C$_\alpha$H), 1.38 (6H, d, J = 6.0 Hz, 2-C$_\beta$H$_3$ and 2-C$_\delta$H$_3$), 1.02 (6H, d, J = 6.8 Hz, 3-C$_\delta$H$_3$ and 3C$_\gamma$H$_3$); HRMS Found: 536.2399, C$_{29}$H$_{33}$N$_3$O$_7$ [M+H]$^+$ requires 536.2391;

H₂N-[*O*-*ᶦ*Bu(3-HABA)]-[*O*-*ᶦ*Bu (3-HABA)]-[*O*-*ᶦ*Bu(3-HABA)]-COOH, 70
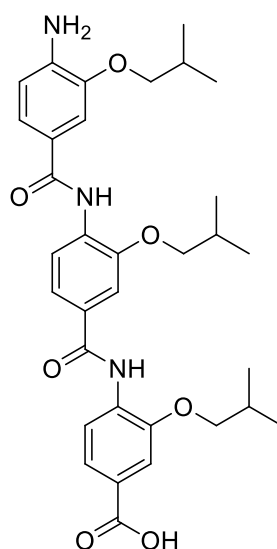
δ_H (500 MHz; d_6-DMSO) 9.31 (1H, s, Amide N-H), 9.00 (1H, s, Amide N-H), 8.15 (1H, d, J = 8.1 Hz, Ar-H), 8.00 (1H, d, J = 8.1 Hz, Ar-H), 7.61 (4H, m, Ar-H), 7.35 (2H, m, Ar-H), 6.71 (1H, d, J = 8.5 Hz, Ar-H), 4.05 (2H, d, J = 6.4 Hz, 1-C$_\alpha$H$_2$), 3.92 (4H, m, 2-C$_\alpha$H$_2$ and 3-C$_\alpha$H$_2$), 2.11 (3H, m, 1-C$_\beta$H, 2-C$_\beta$H and 3-C$_\beta$H), 1.02 (18H, m, 1-C$_\gamma$H$_3$ and C$_\delta$H$_3$, 2-C$_\gamma$H$_3$ and C$_\delta$H$_3$ and 3-C$_\gamma$H$_3$ and C$_\delta$H$_3$); HRMS Found: 614.2840, C$_{33}$H$_{41}$N$_3$O$_7$ [M+Na]$^+$ requires 614.2836.

H$_2$N-[*O-^{i}Bu(3-HABA)*]-[*O-^{i}Pr (3-HABA)*]-[*O-^{i}Bu(3-HABA)*]-COOH, 71
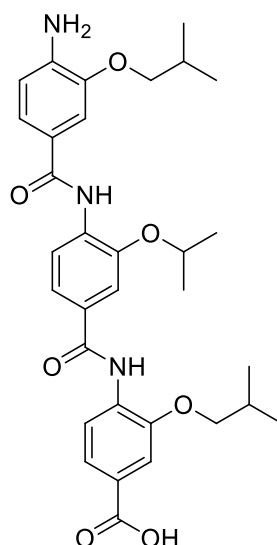
δ_H (500 MHz; d_6-DMSO) 9.35 (1H, s, Amide N-H), 8.96 (1H, s, Amide N-H), 8.28 (1H, d, J = 8.5 Hz, Ar-H), 8.07 (1H, d, J = 8.1 Hz, Ar-H), 7.58 (4H, m, Ar-H), 7.32 (2H, m, Ar-H), 6.72 (1H, d, J = 8.1 Hz, Ar-H), 4.78 (1H, spt, J = 6.0 Hz, 2-C$_\alpha$H$_2$), 3.89 (2H, d, J = 6.8 Hz, 1-C$_\alpha$H$_2$), 3.81 (2H, d, J = 6.8 Hz, 3-C$_\alpha$H), 2.09 (2H, m, 1-C$_\beta$H and 3-C$_\beta$H), 1.38 (6H, d, J = 6.0 Hz, 2-C$_\beta$H$_3$ and 2-C$_\gamma$H$_3$), 1.03 (6H, d, J = 6.8 Hz, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.01 (6H, d, J = 6.8 Hz, 3-C$_\gamma$H$_3$ and 3-C$_\delta$H$_3$); HRMS Found: 578.2863, C$_{32}$H$_{39}$N$_3$O$_7$ [M+H]$^+$ requires 578.286;

H$_2$N-[*O-Bn(3-HABA)*]-[*O-Bn (3-HABA)*]-[*O-Bn(3-HABA)*]-COOH, 72
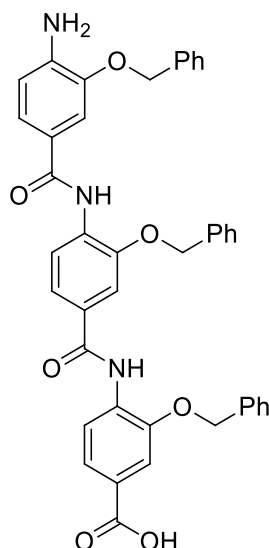
δ_H (500 MHz; d_6-DMSO) 9.54 (1H, m, Amide NH), 9.14 (1H, s, Amide NH), 8.11 (2H, m, Ar-H), 7.63 (10H, m, Ar-H), 7.48 (2H, m, Ar-H), 7.34 (12H, m, Ar-H), 6.76 (1H, d, J = 8.1 Hz, Ar-H), 5.29 (2H, m, Benzylic CH$_2$), 5.25 (2H, s, Benzylic CH$_2$), 5.12 (2H, s, Benzylic CH$_2$); δ_C (125 MHz; d_6-DMSO) 166.8, 165.8, 164.5, 164.4, 149.3, 149.0, 145.9, 136.6, 136.5, 13200, 131.6, 131.2, 130.0, 129.9, 128.4, 128.4, 128.0, 127.9, 127.9, 127.8, 127.5, 127.4, 127.3, 127.2, 125.9, 122.5, 121.8, 121.3,

120.3, 115.4, 113.1, 112.9, 111.6, 111.3, 70.3, 70.2, 70.1; HRMS *m/z* (ESI) Found: 716.2352, $C_{42}H_{35}N_3O_7$ [M+Na]$^+$ requires 716.2367.

H$_2$N-[*O*-^{i}Pr(3-HABA)]-[*O*-^{i}Pr (3-HABA)]-[*O*-^{i}Pr(3-HABA)]-COOH, 73
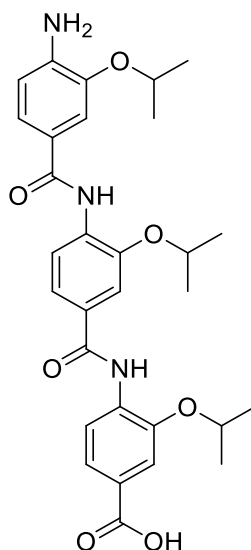
δ_H (500 MHz; d_6-DMSO) 9.28 (1H, s, Amide N-H), 8.94 (1H, s, Amide N-H), 8.3 (1H, d, *J* = 8.5 Hz, Ar-H), 8.18 (1H, d, *J* = 8.5 Hz, Ar-H), 7.60-7.54 (4H, m, Ar-H), 7.35-7.32 (2H, m, Ar-H), 6.72 (1H, d, *J* = 8.1 Hz, Ar-H), 4.79 (1H, spt, *J* = 6.0 Hz, C$_\alpha$H), 4.72 (1H, spt, *J* = 6.0 Hz, C$_\alpha$H), 4.60 (1H, spt, *J* = 6.Hz, C$_\alpha$H), 1.39 (6H, d, *J* = 6.0 Hz, C$_\beta$H$_3$ and C$_\gamma$H$_3$), 1.35 (6H, d, *J* = 6.0 Hz, C$_\beta$H$_3$ and C$_\gamma$H$_3$), 1.32 (6H, d, *J* = 6.0 Hz, C$_\beta$H$_3$ and C$_\gamma$H$_3$); HRMS Found: 550.2558, $C_{30}H_{35}N_3O_7$ [M+H]$^+$ requires 550.2547.

H$_2$N-[*O*-^{i}Bu(3-HABA)]-[*O*-^{i}Pr (3-HABA)]-[*O*-^{i}Pr(3-HABA)]-COOH, 74
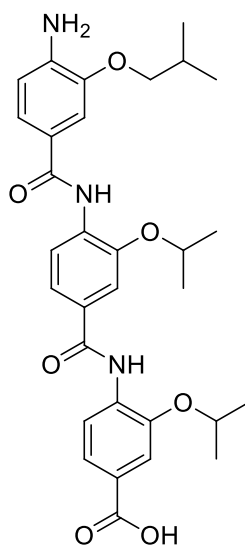
δ_H (500 MHz; d_6-DMSO) 9.30 (1H, s, Amide N-H), 9.08 (1H, s, Amide N-H), 8.25 (1H, d, *J* = 8.5 Hz, Ar-H), 8.17 (1H, d, *J* = 8.1 Hz, Ar-H), 7.59 (4H, m, Ar-H), 7.38 (2H, m, Ar-H), 6.91 (1H, d, *J* = 7.7 Hz, Ar-H), 4.79 (1H, spt, *J* = 6 Hz, 2-C$_\alpha$H), 4.69

(1H, spt, J = 6 Hz, 3-C$_\alpha$H), 3.84 (2H, d, J = 6.8 Hz, 1-C$_\beta$H$_2$), 2.09 (1H, spt, J = 6.8 Hz, 1-CαH), 1.39 (6H, d, J = 6.0 Hz, 2-C$_\beta$H$_3$ and 2-C$_\delta$H$_3$), 1.34 (6H, d, J = 6.0 Hz, 3-C$_\beta$H$_3$ and 3-C$_\delta$H$_3$), 1.02 (6H, d, J = 6.8 Hz, 1-C$_\delta$H$_3$ and 1-C$_\gamma$H$_3$); δ_C (125 MHz; d_6-DMSO)166.9, 164.3, 164.2, 147.6, 147.2, 132.4, 132.0, 129.5, 126.6, 122.1, 121.3, 120.9, 120.6, 120.0, 113.9, 112.2, 110.4, 74.2, 71.4, 71.4, 27.7, 21.8, 21.7; HRMS Found: 564.2716, C$_{31}$H$_{37}$N$_3$O$_7$ [M+H]$^+$ requires 564.2704.

H$_2$N-[*O*-Bn(3-HABA)]-[*O*-^{i}Pr (3-HABA)]-[*O*-^{i}Pr(3-HABA)]-COOH, 75
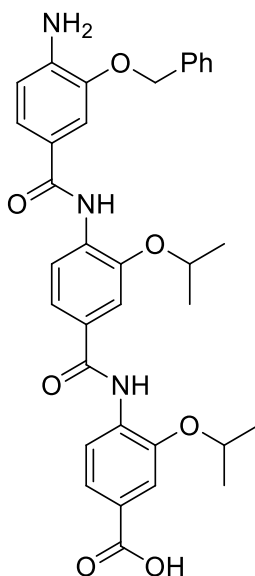
δ_H (500 MHz; d_6-DMSO) 9.31 (1H, s, Amide NH), 9.07 (1H, s, Amide NH), 8.23 (1H, d, J = 8.1 Hz, Ar-H), 8.16 (1H, d, J = 8.1 Hz, Ar-H), 7.60-7.51 (7H, m, Ar-H), 7.41-7.38 (3H, m, Ar-H), 7.34-7.31 (1H, m, Ar-H), 6.92 (1H, d, J = 8.1 Hz, Ar-H), 5.23 (2H, s, 1-C$_\alpha$), 4.78 (1H, spt, J = 6.1 Hz, 2-C$_\alpha$), 4.71 (1H, spt, J = 6.1 Hz, 3-C$_\alpha$), 1.37 (6H, d, J = 6.1 Hz, 2-C$_\beta$ and 2-C$_\gamma$), 1.34 (6H, d, J = 6.1 Hz, 3-C$_\beta$ and 3-Cγ); δ_C (125 MHz; d_6-DMSO) 166.9, 164.3, 164.2, 147.6, 147.3, 136.9, 132.4, 132.0, 129.5, 128.4, 127.7, 127.3, 126.7, 122.1, 121.3, 121.2, 120.8, 119.9, 113.9, 112.3, 111.4, 71.5, 71.4, 69.5, 21.8, 21.7; HRMS Found: 598.2562, C$_{34}$H$_{35}$N$_3$O$_7$ [M+H]$^+$ requires 598.2547.

**H$_2$N-[*O*-2-hydroxyethyl(3-HABA)]-[*O*-^{i}Pr (3-HABA)]-[*O*-^{i}Pr(3-HABA)]-COOH, 76**
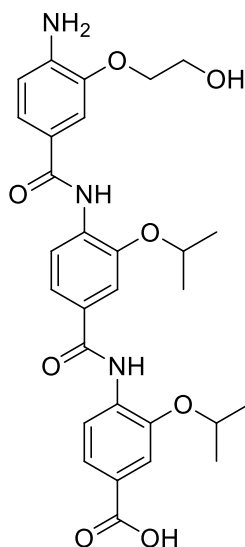
δ_H (500 MHz; d_6-DMSO) 9.27 (1H, s, Amide NH), 8.95 (1H, s, Amide NH), 8.28 (1H, d, *J* = 8.5 Hz, Ar-H), 8.16 (1H, d, *J* = 8.1 Hz, Ar-H), 7.56 (4H, m, Ar-H), 7.35 (2H, m, Ar-H), 6.70 (1H, d, *J* = 8.1 Hz, Ar-H), 4.79 (1H, spt, *J* = 6 Hz, 2-Cα), 4.71 (1H, spt, *J* = 6.0 Hz, 3-Cα), 4.02 (2H, t, *J* = 4.7 Hz, 1-Cα), 3.75 (2H, t, *J* = 4.7 Hz, 1-Cβ), 1.39 (6H, d, *J* = 6.0 Hz, 2-Cβ and 2-Cγ), 1.35 (6H, d, *J* = 6.0 Hz, 3-Cβ and 3-Cγ); HRMS Found: 552.2343, C$_{29}$H$_{33}$N$_3$O$_8$ [M+H]$^+$ requires 552.2340;

5.4.3 Synthesis of KCN-1

4-(2-methylbut-3-yn-2-yloxy)benzaldehyde, 78
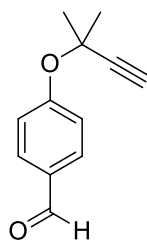
Following a literature procedure,[176] 4-hydroxy benzaldehyde (664 mg, 5.4 mmol), potassium carbonate (1.5 g, 10.8 mmol), potassium iodide (1.5g, 9.2 mmol) and copper iodide (52 mg, 5 mol%) were dissolved in DMF (20 ml) and 3-chloro-3-methyl-1-butyne (1.1 ml, 9.8 mmol) was added. The reaction was heated to 65 °C for 2.5 hours, allowed to cool to r.t. diluted with ether (30 ml) and washed with 1 M aqueous sodium hydroxide (3 × 15 ml) and brine (2 × 15 ml), dried over MgSO$_4$ and concentrated *in vacuo* to yield the *title compound* as an orange oil (1.01g, 99%). ν_{max}/cm^{-1} (solid state) 3291, 2990, 1698, 1601, 1577; δ_H (500 MHz; CDCl$_3$) 9.91 (1H, s, CHO), 7.83 (2H, d, *J* = 8.7 Hz, Ar-H), 7.35 (2H, d, *J* = 8.7 Hz, Ar-H), 2.69 (1H, s, Alkynyl CH), 1.73 (6H, s, 2 × Me); δ_C (125 MHz; CDCl$_3$) 190.9, 161.1,

131.3, 131.2, 119.6, 84.8, 73.4, 72.4, 29.4; HRMS *m/z* (ESI) Found: 188.0842, $C_{12}H_{12}O_2$ [M+H]$^+$ requires 188.0837.

2,2-dimethyl-2H-chromene-6-carbaldehyde, 79
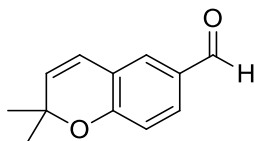
Following a literature procedure,[176] 4-(2-methylbut-3-yn-2-yloxy)benzaldehyde (1 g, 5.3 mmol) was dissolved in toluene (5 ml) and heated to 190 °C under microwave irradiation for 25 minutes, allowed to cool, diluted with ethyl acetate and concentrated *in vacuo* to yield the *title compound* as a dark orange oil (1 g, quant.). v_{max}/cm^{-1} (solid state) 2976, 2648, 1690, 1600; δ_H (500 MHz; CDCl$_3$) 9.83 (1H, s, CHO), 7.66 (1H, dd, *J* = 8.2 Hz, 1.4, Ar-H), 7.52 (1H, dd, *J* = 1.4 Hz, Ar-H), 6.87 (1H, dd, *J* = 8.2 Hz, Ar-H), 6.37 (1H, d, *J* = 10.1 Hz, Alkene), 5.70 (1H, d, *J* = 9.6 Hz, Alkene), 1.48 (6H, s, 2 × Me); δ_C (125 MHz; CDCl$_3$) 190.9, 158.7, 131.9, 131.4, 129.9, 127.8, 121.2, 116.8, 116.0, 77.9, 28.5; HRMS *m/z* (ESI) Found: 188.0830, $C_{12}H_{12}O_2$ [M+H]$^+$ Requires 188.0837.

N-[(2,2-dimethyl-2H-chromen-6-yl)methyl]aniline, 80
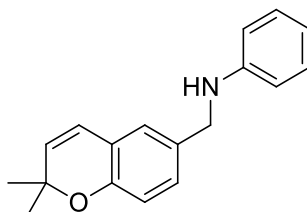
Prepared using a modified literature procedure[177], aniline (1 ml, 10.6 mmol) was added to a solution of 2,2-dimethyl-2H-chromene-6-carbaldehyde (1 g, 5.3 mmol) and *p*-toluenesulfonic acid (200 mg, 1.06 mmol) in methylene chloride (20 ml) and heated to reflux for 3 hours. Reaction mixture concentrated *in vacuo* and NMR used to confirm full conversion to imine, residue dissolved in toluene (30 ml) and cooled to 0 °C (ice bath), diisobutylaluminum hydride (1 M in hexanes, 10.6 ml, 10.6 mmol) was added dropwise and the reaction stirred overnight and allowed to warm to r.t., cooled to 0 °C and quenched by the dropwise addition of saturated aqueous ammonium chloride solution with vigorous stirring. The resulting precipitate was removed by filtration and the filtrate diluted with ethyl acetate (30 ml), washed with water (10 ml), 1M aqueous HCl (3 × 10 ml) and brine (10 ml), dried over MgSO$_4$ and concentrated *in vacuo* to yield the *title compound* as a brown oil (907 mg, 64%). v_{max}/cm^{-1} (solid state) 3417, 2974, 2864, 1603, 1505; δ_H (500 MHz; CDCl$_3$) 7.09 (2H, t, *J* = 7.8 Hz, Aniline Ar-H), 7.00 (1H, dd, *J* = 8.2, 1.8 Hz, Chromene Ar-H), 6.89 (1H, dd, *J* = 1.4 Hz, Chromene Ar-H), 6.67-6.61 (2H, m, Aniline and

Chromene Ar-H), 6.55 (2H, dd, J = 7.8 Hz, Aniline Ar-H), 6.20 (1H, d, J = 9.6 Hz, Alkene), 5.52 (1H, d, J = 9.6 Hz, Alkene), 4.1 (2H, s, Benzylic CH$_2$), 3.88 (1H, s (br), NH), 1.34 (6H, s, 2 × Me); δ_C (125 MHz; CDCl$_3$) 152.2, 148.2, 131.4, 131.0, 129.2, 128.4, 125.6, 122.2, 121.3, 117.5, 116.4, 112.9, 76.2, 47.9, 28.0; HRMS *m/z* (ESI) *m/z* (ESI) Found: 266.1539, C$_{18}$H$_{19}$NO [M+H]$^+$ requires 266.1539.

N-[(2,2-dimethyl-2H-chromen-6-yl)methyl]-3,4-dimethoxy-N-phenylbenzene-1-sulfonamide/KCN-1
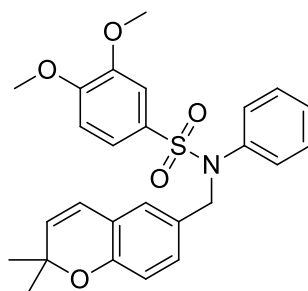
Prepared using a modified literature procedure,[177] 3,4-dimethoxybenzene sulfonyl chloride (890 mg, 3.77 mmol) was added to a solution of N-[(2,2-dimethyl-2H-chromen-6-yl)methyl]aniline (500 mg, 1.88 mmol) and triethylamine (1.3 ml, 9.4 mmol) in methylene chloride (50 ml) and the reaction heated to reflux overnight, allowed to cool to r.t. and diluted with methylene chloride, washed with water, saturated aqueous sodium bicarbonate, 1M aqueous HCl, and brine, dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 20-30% ethyl acetate in petrol to yield the *title compound* as a yellow solid (596 mg, 68%). v_{max}/cm^{-1} (solid state) 2974, 2934, 1586, 1507; δ_H (500 MHz; CDCl$_3$) 7.33 (1H, dd, J = 8.5, 2.1 Hz, Ar-H), 7.23-7.19 (3H, m, Ar-H), 7.01-6.99 (2H, m, Ar-H), 6.96 (1H, dd, J = 1.8 Hz, Ar-H), 6.93-6.86 (3H, m, Ar-H), 6.59 (1H, d, J = 8.2 Hz, Ar-H), 6.21 (1H, d, J = 10.1 Hz, Alkene), 5.56 (1H, d, J = 9.6 Hz, Alkene), 4.62 (2H, s, Benzylic CH$_2$), 3.93 (3H, s, OMe), 3.74 (3H, s, OMe), 1.37 (6H, s, 2 × Me); HRMS *m/z* (ESI) Found: 466.168811, C$_{26}$H$_{27}$NO$_5$S [M+H]$^+$ requires 466.16827.

5.4.4 Synthesis of Helix 2 Mimetics

2-Bromo-*N*-(triphenylmethyl)-acetamide, 87
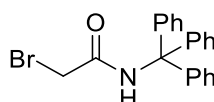
A solution of tritylamine (500 mg, 1.93 mmol) and triethylamine (0.73 mL, 5.25 mmol) in CH$_2$Cl$_2$ (5ml) was added drop-wise to a cooled solution (0 °C) of bromo-acetyl bromide (0.8 mL, 1.75 mmol) and triethylamine (0.73 mL, 5.25 mmol) in DCM (10 mL) with rapid stirring under an inert atmosphere. The reaction was

stirred and allowed to warm to room temperature over 15 hours. The reaction mixture was concentrated *in vacuo* and immediately purified by column chromatography eluting with CH_2Cl_2. Fractions containing product were concentrated *in vacuo* to a black solid which upon trituration with MeOH gave the product as a colourless solid (608 mg, 91%). m.p. 201-203 °C (CH_2Cl_2); R_f 0.35 (CH_2Cl_2); v_{max}/cm^{-1} (solid state) 3261, 3053, 3032, 1660; δ_H (500 MHz; $CDCl_3$) 7.75 (1H, s, NH), 7.33 (10H, m, Ar-H), 7.25 (5H, d, $J = 6.9$ Hz, Ar-H), 3.91 (2H, s, CH_2); δ_C (125 MHz; $CDCl_3$) 164.3 (Carbonyl), 144.1 (trityl), 128.6 (trityl), 128.2 (trityl), 127.3 (trityl), 30.0 (CH_2) quaternary carbon not observed; *m/z* (ES) 404 (100%, MH$^+$), 402 (100%, MH$^+$), 243 (50%, Trt$^+$); HRMS Found: 402.0475; $C_{21}H_{18}NOBr$ [M+Na]$^+$ requires 402.0464.

S-trityl-mercapto-ethanol, 89
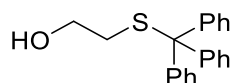
Following a literature procedure,[179] mercaptoethanol (0.23 ml, 3.27 mmol) was added to a solution of trityl chloride (1g, 3.59 mmol) in THF (5 mL) and heated to reflux for 4 hours, the reaction was allowed to cool, concentrated *in vacuo* and triturated with 1:2 EtOAc–Hexane to yield the desired compound as a colourless solid (466 mg, 44%). m.p. 108-110 °C (Hexane); v_{max}/cm^{-1} (film) 3336 (broad), 3063, 2926, 1592; δ_H (500 MHz; $CDCl_3$) 7.50 (6H, d, $J = 7.8$ Hz, trityl), 7.34 (6H, m, trityl), 7.27 (3H, m, trityl), 3.42 (2H, t, $J = 6.2$ Hz, CH_2), 2.53 (2H, t, $J = 6.2$ Hz, CH_2), 1.93 (1H, s, OH); δ_C (125 MHz; $CDCl_3$) 144.8 (trityl), 129.7 (trityl), 128.1 (trityl), 126.8 (trityl), 66.74 (CH_2S), 35.26 (CH_2O), quaternary carbon not observed; *m/z* (ES) [MH]$^+$ 343, [2M+Na]$^+$ 663; HRMS Found: 343.1143; $C_{21}H_{20}OS$ [M+Na]$^+$ requires 343.1127.

Methyl 4-nitro-3-{2-[(triphenylmethyl)sulfanyl]ethoxy}benzoate, 90
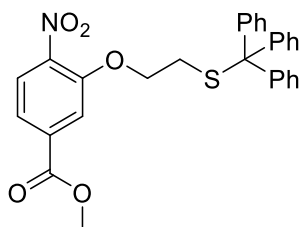
Diisopropyl azodicarboxylate (DIAD) (0.09 mL, 0.45 mmol) was added to a cooled solution (0 °C) of methyl 3-hydroxy-4-nitrobenzoate (88 mg, 0.45 mmol), *S*-trityl mercaptoethanol (200 mg, 0.63 mmol) and triphenylphosphine (116 mg, 0.45 mmol) in anhydrous THF (5 mL) under an inert atmosphere and reaction allowed to warm to room temperature with stirring. After 3 days the reaction was concentrated *in vacuo*, dissolved in CH_2Cl_2 (10 mL) and washed with H_2O (2 × 10 mL), saturated

aqueous sodium bicarbonate (3 × 10 mL) and brine (10 mL), dried over $MgSO_4$ and concentrated *in vacuo* to a yellow oil. The oil was purified by column chromatography eluting with 1:1 CH_2Cl_2–petrol and 7:3 CH_2Cl_2–petrol sequentially to give the product as a colourless oil which crystallized to colourless needles on standing (79 mg, 35%). m.p. 79-81 °C (7:3 CH_2Cl_2, Petrol); R_f 0.35 (1:1, CH_2Cl_2, Petrol); v_{max}/cm^{-1} (Solid State) 3054, 2478, 2254, 2159, 2028, 1974, 1727; δ_H (500 MHz; CDCl$_3$) 7.70 (1H, d, *J* = 8.2 Hz, Ar-H), 7.57 (1H, dd, *J* = 8.5, 1.6 Hz, Ar-H), 7.38 (7H, d, *J* = 7.8 Hz, trityl and Ar-H), 7.22 (6H, t, *J* = 7.8 Hz, trityl), 7.13 (3H, t, *J* = 7.8 Hz, trityl), 3.87 (3H, s, CO$_2$Me), 3.64 (2H, t, *J* = 6.8 Hz, CH$_2$S), 2.63 (2H, t, *J* = 6.8 Hz, OCH$_2$); δ_C (125 MHz; CDCl$_3$) 165.1, 151.4, 144.5, 142.5, 134.7, 129.6, 128.1, 126.9, 125.3, 121.6, 115.7, 68.3, 67.0, 52.8, 33.7; HRMS Found: 522.1339; C$_{29}$H$_{25}$NO$_5$S [M+Na]$^+$ requires 522.1346.

Methyl 3-[2-(methoxymethoxy)ethoxy]-4-nitrobenzoate, 91
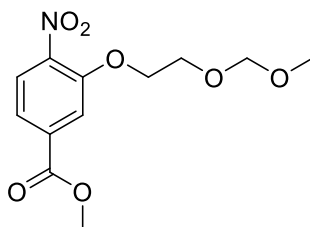
1-Bromo-2-(methoxymethoxy)ethane (0.41 mL, 3.55 mmol) was added to a solution of methyl 3-hydroxy-4-nitrobenzoate (500 mg, 2.54 mmol) and potassium carbonate (1.7 g, 12.7 mmol) in DMF (10 mL) and heated to 50 °C for 24 hours, allowed to cool and partitioned between EtOAc (20 mL) and H$_2$O (20 mL). The organic layer was separated and washed with H$_2$O (20 mL) and brine (20 mL), dried over $MgSO_4$, concentrated *in vacuo* to give a yellow solid and purified by column chromatography eluting with CH_2Cl_2 to give the desired product as a pale yellow solid (209 mg, 29%). m.p. 58-60 °C (CH_2Cl_2); R_f 0.3 (CH_2Cl_2); v_{max}/cm^{-1} (Solid state) 2966, 2942, 2891, 1726; δ_H (500 MHz; CDCl$_3$) 7.83 (1H, d, *J* = 8.2 Hz, Ar-H), 7.77 (1H, s, Ar-H), 7.69 (1H, d, *J* = 8.2 Hz, Ar-H), 4.70 (2H, s, CH$_2$), 4.35 (2H, t, *J* = 4.6 Hz, CH$_2$), 3.96 (3H, s, CO$_2$Me), 3.94 (2H, t, *J* = 4.6 Hz, CH$_2$), 3.38 (3H, s, OMe); δ_C (125 MHz; CDCl$_3$) 165.1, 151.7, 142.6, 134.8, 125.3, 121.6, 115.7, 96.6, 69.4, 65.3, 55.2, 52.8; HRMS Found: 308.0746; C$_{12}$H$_{15}$NO$_7$ [M+Na]$^+$ requires 308.0741.

Methyl 3-[(triphenylmethyl)carbamoyl]methoxy-4-nitrobenzoate, 92
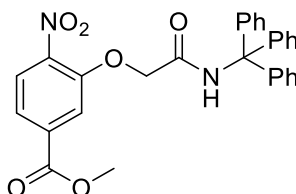
2-Bromo-*N*-(triphenylmethyl)-acetamide (600 mg, 1.58 mmol) was added to a solution of methyl 3-h ydroxy-4-nitrobenzoate (258mg, 1.31 mmol) and potassium carbonate (900 mg, 6.58 mmol) in DMF (10 mL) and heated to 50 °C for 24 hours then allowed to cool and partitioned between EtOAc (20 mL) and H_2O (20 mL). The organic layer was separated and washed with saturated aqueous solution of sodium bicarbonate (20ml), H_2O (20 ml) and brine (20 ml), dried over $MgSO_4$, concentrated *in vacuo* to an orange solid which was recrystallised from methanol to give a white solid (155.4 mg, 23%), m.p. 205 °C (MeOH), v_{max}/cm^{-1} (film) 3407, 3058, 2950, 2512, 2159, 1976, 1728; δH (500 MHz; $CDCl_3$) 8.14 (1H, s (broad), N-H), 8.02 (1H, d, *J* = 8.1 Hz, Ar-H), 7.80 (1H, dd, *J* = 8.5 Hz, 1.7, Ar-H), 7.74 (1H, d, *J* = 1.4 Hz, Ar-H), 7.34-7.26 (15H, m, trityl), 4.68 (2H, s (broad), CH_2), 3.98 (3H, s (broad), OMe); δC (125 MHz; $CDCl_3$) 164.9, 164.6, 150.2, 144.2, 135.8, 128.6, 128.0, 127.2, 126.4, 122.8, 155.5, 91.9, 70.6, 68.3, 53.0; HRMS Found: 519.1533; $C_{29}H_{24}N_2O_6$ [M+Na]$^+$ requires 519.1527.

Methyl 3-{2-[(triphenylmethyl)sulfanyl]ethoxy}-4-aminobenzoate, 93
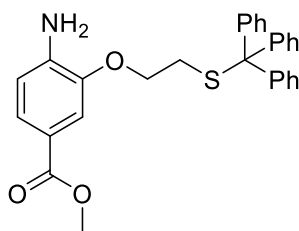
Prepared using procedure E from methyl 3-{2-[(triphenylmethyl)sulfanyl]ethoxy}-4-nitrobenzoate (220 mg, 0.44 mmol) and purified by alumina column chromatography eluting with 1% ethyl acetate–DCM followed by 1% MeOH–DCM to give a pale yellow oil (101 mg, 49%), v_{max}/cm^{-1} (film) 3486, 3374, 3057, 2948, 1704, 1615; δH (300 MHz; $CDCl_3$) 7.52 (1H, dd, *J* = 8.2, 1.6 Hz, Ar-H), 7.45 (6H, m, Trityl and Ar-H), 7.25 (10H, m, Trityl), 6.63 (1H, d, *J* = 8.2 Hz, Ar-H), 3.83 (5H, m, OMe and CH_2), 2.65 (2H, t, J = 6.3 Hz, CH_2); δC (75 MHz; $CDCl_3$) 144.8, 144.6, 141.4, 129.5, 128.0, 126.8, 124.4, 119.3, 113.2, 112.6, 66.99, 51.72, 31.54; HRMS Found: 492.1603; $C_{29}H_{27}NO_3S$ [M+Na]$^+$ requires 492.1604.

Methyl 3-[2-(methoxymethoxy)ethoxy]-4-aminobenzoate, 94
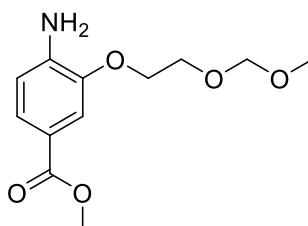
Prepared using procedure E from methyl 3-[2-(methoxymethoxy)ethoxy]-4-nitrobenzoate (416 mg, 1.46 mmol) to give a pale yellow oil (369 mg, 99%),

v_{max}/cm^{-1} (film) 3480, 3365, 2949, 2887, 1704; δ_H (500 MHz; CDCl$_3$) 7.56 (1H, d, J = 8.2 Hz, Ar-H), 7.48 (1H, s, Ar-H), 6.67 (1H, d, J = 7.8 Hz, Ar-H), 4.71 (2H, s, OCH$_2$O), 4.38 (2H, s (broad), NH$_2$), 4.22 (2H, t, J = 4.4 Hz, CH$_2$), 3.92 (2H, t, J = 4.4 Hz, CH$_2$), 3.86 (3H, s, OMe), 3.39 (3H, s, OMe); δ_C (125 MHz; CDCl$_3$) 167.2, 145.0, 141.8, 124.5, 119.2, 113.3, 113.0, 96.5, 68.1, 66.1, 55.3, 51.6; HRMS Found: 278.0999; C$_{12}$H$_{17}$NO$_5$ [M+Na]$^+$ requires 278.0999.

Methyl 3-[(triphenylmethyl)carbamoyl]methoxy-4-aminobenzoate, 95
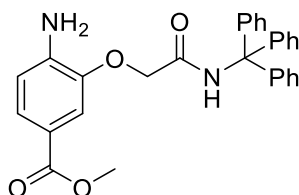
Prepared using procedure E from methyl 3-[(triphenylmethyl)carbamoyl]methoxy-4-nitrobenzoate (220 mg, 0.44 mmol) to give a beige solid (153 mg, 65%), m.p. 132-134 °C, v_{max}/cm^{-1} (film) 3338, 3056, 2949, 2581, 1966, 1671; δ_H (500 MHz; d_6-DMSO) 8.76 (1H, s, NH), 7.42 (1H, dd, J = 8.2, 1.8 Hz, Ar-H), 7.36 (1H, d, J = 1.4 Hz, Ar-H), 7.28-7.17 (15H, m, trityl), 6.68 (1H, d, J = 8.2 Hz, Ar-H), 5.72 (2H, s (broad), NH$_2$), 4.75 (2H, s, CH$_2$), 3.79 (3H, s, OMe); δ_C (125 MHz; d_6-DMSO) 167.1, 166.2, 144.4, 143.8, 128.4, 127.7, 127.5, 126.5, 124.3, 112.9, 112.4, 69.2, 61.8, 51.3, 48.5; HRMS Found: 489.1790, C$_{29}$H$_{26}$N$_2$O$_4$ [M+Na]$^+$ requires 489.1785.

4-Nitro-3-{2-[(triphenylmethyl)sulfanyl]ethoxy}benzoic acid, 96
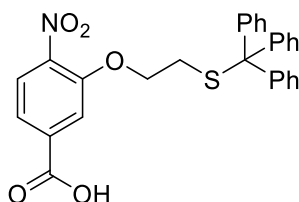
Prepared using procedure C from methyl 4-Nitro-3-{2-[(triphenylmethyl)sulfanyl]ethoxy} benzoate (220 mg, 0.44 mmol) to give a colourless solid (195 mg, 91%), m.p. 194-196 °C v_{max}/cm^{-1} (film) 3057, 2654, 1693; δ_H (500 MHz; CDCl$_3$) 7.94 (1H, d, J = 8.2 Hz, Ar-H), 7.65 (1H, dd, J = 8.2 Hz, 1.4, Ar-H), 7.56 (1H, s, Ar-H), 7.37-7.24 (15H, m, trityl), 3.99 (2H, t, J = 6.2 Hz, CH$_2$), 2.54 (2H, t, J = 6.2 Hz, CH$_2$); δ_C (125 MHz; CDCl$_3$) 165.6, 150.1, 144.1, 142.1, 129.0, 128.0, 126.8, 124.9, 121.6, 115.4, 67.56, 66.19, 30.64; HRMS Found: 508.119; C$_{28}$H$_{23}$NO$_5$S [M+Na]$^+$ requires 508.1189.

3-[2-(Methoxymethoxy)ethoxy]-4-nitrobenzoic acid, 97
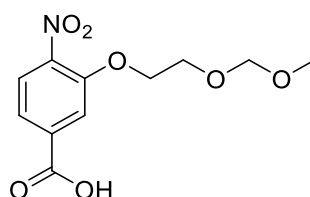
Prepared using procedure C from methyl 3-[2-(methoxymethoxy)ethoxy]-4-nitrobenzoate (400 mg, 1.4 mmol) to give a colourless solid, (262.6 mg, 69%), m.p. 131-133 °C; ν_{max}/cm^{-1} (film) 3063, 2938, 2542, 2159, 2025, 1691; δ_H (500 MHz; CDCl$_3$) 7.90-7.87 (2H, m, Ar-H), 7.81 (1H, dd, J = 8.2, 1.4 Hz, Ar-H), 4.76 (2H, s, OCH$_2$O), 4.41 (2H, t, J = 4.4 Hz, CH$_2$), 4.00 (2H, t, J = 4.4 Hz, CH$_2$), 3.44 (3H, s, OMe); δ_C (125 MHz; CDCl$_3$) 169.4, 151.7, 143.3, 133.8, 125.4, 122.4, 116.3, 96.6, 69.5, 65.4, 55.3; HRMS Found: 294.0571, C$_{11}$H$_{13}$NO$_7$ [M+Na]$^+$ requires 294.0584;

3-[(Triphenylmethyl)carbamoyl]methoxy-4-nitrobenzoic acid, 98
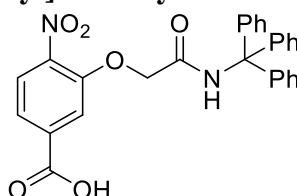
Prepared using procedure C from methyl 3-[(triphenylmethyl)carbamoyl]methoxy-4-nitrobenzoate (220mg, 0.44 mmol) to give an off white solid (157mg, 74%); m.p. 222-224 °C; ν_{max}/cm^{-1} (film) 3398, 3059, 2828, 2581, 1961, 1699; δ_H (500 MHz; d_6-Acetone) 8.21 (1H, s (broad), NH), 8.07 (1H, d, J = 8.2 Hz, Ar-H), 7.93 (1H, d, J = 1.4 Hz, Ar-H), 7.84 (1H, dd, J = 8.5, 1.6 Hz, Ar-H), 7.32-7.25 (15H, m, trityl), 4.98 (2H, s, CH$_2$); δ_C (125 MHz; CDCl3) 166.1, 165.7, 151.4, 145.7, 143.1, 136.5, 129.6, 128.5, 127.7, 126.4, 123.3, 117.0, 70.92, 69.22, 49.6; HRMS Found: 505.1362; C$_{28}$H$_{22}$N$_2$O$_6$ [M+Na]$^+$ requires 505.137.

Methy 3-Hydroxy-4-aminobenzoate, 99
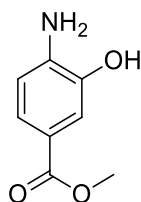
4-amino-3-hydroxybenzoic acid (1 g, 6 mmol) was heated to reflux in methanol (20 ml) and sulphuric acid (2 ml) for 16 hours, allowed to cool to r.t. and sodium bicarbonate added until gas evolution ceased. The resulting solid was removed by filtration and the filtrate concentrated *in vacuo*, suspended in water (20 ml) and extracted with ethyl acetate (3 × 30 ml). The combined organics were washed with brine (20 ml), dried over MgSO$_4$ and concentrated *in vacuo* to yield the *title*

compound as a beige solid (854 mg, 85%). v_{max}/cm^{-1} (solid state) 3398, 3315, 2948, 2583, 1705, 1604; δ_H (500 MHz; d_6-DMSO) 9.40 (1H, s, OH), 7.26 (2H, m, Ar-H), 6.60 (1H, d, J = 8.7 Hz, Ar-H), 5.36 (2H, s, NH), 3.74 (3H, s, OMe); δ_C (125 MHz; d_6-DMSO) 166.4, 142.7, 142.3, 122.5, 116.5, 116.2, 114.5, 112.4, 51.1; HRMS *m/z* (ESI) Found: 190.0493, $C_8H_8NO_3$ [M+Na]$^+$ requires 190.0475.

O$_2$N-[*O*-MOM-Hydroxyethyl(3-HABA)]-[(3-HABA)]-COOMe, 100
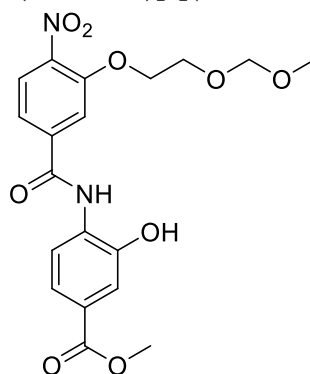
PyBOP (1664 mg, 3.2 mmol) and diisopropylethylamine (1.2 ml, 6.67 mmol) were added sequentially to a solution of 3-[2-(Methoxymethoxy)ethoxy]-4-nitrobenzoic acid (725 mg, 2.67 mmol) in dichloromethane (30 ml) and the reaction stirred for 30 minutes. After which time methy 3-hydroxy-4-aminobenzoate (446 mg, 2.67 mmol) was added and the reaction stirred at r.t. overnight. The reaction mixture was concentrated *in vacuo* and re-dissolved in dimethylformamide (20 ml). Caesium carbonate (3.8 g, 13.55 mmol) was added and the reaction stirred at r.t. for overnight. The reaction mixture was then poured into ethyl acetate (40 ml) and washed copiously with water and brine, dried over MgSO$_4$ and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 5:95 ether/dichloromethane to give the *title compound* as a yellow solid (491 mg, 44%).. v_{max}/cm^{-1} (solid state) 2952, 2869, 1716, 1674, 1524; ^{1}H NMR (500 MHz, d_6-DMSO) δ 10.39 (s, 1H, OH), 9.84 (s, 1H, Amide NH), 8.02 (d, J = 8.3 Hz, 1H, Ar-H), 7.92 (d, J = 8.3 Hz, 1H, Ar-H), 7.89 (s, 1H, Ar-H), 7.66 (dd, J = 8.3, 1.4 Hz, 1H), 7.55 (d, J = 1.9 Hz, 1H, Ar-H), 7.50 (dd, J = 8.3, 1.9 Hz, 1H, Ar-H), 4.64 (s, 2H, O-C*H$_2$*-O), 4.51 – 4.39 (t, 2H, CH$_2$), 3.92 – 3.80 (m, 5H, OMe and CH$_2$), 3.28 (s, 3H, OMe); HRMS found 421.1245, $C_{19}H_{20}N_2O_9$ [M+H]$^+$ requires 421.1241

O$_2$N-[*O*-MOM-hydroxyethyl(3-HABA)]-[*O*-(N-Trt)carbamoylmethoxy-(3-HABA)]-COOMe, 101
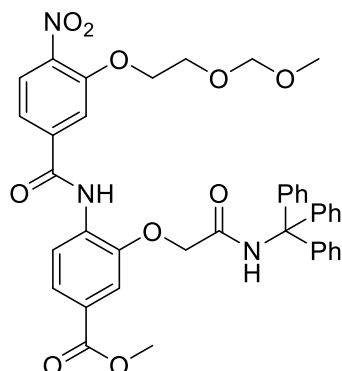
2-Bromo-*N*-(triphenylmethyl)-acetamide (217 mg, 0.57 mmol) was added to a suspension of methyl 3-hydroxy-4-{3-[2-(methoxymethoxy)ethoxy]-4-nitrobenzamido}benzoate (200 mg, 0.48 mmol) and potassium carbonate (331 mg, 2.4 mmol) in dimethylformamide (20 ml) and the reaction heated to 50 °C overnight. The reaction mixture was then poured into water (20 ml) and the resulting precipitate collected by filtration. The precipitate was purified by column chromatography eluting with 5:95 ether/dichloromethane to give the *title compound* as a pale yellow solid (138 mg, 40%). v_{max}/cm^{-1} (solid state) 3381, 2964, 2821, 1969; ^{1}H NMR (500 MHz, CDCl$_3$) δ 9.49 (s, 1H, Amide N-H), 8.53 (d, *J* = 8.5 Hz, 1H, Ar-H), 7.86 (d, *J* = 8.5 Hz, 1H, Ar-H), 7.70 (d, *J* = 8.8 Hz, 1H, Ar-H), 7.63 (s, 1H, Ar-H), 7.44 (d, *J* = 8.4 Hz, 1H, Ar-H), 7.27 (m, *J* = 11.9, 6.7 Hz, 10H, Trityl), 7.20 (d, *J* = 7.1 Hz, 1H, Ar-H), 7.14 − 7.08 (m, 5H, Trityl), 4.68 (s, 4H, O-CH$_2$-O and O-CH$_2$-CONR), 4.21 − 4.13 (t, *J* = 4.2 Hz 2H, CH$_2$), 3.92 (s, 3H, OMe), 3.76 − 3.71 (t, *J* = 4.2 Hz, 2H, CH$_2$), 3.40 (s, 3H, OMe); ^{13}C NMR (125 MHz, CDCl$_3$) δ 167.50, 166.14, 163.51, 152.20, 147.51, 143.83, 141.61, 139.23, 133.67, 128.70, 128.62, 128.50, 128.16, 128.03, 127.48, 126.24, 125.72, 120.71, 119.05, 117.13, 114.03, 96.57, 71.16, 71.07, 69.40, 65.25, 55.27, 52.28; HRMS found 720.2555, C$_{40}$H$_{37}$N$_3$O$_{10}$ requires [M+H]$^+$ 720.2551

H₂N-[*O*-MOM-hydroxyethyl(3-HABA)]-[*O*-(N-Trt)carbamoylmethoxy-(3-HABA)]-COOMe
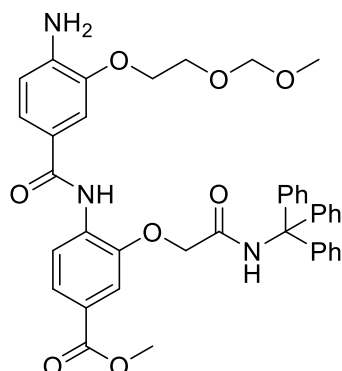
Using standard procedure E on 0.19 mmol scale to give the title compound as a colourless oil (74 mg, 56%). v_{max}/cm^{-1} (solid state) 3343, 2950, 1673; ^{1}H NMR (500 MHz, CDCl₃) δ 8.57 (s, 1H, Amide N-H), 8.49 (d, *J* = 8.5 Hz, 1H, Ar-H), 7.73 (dd, *J* = 8.5, 1.1 Hz, 1H, Ar-H), 7.57 (s, 1H, Ar-H), 7.36 (d, *J* = 1.1, 1H, Ar-H), 7.25 (s, 1H, Ar-H), 7.15 (dd, *J* = 7.4, 2.8 Hz, 10H, Trityl), 7.05 (dd, *J* = 6.7, 2.7 Hz, 5H, Trityl), 6.45 (d, *J* = 8.1 Hz, 1H, Ar-H), 4.62 (s, 2H, O-C*H*₂-O), 4.61 (s, 2H, O-C*H*₂-CON), 4.20 (s, 2H, NH₂), 4.14 – 4.07 (t, *J* = 4.4 Hz, 2H, CH₂), 3.83 (s, 3H, OMe), 3.79 – 3.74 (t, *J* = 4.4 Hz, 2H, CH₂), 3.32 (s, 3H, OMe); ^{13}C NMR (125 MHz, CDCl₃) δ 166.36, 166.26, 165.19, 145.87, 145.79, 144.08, 141.14, 133.17, 128.46, 128.13, 127.31, 125.28, 125.00, 123.46, 120.81, 119.83, 113.51, 113.44, 111.73, 104.21, 96.64, 70.64, 69.46, 68.28, 66.14, 63.78, 55.33, 52.20.

H₂N-[*O*-hydroxyethyl(3-HABA)]-[*O*-carbamoylmethoxy-(3-HABA)]-COOH, 102
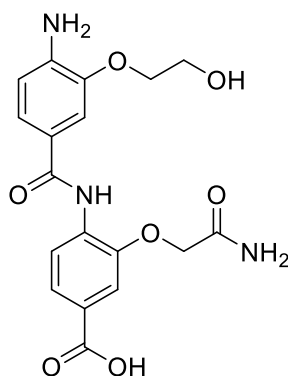
H₂N-[*O*-MOM-hydroxyethyl(3-HABA)]-[*O*-(N-Trt)carbamoylmethoxy-(3-HABA)]-COOMe (35 mg, 0.05 mmol) was dissolved in 1:1 tetrahydrofuran/methanol (5 ml) and treated with 1M sodium hydroxide solution (3 ml) for 2 hours. The reaction mixture was then acidified to pH 1 with concentrated HCl and stirred overnight. The reaction mixture was concentrated *in vacuo* and purified by mass directed preparative HPLC. v_{max}/cm^{-1} (solid state) 3408, 1663; ^{1}H NMR (500 MHz, d_6-DMSO) δ 9.67 (s, 1H, Amide N-H), 8.13 (d, *J* = 8.4 Hz, 1H,

Ar-H), 7.62 (dd, J = 8.4, 1.8 Hz, 1H, Ar-H), 7.51 (d, J = 1.8 Hz, 1H, Ar-H), 7.45 (dd, J = 8.3, 1.9 Hz, 1H, Ar-H), 7.41 (d, J = 1.9 Hz, 1H, Ar-H), 6.67 (d, J = 8.2 Hz, 1H, Ar-H), 4.62 (s, 2H, CH$_2$), 4.01 (t, J = 4.8 Hz, 2H, CH$_2$), 3.75 (t, J = 4.7 Hz, 2H, CH$_2$); HRMS found 390.1301 C$_{18}$H$_{19}$N$_3$O$_7$ requires [M+H]$^+$ 390.1295

Methyl 4-amino-3-(prop-2'-en-1'-yloxy)benzoate
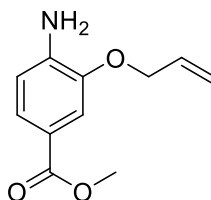
Prepared using standard procedure D on a 2.1 mmol scale (397 mg, 91%). v_{max}/cm^{-1} (solid state) 3367, 2991, 1692; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.55 (dd, J = 8.2, 1.8 Hz, 1H, 6-CH), 7.46 (d, J = 1.7 Hz, 1H, 2-CH), 6.67 (d, J = 8.2 Hz, 1H, 5-CH), 6.08 (ddt, J = 17.2, 10.6, 5.4 Hz, 1H, 2'-CH), 5.42 (dq, J = 17.3, 1.6 Hz, 1H, 3'-CH$_{trans}$)), 5.30 (dq, J = 10.5, 1.4 Hz, 1H, 3'-CH$_{cis}$)), 4.61 (dt, J = 5.4, 1.4 Hz, 2H, 1-CH$_2$), 3.86 (s, 3H, OMe); ^{13}C NMR (125 MHz, CDCl$_3$) δ 167.27, 145.01, 141.38, 133.04, 124.28, 119.46, 117.88, 113.33, 112.68, 69.30, 51.68; HRMS found 208.0969, C$_{11}$H$_{13}$NO$_3$ [M+H]$^+$ requires 208.0968

4-nitro-3-(prop-2'-en-1'-yloxy)benzoic acid
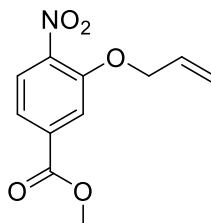
Prepared using standard procedure C on a 2.52 mmol scale (535 mg, 95%). v_{max}/cm^{-1} (solid state) 2824, 2598, 2538, 1683; ^{1}H NMR (500 MHz, d_6-DMSO) δ 13.60 (s, 1H, OH), 7.96 (d, J = 8.3 Hz, 1H, 5-CH), 7.75 (d, J = 1.5 Hz, 1H, 2-CH), 7.63 (dd, J = 8.3, 1.6 Hz, 1H, 6-CH), 6.02 (ddt, J = 17.2, 10.5, 4.9 Hz, 1H, 2'-CH), 5.41 (dq, J = 17.3, 1.7 Hz, 1H, 3'C-H$_{trans}$), 5.29 (dq, J = 10.6, 1.5 Hz, 1H, 3'C-H$_{cis}$), 4.83 (dt, J = 4.9, 1.6 Hz, 2H, 1'-CH$_2$); ^{13}C NMR (125 MHz, d_6-DMSO) δ 165.71, 150.32, 142.18, 135.57, 132.32, 124.98, 121.47, 117.88, 115.73, 69.59; HRMS m/z (ESI) Found: 222.041033, C$_{10}$H$_9$NO$_5$ [M-H]$^-$ Requires 222.040796.

5.4.5 Late Stage Functionalization

The allyl bearing trimer was prepared in parallel using the methods and building blocks described above, checking at pertinent times during the synthesis by crude NMR and LC-MS, to afford the below compounds. Compounds were either pure following final precipitation or purified by preparative HPLC.

O$_2$N-[*O*-Allyl(3-HABA)]-[*O-i*Bu (3-HABA)]-[*O-i*Pr (3-HABA)]-COOMe, 103
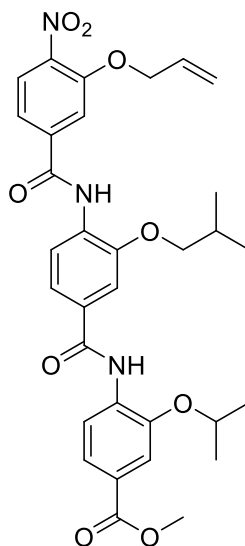
v_{max}/cm^{-1} (solid state) 3421, 2975, 1704, 1680; ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.85 (s, 1H, Amide N-H), 8.76 (s, 1H, Amide N-H), 8.63 (d, *J* = 8.4 Hz, 1H, Ar-H), 8.60 (d, *J* = 8.5 Hz, 1H, Ar-H), 7.93 (d, *J* = 8.3 Hz, 1H, Ar-H), 7.71 (dd, *J* = 8.5, 1.7 Hz, 1H, Ar-H), 7.69 (d, *J* = 1.5 Hz, 1H, Ar-H), 7.59 (m, 2H, Ar-H), 7.42 (m, 2H), 6.06 (ddt, *J* = 17.2, 10.2, 5.0 Hz, 1H, 1-C$_\beta$H), 5.52 (ddd, *J* = 17.1, 2.7, 1.5 Hz, 1H, 1-C$_\gamma$H$_{trans}$), 5.38 (ddd, *J* = 10.7, 2.5, 1.3 Hz, 1H1-C$_\gamma$H$_{cis}$), 4.83 − 4.67 (m, 3H, 1-C$_\alpha$H$_2$ and 3-C$_\alpha$H), 3.98 (d, *J* = 6.5 Hz, 2H, 2- C$_\alpha$H$_2$), 3.90 (s, 3H, OMe), 2.22 (dp, *J* = 13.3, 6.7 Hz, 1H, 2-C$_\beta$H), 1.46 (d, *J* = 6.0 Hz, 6H, 3-C$_\beta$H$_3$ and 3-C$_\gamma$H$_3$), 1.11 (d, *J* = 6.7 Hz, 6H, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 166.74, 164.22, 162.98, 152.15, 147.84, 145.79, 141.96, 139.64, 132.93, 131.13, 130.68, 130.64, 126.02, 125.11, 123.31, 119.01, 118.93, 118.59, 117.71, 114.49, 113.14, 110.62, 75.26, 71.88, 70.31, 52.09, 28.25, 22.23, 19.31; HRMS found 628.2276, C$_{32}$H$_{35}$N$_3$O$_9$ [M+Na]$^+$ requires 628.2265

O$_2$N-[*O-i*Bu(3-HABA)]-[*O*-Allyl (3-HABA)]-[*O-i*Pr (3-HABA)]-COOMe, 104
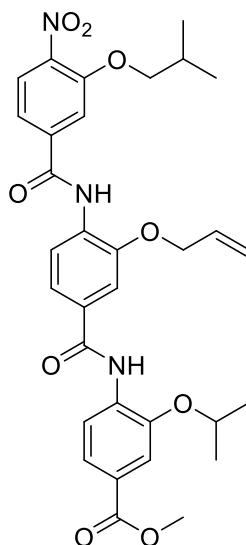
v_{max}/cm^{-1} (solid state) 3428, 2961, 1792, 1714; ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.81 (s, 1H, Amide N-H), 8.73 (s, 1H, Amide N-H), 8.61 (d, *J* = 8.4 Hz, 1H, Ar-H), 8.57 (d, *J* = 8.5 Hz, 1H, Ar-H), 7.88 (d, *J* = 8.3 Hz, 1H, Ar-H), 7.68 (dd, *J* = 8.5, 1.7 Hz, 1H, Ar-H), 7.63 (d, *J* = 1.6 Hz, 1H, Ar-H), 7.59 (d, *J* = 1.8 Hz, 1H, Ar-H), 7.57 (d, *J* = 1.7 Hz, 1H, Ar-H), 7.41 (dd, *J* = 8.5, 1.8 Hz, 1H, Ar-H), 7.38 (dd, *J* = 8.3, 1.7 Hz, 1H, Ar-H), 6.11 (ddt, *J* = 17.2, 10.6, 5.4 Hz, 1H, 2-C$_\beta$H), 5.47 (ddd, *J* = 17.3, 2.8, 1.5 Hz, 1H, 2-C$_\gamma$H$_{trans}$), 5.39 (ddd, *J* = 10.5, 2.3, 1.1 Hz, 1H, 2-C$_\gamma$H$_{cis}$), 4.86 − 4.68 (m, 3H, 2-C$_\alpha$H$_2$ and 3-C$_\alpha$H), 3.94 (d, *J* = 6.5 Hz, 2H, 1-C$_\alpha$H$_2$), 3.88 (s, 3H, OMe), 2.16 (dt, *J* = 13.3, 6.6 Hz, 1H, 1-C$_\beta$H), 1.45 (d, *J* = 6.1 Hz, 6H, 3-C$_\beta$H$_3$ and 3-C$_\gamma$H$_3$), 1.06 (d, *J* = 6.7 Hz, 6H, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 166.69, 164.09, 163.23, 152.75, 147.39, 145.78, 141.70, 139.55, 132.88, 132.03, 130.82, 130.54, 125.81, 125.10, 123.24, 119.22, 119.19, 119.05, 118.58, 117.47, 113.95, 113.12, 111.06, 76.07, 71.87, 69.82, 52.06, 28.20, 22.22, 19.00; HRMS found 606.2456, C$_{32}$H$_{35}$N$_3$O$_9$ [M+H]$^+$ requires 606.2446

O$_2$N-[*O-i*Bu(3-HABA)]-[*O-i*Bu (3-HABA)]-[*O*-Allyl(3-HABA)]-COOMe, 105
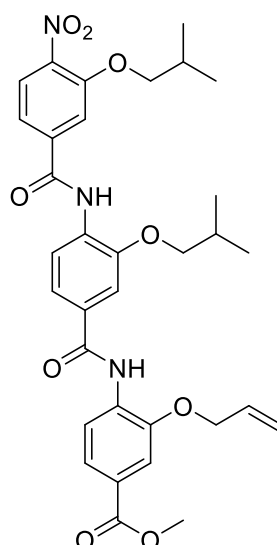
v_{max}/cm^{-1} (solid state) 3435, 2956, 1714, 1679; ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.83 (s, 1H, Amide N-H), 8.78 (s, 1H, Amide N-H), 8.66 (d, J = 8.4 Hz, 1H, Ar-H), 8.64 (d, J = 8.5 Hz, 1H, Ar-H), 7.94 (d, J = 8.3 Hz, 1H, Ar-H), 7.77 (dd, J = 8.5, 1.6 Hz, 1H, Ar-H), 7.66 (d, J = 1.4 Hz, 1H, Ar-H), 7.64 – 7.57 (m, 2H, Ar-H), 7.46 (dd, J = 8.4, 1.6 Hz, 1H, Ar-H), 7.42 (dd, J = 8.3, 1.5 Hz, 1H, Ar-H), 6.16 (ddt, J = 17.2, 10.6, 5.4 Hz, 1H, 3-C$_\beta$H), 5.52 (dd, J = 17.2, 1.2 Hz, 1H, 3-C$_\gamma$H$_{trans}$), 5.43 (dd, J = 10.5, 1.1 Hz, 1H, 3-C$_\gamma$H$_{cis}$), 4.75 (d, J = 5.4 Hz, 2H, 3-C$_\alpha$H$_2$), 4.04 – 3.95 (m, 4H, 1-C$_\alpha$H$_2$ and 2-C$_\alpha$H$_2$), 3.92 (s, 3H, OMe), 2.32 – 2.10 (m, 2H, 1-C$_\beta$H and 2-C$_\beta$H), 1.14 (d, J = 6.7 Hz, 6H, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.10 (d, J = 6.7 Hz, 6H, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 166.63, 164.33, 163.08, 152.82, 147.78, 146.53, 141.72, 139.61, 132.27, 132.25, 130.77, 130.44, 125.92, 125.12, 123.71, 119.07, 118.89, 118.76, 118.62, 117.39, 113.74, 112.05, 110.54, 76.08, 75.20, 69.76, 52.12, 28.28, 28.22, 19.36, 19.03; HRMS found 620.2609, C$_{33}$H$_{37}$N$_3$O$_9$ [M+H]$^+$ requires 620.2602

Standard Proceedure G – Removal of Allyl Groups

Allyl functionalised trimers were treated with palladium tetrakistriphenylphosphine (10 mol%) and sodium toluenesulfinate (1.2 eq) in THF (1ml/mg) overnight, concentrated *in vacuo* and purified by column chromatography eluting with Et$_2$O in DCM.

O₂N-[(3-HABA)]-[*O-i*Bu (3-HABA)]-[*O-i*Pr (3-HABA)]-COOMe, 106
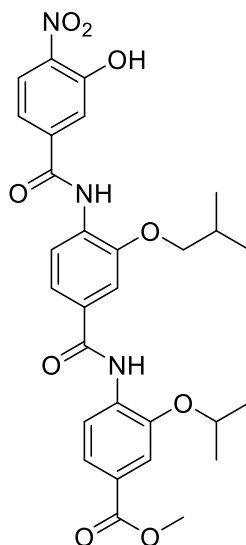
Prepared using standard procedure G on 0.58 mmol scale (267 mg, 81%). ^{1}H NMR (500 MHz, CDCl₃) δ 10.61 (s, 1H, OH), 8.87 (s, 1H, Amide NH), 8.77 (s, 1H, Amide NH), 8.64 (m, 2H, Ar-H), 8.28 (d, J = 8.7 Hz, 1H, Ar-H), 7.74 (dd, J = 8.5, 1.7 Hz, 1H, Ar-H), 7.68 (d, J = 1.9 Hz, 1H, Ar-H), 7.63 (d, J = 1.8 Hz, 1H, Ar-H), 7.61 (d, J = 1.7 Hz, 1H, Ar-H), 7.47 (dd, J = 8.8, 1.9 Hz, 1H, Ar-H), 7.44 (dd, J = 8.5, 1.8 Hz, 1H, Ar-H), 4.78 (hept, J = 6.2 Hz, 1H, 3-C$_α$H), 4.00 (d, J = 6.6 Hz, 2H, 2-C$_α$H₂), 3.92 (s, 3H, OMe), 2.24 (dp, J = 13.4, 6.6 Hz, 1H, 2-C$_β$H), 1.46 (d, J = 6.1 Hz, 6H3-C$_β$H₃ and 3-C$_γ$H₃), 1.12 (d, J = 6.7 Hz, 6H, 2-C$_γ$H₃ and 2-C$_δ$H₃); HRMS Found: 566.212; C₂₉H₃₁N₃O₉ [M+H]$^+$ requires 566.2133

O₂N-[*O-i*Bu(3-HABA)]-[(3-HABA)]-[*O-i*Pr(3-HABA)]-COOMe
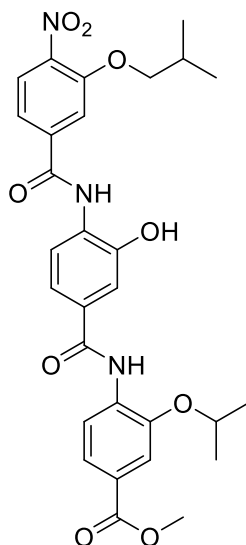
Prepared using standard procedure G on a 0.4 mmol scale (168 mg, 75%). ^{1}H NMR (500 MHz, d_6-DMSO) δ 10.34 (s, 1H, OH), 9.86 (s, 1H, Amide NH), 9.22 (s, 1H, Amide NH), 8.25 (d, J = 8.4 Hz, 1H, Ar-H), 8.00 (d, J = 8.3 Hz, 1H, Ar-H), 7.88 (d, J = 8.3 Hz, 1H, Ar-H), 7.83 (d, J = 1.5 Hz, 1H, Ar-H), 7.66 − 7.60 (m, 2H, Ar-H),

7.58 (d, J = 1.7 Hz, 1H, Ar-H), 7.49 (d, J = 2.0 Hz, 1H, Ar-H), 7.44 (dd, J = 8.3, 2.0 Hz, 1H, Ar-H), 4.82 – 4.69 (m, 1H, 3-C$_\alpha$H), 4.05 (d, J = 6.4 Hz, 2H, 1-C$_\alpha$H$_2$), 3.33 (s, 3H, OMe) 2.06 (sept, 6.6 Hz, 1H, 1-C$_\beta$H), 1.36 (d, J = 6.0 Hz, 6H, 3-C$_\beta$H$_3$ and 3-C$_\gamma$H$_3$), 0.99 (d, J = 6.7 Hz, 6H, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$); HRMS Found 566.2126; C$_{29}$H$_{31}$N$_3$O$_9$ [M+H]$^+$ requires 566.2133.

O$_2$N-[*O-i*Bu(3-HABA)] *O-i*Bu(3-HABA)]-[(3-HABA)]-COOMe
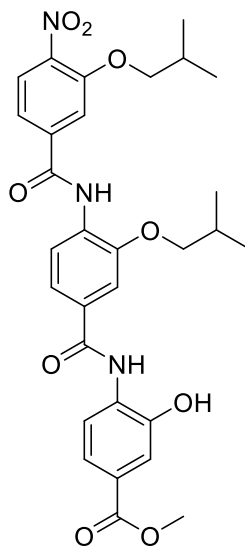
Prepared using standard procedure G on a 1 mmol scale (442 mg, 74%). ^{1}H NMR (500 MHz, d_6-DMSO) δ 10.37 (s, 1H, OH), 9.81 (s, 1H, Amide NH), 9.54 (s, 1H, Amide NH), 8.02 (d, J = 8.3 Hz, 1H, Ar-H), 7.96 (app.t, J = 8.0 Hz, 2H, Ar-H), 7.80 (d, J = 1.5 Hz, 1H, Ar-H), 7.66 (d, J = 1.8 Hz, 1H, Ar-H), 7.64 – 7.59 (m, 2H, Ar-H), 7.53 (d, J = 2.0 Hz, 1H, Ar-H), 7.48 (dd, J = 8.3, 1.9 Hz, 1H, Ar-H), 4.03 (d, J = 6.5 Hz, 2H, 1-C$_\alpha$H$_2$), 3.93 (d, J = 6.4 Hz, 2H, 2-C$_\alpha$H$_2$), 3.83 (s, 3H, OMe), 2.17 – 1.97 (m, 2H, 1-C$_\beta$H and 2-C$_\beta$H), 1.01 (d, J = 6.8 Hz, 6H, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 0.99 (d, J = 6.8 Hz, 6H, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); HRMS Found 580.2288; C$_{30}$H$_{33}$N$_3$O$_9$ [M+H]$^+$ requires 580.2289.

Standard Procedure H – Side chain introduction by Mitsunobu reaction

De-allylated nitro ester trimers were dissolved in THF (1 ml/10 ml substrate) and PPh$_3$ (2 eq.), alcohol (2 eq.) and DIAD (2 eq.) added sequentially. The reaction was stirred overnight at r.t., concentrated *in vacuo*. Amine products were isolated by SCX column and used without further purification.

H₂N-[(3-HABA)]-[*O*--*ᶦ*Bu(3-HABA)]-[*O*-*ᶦ*Pr(3-HABA)]-COOH, 110
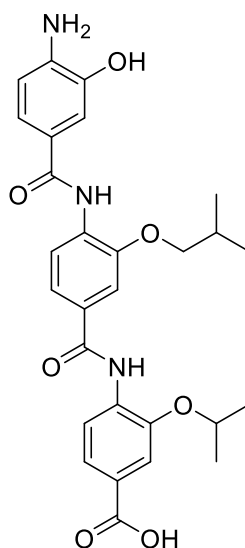
Prepared by standard procedures E and C and purification by mass directed HPLC (14.4 mg, 62%). ¹H NMR (500 MHz, d_6-DMSO) δ 9.30 (s, 1H, Amide NH), 8.86 (s, 1H, Amide NH), 8.28 (d, J = 8.2 Hz, 1H, Ar-H), 8.16 (d, J = 8.3 Hz, 1H, Ar-H), 7.61 − 7.53 (m, 4H), 7.27 (d, J = 2.1 Hz, 1H, Ar-H), 7.19 (dd, J = 8.3, 2.1 Hz, 1H, Ar-H), 6.64 (d, J = 8.2 Hz, 1H, Ar-H), 4.76 − 4.67 (m, 1H, 3-C_αH), 3.97 (d, J = 6.6 Hz, 2H, 2-C_αH₂), 1.90 − 1.79 (m, 1H, 2-C_βH), 1.35 (d, J = 6.0 Hz, 6H, 3-C_βH₃ and 3-C_γH₃), 1.03 (d, J = 6.7 Hz, 6H, 2-C_γH₃ and 2-C_δH₃); HRMS found 522.2244 $C_{28}H_{31}N_3O_7$ [M+H]⁺ requires 522.2234

H₂N-[2-(pyrrolidin-1-yl)ethoxy(3-HABA)]-[*O*--*ᶦ*Bu(3-HABA)]-[*O*—*ᶦ*Pr(3-HABA)]-COOH, 111
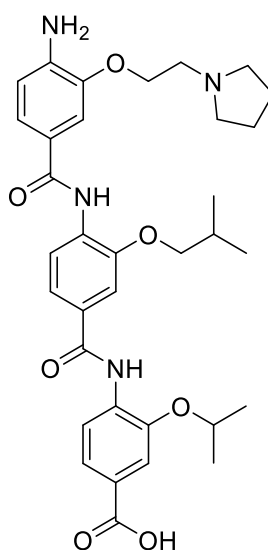
Prepared by standard procedure H followed by standard procedures E and C and purification by mass directed HPLC (7.8 mg, 36%). ¹H NMR (500 MHz, d_6-DMSO)

δ 9.16 (s, 1H, Amide NH), 9.00 (s, 1H, Amide NH), 8.50 (s, 1H, Ar-H), 8.20 (d, J = 8.4 Hz, 1H, Ar-H), 7.89 (d, J = 7.8 Hz, 1H, Ar-H), 7.56 (m, 3H, Ar-H), 7.46 (s, 1H, Ar-H), 7.37 (d, J = 7.2 Hz, 1H, Ar-H), 6.70 (d, J = 8.6 Hz, 1H, Ar-H), 4.61 (dt, J = 12.1, 6.1 Hz, 1H, 3-C_αH), 4.10 (t, J = 6.0 Hz, 2H, 1-C_βH^2), 3.95 (d, J = 6.5 Hz, 2H, 2-C_βH$_2$), 2.83 (t, J = 6.0 Hz, 1H, 1-C_αH$_2$), 2.58 – 2.51 (m, 4H, pyrrolidine 2 × CH$_2$) 2.14 (dt, J = 13.1, 6.5 Hz, 1H, 2-C_βH), 1.73 – 1.65 (m, 4H, pyrrolidine 2 × CH$_2$) 1.31 (d, J = 6.0 Hz, 6H, 3-C_βH$_3$ and 3-C_γH$_3$), 1.25 – 1.09 (m, 5H), 1.08 – 0.97 (d, J = 6.7 Hz, 6H, 2-C_γH$_3$ and 2-C_δH$_3$). HRMS Found 619.3136 C$_{34}$H$_{42}$N$_4$O$_7$ [M+H]$^+$ requires 619.3126

H$_2$N-[O-^{i}Bu(3-HABA)]-[(3-HABA)]-[O-^{i}Pr(3-HABA)]-COOH, 112
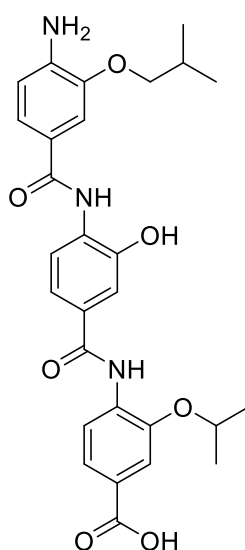
Prepared by standard procedures E and C and purification by mass directed HPLC (11.8 mg, 51%). ^{1}H NMR (500 MHz, d_6-DMSO) δ 9.26 (s, 1H, Amide NH), 9.15 (s, 1H, Amide NH), 8.22 (d, J = 8.3 Hz, 1H, Ar-H), 7.99 (d, J = 8.4 Hz, 1H), 7.63 – 7.52 (m, 4H) 7.47 (d, J = 2.0 Hz, 1H), 7.42 (dd, J = 8.1, 1.7 Hz, 1H), 6.70 (d, J = 8.0 Hz, 1H), 4.77 – 4.69 (m, 1H, 3-C_αH), 3.81 (d, J = 6.5 Hz, 2H, 1-C_αH$_2$), 2.13 – 2.02 (m, 1H, 1-C_βH), 1.36 (d, J = 6.0 Hz, 6H, 3-C_βH$_3$ and 3-C_γH$_3$), 1.02 (d, J = 6.7 Hz, 6H, 1-C_γH$_3$ and 1-C_δH$_3$); HRMS found 522.2241 C$_{28}$H$_{31}$N$_3$O$_7$ [M+H]$^+$ requires 522.2234

H₂N-[*O*--ⁱBu(3-HABA)]-[3-(dimethylamino)propoxy (3-HABA)]-[*O*-ⁱPr(3-HABA)]-COOH, 113

Prepared by standard procedure H followed by standard procedures E and C and purification by mass directed HPLC (11 mg, 49%). ^{1}H NMR (500 MHz, d_6-DMSO) δ 9.28 (s, 1H, Amide N-H), 9.04 (s, 1H, Amide N-H), 8.22 (d, J = 8.3 Hz, 1H, Ar-H), 8.11 (d, J = 8.5 Hz, 1H, Ar-H), 7.64 − 7.55 (m, 5H, Ar-H), 7.37 − 7.33 (m, 1H, Ar-H) 6.70 (d, J = 8.0 Hz, 1H, Ar-H), 4.74 − 4.66 (m, 1H, 3-C$_α$H), 4.20 (t, J = 6.2 Hz, 2H, 2-C$_α$H$_2$), 3.80 (t, J = 5.4 Hz, 2H, 2-C$_γ$H$_2$), 2.41 (m, 2H, 2-C$_β$H$_2$), 2.14 (s, 6H, 2 × NMe), 1.97 (m, 1H, 1-C$_β$H) 1.35 (dd, J = 6.0, 1.8 Hz, 6H, 3-C$_β$H$_3$ and 3-C$_γ$H$_3$), 1.02 (dd, J = 6.7, 3.1 Hz, 6H, 1-C$_γ$H$_3$ and 1-C$_δ$H$_3$); HRMS Found 607.3138, C$_{33}$H$_{42}$N$_4$O$_7$ [M+H]$^{+}$ requires 607.3126.

H₂N-[*O*--ⁱBu(3-HABA)]-[2-(pyrrolidin-1-yl)ethoxy (3-HABA)] -[*O*-ⁱPr(3-HABA)]-COOH, 114

Prepared by standard procedure H followed by standard procedures E and C and purification by mass directed HPLC (6.2 mg, 29%). ^{1}H NMR (500 MHz, d_6-DMSO)

δ 9.28 (s, 1H, Amide NH), 9.26 (s, 1H, Amide NH), 8.39 – 8.25 (m, 2H, Ar-H), 8.08 (d, J = 9.6 Hz, 1H, Ar-H), 7.72 – 7.50 (m, 5H, Ar-H), 7.33 (m, 1H, Ar-H), 4.74 – 4.63 (m, 1H, 3-C_αH), 4.29 (m, 2H, 2-C_βH$_2$), 3.80 (d, J =6.5 Hz, 2H, 1-C_αH$_2$), 2.84 (m, 2H, 2-C_αH$_2$), 2.09 (m, 1H, 1-C_βH), 1.60 – 1.50 (m, 4H, pyrrolidine 2 × CH$_2$), 1.35 (d, J = 6.0 Hz, 6H, 3-C_βH$_3$ and 3-C_γH$_3$), 1.03 (d, J = 6.7 Hz, 6H, 1-C_γH$_3$ and 1-C_δH$_3$); HRMS Found 619.3142, C$_{34}$H$_{42}$N$_4$O$_7$ [M+H]$^+$ requires 619.3126j

H$_2$N-[O-^{i}Bu(3-HABA)]-[O-^{i}Bu(3-HABA)]-[(3-HABA)]-COOH, 115
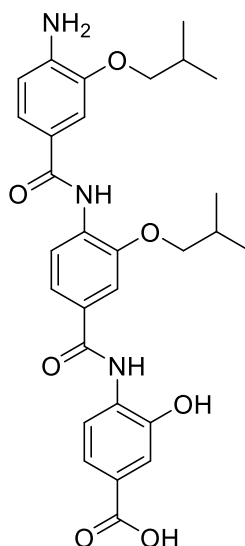
Prepared by standard procedures E and C and purification by mass directed HPLC (5.5 mg, 20%). ^{1}H NMR (500 MHz, d_6-DMSO) δ 9.47 (s, 1H, Amide NH), 8.99 (s, 1H, Amide NH), 8.24 (d, J = 8.3 Hz, 1H, Ar-H), 7.92 (d, J = 8.3 Hz, 1H, Ar-H), 7.64 – 7.58 (m, 2H, Ar-H), 7.50 (d, J = 1.9 Hz, 1H, Ar-H), 7.45 (dd, J = 8.3, 1.9 Hz, 1H, Ar-H), 7.35 (dd, J = 8.2, 1.9 Hz, 1H, Ar-H), 7.31 (d, J = 1.9 Hz, 1H, Ar-H), 6.70 (d, J = 8.2 Hz, 1H, Ar-H), 3.96 (d, J = 6.3 Hz, 2H, 1-C_αH$_2$), 3.79 (d, J = 6.5 Hz, 2H, 2-C_αH$_2$), 2.18 – 2.01 (m, 2H, 1-C_βH and 2-C_βH), 1.05 (d, J = 6.7 Hz, 6H, 1-C_γH$_3$ and 1-C_δH$_3$), 1.02 (d, J = 6.7 Hz, 6H, 2-C_γH$_3$ and 2-C_δH$_3$); HRMS found 536.23994 C$_{29}$H$_{33}$N$_3$O$_7$ [M+H]$^+$ requires 536.2391

H₂N-[*O*-*ⁱ*Bu (3-HABA)]- [*O*--*ⁱ*Bu(3-HABA)]- [3-(dimethylamino)propoxy (3-HABA)]-COOH, 116
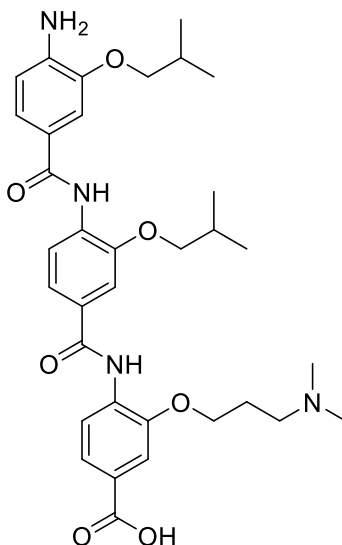
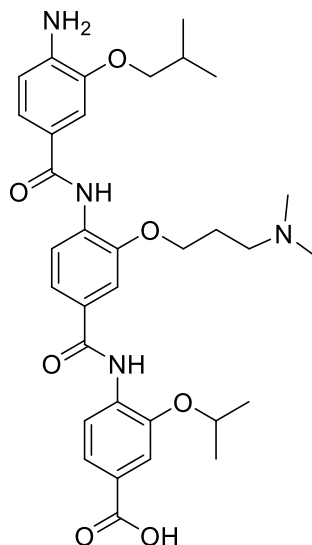Prepared by standard procedure H followed by standard procedures E and C and purification by mass directed HPLC (15 mg, 70%). ^{1}H NMR (500 MHz, d_6-DMSO) δ 9.50 (s, 1H, Amide N-H), 9.00 (s, 1H, Amide N-H), 8.25 (d, J = 8.3 Hz, 1H, Ar-H), 8.00 (t, J = 9.2 Hz, 1H, Ar-H), 7.64 – 7.59 (m, 3H, Ar-H), 7.58 (d, J = 1.6 Hz, 1H, Ar-H), 7.35 (dd, J = 8.2, 1.9 Hz, 1H, Ar-H), 7.29 (d, J = 1.8 Hz, 1H, Ar-H), 6.69 (d, J = 8.2 Hz, 1H, Ar-H), 4.19 (t, J = 5.9 Hz, 2H, 3-C$_\alpha$H$_2$), 3.10 – 2.99 (m, 3H, 3-C$_\gamma$H$_2$), 2.64 – 2.57 (s, 6H, 2 × NMe), 2.19 – 2.01 (m, 4H, 1-C$_\beta$H, 2-C$_\beta$H and 3-C$_\beta$H$_2$), 1.05 (t, J = 5.5 Hz, 6H, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.01 (d, J = 4.8 Hz, 6H, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); HRMS Found 621.3293 C$_{34}$H$_{44}$N$_2$O$_7$ [M+H]$^+$ requires 621.3288.

H₂N-[*O*-*ⁱ*Bu (3-HABA)] -[*O*-*ⁱ*Bu(3-HABA)] -[2-(pyrrolidin-1-yl) ethoxy(3-HABA)]-COOH, 109
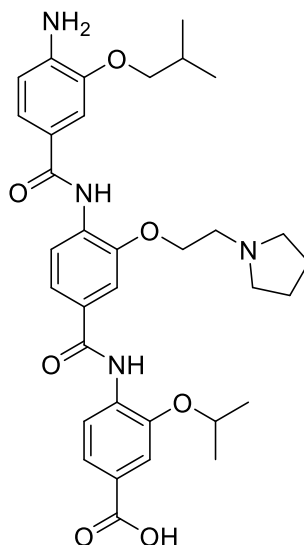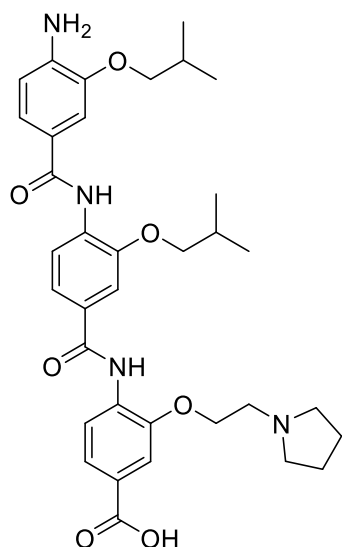

Prepared by standard procedure H followed by standard procedures E and C and purification by mass directed HPLC (16 mg, 72%). ^{1}H NMR (500 MHz, d_6-DMSO)

δ 9.83 (s, 1H, Amide-NH), 8.99 (s, 1H, Amide-NH), 8.25 (d, J = 8.4 Hz, 1H, Ar-H), 8.04 (d, J = 8.3 Hz, 1H, Ar-H), 7.77 (s, 1H, Ar-H), 7.69 (dd, J = 8.4, 1.7 Hz, 1H, Ar-H), 7.64 (dd, J = 8.3, 1.4 Hz, 1H, Ar-H), 7.61 (d, J = 1.5 Hz, 2H, Ar-H), 7.35 (dd, J = 8.3, 1.6 Hz, 1H, Ar-H), 7.30 (d, J = 1.7 Hz, 1H, Ar-H), 6.71 (d, J = 8.2 Hz, 1H, Ar-H), 4.45 (t, J = 4.4 Hz , 2H, 3-C$_\alpha$H$_2$), 4.01 (d, J = 6.3 Hz, 2H, 1-C$_\alpha$H$_2$), 3.79 (d, J = 6.5 Hz, 2H, 2-C$_\alpha$H$_2$), 3.59 (m, 2H, 3-C$_\beta$H$_2$), 2.15 (dt, J = 13.5, 6.6 Hz, 1H, 1-C$_\beta$H), 2.08 (dt, J = 13.2, 6.7 Hz, 1H, 2-C$_\beta$H),1.93-1.86 (m, 4H, pyrrolidine 2 × CH$_2$), 1.27 − 1.22 (m, 4H, pyrrolidine 2 × CH$_2$), 1.05 (d, J = 6.7 Hz, 6H, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.02 (d, J = 6.7 Hz, 6H, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); HRMS Found 633.3294 C$_{35}$H$_{44}$N$_2$O$_7$ [M+H]$^+$ requires 633.3282

H$_2$N-[O-^{i}Bu (3-HABA)]- [O--^{i}Bu(3-HABA)]- [pyridin-4-ylmethoxy (3-HABA)]-COOH, 117
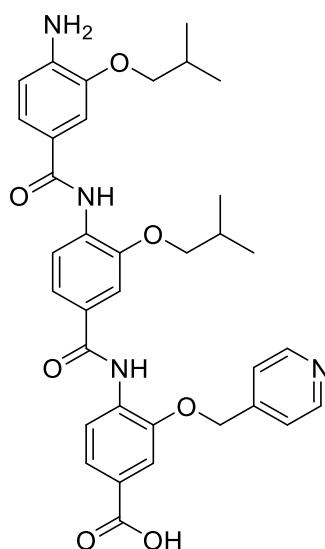
Prepared by standard procedure H followed by standard procedures E and C and purification by mass directed HPLC (8.7 mg, 40%). ^{1}H NMR (500 MHz, d_6-DMSO) δ 9.57 (s, 1H, Amide NH), 8.99 (s, 1H, Amide NH), 8.22 (d, J = 8.4 Hz, 1H, Ar-H), 8.04 (d, J = 8.3 Hz, 1H, Ar-H), 8.01 (d, J = 7.9 Hz, 1H, Ar-H), 7.71 (d, J = 1.7 Hz, 1H, Ar-H), 7.66 − 7.52 (m, 4H, Ar-H), 7.51 (d, J = 1.9 Hz, 1H, Ar-H), 7.44 (dt, J = 7.6, 3.8 Hz, 1H, Ar-H), 7.35 (dd, J = 8.2, 1.9 Hz, 1H, Ar-H), 7.31 (t, J = 2.0 Hz, 1H, Ar-H), 6.71 (d, J = 8.2 Hz, 1H, Ar-H), 5.41 (s, 2H, 3-C$_\alpha$H$_2$), 3.90 (d, J = 6.2 Hz, 2H, , 1-C$_\alpha$H$_2$), 3.79 (d, J = 6.5 Hz, 2H, , 2-C$_\alpha$H$_2$), 2.19 − 2.03 (m, 2H, 1-C$_\beta$H and 2-C$_\beta$H), 1.04 (d, J = 6.7 Hz, 6H, 1-C$_\gamma$H$_3$ and 1-C$_\delta$H$_3$), 1.02 (d, J = 6.7 Hz, 6H, 2-C$_\gamma$H$_3$ and 2-C$_\delta$H$_3$); HRMS Found 627.2820 C$_{35}$H$_{38}$N$_4$O$_7$ [M+H]$^+$ requires 627.2813

H$_2$N-[*O--^{i}Bu(3-HABA)]-[O-TFD* (3-HABA)]-[*O-^{i}Pr(3-HABA)]*-COOH, 128
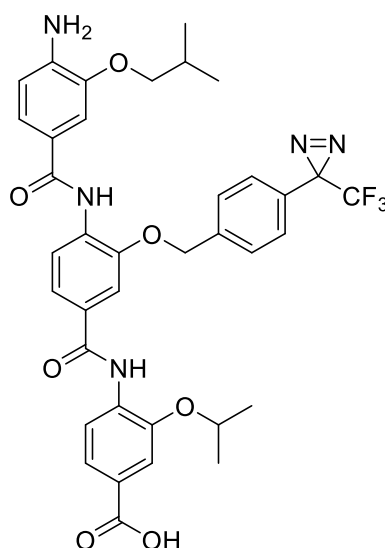
Prepared following standard procedure G followed by standard procedure E. The resulting material was dissolved in THF and cooled to 0 °C and sodium hydride (1 eq.) added. The reaction mixture was stirred for 30 minutes at 0 °C before 3-[4-(iodomethyl)phenyl]-3-(trifluoromethyl)-3H-diazirine (1.1 eq.) was added. The reaction mixture was stirred overnight and allowed to warm to r.t. The reaction mixture was then diluted with methanol and treated with 1M NaOH (aq.). Finally the reaction mixture was concentrated *in vacuo* and purified by mass directed HPLC (2.05 mg, 16%). ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.82 (s, 1H, Amide NH), 8.64 (d, *J* = 8.3 Hz, 1H, Ar-H), 8.57 (m, 2H, Amide NH and Ar-H), 7.71 (d, *J* = 8.5 Hz, 1H, Ar-H), 7.63 (s, 1H, Ar-H), 7.55 (s, 1H, Ar-H), 7.44 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.37 (d, *J* = 9.1 Hz, 1H, Ar-H), 7.33 (s, 1H, Ar-H), 7.08 (d, *J* = 7.7 Hz, 1H, Ar-H), 6.60 (d, *J* = 8.1 Hz, 1H, Ar-H), 5.21 (s, 2H, 2-C$_α$H$_2$), 4.73 – 4.64 (m, 1H, 3-C$_α$H), 3.75 (d, *J* = 6.4 Hz, 2H, 1-C$_α$H$_2$), 2.11 – 2.04 (m, 1H, 1-C$_β$H), 1.39 (d, *J* = 6.0 Hz, 6H, 3-C$_β$H$_3$ and 3-C$_γ$H$_3$), 0.97 (t, *J* = 9.0 Hz, 6H, 2-C$_γ$H$_3$ and 2-C$_δ$H$_3$); HRMS Found: 720.2639; C$_{37}$H$_{36}$F$_3$N$_5$O$_7$ [M+H]$^+$ requires 720.2639.

H₂N-[*O*--*i*Bu(3-HABA)]-[*O*-*i*Bu(3-HABA)]-[*O*-*TFD*(3-HABA)]-COOH, 129
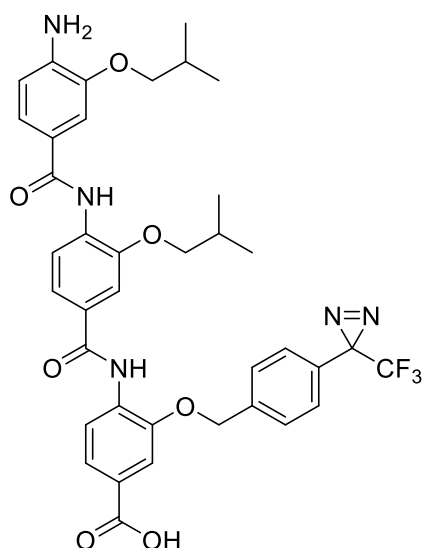
Prepared following standard procedure G followed by standard procedure E. The resulting material was dissolved in THF and cooled to 0 °C and sodium hydride (1 eq.) added. The reaction mixture was stirred for 30 minutes at 0 °C before 3-[4-(iodomethyl)phenyl]-3-(trifluoromethyl)-3H-diazirine (1.1 eq.) was added. The reaction mixture was stirred overnight and allowed to warm to r.t. The reaction mixture was then diluted with methanol and treated with 1M NaOH (aq.). Finally the reaction mixture was concentrated *in vacuo* and purified by mass directed HPLC (4.6 mg, 44%). ^{1}H NMR (500 MHz, CDCl₃) δ 8.79 (s, 1H, Amide NH), 8.76 (s, 1H, Amide NH), 8.72 (d, *J* = 8.5 Hz, 1H, Ar-H), 8.68 (d, *J* = 8.4 Hz, 1H, Ar-H), 7.88 (d, *J* = 8.6 Hz, 1H, Ar-H), 7.72 (s, 1H, Ar-H), 7.58 (s, 1H, Ar-H), 7.53 (d, *J* = 8.2 Hz, 2H, Ar-H), 7.44 (s, 1H, Ar-H), 7.40 – 7.33 (m, 2H, Ar-H), 7.30 (d, *J* = 3.8 Hz, 2H, Ar-H), 6.78 (d, *J* = 8.1 Hz, 1H, Ar-H), 5.31 (s, 2H, 3-C$_\alpha$H₂), 3.96 (d, *J* = 6.4 Hz, 2H, 1-C$_\alpha$H₂), 3.90 (d, *J* = 6.5 Hz, 2H, 2-C$_\alpha$H₂), 2.31 – 2.04 (m, 2H, 1-C$_\beta$H and 2-C$_\beta$H), 1.14 (d, *J* = 6.7 Hz, 6H, 1-C$_\gamma$H₃ and 1-C$_\delta$H₃), 1.10 (d, *J* = 6.7 Hz, 6H, 2-C$_\gamma$H₃ and 2-C$_\delta$H₃); HRMS Found: 734.2780; C₃₈H₃₈F₃N₅O₇ [M+H]$^+$ requires 734.2796.

5.4.6 Protein Digestion and LC-MS/MS

p300 protein was dialysed into 50 mM ammonium bicarbonate buffer overnight and concentrated to ~ 1mg/ml. 1 ml of protein solution was heated to 40 °C with shaking for 10 minutes. The solution was then treated with DTT (final concentration 10 mM) and heated to 80 °C with shaking for 15 minutes. The solution was then allowed to cool, and treated with iodoacetamide (final concentration 20 mM) in the dark for 30 minutes. Trypsin was then added (1:50 trypsin:protein) and the reaction agitated at 37 °C overnight. The reaction was then quenched with TFA and MeCN and analysed by LC-MS/MS on a Bruker Maxis II instrument. Data was analysed using BioTools (Bruker Daltronics).

5.4.7 Photo-Crosslinking Experiments

A solution containing the photo-crosslinking compound (150 μM) and p300 (100 μM) in assay buffer was prepared and analysed by LC-MS on a Bruker HCT Ultra. Separate solutions containing only compound or only protein were treated identically. The solutions were then irradiated with UV light (365 nm) for 1 hour whilst cooled in an ice bath. The LC-MS analysis was then repeated. The protein signals were examined for any increase of adduct formation.

5.4.8 Synthesis of glycine linked monomer

Prop-2'-en-1'-yl 3-hydroxy-4-nitrobenzoate, 132
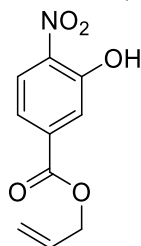
Para-toluene sulfonic acid (2 g, 11 mmol) was added to a solution of 3-hydroxy-4-nitrobenzoic acid (10 g, 55 mmol) in allyl alcohol (30 ml) and the reaction heated to 80 °C overnight then allowed to cool to r.t. The reaction mixture was diluted with ethyl acetate (50 ml) and washed with water (30 ml), saturated aqueous sodium bicarbonate (2 × 30 ml) and brine (30 ml), dried over $MgSO_4$ and concentrated *in vacuo* to a yellow amorphous solid (8.9 g, 73%) v_{max}/cm^{-1} (solid state) 3316, 1722; ^{1}H NMR (500 MHz, CDCl$_3$) δ 10.53 (s, 1H, phenol OH), 8.20 (d, *J* = 8.8 Hz, 1H, 6-CH), 7.87 (s, 1H, 2-CH), 7.66 (d, *J* = 8.8 Hz, 1H, 5-CH), 6.06 (ddd, *J* = 16.7, 11.2, 5.6 Hz, 1H, 2'C-H), 5.46 (d, *J* = 17.2 Hz, 1H, 3'C-H$_{trans}$), 5.36 (d, *J* = 10.4 Hz, 1H, 3'C-H$_{cis}$), 4.88 (d, *J* = 5.6 Hz, 2H, 1'-CH$_2$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 164.01, 154.67, 138.05, 135.83, 131.43, 125.29, 121.71, 120.65, 119.22, 66.55; HRMS Found 222.0411 C$_{10}$H$_9$NO$_5$ [M-H]$^-$ requires 222.0407

Prop-2'-en-1'-yl 3-(2''-methylpropoxy)-4-nitrobenzoate, 133
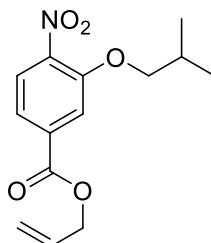
1-bromo-2-methylpropane (3.6 ml, 33.6 mmol) was added to a suspension of Prop-2'-en-1'-yl 3-hydroxy-4-nitrobenzoate (5 g, 22.4 mmol) and potassium carbonate (15.4 g, 112 mmol) in dimethylformamide (50 ml) and the reaction heated to 50 °C with stirring overnight. The reaction mixture was then diluted with ethyl acetate

(100 ml), washed copiously with water and brine, dried over MgSO$_4$ and concentrated *in vacuo* to give the title compound as a red oil (3.55 g, 60%). v_{max}/cm^{-1} (solid state) 2961, 2875, 1718; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.80 (d, *J* = 8.4 Hz, 1H, 5-CH), 7.73 (s, 1H, 2C-H), 7.67 (dd, *J* = 8.4, 1.3 Hz, 1H, 6-CH), 6.04 (ddd, *J* = 16.3, 11.0, 5.8 Hz, 1H, 2'-CH), 5.43 (dd, *J* = 17.2, 1.2 Hz, 1H, 3'-CH$_{trans}$), 5.33 (dd, *J* = 10.3, 0.8 Hz, 1H, 3'-CH$_{cis}$), 4.86 (d, *J* = 5.8 Hz, 2H, 1'-CH$_2$), 3.93 (d, *J* = 6.4 Hz, 2H, 1''-CH$_2$), 2.15 (spt, *J* = 6.6 Hz, 1H, 2''-CH), 1.06 (d, *J* = 6.9 Hz, 6H, 3''-CH$_3$ and 4''CH$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 164.45, 152.05, 142.51, 134.75, 131.64, 125.15, 121.06, 118.98, 115.42, 75.95, 66.39, 28.19, 18.98; HRMS Found: 302.0998, C$_{14}$H$_{17}$NO$_5$ [M+Na]$^+$ requires 302.0998.

Prop-2'-en-1'-yl 4-amino-3-(2''-methylpropoxy)benzoate, 134
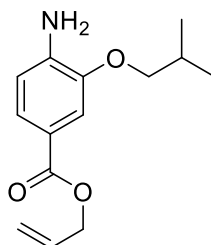
Tin (II) chloride dihydrate (4 g, 17.9 mmol) was added to a solution of prop-2'-en-1'-yl 3-(2''-methylpropoxy)-4-nitrobenzoate (1 g, 3.58 mmol) in ethyl acetate (70 ml) and the reaction heated to 50 °C overnight. The reaction was allowed to cool and poured into 1M NaOH solution (50 ml). The organic layer was separated and washed with 1M NaOH (3 × 30 ml), dried over MgSO$_4$ and concentrated *in vacuo* to a pale yellow oil (880 mg, 98%). v_{max}/cm^{-1} (solid state) 3588, 2914, 1742, 1635; ^{1}H NMR (500 MHz, CDCl$_3$) δ 7.60 (dd, *J* = 8.2, 1.6 Hz, 1H, 6-CH), 7.48 (d, *J* = 1.4 Hz, 1H, 2C-H), 6.69 (d, *J* = 8.2 Hz, 1H, 5-CH), 6.07 (ddd, *J* = 17.2, 10.4, 1.4 Hz, 1H, 2'-CH), 5.42 (dd, *J* = 17.2, 1.4 Hz, 1H, 3'-CH$_{trans}$), 5.29 (dd, *J* = 10.4, 1.4 Hz, 1H, 3'-CH$_{cis}$), 4.81 (d, *J* = 5.6 Hz, 2H, 1'-CH$_2$), 3.85 (d, *J* = 6.5 Hz, 2H, 1''-CH$_2$), 2.25 – 1.98 (sept, 6.6 Hz, 1H, , 2''-CH), 1.08 (d, *J* = 6.6 Hz, 6H, 3''-CH$_3$ and 4''CH$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 166.54, 145.60, 141.36, 132.82, 123.99, 119.40, 117.70, 113.08, 112.15, 74.72, 65.07, 28.30, 19.34; HRMS Found: 250.1438, C$_{14}$H$_{19}$NO$_3$ [M+H]$^+$ requires 250.1437

Prop-2'-en-1'-yl4-(2-{[(9H-fluoren-9-ylmethoxy)carbonyl]amino}acetamido)-3-(2''-methylpropoxy)benzoate, 135
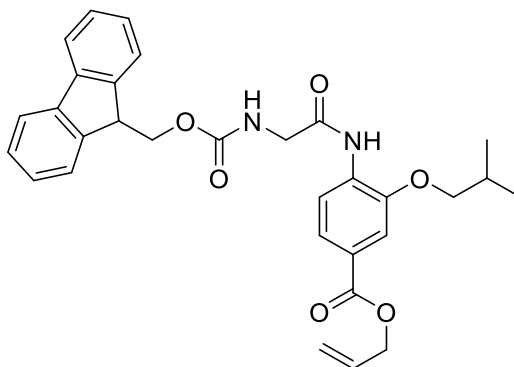
Dichlorotriphenylphosphorane (5.8 g, 17.6 mmol) was added to a solution of Fmoc-glycine (2 g, 7.06 mmol) and prop-2'-en-1'-yl 4-amino-3-(2''-methylpropoxy)benzoate (880 mg, 3.53 mmol) in chloroform (50 ml) and the reaction heated to reflux overnight. The reaction was allowed to cool to r.t. and concentrated *in vacuo*. The residue was purified by column chromatography eluting with dichloromethane to give the *title compound* as a colourless amorphous solid (1.3g, 71%). v_{max}/cm^{-1} (solid state) 3398, 2958, 2873, 1697; ^{1}H NMR (500 MHz, CDCl$_3$) δ 8.59 (s, 1H, Amide NH), 8.48 (d, J = 8.4 Hz, 1H, Ar-H), 7.80 (d, J = 6.3 Hz, 1H, Ar-H), 7.74 (d, J = 8.4 Hz, 1H, Ar-H), 7.62 (s, 1H, Ar-H), 7.56 (d, J = 1.2 Hz, 1H, Ar-H), 7.43 (s, 1H, Ar-H), 7.33 (s, 1H, Ar-H), 6.08 (ddd, J = 17.2, 10.4, 5.6 Hz, 1H, 2'-CH), 5.44 (d, J = 17.2 Hz, 1H, 3'-CH$_{trans}$), 5.33 (d, J = 10.4 Hz, 1H, 3'-CH$_{cis}$), 4.85 (d, J = 5.6 Hz, 2H, 1'-CH$_2$), 4.48 (d, J = 7.0 Hz, 2H, gly-CH$_2$), 4.26 (t, J = 6.9 Hz, 1H, Fmoc-CH), 4.10 (s, 1H br s, 1H, carbamate NH), 3.85 (d, J = 6.4 Hz, 2H, Fmoc-CH$_2$), 2.12 (sept, J = 6.7 Hz, 1H, 2''-CH), 1.04 (d, J = 6.7 Hz, 6H, 3''-CH$_3$ and 4''CH$_3$); ^{13}C NMR (125 MHz, CDCl$_3$) δ 167.19, 165.92, 146.97, 143.62, 141.33, 132.32, 131.40, 127.84, 127.12, 125.49, 124.99, 123.21, 120.07, 118.67, 118.28, 111.66, 75.15, 67.58, 65.63, 47.08, 45.90, 28.20, 19.25; HRMS found: 529.234309, C$_{31}$H$_{32}$N$_2$O$_6$ [M+H]$^+$ Requires 529.233313.

4-(2-{[(9H-fluoren-9-ylmethoxy)carbonyl]amino}acetamido)-3-(2-methylpropoxy)benzoic acid, 136
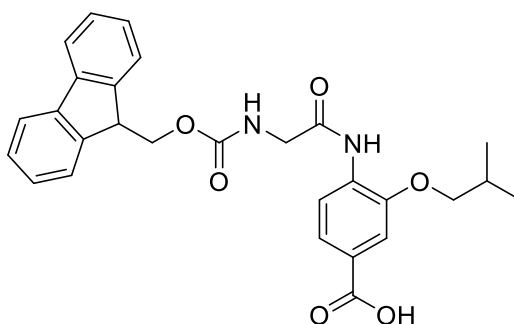
Palladium tetrakis(triphenylphosphine) (110 mg, 0.095 mmol, 5 mol%) was added to a solution of Prop-2'-en-1'-yl4-(2-{[(9H-fluoren-9-ylmethoxy)carbonyl]amino}acetamido)-3-(2''-methylpropoxy)benzoate (1 g, 1.89 mmol) and sodium p-toluenesulfinate (505 mg, 2.83 mmol) in tetrahydrofuran (50 ml) and the reaction stirred at r.t. for 24 hours. The reaction mixture was filtered through Celite and concentrated *in vacuo*. The residue was purified by column chromatography eluting with 5-10% methanol in dichloromethane to give the *title compound* as a yellow solid (741 mg, 80%). v_{max}/cm^{-1} (solid state) 3395, 3302, 2963, 1697; ^{1}H NMR (500 MHz, d_6-DMSO) δ 8.98 (s, 1H, Amide N-H) 8.24 (d, J = 8.0 Hz, 1H, Ar-H), 7.95 (br. s, 1H, carbamate-NH), 7.90 (d, J = 7.4 Hz, 2H, Ar-H), 7.72 (d, J = 7.4 Hz, 2H, Ar-H), 7.57 (d, J = 7.4 Hz, 1H, Ar-H), 7.52 (s, 1H, Ar-H), 7.42 (t, J = 7.2 Hz, 2H, Ar-H), 7.33 (t, J = 7.3 Hz, 2H, Ar-H), 4.34 (d, J = 6.6 Hz, 2H, Fmoc-CH$_2$), 4.25 (t, J = 6.6 Hz, 1H, Fmoc-CH), 3.86 (d, J = 6.0 Hz, 2H, Gly-CH$_2$), 3.83 (d, J = 6.6 Hz, 2H, 1''-CH$_2$), 2.07 – 1.94 (m, 1H, 2''-CH), 0.95 (d, J = 6.7 Hz, 6H, 3''-CH$_3$ and 4''CH$_3$); ^{13}C NMR (125 MHz, d_6-DMSO) δ 168.27, 147.24, 143.68, 140.72, 130.81, 127.65, 127.05, 125.14, 122.29, 120.11, 118.80, 112.00, 74.62, 66.01, 46.57, 27.59, 18.94; HRMS found 489.2032 C$_{28}$H$_{28}$N$_2$O$_6$ requires [M+H]$^+$ 489.2020

5.4.9 Preparation of Peptide-Oligobenzamide Hybrids

Prepared following an adapted literature method.[152] Fmoc-monomer acid chlorides were prepared as previously described and loaded onto a CEM liberty peptide synthesiser. Gly-loaded Wang resin was swelled in NMP for 30 minutes prior to loading onto the synthesiser. The previously described methods were used to synthesise the oligobenzamide portion of the molecule, ending with the Fmoc-Gly-monomer. The peptide portion was then extended using standard Fmoc peptide synthesis as described above.

H₂N-QLTSYCEVNAPIQSRNLLQG--[*O*-*ⁱ*Bu(3-HABA)]-[*O*-*ⁱ*Bu(3-HABA)]-[*O*-*ⁱ*Pr(3-HABA)]-G-NH₂, 137
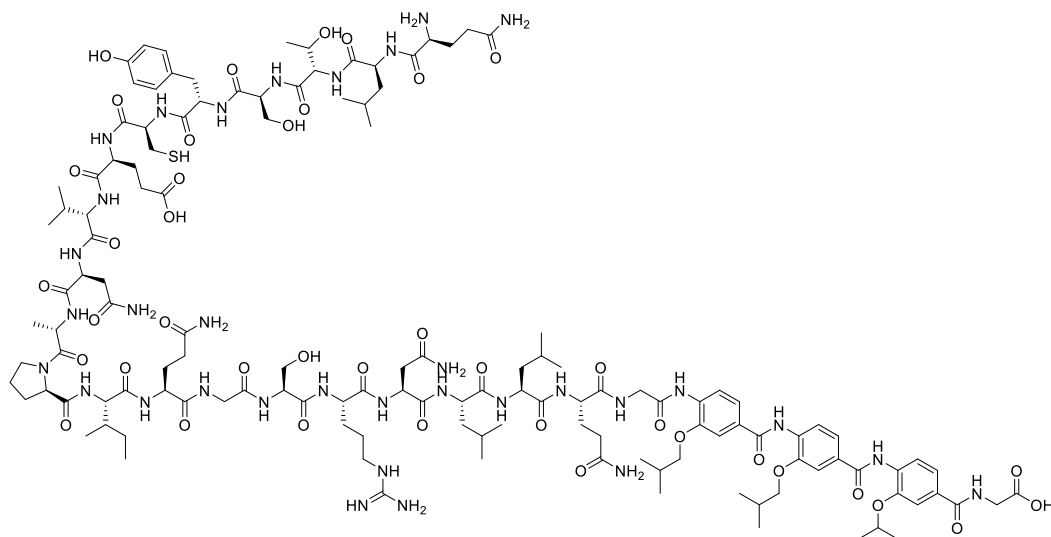
LC-MS *m/z* (ES) 1455.3 [M+2H]²⁺, 970.4 [M+3H]³⁺ HRMS Found: 1454.2223 C₁₃₁H₁₉₉N₃₃O₄₀S requires [M+2H]²⁺ 1454.2209
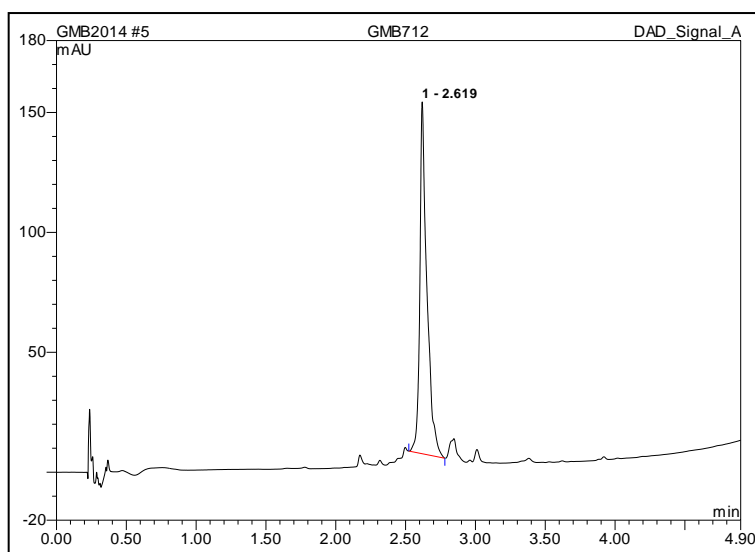
**Ac-QLTSYSEVNAPIQSRNLLQG--[*O-*^{i}Bu(3-HABA)]-[*O-*^{i}Bu(3-HABA)]-[*O-*
^{i}Pr(3-HABA)]-G-NH$_2$, 138**
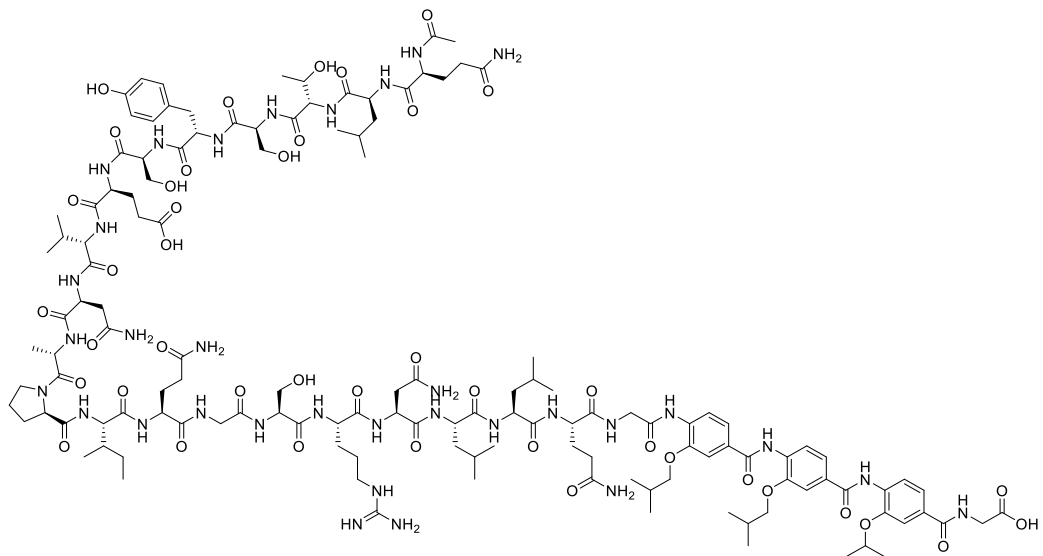
LC-MS *m/z* (ES) 1468.2 [M+2H]$^{2+}$ HRMS Found: 1467.7340, C$_{133}$H$_{201}$N$_{33}$O$_{42}$
requires [M+2H]$^{2+}$ 1467.7393
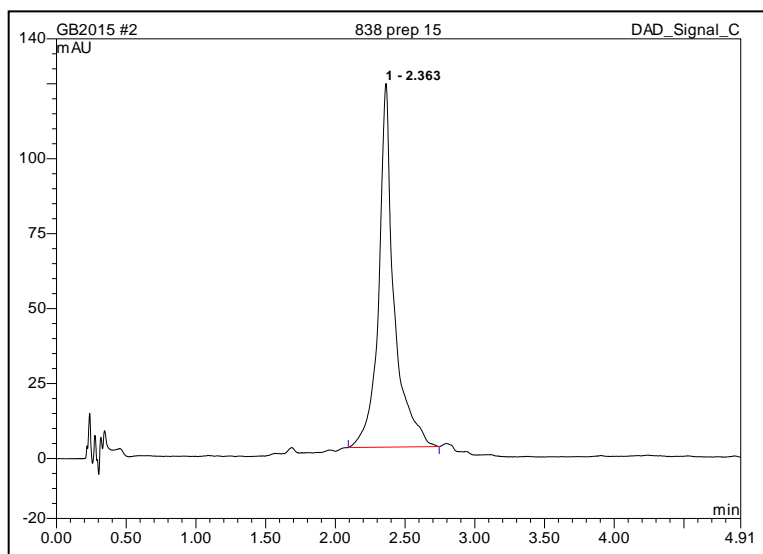
Appendix I – Alternative Peptide PEG Linking Strategies

Initially, a one pot conjugation procedure was developed to allow the modular assembly of peptidic regions with linkers of various lengths thus negating the requirement to perform the peptide synthesis in triplicate. A route combining copper catalysed azide alkyne cycloaddition and thiol-ene reactions was envisioned. This was trialled with a model system as shown below. This reaction sequence progressed successfully in one pot as followed by LC-MS. By treating the commercially available NHS-PEG-Maleimide compound with a short azido-alkylamine followed by a thiol, the click handle and the first component can be introduced regioselectively. Finally, addition of an alkyne along with a copper catalyst affords the PEG linked compound.
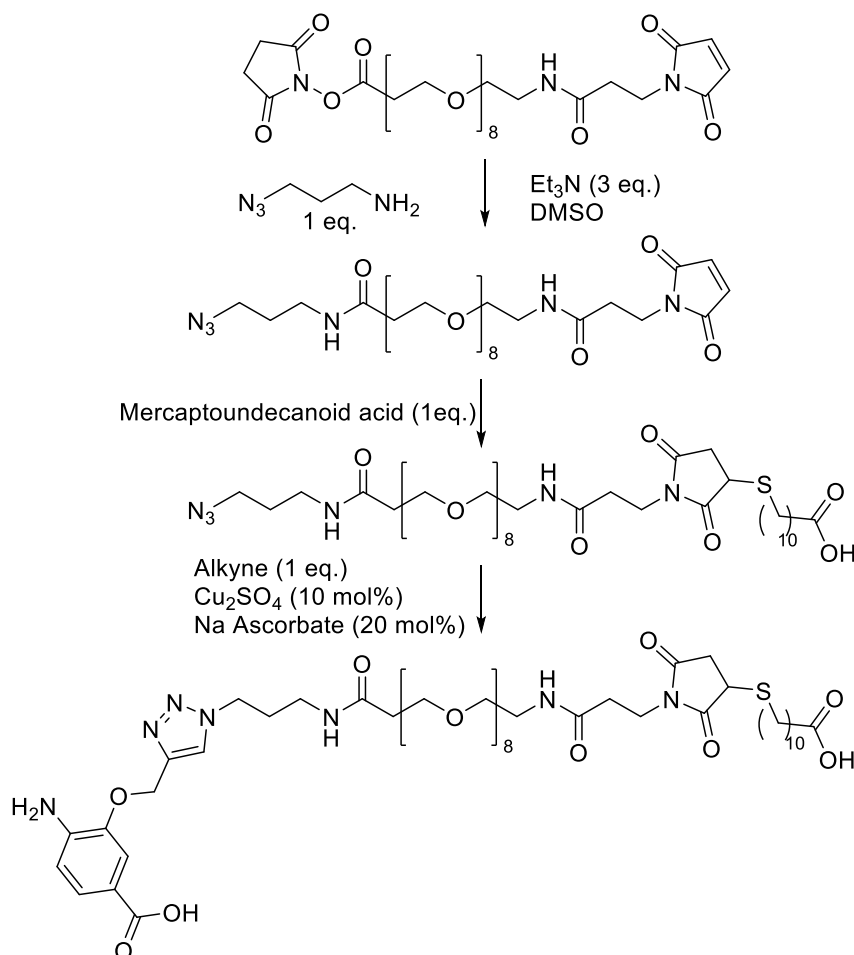
This one pot procedure was then progressed onto using a peptide containing and *C*-terminal propargyl glycine and another with an *N*-terminal thiol as shown below and monitored by LC-MS. The final click reaction progressed with small molecule alkynes but not the alkynyl peptide. The peptide appeared to be unstable to the reaction conditions.
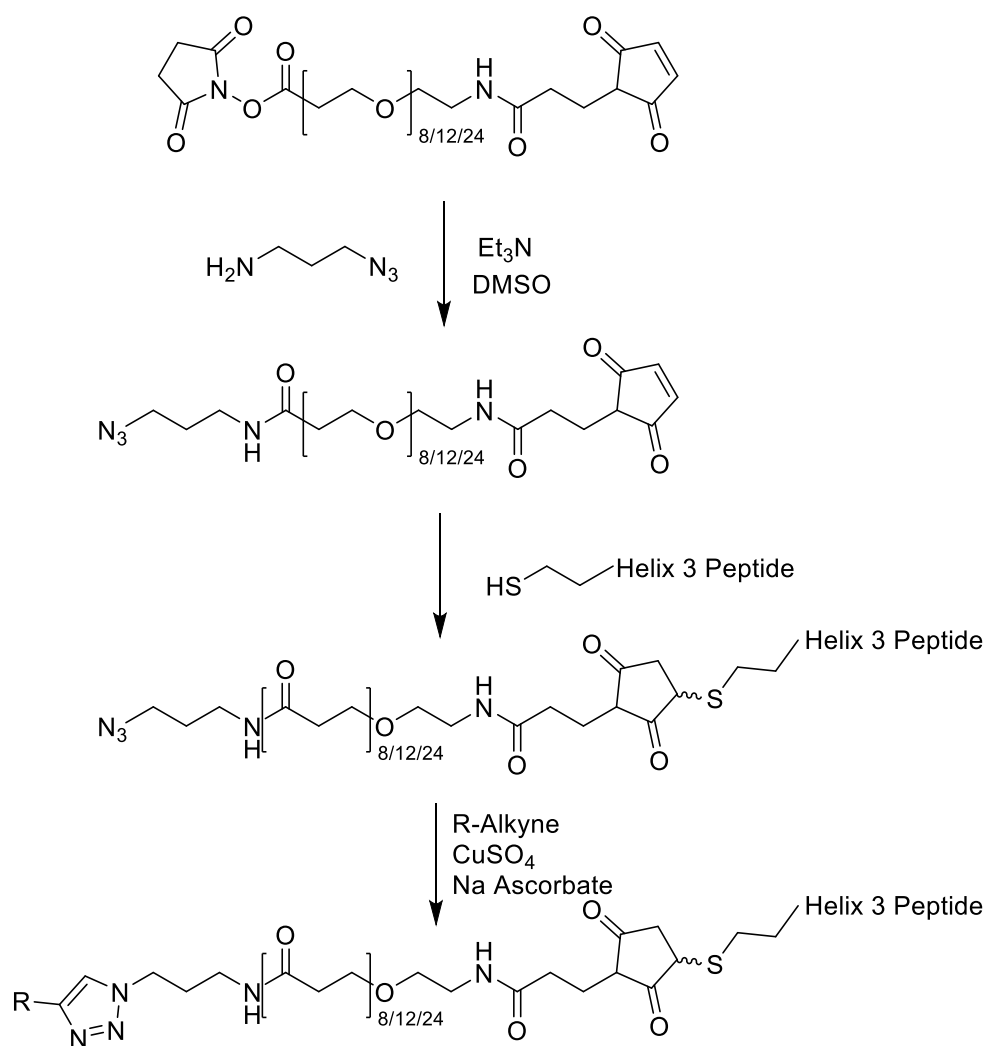
A shorter alkynyl peptide was prepared and used to optimise the click reaction as shown in the table below. Using these optimised conditions, the full length peptide could cleanly undergo click reactions. However, when the complete one pot procedure was attempted using the optimised conditions, no product was observed.
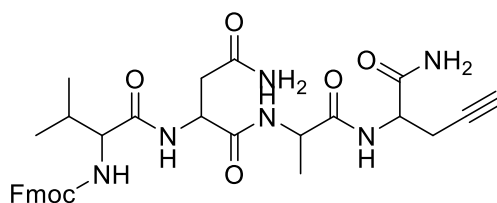
Conditions	Outcome
1 eq. Cu, 1 eq. Ascorbate, DMSO	No Reaction
1.1 eq. Cu, 1 eq. Ascorbate, DMSO	Conversion to product and degradation
1.1 eq. Cu, 10 eq. Ascorbate, DMSO	Conversion to product and degradation
1 eq. Cu, 1 eq. TBTA, 10 eq. Ascorbate, DMSO	Conversion to product and degradation
1.1 eq. Cu, 10 eq. TCEP, DMSO	No Reaction
1 eq. Cu, 10 eq. Ascorbate, 1:1 THF/Water	Conversion to product
1 eq. Cu, 10 eq. Ascorbate, 10 eq. TBTA, 1:1 THF/Water	Conversion to product and stable o/n
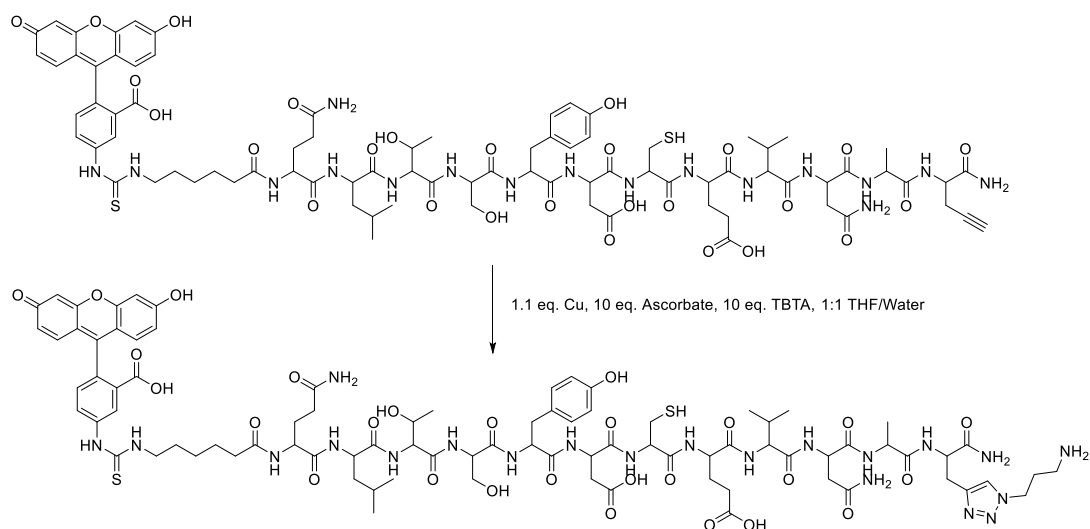
Clean progression by LC-MS

At this stage it was decided to prepare the PEG-Peptide conjugates via linear peptide synthesis incorporating the PEG chain as with any other monomer.

Appendix II – Alternative Synthetic Routes to Designed Compounds

Alternative Route via cyclic sulfamidate:
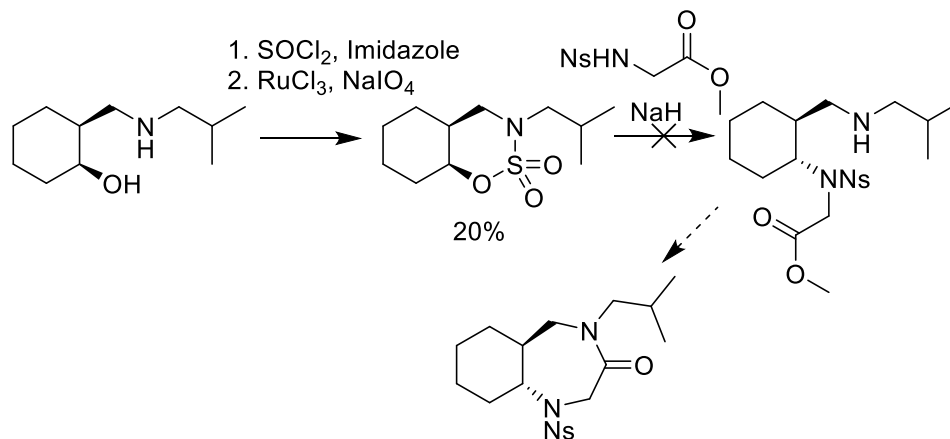
No sulfamidate ring opening was observed with the sulfonamide anion under forcing conditions.

Alternative Route via aminolysis:
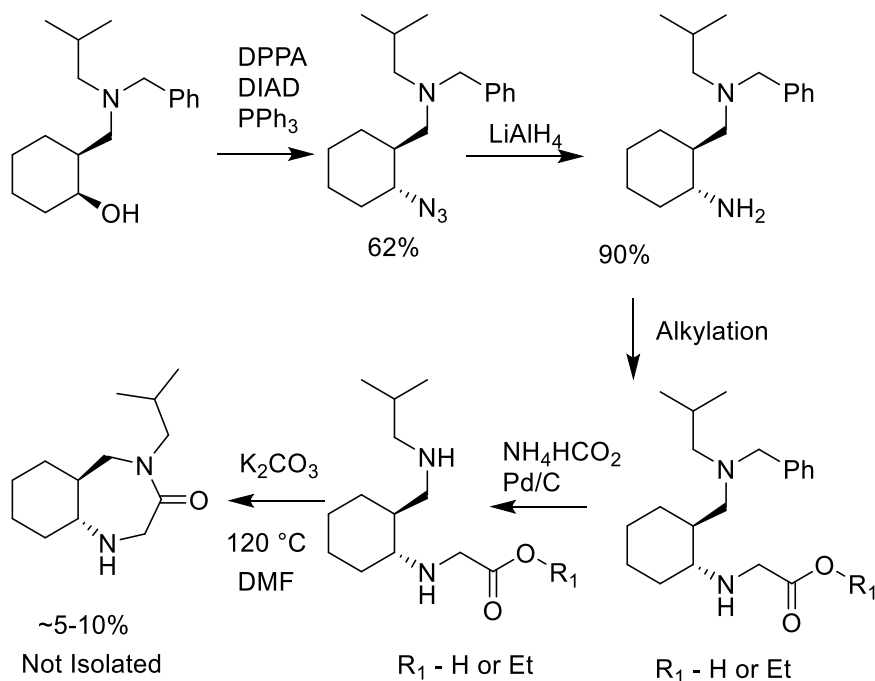
All attempts to improve the conversion in the final step lead to a complex mixture of undesired compounds.

Alternative Route via Mitsunobu ring closing:
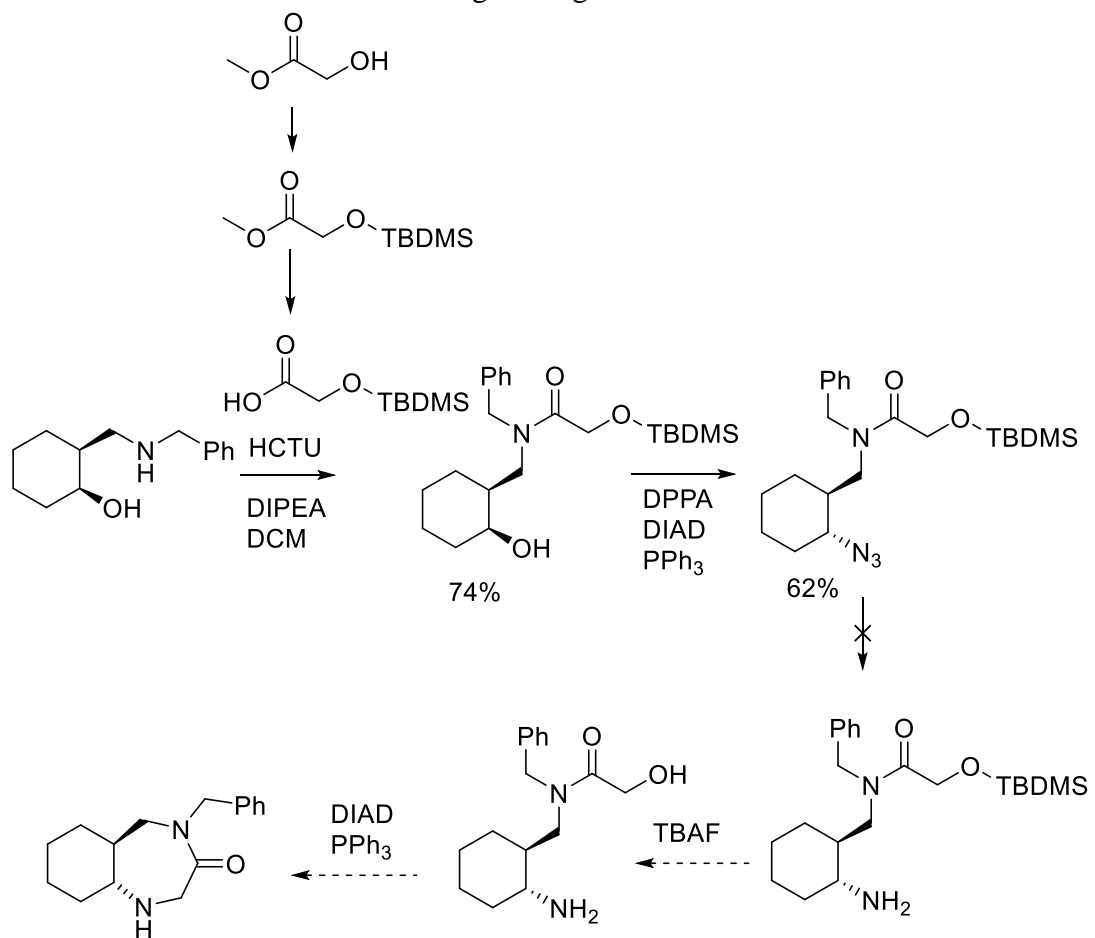
Conditions to cleanly reduce the azide could not be identified.

Appendix III – Crystallographic Data

The diffraction data were collected and solved by Dr Chris Pask at the School of Chemistry X-ray facility.

Measurements were carried out at 120K on an Agilent SuperNova diffractometer equipped with an Atlas CCD detector and connected to an Oxford Cryostream low temperature device using mirror monochromated Cu K_α radiation (λ= 1.54184 Å) from a Microfocus Nova X-ray source. The structure was solved by direct methods using SHELXS[203] and refined by a full matrix least squares technique based on F2 using SHELXL97.[203]

The compound crystallised as colourless needles by slow evaporation of 1:1 $CHCl_3$ and Et_2O. The compound crystallised in an orthorhombic cell and was solved in the $P2_12_12_1$ space group, with one molecule in the asymmetric unit.

All non-hydrogen atoms were located in the Fourier Map and refined anisotropically. All hydrogen atoms were placed in calculated positions and refined isotropically using a "riding model".

Empirical formula	$C_{13}H_{24}N_2O$
Formula weight	224.34
Temperature/K	120.0(2)
Crystal system	orthorhombic
Space group	$P2_12_12_1$
a/Å	5.06544(18)
b/Å	11.4671(4)
c/Å	22.0794(8)
α/°	90.00
β/°	90.00
γ/°	90.00
Volume/Å^3	1282.51(8)
Z	4
ρ_{calc}g/cm^3	1.162
μ/mm^{-1}	0.572
F(000)	496.0
Crystal size/mm^3	$0.16 \times 0.07 \times 0.04$
Radiation	CuKα (λ = 1.54184)
2Θ range for data collection/°	8 to 148.36
Index ranges	-6 ≤ h ≤ 6, -14 ≤ k ≤ 13, -27 ≤ l ≤ 26
Reflections collected	13058
Independent reflections	2564 [R_{int} = 0.0468, R_{sigma} = 0.0292]
Data/restraints/parameters	2564/0/150
Goodness-of-fit on F^2	1.046
Final R indexes [I>=2σ (I)]	R_1 = 0.0344, wR_2 = 0.0849
Final R indexes [all data]	R_1 = 0.0375, wR_2 = 0.0871
Largest diff. peak/hole / e Å^{-3}	0.15/-0.23
Flack parameter	-0.2(3)

Appendix IV

Reducing agent screen on the reaction of **40** to **42** – ratios determined by crude NMR with reference to authentic separated samples.

Reducing Agent	dr / cis:trans
$NaBH_4$	60:40
$NaBH(OAc)_3$	80:20
Pd/C, Ammonium Formate, Reflux	92:8
Borane•THF	58:42
Ammonium Formate, Reflux	22:78

Bibliography

1. R. E. Babine and S. L. Bender, *Chem. Rev.*, 1997, **97**, 1359-1472.
2. A. H. C. W. Bertoncini, X. Salvatella in *Protein Surface Recognition: Approaches for Drug Discovery*, ed. A. H. C. W. Bertoncini, X. Salvatella Wiley, New York, 1st edn., 2010.
3. J. A. Wells and C. L. McClendon, *Nature*, 2007, **450**, 1001-1009.
4. T. Clackson and J. Wells, *Science*, 1995, **267**, 383-386.
5. A. A. Bogan and K. S. Thorn, *J. Mol. Biol.*, 1998, **280**, 1-9.
6. T. A. Edwards and A. J. Wilson, *Amino Acids*, 2011, **41**, 743-754.
7. M. P. H. Stumpf, T. Thorne, E. de Silva, R. Stewart, H. J. An, M. Lappe and C. Wiuf, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**, 6959-6964.
8. B. N. Bullock, A. L. Jochim and P. S. Arora, *J. Am. Chem. Soc.*, 2011, **133**, 14220-14223.
9. V. Azzarito, K. Long, N. S. Murphy and A. J. Wilson, *Nature Chemistry*, 2013, **5**, 161-173.
10. A. A. Ivanov, F. R. Khuri and H. Fu, *Trends Pharmacol. Sci.*, 2013, **34**, 393-400.
11. C. C. Harris, *Proc. Natl. Acad. Sci. U.S.A.*, 2006, **103**, 1659-1660.
12. A. Burlacu, *J. Cell. Mol. Med.*, 2003, **7**, 249-257.
13. M. Ivan, K. Kondo, H. Yang, W. Kim, J. Valiando, M. Ohh, A. Salic, J. M. Asara, W. S. Lane and W. G. Kaelin Jr., *Science*, 2001, **292**, 464-468.
14. Z. Arany, L. E. Huang, R. Eckner, S. Bhattacharya, C. Jiang, M. A. Goldberg, H. F. Bunn and D. M. Livingston, *Proc. Natl. Acad. Sci. U.S.A.*, 1996, **93**, 12969-12973.
15. F. Wang, C. Marshall and M. Ikura, *Cell. Mol. Life Sci.*, 2013, **70**, 3989-4008.
16. R. Chowdhury, A. Hardy and C. J. Schofield, *Chem. Soc. Rev.*, 2008, **37**, 1308-1319.
17. G. L. Semenza, P. H. Roth, H. M. Fang and G. L. Wang, *J. Biol. Chem.*, 1994, **269**, 23757-23763.
18. A. L. Kung, S. Wang, J. M. Klco, W. G. Kaelin and D. M. Livingston, *Nat Med*, 2000, **6**, 1335-1340.
19. I. K. Nordgren and A. Tavassoli, *Chem. Soc. Rev.*, 2011, **40**, 4307-4317.
20. P. Jaakkola, D. R. Mole, Y. M. Tian, M. I. Wilson, J. Gielbert, S. J. Gaskell, A. von Kriegsheim, H. F. Hebestreit, M. Mukherji, C. J. Schofield, P. H. Maxwell, C. W. Pugh and P. J. Ratcliffe, *Science*, 2001, **292**, 468-472.
21. P. C. Mahon, K. Hirota and G. L. Semenza, *Genes & Development*, 2001, **15**, 2675-2686.
22. D. Lando, D. J. Peet, J. J. Gorman, D. A. Whelan, M. L. Whitelaw and R. K. Bruick, *Genes & Development*, 2002, **16**, 1466-1471.
23. M. S. Wiesener, J. S. Jürgensen, C. Rosenberger, C. Scholze, J. H. Hörstrup, C. Warnecke, S. Mandriota, I. Bechmann, U. A. Frei, C. W. Pugh, P. J. Ratcliffe, S. Bachmann, P. H. Maxwell and K.-U. Eckardt, *The FASEB Journal*, 2002.
24. G. L. Wang, B. H. Jiang, E. A. Rue and G. L. Semenza, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, **92**, 5510-5514.
25. Z. J. Huang, I. Edery and M. Rosbash, *Nature*, 1993, **364**, 259-262.

26. Y. Guo, T. H. Scheuermann, C. L. Partch, D. R. Tomchick and K. H. Gardner, *J. Biol. Chem.*, 2015, **290**, 7707-7721.
27. B.-H. Jiang, J. Z. Zheng, S. W. Leung, R. Roe and G. L. Semenza, *J. Biol. Chem.*, 1997, **272**, 19253-19260.
28. P. J. Kallio, K. Okamoto, S. O'Brien, P. Carrero, Y. Makino, H. Tanaka and L. Poellinger, *The EMBO Journal*, 1998, **17**, 6573-6586.
29. S. J. Freedman, Z.-Y. J. Sun, F. Poy, A. L. Kung, D. M. Livingston, G. Wagner and M. J. Eck, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 5367-5372.
30. S. A. Dames, M. Martinez-Yamout, R. N. De Guzman, H. J. Dyson and P. E. Wright, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 5271-5276.
31. S. Bhattacharya, C. L. Michels, M.-K. Leung, Z. P. Arany, A. L. Kung and D. M. Livingston, *Genes & Development*, 1999, **13**, 64-75.
32. J. Bragança, T. Swingler, F. I. R. Marques, T. Jones, J. J. Eloranta, H. C. Hurst, T. Shioda and S. Bhattacharya, *J. Biol. Chem.*, 2002, **277**, 8559-8565.
33. J. L. Ruas, L. Poellinger and T. Pereira, *J. Biol. Chem.*, 2002, **277**, 38723-38730.
34. J. Gu, J. Milligan and L. E. Huang, *J. Biol. Chem.*, 2001, **276**, 3550-3554.
35. H. Cho, D. R. Ahn, H. Park and E. G. Yang, *FEBS Lett.*, 2007, **581**, 1542-1548.
36. I. M. Yasinska and V. V. Sumbayev, *FEBS Lett.*, 2003, **549**, 105-109.
37. M. H. Rabinowitz, *J. Med. Chem.*, 2013, **56**, 9369-9402.
38. H. J. Choi, B. J. Song, Y. D. Gong, W. J. Gwak and Y. Soh, *Br. J. Pharmacol.*, 2008, **154**, 114-125.
39. D. L. Buckley, J. L. Gustafson, I. Van Molle, A. G. Roth, H. S. Tae, P. C. Gareiss, W. L. Jorgensen, A. Ciulli and C. M. Crews, *Angew. Chem. Int. Ed.*, 2012, **51**, 11463-11467.
40. D. L. Buckley, I. Van Molle, P. C. Gareiss, H. S. Tae, J. Michel, D. J. Noblin, W. L. Jorgensen, A. Ciulli and C. M. Crews, *J. Am. Chem. Soc.*, 2012, **134**, 4465-4468.
41. K. Lee, H. Zhang, D. Z. Qian, S. Rey, J. O. Liu and G. L. Semenza, *Proc Natl Acad Sci U S A*, 2009, **106**, 17910-17915.
42. E. Miranda, I. K. Nordgren, A. L. Male, C. E. Lawrence, F. Hoakwie, F. Cuda, W. Court, K. R. Fox, P. A. Townsend, G. K. Packham, S. A. Eccles and A. Tavassoli, *J. Am. Chem. Soc.*, 2013, **135**, 10418-10425.
43. T. H. Scheuermann, Q. Li, H.-W. Ma, J. Key, L. Zhang, R. Chen, J. A. Garcia, J. Naidoo, J. Longgood, D. E. Frantz, U. K. Tambar, K. H. Gardner and R. K. Bruick, *Nat Chem Biol*, 2013, **9**, 271-276.
44. N. G. Nickols, C. S. Jacobs, M. E. Farkas and P. B. Dervan, *ACS Chem. Biol.*, 2007, **2**, 561-571.
45. D. Kong, E. J. Park, A. G. Stephen, M. Calvani, J. H. Cardellina, A. Monks, R. J. Fisher, R. H. Shoemaker and G. Melillo, *Cancer Research*, 2005, **65**, 9047-9055.
46. Y. Guo, C. L. Partch, J. Key, P. B. Card, V. Pashkov, A. Patel, R. K. Bruick, H. Wurdak and K. H. Gardner, *ACS Chem. Biol.*, 2013, **8**, 626-635.
47. A. L. Kung, S. D. Zabludoff, D. S. France, S. J. Freedman, E. A. Tanner, A. Vieira, S. Cornell-Kennon, J. Lee, B. Wang, J. Wang, K. Memmert, H.-U. Naegeli, F. Petersen, M. J. Eck, K. W. Bair, A. W. Wood and D. M. Livingston, *Cancer Cell*, 2004, **6**, 33-43.
48. K. M. Cook, S. T. Hilton, J. Mecinovic, W. B. Motherwell, W. D. Figg and C. J. Schofield, *J. Biol. Chem.*, 2009, **284**, 26831-26838.

49. R. Dubey, M. D. Levin, L. Z. Szabo, C. F. Laszlo, S. Kushal, J. B. Singh, P. Oh, J. E. Schnitzer and B. Z. Olenyuk, *J. Am. Chem. Soc.*, 2013, **135**, 4537-4549.

50. M. K. Jayatunga, S. Thompson, T. C. McKee, M. C. Chan, K. M. Reece, A. P. Hardy, R. Sekirnik, P. T. Seden, K. M. Cook, J. B. McMahon, W. D. Figg, C. J. Schofield and A. D. Hamilton, *Eur. J. Med. Chem.*, 2015, **94**, 509-516.

51. C. Tan, R. G. de Noronha, N. S. Devi, A. A. Jabbar, S. Kaluz, Y. Liu, S. R. Mooring, K. C. Nicolaou, B. Wang and E. G. Van Meir, *Bioorg. Med. Chem. Lett.*, 2011, **21**, 5528-5532.

52. S. Yin, S. Kaluz, N. S. Devi, A. A. Jabbar, R. G. de Noronha, J. Mun, Z. Zhang, P. R. Boreddy, W. Wang, Z. Wang, T. Abbruscato, Z. Chen, J. J. Olson, R. Zhang, M. M. Goodman, K. C. Nicolaou and E. G. Van Meir, *Clinical Cancer Research*, 2012, **18**, 6623-6633.

53. G. M. Burslem, H. F. Kyle, A. L. Breeze, T. A. Edwards, A. Nelson, S. L. Warriner and A. J. Wilson, *ChemBioChem*, 2014, **15**, 1083-1087.

54. L. K. Henchey, S. Kushal, R. Dubey, R. N. Chapman, B. Z. Olenyuk and P. S. Arora, *J. Am. Chem. Soc.*, 2009, **132**, 941-943.

55. S. Kushal, B. B. Lao, L. K. Henchey, R. Dubey, H. Mesallati, N. J. Traaseth, B. Z. Olenyuk and P. S. Arora, *Proc. Natl. Acad. Sci. U.S.A.*, 2013.

56. B. B. Lao, I. Grishagin, H. Mesallati, T. F. Brewer, B. Z. Olenyuk and P. S. Arora, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 7531-7536.

57. B. B. Lao, K. Drew, D. A. Guarracino, T. F. Brewer, D. W. Heindel, R. Bonneau and P. S. Arora, *J. Am. Chem. Soc.*, 2014, **136**, 7877-7888.

58. B. J. Lamphear, R. Kirchweger, T. Skern and R. E. Rhoads, *J. Biol. Chem.*, 1995, **270**, 21975-21983.

59. N. Sonenberg and A.-C. Gingras, *Curr. Opin. Cell Biol.*, 1998, **10**, 268-275.

60. J. Pelletier and N. Sonenberg, *Nature*, 1988, **334**, 320-325.

61. M. LÓPEZ-LASTRA, A. RIVAS and M. I. BARRÍA, *Biological Research*, 2005, **38**, 121-146.

62. J. Marcotrigiano, A.-C. Gingras, N. Sonenberg and S. K. Burley, *Molecular Cell*, 1999, **3**, 707-716.

63. S. Mader, H. Lee, A. Pause and N. Sonenberg, *Mol. Cell. Biol.*, 1995, **15**, 4990-4997.

64. A. Bah, R. M. Vernon, Z. Siddiqui, M. Krzeminski, R. Muhandiram, C. Zhao, N. Sonenberg, L. E. Kay and J. D. Forman-Kay, *Nature*, 2015, **519**, 106-109.

65. C. Igreja, D. Peter, C. Weiler and E. Izaurralde, *Nat Commun*, 2014, **5**.

66. Y. Mamane, E. Petroulakis, L. Rong, K. Yoshida, L. W. Ler and N. Sonenberg, *Oncogene*, 0000, **23**, 3172-3179.

67. J. R. Graff, B. W. Konicek, J. H. Carter and E. G. Marcusson, *Cancer Research*, 2008, **68**, 631-634.

68. C. J. Brown, C. S. Verma, M. D. Walkinshaw and D. P. Lane, *Cell Cycle*, 2009, **8**, 1905-1911.

69. C. J. Brown, I. McNae, P. M. Fischer and M. D. Walkinshaw, *J. Mol. Biol.*, 2007, **372**, 7-15.

70. A. Niedzwiecka, J. Marcotrigiano, J. Stepinski, M. Jankowska-Anyszka, A. Wyslouch-Cieszynska, M. Dadlez, A.-C. Gingras, P. Mak, E. Darzynkiewicz, N. Sonenberg, S. K. Burley and R. Stolarski, *J. Mol. Biol.*, 2002, **319**, 615-635.

71. E. Papadopoulos, S. Jenni, E. Kabha, K. J. Takrouri, T. Yi, N. Salvi, R. E. Luna, E. Gavathiotis, P. Mahalingam, H. Arthanari, R. Rodriguez-Mias, R.

Yefidoff-Freedman, B. H. Aktas, M. Chorev, J. A. Halperin and G. Wagner, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, E3187-E3195.

72. C. J. Brown, J. J. Lim, T. Leonard, H. C. A. Lim, C. S. B. Chia, C. S. Verma and D. P. Lane, *J. Mol. Biol.*, 2011, **405**, 736-753.

73. L. Beretta, A.-C. Gingras, Y. V. Svitkin, M. N. Hall and N. Sonenberg, *The EMBO Journal*, 1996, **15**, 658-664.

74. S. Faivre, G. Kroemer and E. Raymond, *Nat Rev Drug Discov*, 2006, **5**, 671-688.

75. C. M. Chresta, B. R. Davies, I. Hickson, T. Harding, S. Cosulich, S. E. Critchlow, J. P. Vincent, R. Ellston, D. Jones, P. Sini, D. James, Z. Howard, P. Dudley, G. Hughes, L. Smith, S. Maguire, M. Hummersone, K. Malagu, K. Menear, R. Jenkins, M. Jacobsen, G. C. M. Smith, S. Guichard and M. Pass, *Cancer Research*, 2010, **70**, 288-298.

76. A. Kentsis, I. Topisirovic, B. Culjkovic, L. Shao and K. L. B. Borden, *Proc. Natl. Acad. Sci. U.S.A.*, 2004, **101**, 18105-18110.

77. T. P. Herbert, R. Fåhraeus, A. Prescott, D. P. Lane and C. G. Proud, *Current Biology*, 2000, **10**, 793-796.

78. D. Lama, S. T. Quah, C. S. Verma, R. Lakshminarayanan, R. W. Beuerman, D. P. Lane and C. J. Brown, *Sci. Rep.*, 2013, **3**.

79. W. Zhou, S. T. Quah, C. S. Verma, Y. Liu, D. P. Lane and C. J. Brown, *PLoS ONE*, 2012, **7**, e47235.

80. N. J. Moerke, H. Aktas, H. Chen, S. Cantel, M. Y. Reibarkh, A. Fahmy, J. D. Gross, A. Degterev, J. Yuan, M. Chorev, J. A. Halperin and G. Wagner, *Cell*, 2007, **128**, 257-267.

81. P. Mahalingam, K. Takrouri, T. Chen, R. Sahoo, E. Papadopoulos, L. Chen, G. Wagner, B. H. Aktas, J. A. Halperin and M. Chorev, *J. Med. Chem.*, 2014, **57**, 5094-5111.

82. R. Cencic, D. R. Hall, F. Robert, Y. Du, J. Min, L. Li, M. Qui, I. Lewis, S. Kurtkaya, R. Dingledine, H. Fu, D. Kozakov, S. Vajda and J. Pelletier, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 1046-1051.

83. J. B. Baell and G. A. Holloway, *J. Med. Chem.*, 2010, **53**, 2719-2740.

84. J. B. Baell, L. Ferrins, H. Falk and G. Nikolakopoulos, *Aust. J. Chem.*, 2013, **66**, 1483-1494.

85. Y. Suzuki, M. Minami, M. Suzuki, K. Abe, S. Zenno, M. Tsujimoto, K. Matsumoto and Y. Minami, *J. Biol. Chem.*, 2009, **284**, 35597-35604.

86. I. D. Campbell, *Biophysical techniques*, Oxford University Press, Oxford; New York, 2012.

87. J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards and A. J. Wilson, *Chem Commun (Camb)*, 2009, 5091-5093.

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

89. 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.

90. A. Barnard, J. A. Miles, G. M. Burslem, A. M. Barker and A. J. Wilson, *Org. Biomol. Chem.*, 2015, **13**, 258-264.

91. R. E. Moellering, M. Cornejo, T. N. Davis, C. D. 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.

92. T. Heyduk, Y. Ma, H. Tang and R. H. Ebright, in *Methods Enzymol.*, ed. A. Sankar, Academic Press, 1996, vol. Volume 274, pp. 492-503.

93. R. Z. Cer, U. Mudunuri, R. Stephens and F. J. Lebeda, *Nucleic Acids Research*, 2009, **37**, W441-W445.
94. J. Chen, N. Sawyer and L. Regan, *Protein Sci.*, 2013, **22**, 510-515.
95. W. Guo, J. A. Wieniewski and H. Ji, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 3652.
96. S. Surade and Tom L. Blundell, *Chemistry & Biology*, 2012, **19**, 42-50.
97. J. T. Ernst, J. Becerril, H. S. Park, H. Yin and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2003, **42**, 535-539.
98. J. M. Rodriguez, N. T. Ross, W. P. Katt, D. Dhar, G.-i. Lee and A. D. Hamilton, *ChemMedChem*, 2009, **4**, 649-656.
99. S. Melkko, C. E. Dumelin, J. Scheuermann and D. Neri, *Chemistry & Biology*, **13**, 225-231.
100. W. B. Turnbull, B. L. Precious and S. W. Homans, *J. Am. Chem. Soc.*, 2004, **126**, 1047-1054.
101. Z. E. Reinert, E. D. Musselman, A. H. Elcock and W. S. Horne, *ChemBioChem*, 2012, **13**, 1107-1111.
102. M. Kanai, K. H. Mortell and L. L. Kiessling, *J. Am. Chem. Soc.*, 1997, **119**, 9931-9932.
103. C. Jeppesen, J. Y. Wong, T. L. Kuhl, J. N. Israelachvili, N. Mullah, S. Zalipsky and C. M. Marques, *Science*, 2001, **293**, 465-468.
104. S. Oelmeier, F. Dismer and J. Hubbuch, *BMC Biophysics*, 2012, **5**, 14.
105. F. Kienberger, V. P. Pastushenko, G. Kada, H. J. Gruber, C. Riener, H. Schindler and P. Hinterdorfer, *Single Molecules*, 2000, **1**, 123-128.
106. D. J. Craik, D. P. Fairlie, S. Liras and D. Price, *Chemical Biology & Drug Design*, 2013, **81**, 136-147.
107. G. Klebe, *Nat Rev Drug Discov*, 2015, **14**, 95-110.
108. S. A. Dames, M. Martinez-Yamout, R. N. De Guzman, H. J. Dyson and P. E. Wright, *Proc Natl Acad Sci U S A*, 2002, **99**, 5271-5276.
109. A. Modrak-Wojcik, M. Gorka, K. Niedzwiecka, K. Zdanowski, J. Zuberek, A. Niedzwiecka and R. Stolarski, *FEBS Lett.*, 2013.
110. W. L. Jorgensen, *Science*, 2004, **303**, 1813-1818.
111. B. K. Shoichet, S. L. McGovern, B. Wei and J. J. Irwin, *Curr. Opin. Chem. Biol.*, 2002, **6**, 439-446.
112. G. Jones and P. Willett, *Curr. Opin. Biotechnol.*, 1995, **6**, 652-656.
113. G. Schneider and U. Fechner, *Nat Rev Drug Discov*, 2005, **4**, 649-663.
114. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J. Med. Chem.*, 2004, **47**, 1739-1749.
115. T. A. Halgren, R. B. Murphy, R. A. Friesner, H. S. Beard, L. L. Frye, W. T. Pollard and J. L. Banks, *J. Med. Chem.*, 2004, **47**, 1750-1759.
116. P. C. D. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson and M. T. Stahl, *J. Chem. Inf. Mod.*, 2010, **50**, 572-584.
117. *LigPrep version 2.9*, (2014) Schrödinger, LLC, New York.
118. N. Sauton, D. Lagorce, B. Villoutreix and M. Miteva, *BMC Bioinformatics*, 2008, **9**, 184.
119. S.-Y. Yang, *Drug Discovery Today*, 2010, **15**, 444-450.
120. G. Schneider, W. Neidhart, T. Giller and G. Schmid, *Angew. Chem. Int. Ed.*, 1999, **38**, 2894-2896.
121. P. C. D. Hawkins, A. G. Skillman and A. Nicholls, *J. Med. Chem.*, 2006, **50**, 74-82.

122. E. Ko, A. Raghuraman, L. M. Perez, T. R. Ioerger and K. Burgess, *J. Am. Chem. Soc.*, 2012, **135**, 167-173.
123. D. Xin, E. Ko, L. M. Perez, T. R. Ioerger and K. Burgess, *Org. Biomol. Chem.*, 2013, **11**, 7789-7801.
124. *US Pat.*, US 20100144805 A1, 2010.
125. S. Zheng, Q. Zhong, Q. Jiang, M. Mottamal, Q. Zhang, N. Zhu, M. E. Burow, R. A. Worthylake and G. Wang, *ACS Med. Chem. Lett.*, 2013, **4**, 191-196.
126. H. M. Meshram, D. A. Kumar and B. R. Vara Prasad, *Synth. Commun.*, 2009, **39**, 2317-2320.
127. N. Miyaura and A. Suzuki, *Chem. Rev.*, 1995, **95**, 2457-2483.
128. S. M. Devine, M. D. Mulcair, C. O. Debono, E. W. W. Leung, J. W. M. Nissink, S. S. Lim, I. R. Chandrashekaran, M. Vazirani, B. Mohanty, J. S. Simpson, J. B. Baell, P. J. Scammells, R. S. Norton and M. J. Scanlon, *J. Med. Chem.*, 2015, **58**, 1205-1214.
129. *Pipeline Pilot v8.5.0.200*, (2011) Accelrys© Software Inc.
130. L. H. Sternbach, *J. Med. Chem.*, 1979, **22**, 1-7.
131. E. Vitaku, D. T. Smith and J. T. Njardarson, *J. Med. Chem.*, 2014, **57**, 10257-10274.
132. A. R. Katritzky, I. V. Shcherbakova, R. D. Tack and X.-Q. Dai, *Tetrahedron*, 1993, **49**, 3907-3918.
133. F. Fülöp, L. Simon, G. Simon-Talpas and G. Bernáth, *Synth. Commun.*, 1998, **28**, 2303-2309.
134. O. V. Larionov and E. J. Corey, *J. Am. Chem. Soc.*, 2008, **130**, 2954-2955.
135. D. B. Dess and J. C. Martin, *J. Org. Chem.*, 1983, **48**, 4155-4156.
136. A. Nadin, C. Hattotuwagama and I. Churcher, *Angew. Chem. Int. Ed.*, 2012, **51**, 1114-1122.
137. J.-L. Wang, D. Liu, Z.-J. Zhang, S. Shan, X. Han, S. M. Srinivasula, C. M. Croce, E. S. Alnemri and Z. Huang, *Proc. Natl. Acad. Sci. U.S.A.*, 2000, **97**, 7124-7129.
138. S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173-180.
139. B. P. Orner, J. T. Ernst and A. D. Hamilton, *J. Am. Chem. Soc.*, 2001, **123**, 5382-5383.
140. H. Yin and A. D. Hamilton, *Bioorg. Med. Chem. Lett.*, 2004, **14**, 1375-1379.
141. 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.
142. J. M. Rodriguez and A. D. Hamilton, *Tetrahedron Lett.*, 2006, **47**, 7443-7446.
143. J. M. Rodriguez and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2007, **46**, 8614-8617.
144. J. Plante, F. Campbell, B. Malkova, C. Kilner, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2008, **6**, 138-146.
145. F. Campbell, J. P. Plante, T. A. Edwards, S. L. Warriner and A. J. Wilson, *Org. Biomol. Chem.*, 2010, **8**, 2344-2351.
146. 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.
147. I. Saraogi, C. D. Incarvito and A. D. Hamilton, *Angew. Chem. Int. Ed. Engl.*, 2008, **47**, 9691-9694.

148. A. Barnard, K. Long, H. L. Martin, J. A. Miles, T. A. Edwards, D. C. Tomlinson, A. Macdonald and A. J. Wilson, *Angew. Chem. Int. Ed.*, 2015, **54**, 2960-2965.
149. B. Mensa, Y. H. Kim, S. Choi, R. Scott, G. A. Caputo and W. F. DeGrado, *Antimicrob. Agents Chemother.*, 2011, **55**, 5043-5053.
150. F. Lu, S.-W. Chi, D.-H. Kim, K.-H. Han, I. D. Kuntz and R. K. Guy, *J. Comb. Chem.*, 2006, **8**, 315-325.
151. T. Shahian, G. M. Lee, A. Lazic, L. A. Arnold, P. Velusamy, C. M. Roels, R. K. Guy and C. S. Craik, *Nat. Chem. Biol.*, 2009, **5**, 640-646.
152. 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.
153. T.-K. Lee and J.-M. Ahn, *ACS Combinatorial Science*, 2010, **13**, 107-111.
154. 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.
155. M. J. Adler, R. T. W. Scott and A. D. Hamilton, *Chem. Eur. J.*, 2012, **18**, 12974-12977.
156. M. J. Adler and A. D. Hamilton, *J. Org. Chem.*, 2011, **76**, 7040-7047.
157. R. P. Cheng, S. H. Gellman and W. F. DeGrado, *Chem. Rev.*, 2001, **101**, 3219-3232.
158. L. A. Pilsl and O. Reiser, *Amino Acids*, 2011, **41**, 709-718.
159. J. Frackenpohl, P. I. Arvidsson, J. V. Schreiber and D. Seebach, *ChemBioChem*, 2001, **2**, 445-455.
160. E. A. Harker, D. S. Daniels, D. A. Guarracino and A. Schepartz, *Biorg. Med. Chem.*, 2009, **17**, 2038-2046.
161. L. D. Walensky and G. H. Bird, *J. Med. Chem.*, 2014.
162. D. Wang, W. Liao and P. S. Arora, *Angew. Chem. Int. Ed.*, 2005, **44**, 6525-6529.
163. A. Glas, D. Bier, G. Hahne, C. Rademacher, C. Ottmann and T. N. Grossmann, *Angew. Chem. Int. Ed.*, 2014, **53**, 2489-2493.
164. A. Patgiri, A. L. Jochim and P. S. Arora, *Acc. Chem. Res.*, 2008, **41**, 1289-1300.
165. C. E. Schafmeister, J. Po and G. L. Verdine, *J. Am. Chem. Soc.*, 2000, **122**, 5891-5892.
166. H. E. Blackwell and R. H. Grubbs, *Angew. Chem. Int. Ed.*, 1998, **37**, 3281-3284.
167. 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.
168. S. J. Freedman, Z. Y. Sun, F. Poy, A. L. Kung, D. M. Livingston, G. Wagner and M. J. Eck, *Proc Natl Acad Sci U S A*, 2002, **99**, 5367-5372.
169. F. Lu, S.-W. Chi, D.-H. Kim, K.-H. Han, I. D. Kuntz and R. K. Guy, *J. Comb. Chem.*, 2006, **8**, 315-325.
170. 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.
171. J. L. Yap, X. Cao, K. Vanommeslaeghe, K. Y. Jung, C. Peddaboina, P. T. Wilder, A. Nan, A. D. MacKerell, Jr., W. R. Smythe and S. Fletcher, *Org Biomol Chem*, 2012, **10**, 2928-2933.
172. J. C. Fuller, R. M. Jackson, T. A. Edwards, A. J. Wilson and M. R. Shirts, *PLoS ONE*, 2012, **7**, e43253.
173. I. Azumaya, T. Okamoto, F. Imabeppu and H. Takayanagi, *Tetrahedron*, 2003, **59**, 2325-2331.

174. O. V. Kulikov, C. Incarvito and A. D. Hamilton, *Tetrahedron Lett.*, 2011, **52**, 3705-3709.
175. M. Colombo, S. Bossolo and A. Aramini, *J. Comb. Chem.*, 2009, **11**, 335-337.
176. R. L. Farmer, M. M. Biddle, A. E. Nibbs, X. Huang, R. C. Bergan and K. A. Scheidt, *ACS Med. Chem. Lett.*, 2010, **1**, 400-405.
177. J. Mun, A. A. Jabbar, N. S. Devi, Y. Liu, E. G. Van Meir and M. M. Goodman, *Biorg. Med. Chem.*, 2012, **20**, 4590-4597.
178. K. Long, T. A. Edwards and A. J. Wilson, *Bioorg. Med. Chem.*, 2013, **21**, 4034-4040.
179. C. B. Cooley, B. M. Trantow, F. Nederberg, M. K. Kiesewetter, J. L. Hedrick, R. M. Waymouth and P. A. Wender, *J. Am. Chem. Soc.*, 2009, **131**, 16401-16403.
180. J.-M. Ahn and S.-Y. Han, *Tetrahedron Lett.*, 2007, **48**, 3543-3547.
181. A. Sharma and J. F. Hartwig, *Nature*, 2015, **517**, 600-604.
182. J. L. Jeffrey and R. Sarpong, *Chem. Sci.*, 2013, **4**, 4092-4106.
183. M. Honda, H. Morita and I. Nagakura, *J. Org. Chem.*, 1997, **62**, 8932-8936.
184. O. Mitsunobu, *Synthesis*, 1981, **1981**, 1-28.
185. D. S. Palacios, I. Dailey, D. M. Siebert, B. C. Wilcock and M. D. Burke, *Proc. Natl. Acad. Sci. U.S.A.*, 2011, **108**, 6733-6738.
186. G. W. Preston and A. J. Wilson, *Chem. Soc. Rev.*, 2013, **42**, 3289-3301.
187. G. W. Preston, S. E. Radford, A. E. Ashcroft and A. J. Wilson, *ACS Chem. Biol.*, 2013, **9**, 761-768.
188. J. T. Bush, L. J. Walport, J. F. McGouran, I. K. H. Leung, G. Berridge, S. S. van Berkel, A. Basak, B. M. Kessler and C. J. Schofield, *Chem. Sci.*, 2013, **4**, 4115-4120.
189. D. P. Smith, J. Anderson, J. Plante, A. E. Ashcroft, S. E. Radford, A. J. Wilson and M. J. Parker, *Chem. Commun.*, 2008, 5728-5730.
190. K.-H. Altmann, B. Johannes, H. Kessler, F. Diederich, B. Krautler, S. Lippard, R. Liskamp, K. Muller, E. M. Nolan, B. Samori, G. Schneider, S. L. Schreiber, H. Schwalbe, C. Toniolo, C. A. A. van Boeckel, H. Waldmann and C. T. Walsh, *ChemBioChem*, 2009, **10**, 16-29.
191. J. M. Chalker, L. Lercher, N. R. Rose, C. J. Schofield and B. G. Davis, *Angew. Chem. Int. Ed.*, 2012, **51**, 1835-1839.
192. N. Timms, C. L. Windle, A. Polyakova, J. R. Ault, C. H. Trinh, A. R. Pearson, A. Nelson and A. Berry, *ChemBioChem*, 2013, **14**, 474-481.
193. A. A. Vinogradov, E. D. Evans and B. L. Pentelute, *Chem. Sci.*, 2015.
194. U. Arnold, M. P. Hinderaker, B. L. Nilsson, B. R. Huck, S. H. Gellman and R. T. Raines, *J. Am. Chem. Soc.*, 2002, **124**, 8522-8523.
195. U. Arnold, M. P. Hinderaker, J. Köditz, R. Golbik, R. Ulbrich-Hofmann and R. T. Raines, *J. Am. Chem. Soc.*, 2003, **125**, 7500-7501.
196. G. J. Williams, T. Woodhall, L. M. Farnsworth, A. Nelson and A. Berry, *J. Am. Chem. Soc.*, 2006, **128**, 16238-16247.
197. A. Tam, U. Arnold, M. B. Soellner and R. T. Raines, *J. Am. Chem. Soc.*, 2007, **129**, 12670-12671.
198. R. David, R. Günther, L. Baumann, T. Lühmann, D. Seebach, H.-J. Hofmann and A. G. Beck-Sickinger, *J. Am. Chem. Soc.*, 2008, **130**, 15311-15317.
199. Z. E. Reinert, G. A. Lengyel and W. S. Horne, *J. Am. Chem. Soc.*, 2013, **135**, 12528-12531.
200. B. Eckhardt, W. Grosse, L.-O. Essen and A. Geyer, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 18336-18341.

201. C. Mayer, M. M. Müller, S. H. Gellman and D. Hilvert, *Angew. Chem. Int. Ed.*, 2014, **53**, 6978-6981.
202. V. P. Kumar, M. Narender, R. Sridhar, Y. V. D. Nageswar and K. R. Rao, *Synth. Commun.*, 2007, **37**, 4331-4336.
203. G. Sheldrick, *Acta Crystallographica Section A*, 2008, **64**, 112-122.