Chemical biology tools to explore the activities and mechanisms of Class I HDAC corepressor complexes

Thesis submitted for the degree of

Doctor of Philosophy

at the University of Leicester

by

Naomi Stephanie Robertson MSci (Glasgow)

Department of Chemistry

University of Leicester

May 2016

For Dad

(1928-2015)

Abstract

Chemical biology tools to explore the activities and mechanisms of Class I HDAC corepressor complexes

Naomi Stephanie Robertson MSci (University of Glasgow)

Isoform selective histone deacetylase (HDAC) inhibitors are urgently required to investigate HDAC isoform phenotypes and provide targeted therapeutics. Class 1 HDACs (1, 2 and 3) are recruited into large multi-subunit co-repressor complexes for maximal activity. However, little attention has been paid to targeting the disruption of these protein-protein interactions (PPIs) and so represents a novel approach to develop more specific therapeutics and chemical probes.

This thesis describes the design, synthesis and evaluation of stapled peptide analogues of the SMRT-DAD protein. A systematic, structure based approach was taken to design the second generation stapled peptides containing phosphotyrosine residues. These were shown to be super activators of the HDAC3:SMRT-DAD corepressor complex.

The synthesis of a hydroxamic acid amino acid is also described and an improvement on the current synthetic scheme is discussed. A novel hydroxamic acid containing histone H4 tail peptide was then prepared which was shown to be an inhibitor of the HDAC1:MTA1 corepressor complex with an IC_{50} of 366 nM. A crystal structure of this inhibitor peptide bound in the active site of HDAC1 was also obtained.

The challenging synthesis of a series of histone H3 (1-21) tail sequence peptides has also been described. NMR experiments were carried out using these peptides to study both the demethylase and deacetylase activities of the ternary LSD1:CoREST:HDAC1. The deacetylase activity of HDAC1 has been shown to be much faster than the demethylase activity of LSD1. In order to obtain a crystal structure of the ternary LSD1:CoREST:HDAC1 complex and to determine the distance between the LSD1 and the HDAC1 components, three

dual peptide inhibitors based on SNAIL have been designed and synthesised. These crystallisation experiments and biological studies are in progress.

Acknowledgements

I would like to express my appreciation and thanks to my PhD supervisors Dr Andrew Jamieson and Prof John Schwabe. Without them, I would not have been able to work on this project. Their help, support and mentorship throughout the last four years has been invaluable. I would also like to thank Dr Peter Watson for taking me under his wing and teaching me biochemistry for the chemist. I wish to extend my thanks to the rest of the Jamieson and Schwabe groups for valued discussions and a good working atmosphere.

I would like to thank those who assisted in carrying out this research, including Dr Gerry Griffith for his help with NMR, Mick Lee for assistance with mass spectrometry and Kuldip Singh for X-ray crystallography.

A special thanks to my family for all the ongoing support. Most of all I would like to thank my loving, supportive, encouraging, and patient husband Craig. He was always there to cheer me up in the moments when the PhD was getting to me. I could not have done this without him!

List of abbreviations

˚C	Degrees Celcius
µL	Microliter
µM	Micromolar
Å	Angstrom
All	Allyl
Boc	*Tert*-butyloxycarbonyl
CD	Circular dichroism
COSY	Correlation spectroscopy
d	Doublet
DCC	*N*,*N*'-Dicyclohexylcarbodiimide
DCM	Dichloromethane
dd	doublet of doublets
ddt	doublet of doublet of triplets
DIC	*N*,*N*'-Diisopropylcarbodiimide
DIPEA	*N*,*N*'-diisopropylethyl amine (Hunig's base)
DMAP	4-(Dimethylamino)pyridine
DMF	*N*,*N'*-dimethylformamide
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dt	doublet of triplets
Fmoc	Fluorenylmethyloxycarbonyl
h	Hours
HBTU	*O*-(Benzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate
HCTU	*O*-(6-Chlorobenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate
HDAC	Histone deacetylase
HDACi	Histone deacetylase inhibitor
HPLC	High performance liquid chromatography
HRMS-ESI	High resolution mass spectrometry – electrospray ionisation
J	Coupling constant
LC-MS	Liquid chromatography – mass spectrometry
M	Molar
M	Multiplet
min	Minutes
mmol	Millimoles
mol	Moles
mp	Melting point

nL	Nanolitre
nm	Nanometre
nM	Nanomolar
NMP	*N*-methylpyrrolidine
NMR	Nuclear magnetic resonance
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl
PEG	Polyethylene glycol
ppm	Parts per million
q	Quartet
quint	Quintet
R_8	(*R*)-2-(4-pentenyl)alanine
S_5	(*S*)-2-(4-pentenyl)alanine
SPPS	Solid phase peptide synthesis
t	Triplet
^{t}Bu	*Tert*-butyl
TFA	Trifluoroacetic acid
TFE	Trifluoroethanol
THF	Tetrahydrofuran
Trt	Trityl
Δ, δ	Delta

Contents

Chapter 1: Introduction

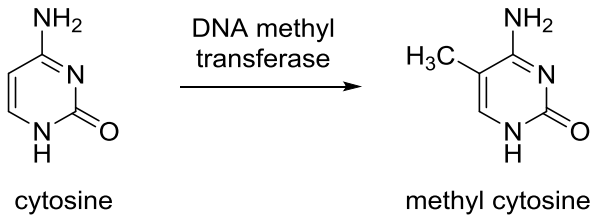
1.1. Epigenetics

The term epigenetics was first introduced by Conrad Waddington in the 1940s.[1] It can be described as the study of heritable phenotypical changes that happen during the cell cycle which is not a consequence of altering the DNA sequence.[2]

The life cycle of a butterfly can be thought of as an epigenetic phenomenon, specifically the metamorphosis from a caterpillar to a butterfly (Figure 1). The DNA of the caterpillar has not changed, but the changes in its phenotype to a butterfly have been caused by the difference in the expression of its genes.

Figure 1: The transformation from a caterpillar to a butterfly demonstrates an epigenetic phenomenon.

In epigenetics, there are two main molecular mechanisms involved. In humans, the most well studied epigenetic modification is DNA methylation or more specifically the methylation of the base cytosine by DNA methyl transferases (Dnmts) (Scheme 1). This has the effect of preventing transcription as the methylated DNA is a less efficient substrate for transcription factors and can also have a high affinity to other repressor proteins.[3]



Scheme 1: DNA methylation of cytosine is achieved using DNA methyl transferases (Dnmts).

The second epigenetic mechanism involves the post-translational modification (PTM) of histone proteins and is much more dynamic than DNA methylation. Histones play an important role in packaging the DNA within cells. Each eukaryotic cell has approximately 1.6 metres of DNA packaged into the cell's nucleus tightly wrapped around a group of small basic proteins called histones. This complex formed by the DNA and histones is called chromatin. In fact, chromatin constitutes about half the mass of eukaryotic chromosomes.[4]

In the nucleosome, there are five major histone proteins (Figure 2). Four of these (H2A, H2B, H3 and H4) are associated with one another and known as the core histones. These core histones make up the histone octamer. There are two copies of each of the core histone proteins in the nucleosome to give the histone octamer. Wrapped around the histone octamer are approximately 146 base pairs of DNA in 1.7 superhelical coils to complete the nucleosome core.[5–8]

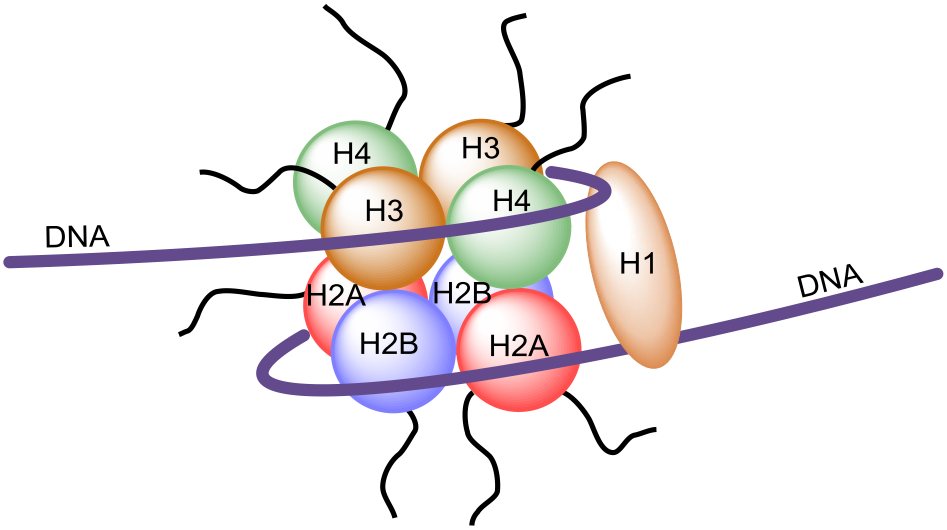


Figure 2: Structure of the nucleosome showing DNA wrapping around a histone octamer, with protruding histone tails coming from each histone protein.

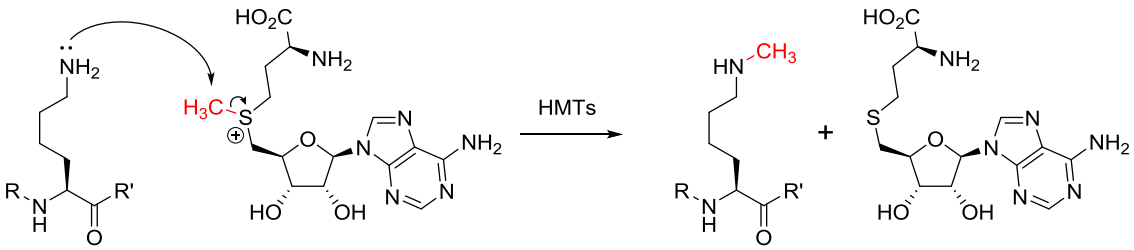
The other major histone protein is H1 and its variants, which are termed linker histones because they do not form part of the nucleosome core (Figure 2).[8] Additionally, the binding of H1 proteins to chromatin is known to be dynamic in nature.[9] However, it has been proposed that the H1 proteins are implicated in

the formation of higher order chromatin structures. This is done by stabilising the nucleosome core and helping the chromatin fold through the binding of H1 proteins to the nucleosomal dyad and the linker DNA entering and exiting the nucleosomal core.[9]

1.1.1. Post-translational modifications

Many different PTMs of histone proteins are possible, including methylation, phosphorylation,[10] ubiquitination, citrullination,[11,12] and acetylation.

The methylation of histone proteins occurs primarily on lysine or arginine residues.[13] This is achieved through an S_N2 reaction of the lysine or arginine residue with the cofactor *S*-adenosylmethionine (SAM), resulting in the formation of the *N*-methylated amino acid residue and *S*-adenosylhomocysteine (SAH) (Scheme 2). The methylation reaction is catalysed by histone methyl transferase (HMT) enzymes. Lysine residues can be mono-, di- or trimethylated, while arginine residues can be mono- or dimethylated and the dimethylation further divided into symmetrical and unsymmetrical dimethylation (Figure 3).[14,15,13] Methylation of lysine and arginine residues does not alter the charge of the histone proteins.[13]



Scheme 2: The mechanism of the SN2 methylation reaction of lysine residues on peptide chains is catalysed by histone methyl transferases (HMTs), using the co-factor *S*-adenosylmethionine, to give the methylated lysine residue and *S*-adenosylhomocysteine.

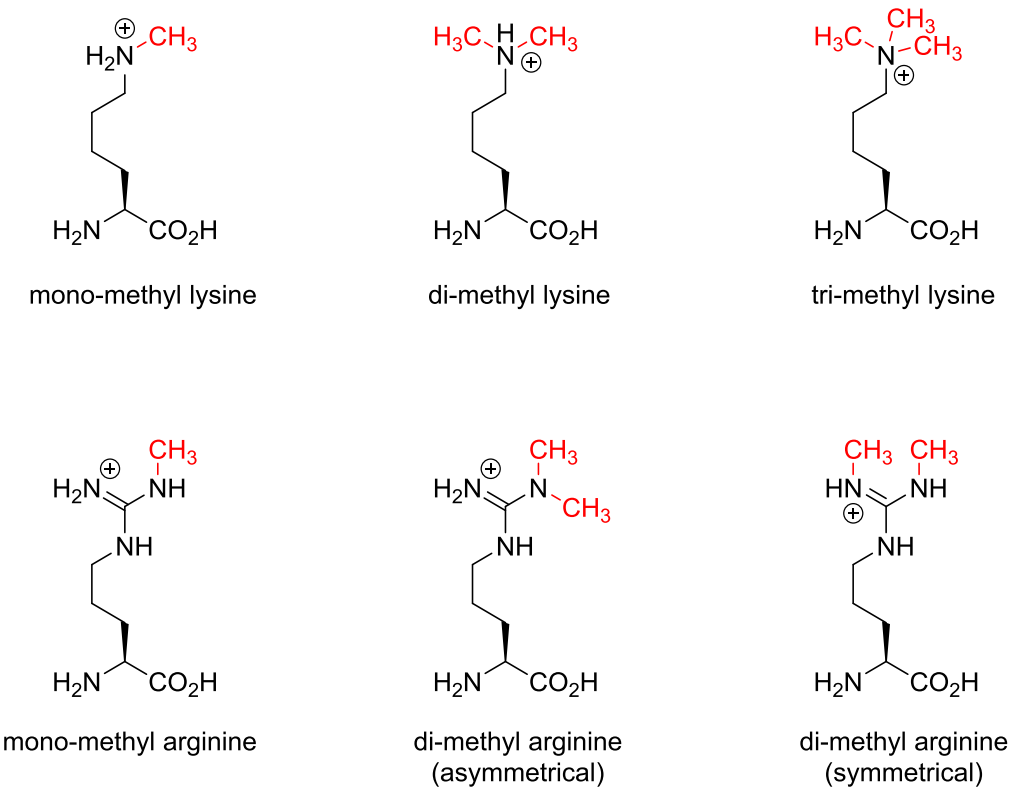
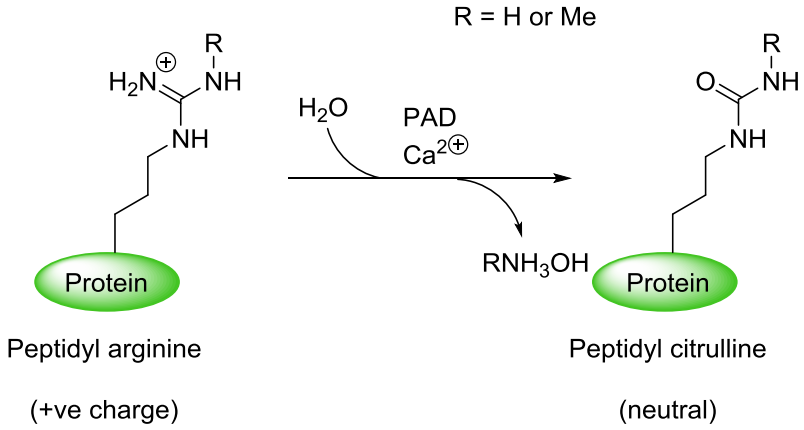


Figure 3: Lysine residues can be methylated up to three times, while arginine residues can be mono- or di-methylated, with the di-methylation being either asymmetric or symmetric.

Histone phosphorylation is a dynamic process controlled by kinases and phosphatases[13] and is involved in DNA damage repair pathways including non-homologous end joining, homologous recombination and replication-coupled DNA repair.[10] Phosphorylation occurs mainly on tyrosine, serine and threonine residues, but histidine phosphorylation is also known.[16]

Ubiquitination is one of the least understood of the histone modifications and occurs at lysine residues.[17] The most abundantly ubiquitinated proteins in the nucleus are H2A and H2B, and are generally monoubiquitinated.[18,19] Monoubiquitination is a signalling factor affecting cellular processes which include membrane trafficking and endocytosis.[20,21] However, polyubiquitination can also occur, with H2A known to have tetra- to hexaubiquitinated forms.[22] Where there are at least four ubiquitin subunits in a chain, polyubiquitination also leads to protein degradation as it can be recognised by the 26S proteasome.[23]

Arginine residues can also undergo deimination as a PTM. The deimination of arginine by peptidylarginine deiminases (PADs) leads to the transformation of the arginine ketimine group to a urea, forming the non-coded amino acid citrulline (Scheme 3). Thus, this PTM is also known as citrullination. Peptidyl monomethyl arginine residues can also undergo citrullination.[24,25] Arginine is positively charged at physiological pH, however, citrulline is not. Therefore, arginine citrullination can lead to a change in protein folding.



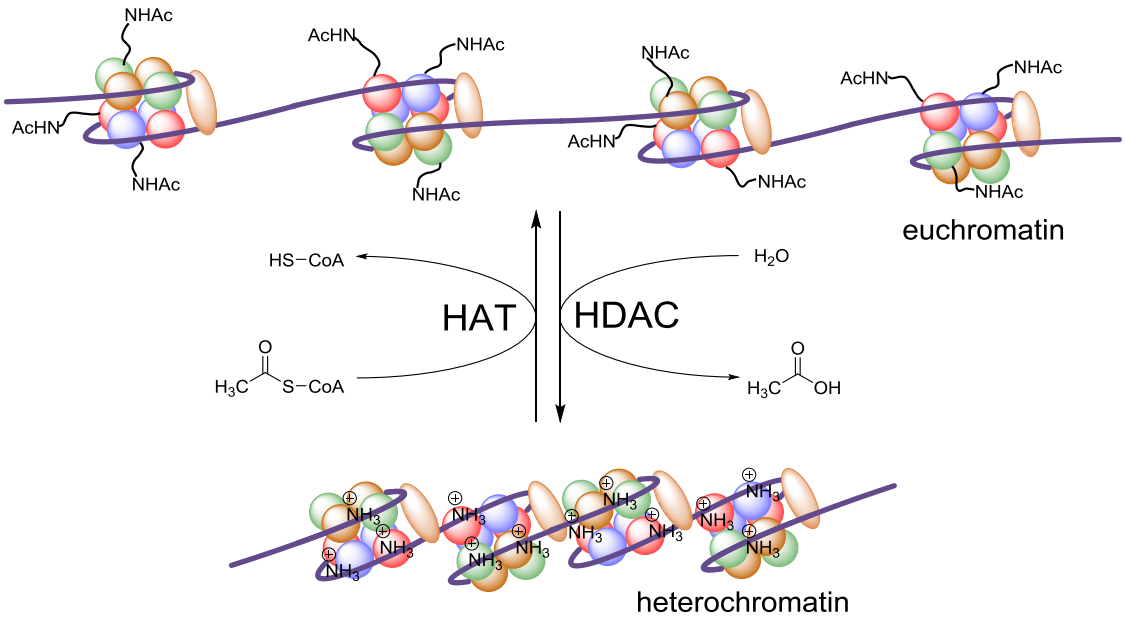
Scheme 3: Citrullination of arginine residues by peptidylarginine deiminases (PADs) leads to the formation of non-coded amino acid citrulline and a change in the overall charge of peptide.

1.1.2. Histone acetylation

Histone acetylation was first reported by Allfrey and co-workers in 1964[26] and is one of the most widespread PTMs. Acetylation can occur at either arginine or lysine residues, although lysine acetylation is the more common of the two. In fact the acetylation of lysine residues have been found at 3600 different sites, in 1750 different proteins, which include 61 histone proteins.[27] Acetylation of histones promote the recruitment of effector proteins, relaxation of chromatin and an increase in transcription. Acetylation is a very dynamic process regulated by the interplay between three classes of proteins that interact with acetylated proteins; histone acetyl transferases (HATs), bromodomains and

histone deacetylases (HDACs). HATs install the acetyl mark on the histone tails; bromodomains recognise the histone modifications and HDACs remove the epigenetic acetyl marks.

HATs catalyse the transfer of an acetyl group, from acetyl coenzyme A, onto the ε-amino group of a lysine residue on histone tails.[13] As a result of acetylation, the positive charge on the lysine residue is removed and so the residue becomes neutral. Thus, the chromatin forms a more opened and relaxed structure, termed euchromatin (Scheme 4). In this form, transcription factors can be recruited to the chromatin so that the DNA can be copied and gene expression can occur.



Scheme 4: The dynamic process by which histone tails on chromatin are acetylated, by histone acetyl transferases (HATs), and deactylated be histone deacetylases (HDACs)

Bromodomains consist of a sequence of about 110 amino acid residues that are part of much larger protein structures.[28,29] Indeed, there are 61 bromodomains across 46 unique proteins.[30] They act as the readers in the histone acetylation story recognising the acetylated lysine residues on histone tails and initiating the recruitment of transcription factors to the chromatin.[31]

HDAC enzymes catalyse the removal of acetyl groups from the lysine residues of histone tails (Scheme 4). This is achieved through the use of a water molecule buried within the active site of the HDAC enzyme to give an acetate side product. The deacetylation leaves a positively charged ammonium salt, which associates with the negatively charged phosphate groups on the DNA backbone creating a tightly packed chromatin structure, termed heterochromatin. As a result of the DNA being packed into the heterochromatin form, gene repression occurs.

The inhibition of HDAC enzymes is a useful tool in cancer research because HDAC inhibitors allow chromatin to be in its euchromatin form. This allows the expression of tumour suppressor genes, which leads to the slowing of cell division, DNA repair mechanisms to be activated and apoptosis to occur.[32]

1.2. Histone deacetylase enzymes

1.2.1. Function of HDACs

As their name suggests, HDACs catalyse the deacetylation of lysine residues on histone tails. However, it is known that HDACs do not only deacetylate histone proteins, but also many other lysine residues that are part of other protein structures. Thus, HDACs have also been described as lysine deacetylases (KDACs).[27,33]

1.2.2. HDACs as a superfamily

Currently, 18 mammalian HDAC enzymes have been identified[34–36] and, along with acetylpolyamine aminohydrolases and acetoin utilisation proteins, are members of an ancient protein superfamily known as the histone deacetylase superfamily.[37] HDACs themselves are subdivided into four main groups based on their homology to yeast proteins (Figure 4).[34,38]

Class I HDACS, which include HDAC1, HDAC2, HDAC3 and HDAC8 are homologous to Rpd3 in yeast.[35]

Class II HDACs are homologous to yeast Hda1 and can be further subdivided into two groups; Class IIa and Class IIb. In Class IIa, reside HDAC4, HDAC5, HDAC7 and HDAC9, which are all localised within the nucleus and cytoplasm of the cell, whereas Class IIa HDACs, HDAC6 and HDAC10 are localized mostly in the cell cytoplasm.[34]

Class III HDACs are the sirtuins which are homologous to Sir2 in yeast. There is only one Class IV HDAC, HDAC11, which has conserved residues in its catalytic domain shared by both the Class I and Class II HDACs.[34]

Class I and Class II HDACs are considered to be the classical HDACs and have zinc bound in their active site. HDAC11, the only known member of the Class IV HDACs also contain zinc. These three classes are inhibited by small molecule compounds such as Trichostatin A (TSA) (Figure 5).

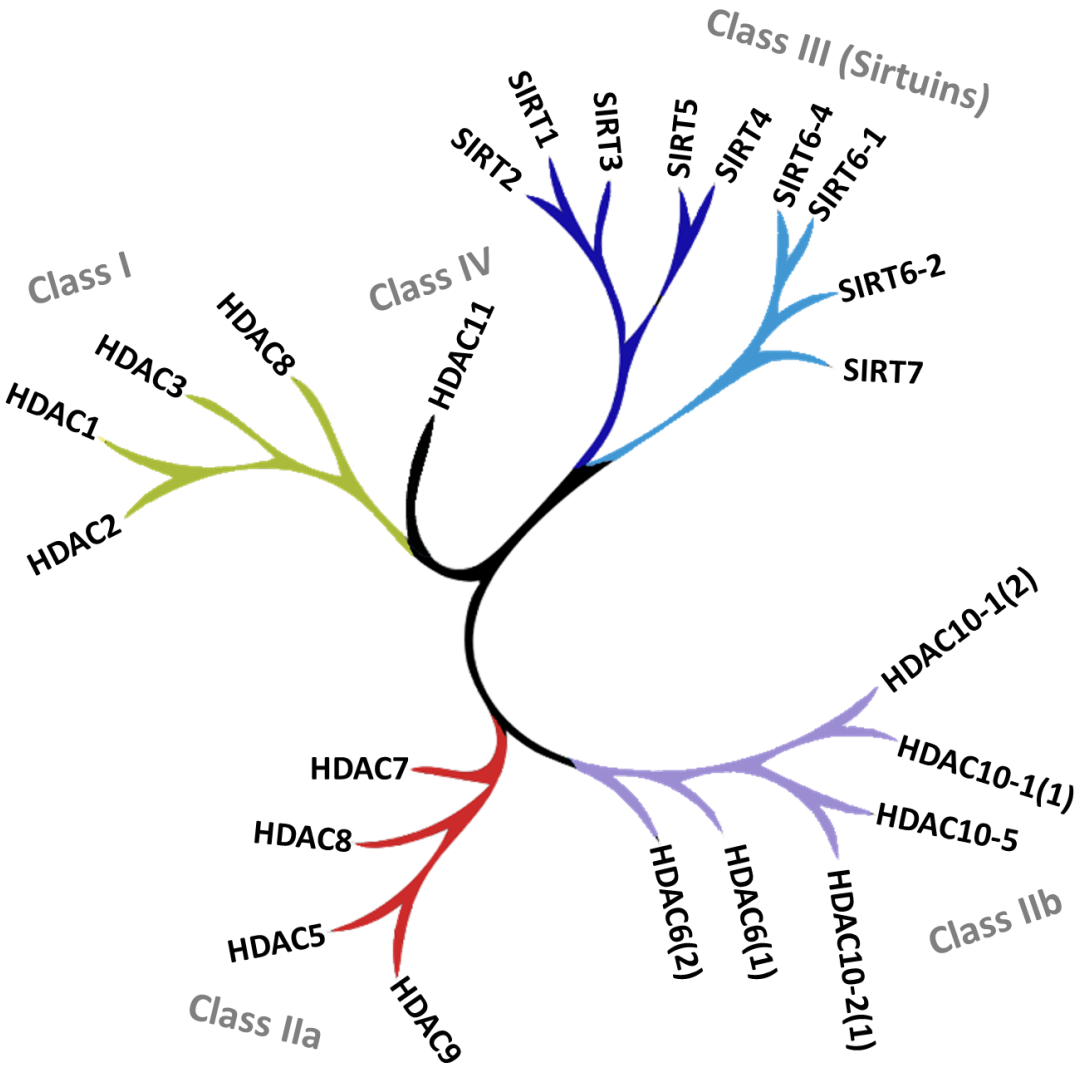


Figure 4: Phylogenetic tree of the HDAC family of proteins showing how they are subdivided into the different classes: Class I, IIa, IIb, III and IV (adapted from Arrowsmith *et al.*).[38]

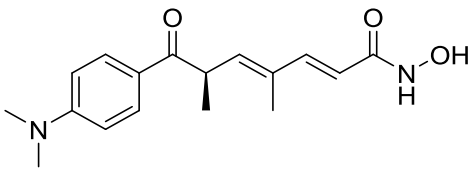


Figure 5: Structure of Trichostatin A, a known inhibitor of Class I, II and IV HDAC enzymes, through its binding with the zinc atom in the HDAC active site pocket.

Class I HDAC enzymes are known to form multi domain complexes with a variety of different proteins in order to be active (Figure 6).[39] The exception to the rule is HDAC8, which is the only HDAC enzyme known to be active without the need to recruit co-repressor proteins. HDAC1 and HDAC2 are fairly promiscuous enzymes, forming complexes with a variety of different proteins. Examples of which include the Sin3A, Nucleosome Remodelling and Deacetylase (NuRD) and CoREST complexes.[39]

On the contrary, HDAC3 is very specific and will only form a complex with SMRT.

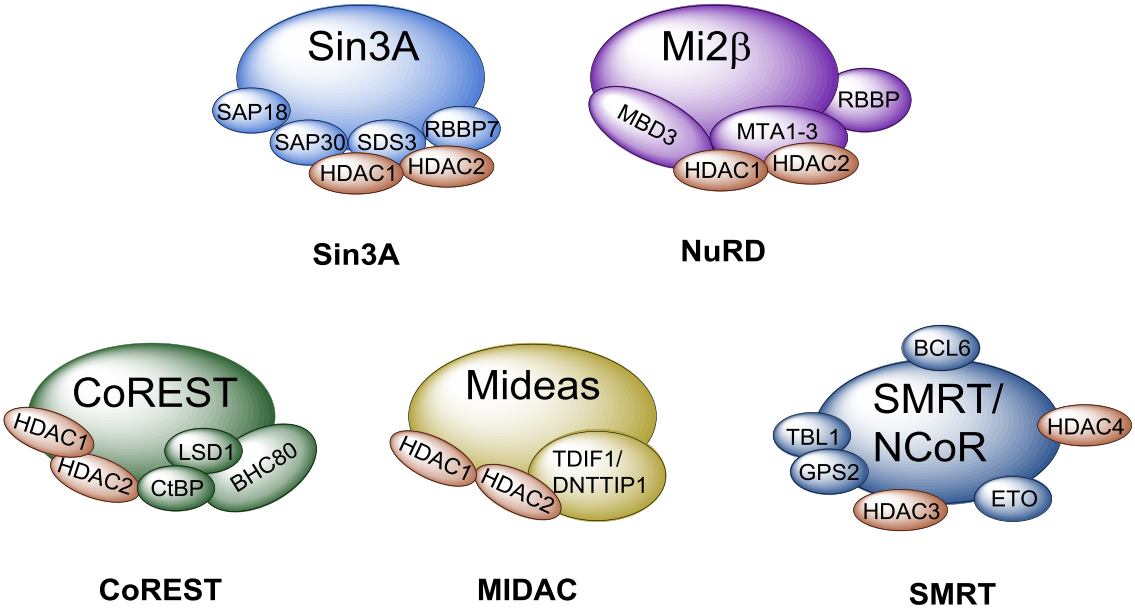


Figure 6: Schematic diagrams of complexes containing histone deacetylases. HDAC1 and 2 are located together in three co-repressor complexes: Sin3A, Nucleosome Remodelling and Deacetylase (NuRD) and CoREST complexes. HDAC3 only associates with SMRT complex. HDAC4 also associates with SMRT. (adapted from Watson *et al.*[39])

1.2.3. Structure of HDACs

Histone deacetylase-like protein (HDLP) was the first crystal structure elucidated for HDAC proteins in the late 90s.[40] It is an HDAC homologue from *Aquifex aeolicus*, a hyperthermophilic bacterium thought to be part of a group of some of the oldest species of bacteria. HDLP from *A. aeolicus* shares 35.2% identity with human HDAC1, which equates to over 375 conserved residues.[40]

Finnin and colleagues crystallised HDLP alone and with HDAC inhibitors SAHA and TSA bound in the active site (Figure 7).[40] As with all Class I HDAC proteins, buried deep in the active site of HDLP, is a zinc atom to which both SAHA and TSA coordinate to provide their inhibiting affinity.

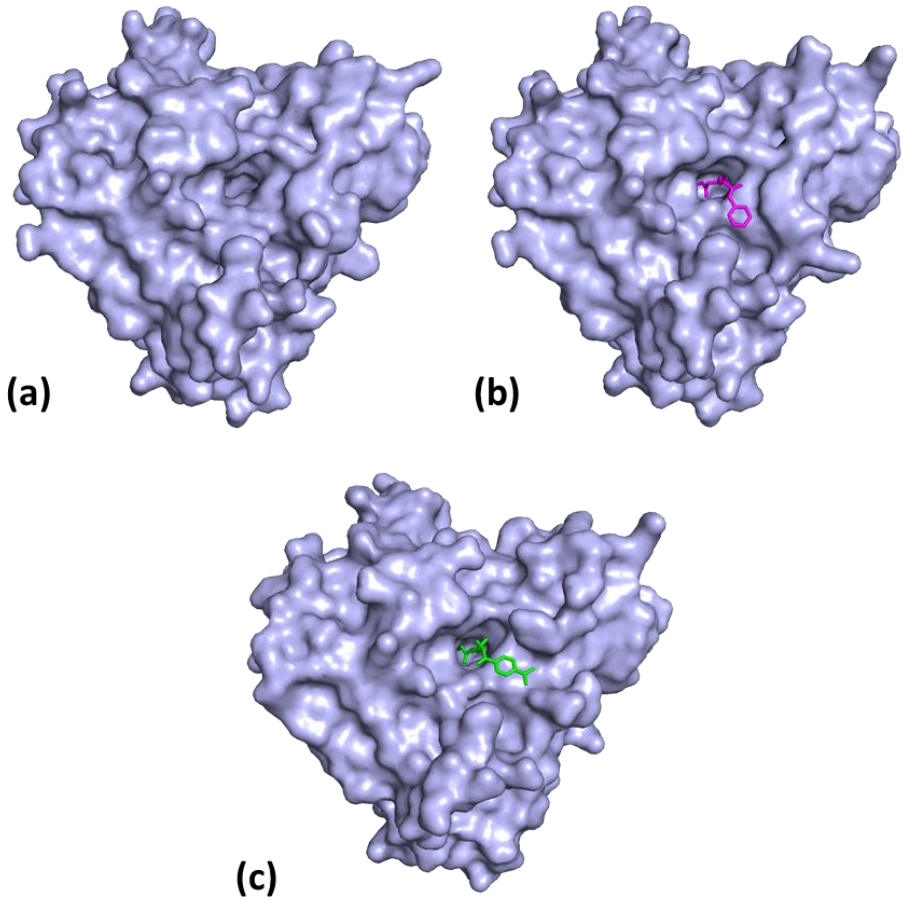


Figure 7: Crystal structures of (a) HDLP (pdb: 1C3P), (b) HDLP bound to SAHA (magenta) (pdb: 1C3S) and (c) HDLP bound to TSA (green) (pdb: 1C3R).[40]

The first human HDAC to be crystallized was HDAC8 (Figure 8). It was crystallized in 2004 by Somoza and colleagues, bound to a selection of HDAC inhibitors (TSA, SAHA, MS-344 and CRA-A). The binding of these inhibitors to HDAC8 showed the active site to be extremely malleable. This can be seen by the differences in the structure of HDAC8 around it's active site when inhibitors are bound (Figure 9).[41]

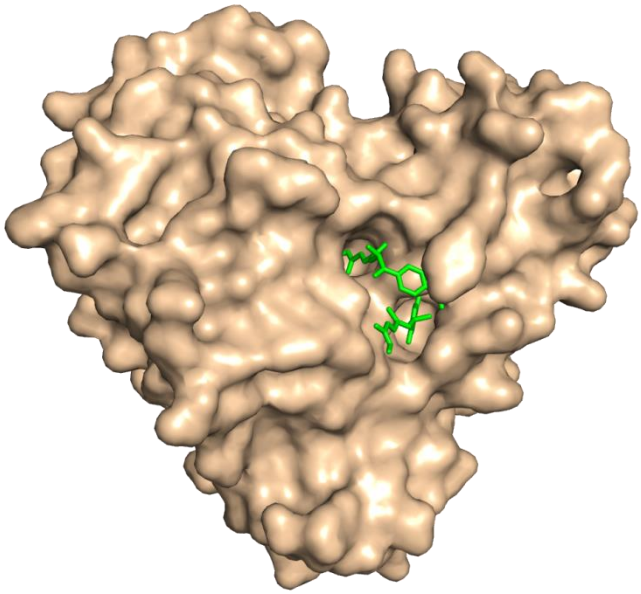


Figure 8: Crystal structure of human HDAC8 (surface) complexed with Trichostatin A (green sticks) (pdb: 1T64), which was the first human HDAC enzyme to be crystallized.[41]

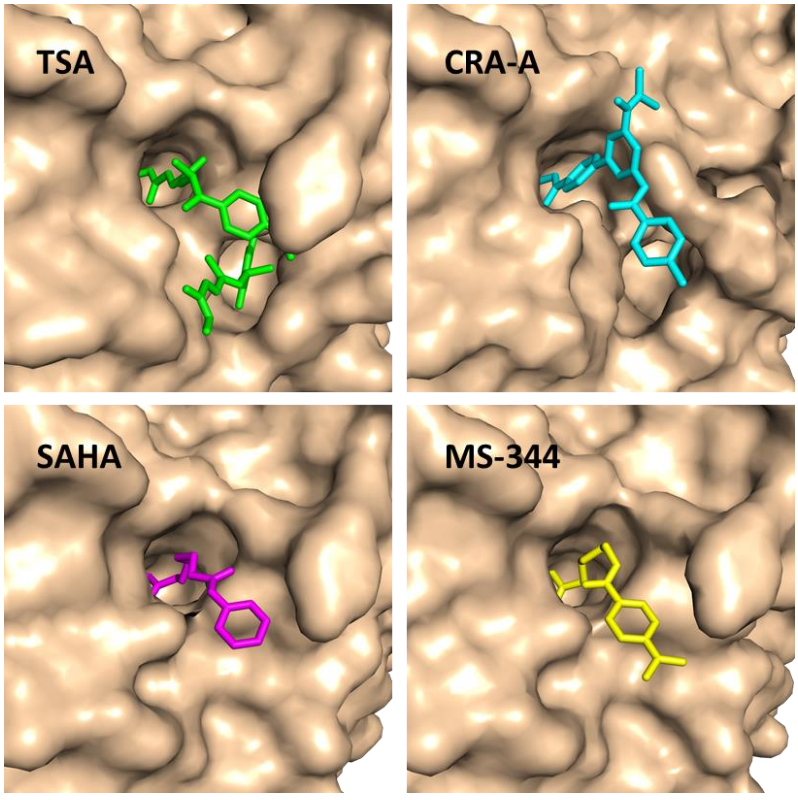


Figure 9: HDAC inhibitors TSA (green) (pdb: 1T64), CRA-A (cyan) (pdb: 1VKG), SAHA (magenta) (pdb: 1T69) and MS-344 (yellow) (pdb: 1T67), bound in the active site of HDAC8 (surface) showing the flexibility of the HDAC enzyme around the active site pocket when different inhibitors are bound.[41]

In comparison with HDLP, they found that the structure overlay of the α-carbons of HDAC8 and HDLP had a root-mean-square deviation of 2.2 Å, which is over 90% of the residues. However, even with this, the two proteins differ in their outer portions of their active sites.

In 2012, Watson and colleagues elucidated the structure of HDAC3 bound to its corepressor complex (deacetylase activation domain, DAD) from the human silencing mediator for retinoid or thyroid-hormone receptor (SMRT) corepressor (also called NCOR2) and a molecule of inositol tetraphosphate (D-myo-inositol-(1,4,5,6)-tetrakisphosphate, Ins(1,4,5,6)P_4) (Figure 10).[42]

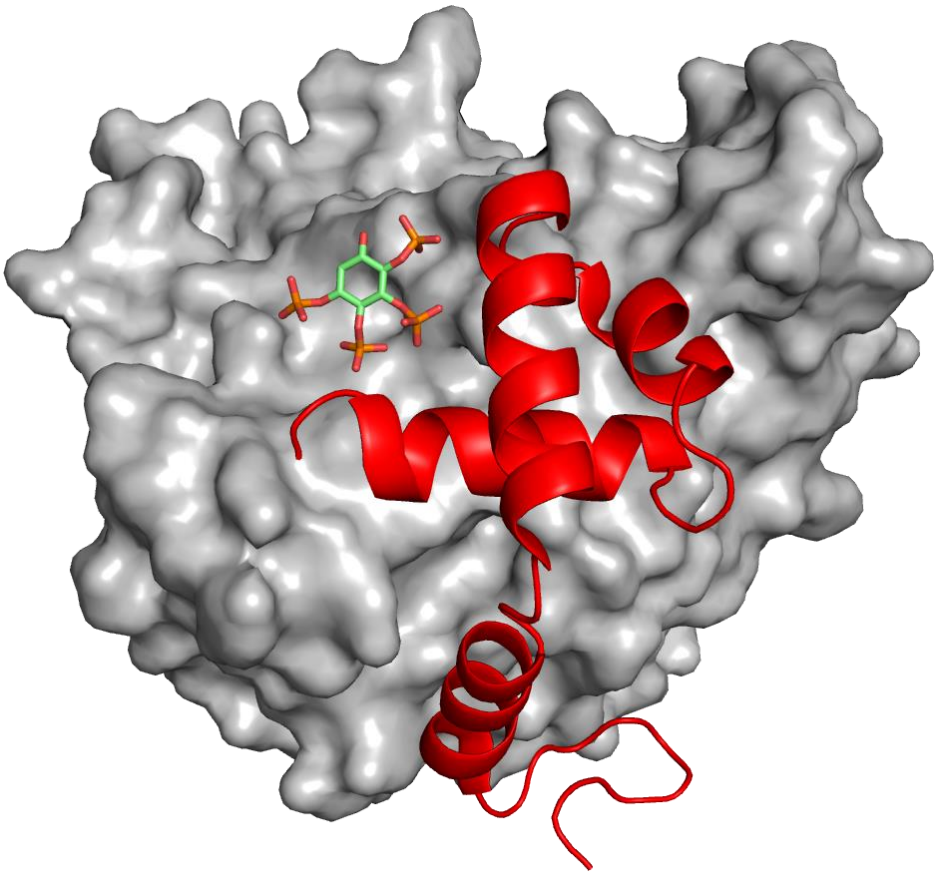


Figure 10: Crystal structure of HDAC3 (surface) bound to its co-repressor (DAD, red ribbon) and inositol tetraphosphate (sticks) (pdb: 4A69).[42]

From the structure, they found that the SMRT-DAD undergoes a large structural rearrangement upon binding to HDAC3 (Figure 11). In the unbound form, the DAD specific motif is associated closely with the rest of the three helices in the bundle. However, upon binding, the DAD specific motif undergoes a sharp change in its structure, acting as a hinge for DAD. This structural change allows helices 1, 2 and 3 the ability to dissociate from HDAC3.

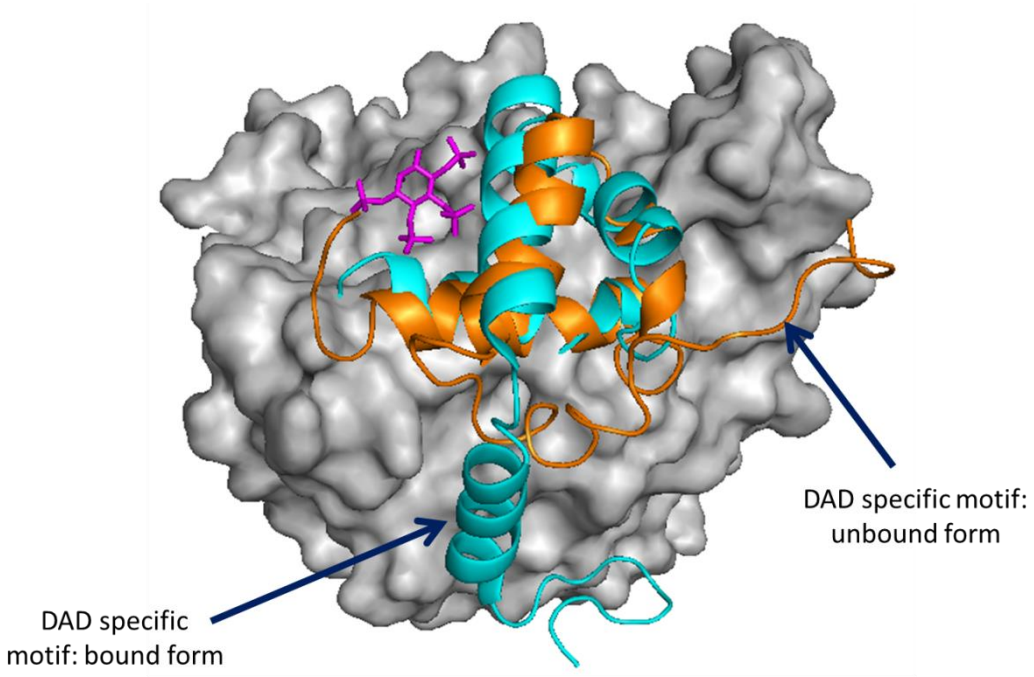


Figure 11: NMR solution structure of unbound DAD (orange ribbon) (pdb: 1XCR)[43] overlaid onto the crystal structure of HDAC3 (grey surface) bound to Ins(1,4,5,6)P_4 (magenta sticks) and DAD (cyan ribbon) (pdb: 4A69).[42]

In addition, they suggested that the Ins(1,4,5,6)P_4 may act as a regulator or 'intermolecular glue' between HDAC3 and SMRT-DAD, as it was found to be vital in the assembly of the HDAC3/DAD complex.

Both of these (the DAD and Ins(1,4,5,6)P_4) combined are thought to be important in the inducement of HDAC3 function. It is hypothesised that when either the DAD or the Ins(1,4,5,6)P_4 molecule is missing from the complex, then HDAC3 will not be able to perform its function.

HDAC3 is one of the class 1 HDAC proteins and therefore like other class I HDACs, includes a zinc atom bound in its active site. As the structure shows, both DAD and Ins(1,4,5,6)P_4 are bound away from the catalytic site at an allosteric site of the protein (Figure 12). However, unlike the central dogma of allostery, the binding of the corepressor and Ins(1,4,5,6)P_4 does not induce a structural change to mediate the thermodynamics or kinetics of an enzyme mechanism.[44]

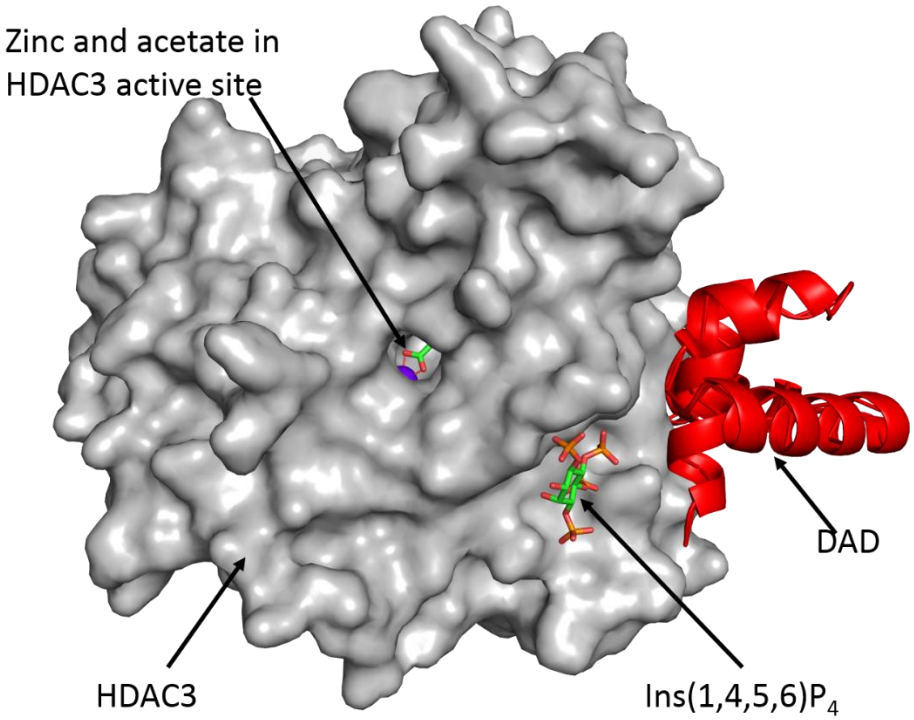


Figure 12: Structure of HDAC3 (grey surface) showing the relative positions of its corepressor, DAD (red cartoon) and Ins(1,4,5,6)P_4 (sticks) with regard to the HDAC3 active site. The active site is occupied by a zinc atom (purple sphere) and an acetate molecule (stcks). (pdb: 4A69)[42]

Following on from HDAC3, the crystal structure of HDAC1 was elucidated by Millard and colleagues. They crystallized HDAC1 in complex with the dimeric ELM2-SANT domain of MTA1 from the NuRD complex (Figure 13).[45] In comparison with the elucidated HDAC3 complex, the HDAC1 complex was not crystallized with Ins(1,4,5,6)P_4. Three sulfate ions were instead found in the

position that the inositol phosphate occupied in HDAC3. However, in parallel with HDAC3, it was found that Ins(1,4,5,6)P_4, was an important factor in the activity of HDAC1. Additionally, in the crystallisation of HDAC1/MTA1, a longer construct was used for the SMRT. The longer tail allowed for the anchoring of SMRT to the complex.

All three of these HDAC complexes are class 1 HDACs, and unsurprisingly their structure is mostly conserved (Figure 14). It is the small differences between them that potentially allow for specificity and hence a way to target just one HDAC protein over another.

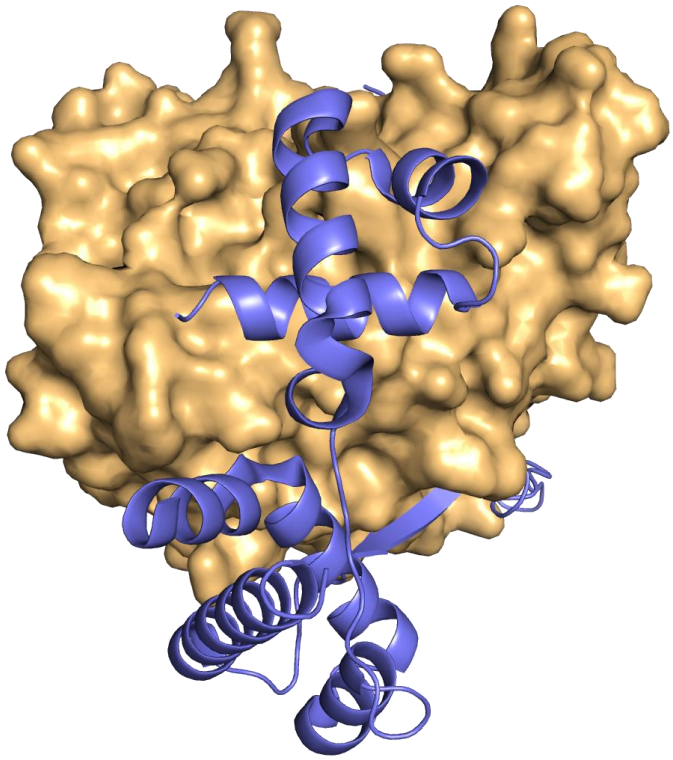


Figure 13: Crystal structure of HDAC1 (surface) in complex with the dimeric ELM2-SANT domain of MTA1 (blue ribbon) from the NuRD complex (pdb: 4BKX).[45]

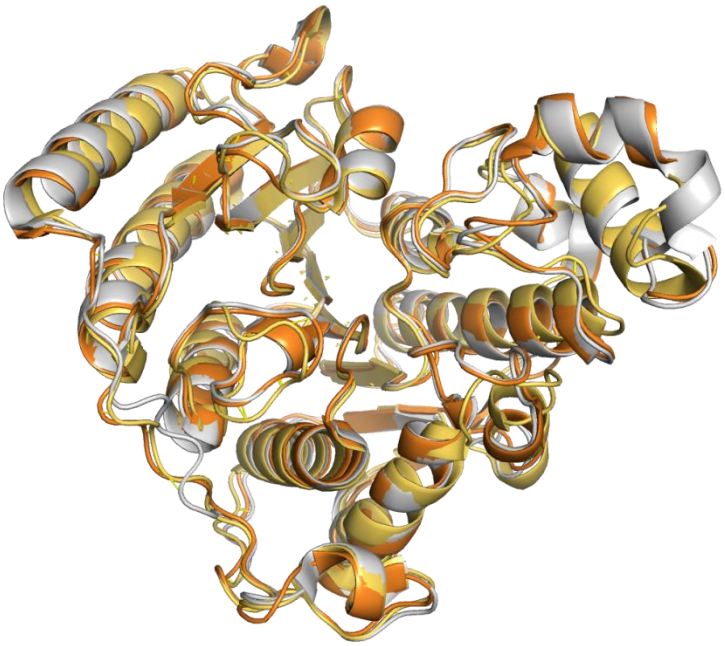


Figure 14: Overlaid crystal structures of HDAC1 (orange), HDAC3 (grey) and HDAC8 (yellow) showing the conservation of structure between the Class I HDAC enzymes.[41,42,45]

1.2.4. HDAC deacetylation mechanism

There are a number of mechanisms that have been proposed for the deacetylation of acetylated lysine residues. One of the most readily accepted mechanisms will be described here using HDAC8 (Scheme 5).

In parallel with other Zn^{2+}-metalloenzymes such as thermolysin[46] and carboxypeptidase A,[47] decetylation by HDACs proceed through a promoted water mechanism.[48]

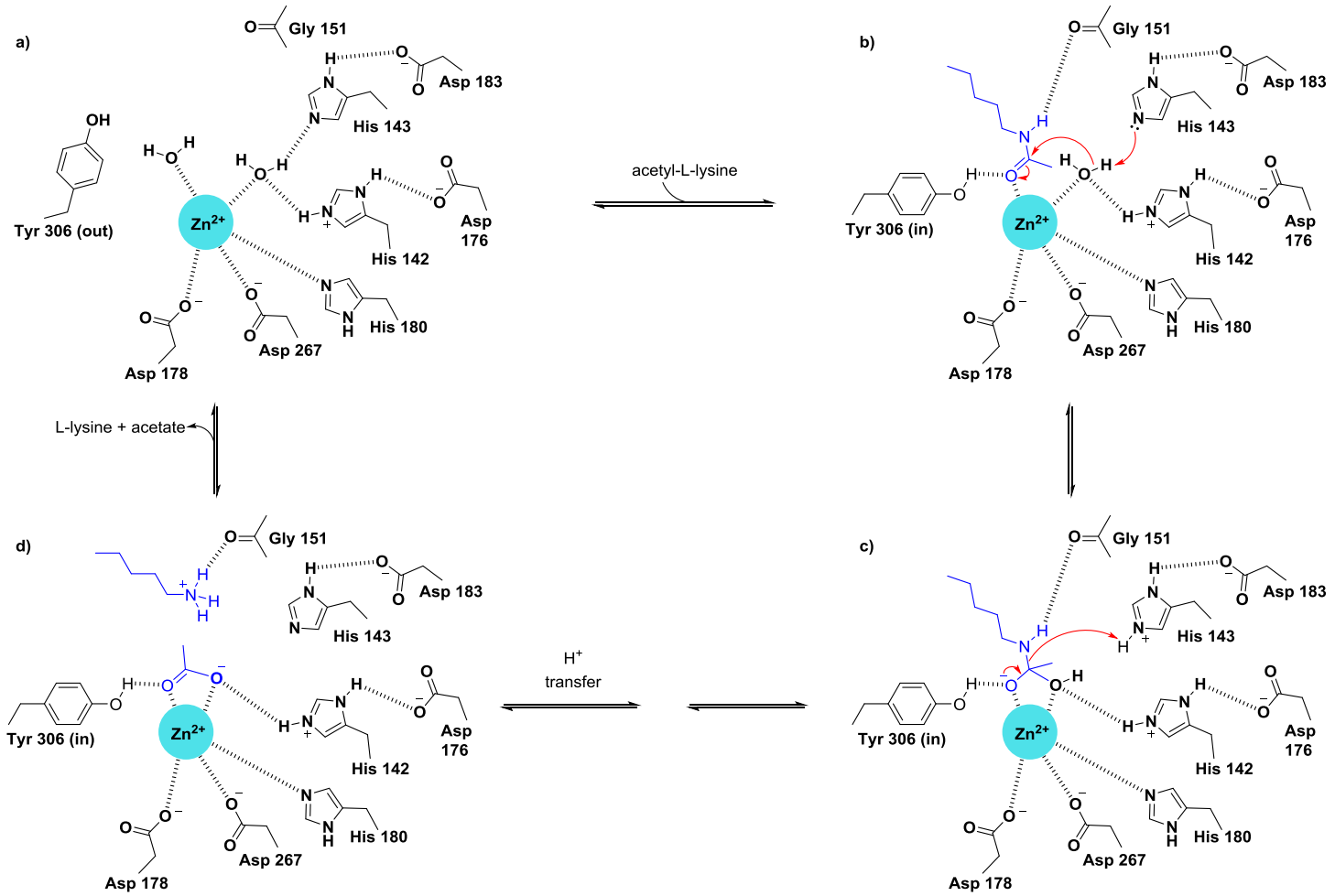
In the HDAC8 active site, the Zn^{2+} ion has a 5-coordinate square pyramidal geometry. The zinc is coordinated to three residues from HDAC8 (D178, H180 and D267) and two water molecules (Scheme 5a).

The catalytic mechanism also involves a gatekeeper tyrosine residue (Y306). This is thought to undergo a conformational change from ‘out’ to ‘in’ upon the binding of the acetylated lysine substrate into the enzyme active site (Scheme 5). As a result of this movement, the binding of the acetylated lysine substrate suggests an induced-fit mechanism.[48] Y306 has also been shown to be

important for the activity of the enzyme through a Y306F mutation which inactivates HDAC8.[49,50]

When the acetylated lysine substrate enters the active site pocket, one of the water molecules in the active site is displaced to allow the substrate carbonyl to coordinate the zinc. Additionally, the substrate carbonyl also coordinates the hydroxyl of Y306, thus stabilising the substrate and polarising the carbonyl ready for deacetylation.

A histidine residue in the HDAC active site (H143 in HDAC8) then acts as a base to begin the catalytic cycle. H142 was also thought to act as a base; however, with a dramatic 10^5-fold decrease in enzyme activity upon a H143A mutation,[49] H143 has been proven to act as the base. The deprotonation of the remaining water molecule coordinated to the zinc by H143 thus activates the hydroxyl to initiate a nucleophilic attack onto the lysine carbonyl (Scheme 5b).[48]



Scheme 5: Proposed mechanism of HDAC enzymes, more particularly HDAC8. (adapted from Lombardi *et al*.[48])

The protonated H143 then acts as an acid to enable the collapse of the tetrahedral intermediate and reforms the carbonyl from the oxyanion (Scheme 5c). Proton transfer then occurs, which protonates the amino side chain of the lysine and leaves the acetate product coordinated to Y306, H142 and twice to zinc (Scheme 5c).

Both the deacetylated product and the acetate can then leave the active site pocket to allow the HDAC enzyme to deacetylate another substrate.

1.2.5. HDACs as oncology targets

HDAC enzymes are involved in many biological functions from cell differentiation to apoptosis and protein-DNA interactions to post translational modifications.[51] Unsurprisingly, HDAC enzymes have therefore been linked to many cancers, in particular changes in their cellular protein expression. Overexpression of HDAC1 has been reported in gastric,[52] colon,[53] prostate[54] and breast[55] cancers, whereas low expression of HDAC1 has been associated with a more invasive phenotype in patients with oesophageal squamous cell carcinomas.[56] Overexpression of HDAC2 is seen in cervical,[57] gastric[58,59] and colon[53] cancers. In addition, increased expression of HDAC3 is also associated with colon cancer,[53] whereas overexpression of HDAC6 is seen in breast cancer.[60]

As a result of the involvement of HDAC enzymes in these various cancers, it is clear to see that the changes in expression of HDAC enzymes could have an active role in tumour onset and progression.[61] Consequently, HDAC enzymes can be seen as valid cancer therapeutic targets and indeed many small molecule drugs are being investigated as inhibitors of HDACs.

1.2.6. HDAC inhibitors

HDACs have been validated as suitable drug targets for cancer. As a result, four HDAC inhibitors, vorinostat (SAHA), belinostat, romidepsin and chidamide,

have recently been approved by the FDA for the treatment of T-cell lymphomas.[62,63] Additionally, panobinostat has been approved for treatment of multiple myeloma.[64] There are also a large number of different small molecule compounds that are currently in clinical trials as drugs that target HDACs,[65,66] These small molecules include abexinostat (PCI-24781),[67] pracinostat (SB939),[68,69] resminostat,[70] givinostat[71–73] and CUDC-101,[74–76] which are hydroxamic acids, mocetinostat (MGCD0103) and entinostat,[77] which are benzamides and short chain fatty acids like valproic acid.[51,78]

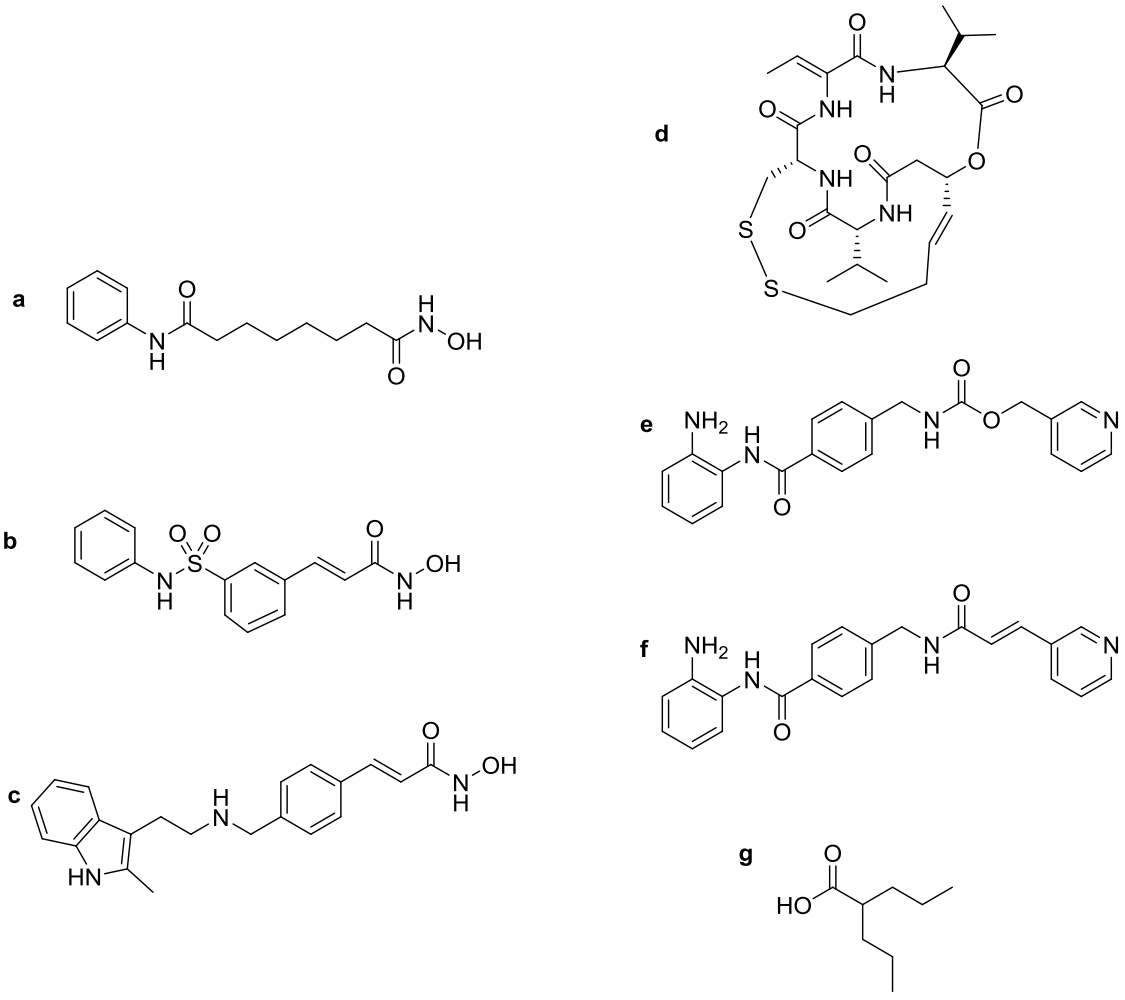


Figure 15: Small molecule HDAC inhibitors a) Vorinostat, b) Belinostat, c) Panobinostat, d) Romidepsin, e) Entinostat, f) Chidamide and g) Valproic acid.

These HDAC inhibitors are generally termed *pan inhibitors*. They are small molecule inhibitors that are not isoform selective. Thus, new methods of

inhibiting HDAC enzymes are needed which have selectivity for one isoform over the others.

Current HDAC inhibitors can be broken down into four different types: hydroxamic acids, cyclic and depsipeptides, benzamides and aliphatic acids.

Vorinostat (suberoylanilide hydroxamic acid, SAHA) is an example of a hydroxamic acid HDAC inhibitor (Figure 15a) which is classed as a pan-HDAC inhibitor because it is not specific for just one HDAC protein.[79] It is marketed by Merck as Zolinza. Vorinostat works by chelating the zinc atom that sits in the active site of HDAC enzymes and is a potent inhibitor of HDAC1, HDAC2, HDAC3 and HDAC6. It was the first HDACi to be approved by the FDA for clinical use in patients with cutaneous T-cell lymphoma (CTCL) in October 2006.[62]

Belinostat (Figure 15b) is another hydroxamic acid containing HDACi and in particular a sulphonamide hydroxamic acid. It is an inhibitor of class I, II and IV HDACs. It was developed by TopoTarget and Spectrum Pharmaceuticals for the treatment of relapsed or refractory peripheral T-cell lymphoma (PTCL) and was granted FDA approval in 2014.[63,80,81]

Panobinostat (Figure 15c), also known as Farydak, and developed by Novartis, is another hydroxamic acid based HDACi. In February 2015, panobinostat was granted an FDA accelerated approval for use in combination with bortezomib and dexamethasone for the treatment of multiple myeloma.[64] This treatment was approved for patients who have already undergone at least two previous treatments.[64]

Romidepsin (Figure 15d), marketed as Istodax by Celgene, is a depsipeptide and an example of a cyclic peptide used as an HDAC inhibitor. It consists of three fused peptide rings and is a natural product first found in the fermentation broth of the bacteria *Chromobacterium violaceum*. In November 2009, romidepsin was FDA approved for use in the clinic for patients with CTCL, the fourth drug to be approved for CTCL and only the second HDAC inhibitor to ever be approved. However, although romidepsin has been approved by the FDA, it is only approved for use in the United States. In Europe, the European

Medical Agency (EMA) and the Committee for Medical Products for Human Use (CHMP) recommended the refusal of the marketing authorisation for romidepsin for the treatment of PTCL. They concluded that although the clinical study showed the anti-tumour capabilities of romidepsin, it wasn't compared to any other treatment. Additionally, they could not determine survival and progression-free survival compared to treatments currently used for PTCL.[82]

Romidepsin is different to vorinostat, belinostat and panobinostat in that it is a prodrug and requires the disulphide bond to be reduced by glutathione in the cell to allow the release of a zinc binding thiol (Figure 16).[83–85]

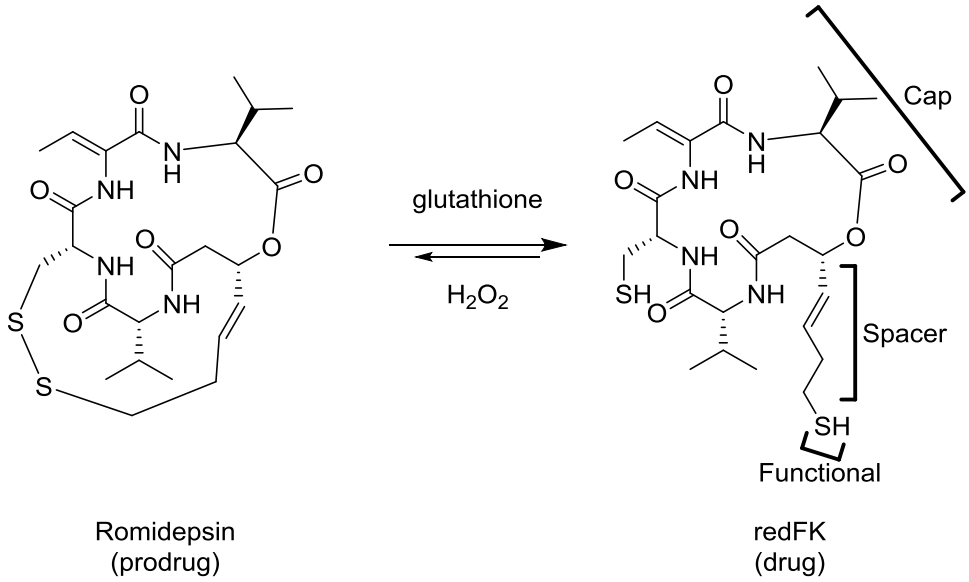


Figure 16: Reduction in cells of the depsipeptide prodrug Romidepsin with glutathione gives the active HDAC drug redFK, a monocyclic dithiol.

Benzamides are another class of HDACi. Included in this group is entinostat (Figure 15e), an oral isoform selective HDACi.[77] It is currently undergoing Phase III drug trials in combination with the aromatase inhibitor exemestane versus exemestane alone in targeting resistance to hormonal therapies in oestrogen receptor-positive (ER+) breast cancer. Chidamide (Figure 15f) is another benzamide HDACi and has recently been approved by the Chinese FDA for relapsed or refractory PTCL.[86]

Valproic acid (Figure 15g) is an example of an aliphatic acid HDACi which primarily targets class I and class IIa HDACs. It is currently marketed as an anti-seizure drug that has been widely available for over 30 years.[51] However, as an HDACi, aliphatic acids, including valproic acid are generally weaker in comparison to hydroxamic acids or cyclic peptides.[34] Even so, as a monotherapy, valproic acid was shown to have some therapeutic effect in myelodysplastic syndromes.[78]

1.3. Protein-protein interactions

Protein-protein interactions (PPIs) are vitally important in many processes in biological systems. In fact, it has been estimated that the human protein-protein interactome consists of between 130,000 and 650,000 interactions.[87–89]

As a consequence, in recent years, PPIs have become extremely attractive drug targets because they have been linked to the progression and development of many disease states.[90–92] Additionally, the chemical space occupied by PPIs is underexplored and so is a potential new area for drug research.

However, because of the relatively large surface areas that PPIs encompass (~1600 $Å^2$), and their dynamic nature, the development of new drugs is a challenge in medicinal and chemical biology.[93] Although a typical PPI can cover a large area of a protein surface, typically, only a few key amino acid residues, which can be described as interaction 'hot spots', are important in delivering the majority of the binding affinity and specificity of a PPI.[93–95] These hot spots have been found to be secondary structure motifs such as the α-helix and β-sheets. Of the secondary structures involved in PPIs, the α-helix has received a lot of interest because α-helices comprise about 62% of the secondary structures in protein complexes[96] and they also mediate a large number of key PPIs.[97,98]

As part of larger proteins, the peptide sequences at the hot spots are thermodynamically folded. However, when the sequences are isolated away from the larger protein, they lack the ability to spontaneously adopt their bioactive conformation. Furthermore, isolated peptides also have many other

unattractive properties such as poor bioavailability, poor proteolytic stability and lack of membrane permeability. For this reason, the use of peptides as drugs has been very limited.

There is therefore an ongoing challenge to design molecules which can target the chemical space occupied by PPIs while at the same time overcoming the inherent undesirable properties of isolated peptides. It is for this reason that the field of proteomimetics has significantly grown in recent years. One major reason is that proteomimetic molecules can mimic the structure and function of proteins, while being a fraction of the size of a full protein.

1.3.1 Peptidomimetics

Peptides make a number of key intra- and intermolecular interactions which help to stabilise their secondary structure. In solution, α-helices form intramolecular hydrogen bond between the carbonyl of the residue *i* and the NH group of the *i+4* residue (Figure 17a). Additionally, intermolecular hydrogen bonds can also form, for example with side chains of other helices (Figure 17b), or even with the solvent. Hydrophobic interactions between different helices can also form and help stabilise the formation of an α-helix.

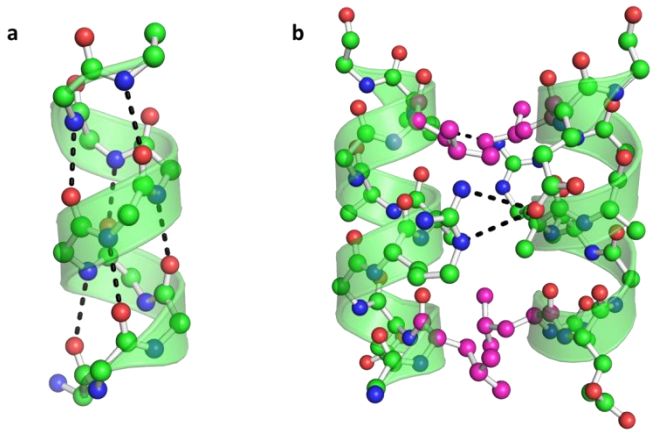


Figure 17: Ribbon and ball-and-stick diagram of α-helix peptides showing a) intramolecular hydrogen bonds (dashed lines) between NH and CO groups on the same helix and b) intermolecular hydrogen bonds (dashed lines) between side chain residues of different helices and hydrophobic interactions (magenta balls) between different helices.

Many types of peptide mimics are known specifically for the α-helix structure. These include small molecules,[99–102] mixed α/β-peptides,[103,104] peptoids,[105] and proteomimetics.

Small molecule peptidomimetics have been successful in inhibiting PPIs and have been taken on by the pharmaceutical industry. In fact, Abbott Laboratories have developed ABT-737 (Figure 18a) and ABT-236 (Figure 18b) which are small molecule inhibitors of the Bcl-x_L/Bak PPI.[99,100] Additionally, Hoffmann-La Roche have developed the Nutlin family of small molecule compounds (Figure 18c), while Johnson & Johnson Pharmaceuticals have developed the benzodiazapinedione family (Figure 18d), both of which inhibit the p53/mDM2 PPI.[101,102]

In the field of mixed α/β-peptides, the Gellman group designed a chimeric inhibitor of the Bcl-2 family of PPIs using an unnatural 14/15-helical α/β-peptide. It was designed to mimic the N-terminal segment of Bak and an α-segment of the mixed peptide mimicked key residues of the C-terminal to give a nanomolar ligand of Bcl-x_L (K_i = 2 nM).[103] Following on from this, the group also crystallized the most potent α/β-peptide mimicking the BIM-BH3 peptide bound to Bcl-x_L (Figure 19).[104]

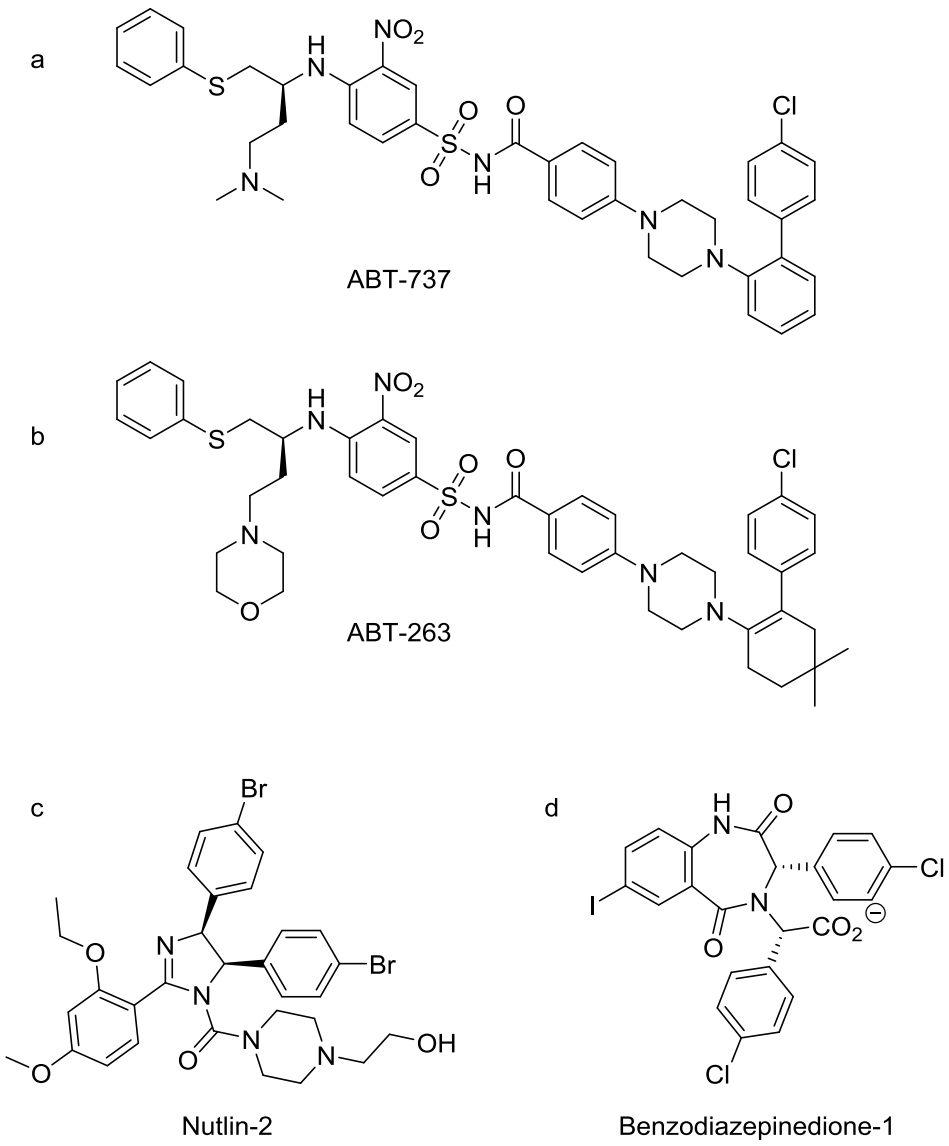


Figure 18: A series of small molecule peptidomimetics which target specific protein-protein interactions.

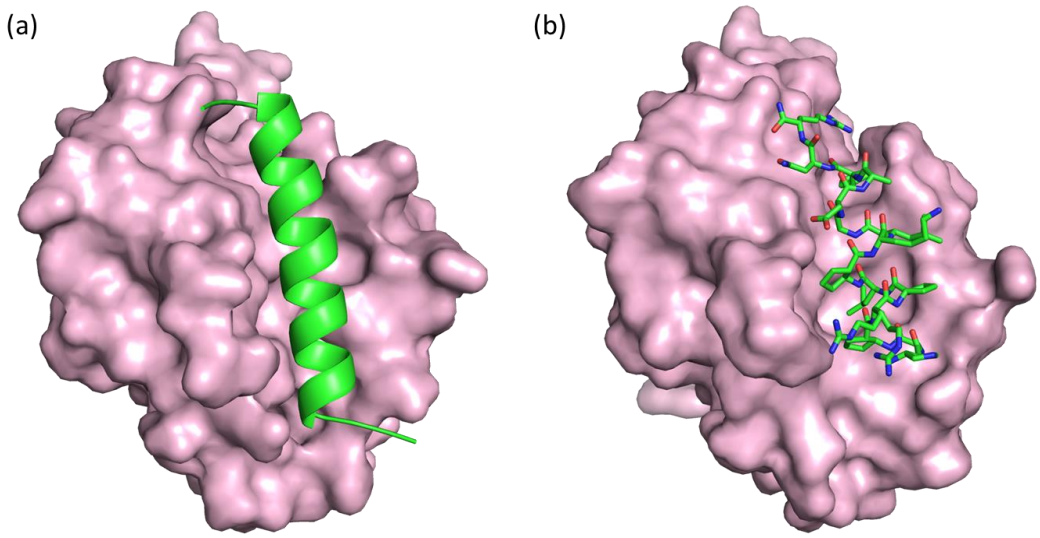


Figure 19: Crystal structures of (a) the BIM-BH3 peptide (green helix) in complex with Bcl-x_L (pink surface) (pdb: 3FDL) and (b) an α/β-peptide foldamer (sticks) mimicking the BIM-BH3 bound to Bcl-x_L (pink surface) (pdb: 3FDM).[104]

Polyproline helices have also been used as peptide mimics particularly by Appella and colleagues, who used peptoids as inhibitors of p53/hDM2.[105] They introduced N-alkylation and used different groups at that position from carboxylic acids to phosphonates and indoles (Figure 20).[105]

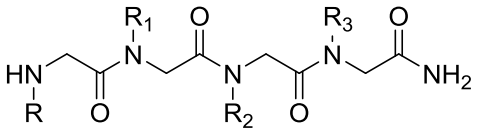


Figure 20: Schematic of a peptoid where the R groups can vary from carboxylic acids to phosphonates to indoles.

Unlike α/β-peptides and peptoids, proteomimetics are suitable α-helix mimics through the way they mimic not the helix itself, but the side chains which protrude from the helix. To this end, in the early to mid-90s, the Rees and Willems groups synthesized trisubstituted indanes to act as inhibitors of tachykinnin receptors (Figure 21).[106,107]

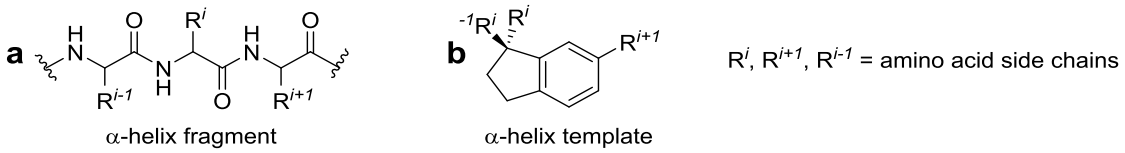


Figure 21: a) An α-helix fragment and b) an α-helix template to mimic the side chain residues of the α-helix structure for use as proteomimetics.

Developed by Hamilton and colleagues, terphenyls are another example of proteomimetics, which mimic an α-helix.[108,109] They synthesized a trisubstituted series of 3,2',2''-terphenyl compounds in which the aryl cores adopts a staggered conformation, which mimicked the *i, i+3(4)* and *i+7* residues of a peptide through the *ortho* positions of the terphenyl core[110] (Figure 22a). Following on from this, the Hamilton group also developed other scaffolds which mimic the protruding side chains on the α-helix, including terephthalamide[111] (Figure 22b), 4,4'-dicarboxamine[112] (Figure 22c), 5-6-5-imidazole-phenyl-thiazole[113] (Figure 22d) trispyridylamine[114] (Figure 22e) and enaminone[115] (Figure 22f) templates.

Many other groups have developed different types of helix proteomimetics. An oxazole-pyridazine-piperazine mimic (Figure 23a) was developed by the Rebek group incorporating hydrophobic side chain residues to mimic the *i, i+4* and *i+7* residues.[116] Fletcher and colleagues also reported a series of mimetics that control the extent of hydrogen bonding and therefore flexibility of the mimetic by using a distribution of benzene and pyridine rings (Figure 23b),[117] similar to Hamilton's work with trispyridylamine based mimics.[114]

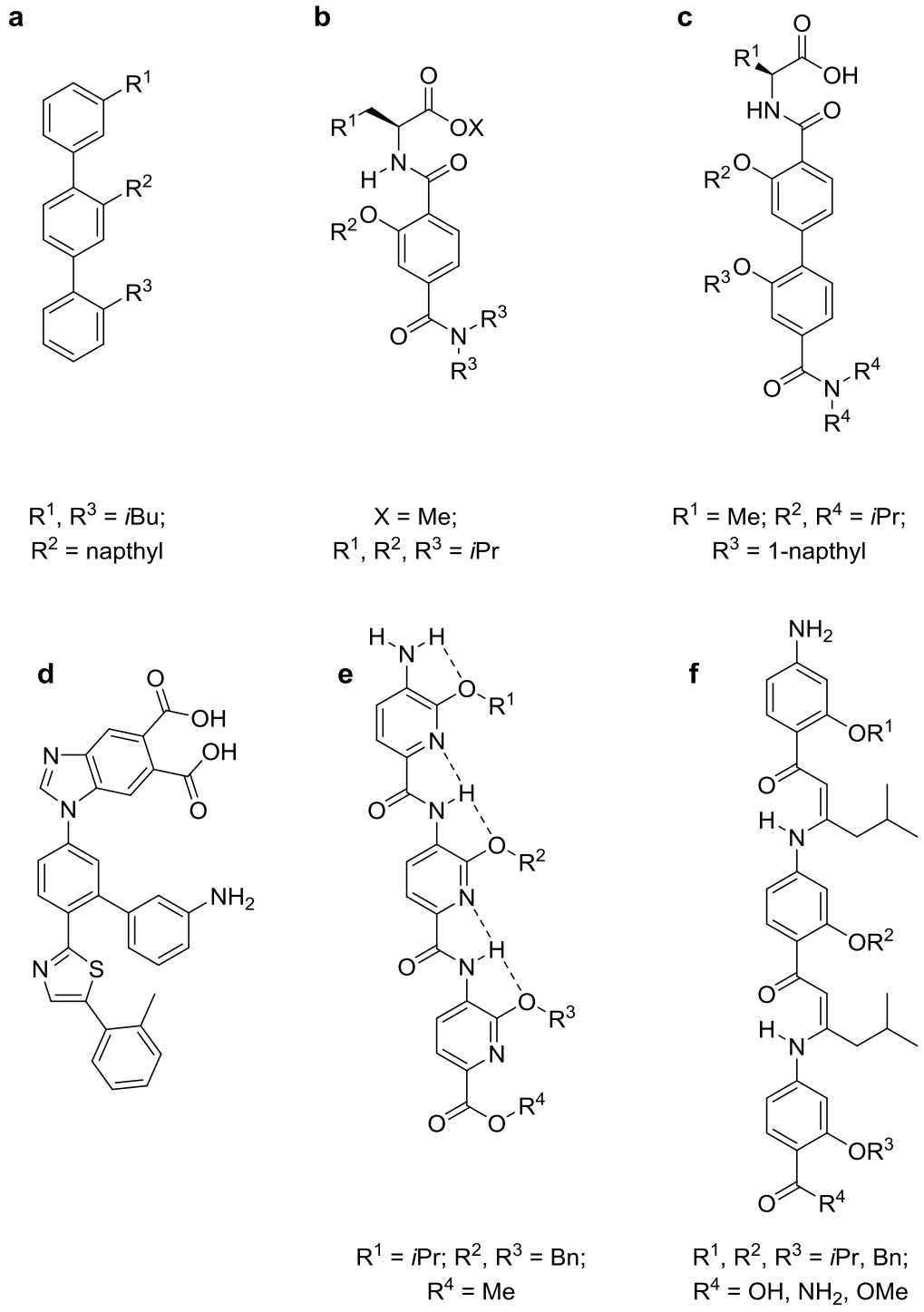


Figure 22: Proteomimetic scaffolds developed by the Hamilton group which have been used to target protein-protein interactions.[110–115]

Following on from these examples, the Wilson group have been able to show that both 3-*O*-alkylated oligoamides (Figure 23c) and 2-*O*-alkylated oligoamides (Figure 23d) can act as inhibitors of p53/hDM2.[118,119] The Wilson group were

also the first to describe a solid-phase synthesis for an α-helix mimetic with *N*-alkylated oligobenzamides (Figure 23e) which acted as inhibitors of p53/hDM2.[120,121]

However, even with all the different variations of α-helix proteomimetics, so far, none have been able to mimic more than one face of the helix. Consequently, conformationally constrained peptides have emerged as desirable molecules to mediate PPIs.

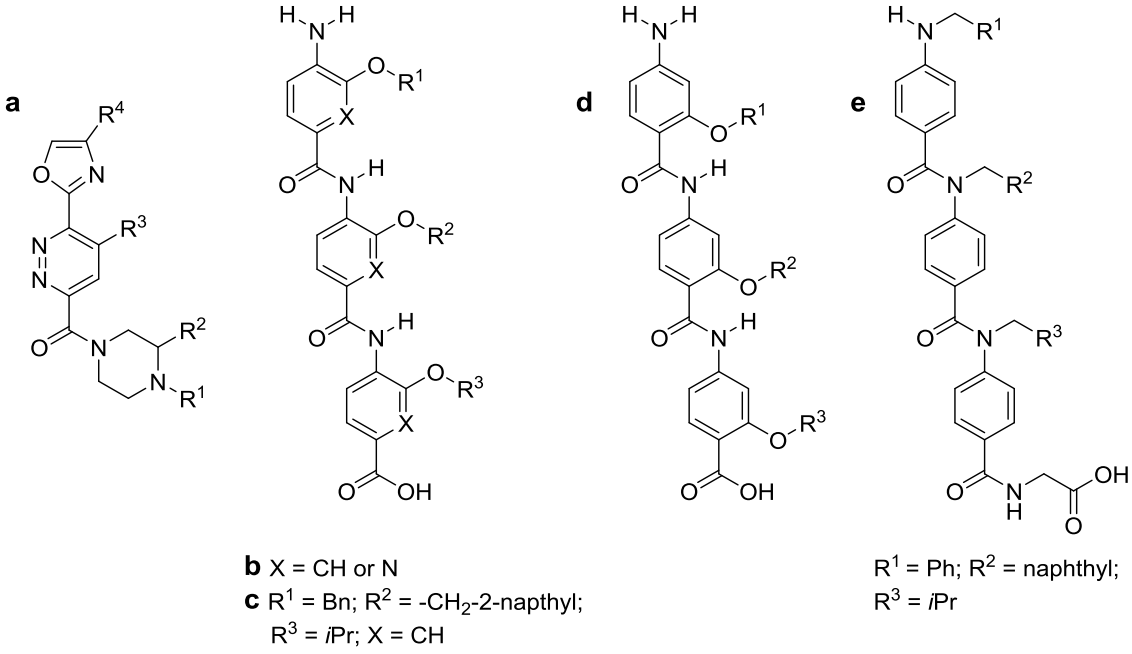


Figure 23: Proteomimetic scaffolds used to mimic one face of a helix to target protein-protein interactions.[116–121]

1.3.2 Conformationally constrained peptides

Short isolated peptides generally do not fold into their bioactive conformation away from their parent proteins. Thus, the stabilisation of short peptide sequences into an α-helical conformation has been studied extensively. This stabilisation of the α-helix structure has been achieved through the addition of helical constraints. As a result of the need to constrain α-helices, many different constraints have been developed. These include salt bridges between charged

amino acid side chain residues[122,123] (Figure 24a), lactam bridges[124–128] (Figure 24b), disulphide bridges[129] (Figure 24c), hydrophobic interactions[108] (Figure 24d), metal ligation[130,131] (Figure 24e), hydrogen-bonding surrogates,[132,133] triazole staples[134] (Figure 24f), photocontrollable macrocycles[135] (Figure 24g), introduction of α,α-disubstituted amino acids[136,137] and hydrocarbon stapling[138] (Figure 24h).

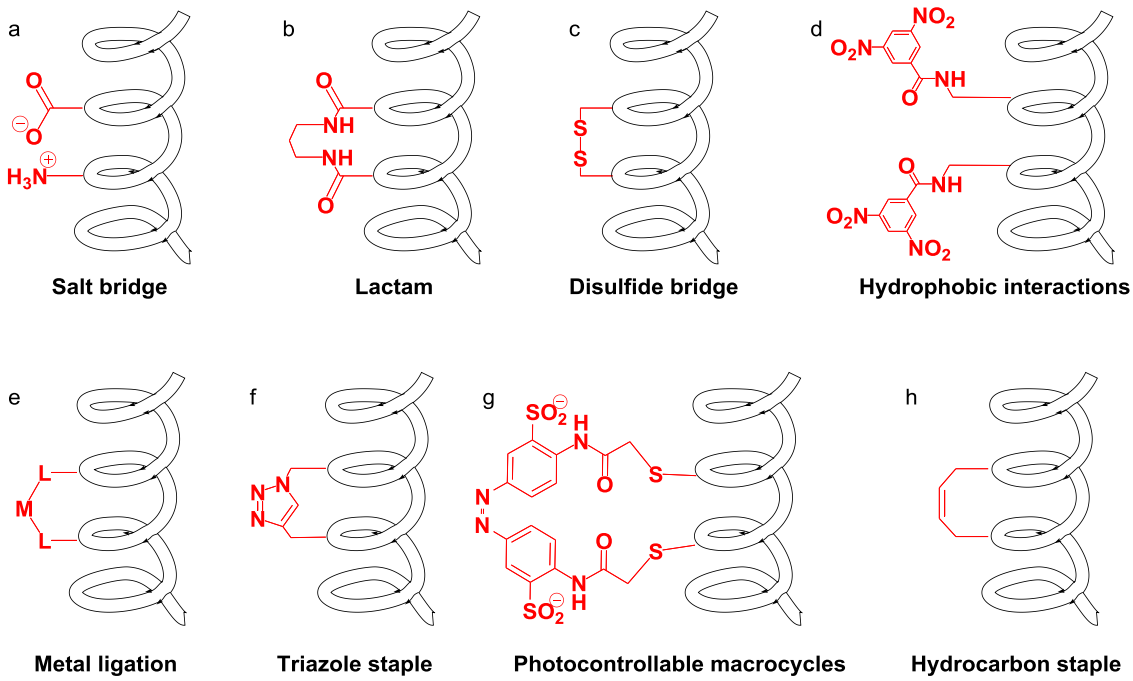


Figure 24: A selection of methods used to constrain short peptides into an α-helical conformation.

Each of these methods of constraint, exhibit their own set of advantages and disadvantages. Nonetheless, they can improve the binding affinity of an isolated peptide sequence to its binding partner as a result of overcoming the entropic penalty required for folding. Furthermore, peptidases are prevented from recognising the isolated peptide because peptidases require the peptides to be in their extended form in order to hydrolyse the amide linkages between the amino acids.

Salt bridges

The use of salt bridges was one of the first methods of constraint and involves the interactions between the *i,i+3, i,i+4* or *i,i+7* residues, in particular lysine and glutamic acid residues.[123] Other amino acid pairings are also possible including that of aspartic acid and the guanidinium ion of arginine.

Salt bridges consist of electrostatic interactions which although fairly weak (a maximum Glu^{-}-Lys^{+} ion-pair interaction of ~0.45 kcal mol^{-1}),[123] these interactions can add up forming much stronger stabilizations. The distance required for a salt bridge to form is deemed to be less than 4 Å[139] and a distance greater than this between two amino acid side chains, is not classed as a salt bridge.

Positioning of lysine and glutamic acid residues on the peptide also affects the stabilization of the α-helix structure of the peptide. For instance, having a glutamic acid residue at the *i* position and a lysine residue at the *i+4* position, increases the helical properties of a peptide compared with that with the residues in reversed configuration.[122] This difference in helicity is even more pronounced when the amino acids (E,K → K,E) are in the *i, i+3* orientation.[122]

The environment which the peptide occupies also plays a role in the formation and stabilisation of the α-helix induced by the salt bridge. The helicity of the peptides can be affected by pH and salt concentration.[122] Marqusee and Baldwin found that at 1°C, a pH titration showed the most helical peptide conformation to be at around pH 7 in 0.01 M NaCl. For the two extremes, high and low pH, the helicity of the peptide is greatly decreased. Changing the concentration to 1.0 M NaCl also had an effect on the helical structure of the peptide. At high salt concentration, the peptide had low helical character in comparison with that at 0.01 M NaCl.[122]

As a result of the need for pH and salt control in constraining α-helices using salt bridges, many other constraint methods are generally preferred.

Lactams

The first example of lactam bridge used for stabilizing a protein was reported by Rosenblatt and co-workers.[124] They utilized a lactam bridge between a lysine and aspartic acid residue at the *i* and *i+4* positions of their protein of interest (Figure 25), parathyroid-hormone-related protein (PTHrP). With this they elucidated that the covalently bonded lactam bridge allowed PTHrP to be in predominantly in its bioactive form. As a result of the lactam bridge, the peptide displayed higher potency in comparison with the native sequence, which is stabilized at these same positions by a non-covalent salt bridge, which can be disrupted.[124]

McDowell and colleagues have also utilised lactam bridges on a 27-residue peptide to target HIV-1 gp41.[126] They achieved the lactam bridge through the covalent linkage of two glutamic acid residues at the *i* and (*i+7*) positions of the peptide with an α,ω-diaminoalkane group (Figure 25b). In addition to the singular lactam bridge, the group also synthesized a peptide bearing two of these lactam bridges separated by two complete turns of the helix. The doubly lactam bridged peptide showed greater helicity in comparison with the singly lactam bridged peptides, and in addition it also gave a better inhibitory potency.[126]

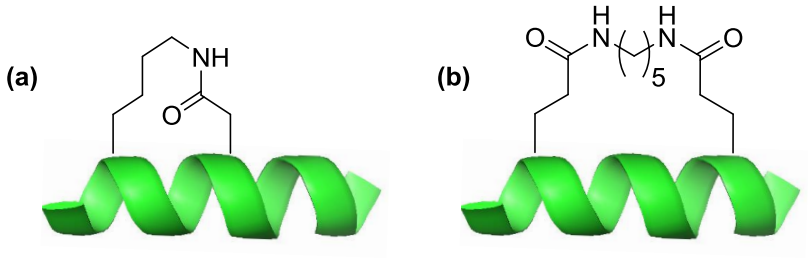


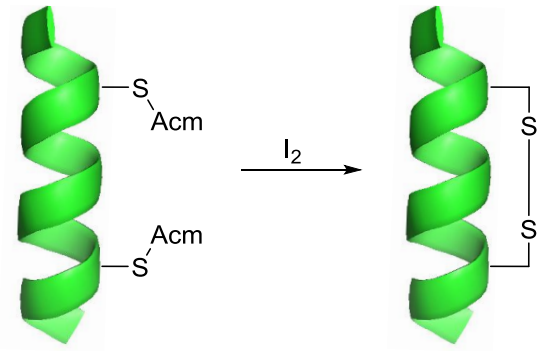
Figure 25: Examples of lactam bridges used in constraining α-helices including (a) an *i,i+4* lactam staple made using lysine and aspartic acid residues and (b) an *i,i+7* lactam staple where two glutamic acid residues are covalently linked with an α,ω-diaminoalkane group.

Following on from this work by McDowell, the Kim group utilized a lactam bridge across an *i* and *i+7* position of their peptide with the aim of stabilizing a 14-residue peptide sequence that once again would bind the hydrophobic pocket of HIV-1 gp41.[127]

Lactam bridges have also been employed to determine the effect of the reduction of hydrophobicity of the non-polar face of a helical peptide. More specifically, the effect on the helical content and the stability of the peptide.[140] To this end, lactam bridges were incorporated with an *i*,*i+4* spacing in 14-residue amphipathic α-helical peptides. The study found that the addition of a lactam bridge into the peptide increased not only the helicity of the peptides but also enhanced the dimerization of the peptides based on a 3,4 hydrophobic heptad repeat.[140]

Disulfide bridges

Many groups have exploited disulphide bridges in order to stabilize the α-helicity of certain peptides.[125,129] The Schultz group synthesized short peptides containing a two turn α-helix stabilized by a single intramolecular disulphide bond bridging the *i* and *i+7* residues (Scheme 6).[129] These peptides showed high helicity in comparison to their native sequence in water at 0 ˚C and 60 ˚C.



Scheme 6: Formation of an *i,i+7* disulfide bridged stapled α-helical peptide through the use of iodine. The sulfur containing residues were protected by acetamidomethyl (Acm) groups.

Hydrophobic interactions

Aromatic side chains such as tyrosine, phenylalanine, histidine and tryptophan are also known to help contribute to the stabilization of α-helices through π-π stacking[141] and contribute free energies of formation of between -0.6 and -1.3 kcal/mole.[142,143] A particular example of hydrophobic interactions stabilising

helices is from the Hamilton group. They designed a series of short α-helix peptides (14-20 amino acid residues) containing ε-(3,5-dinitrobenzoyl)Lys (Dnb) position at the *i and i+4, i+7* and *i+13* positions[144] (Figure 26). The Dnb *i, i+4* and the Dnb *i, i+7* peptides have the aromatic residues on the same face of the helix with a difference of either one or two turns of the helix. However, for the Dnb *i+13* peptide, the aromatic residues are nearly on opposite faces of the helix. They found that the Dnb *i+4* peptide showed the greatest α-helical character of the three peptides due to the relative closeness of the two aromatic residues (~ 6 Å).[144]

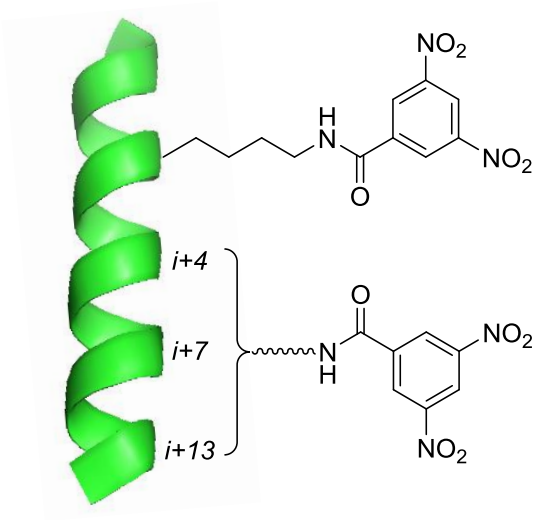


Figure 26: Hydrophobic interactions, such as those using ε-(3,5-dinitrobenzoyl)Lys (Dnb), can act to stabilise α-helices.

Metal ligation

Another important α-helix stabilization technique is the use of metal ligation. This involves the use of a metal ion to bind ligands that are incorporated into the peptide sequence (Figure 27). The thermodynamics of coordinating a metal ion to the ligands on the peptide reduces the entropy required for the peptide to form a helix and therefore should increase the helical content.[130] Ruan and colleagues showed that an *i, i+4* ligand spacing was favoured for α-helix stabilization by metal ligation, enhancing the helical content in 15 out of 20 cases, unlike an *i, i+3* ligand spacing which enhanced helical content in only

four of 20 cases.[130] The metal ions used for this stabilisation include Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+} and Cd^{2+}.

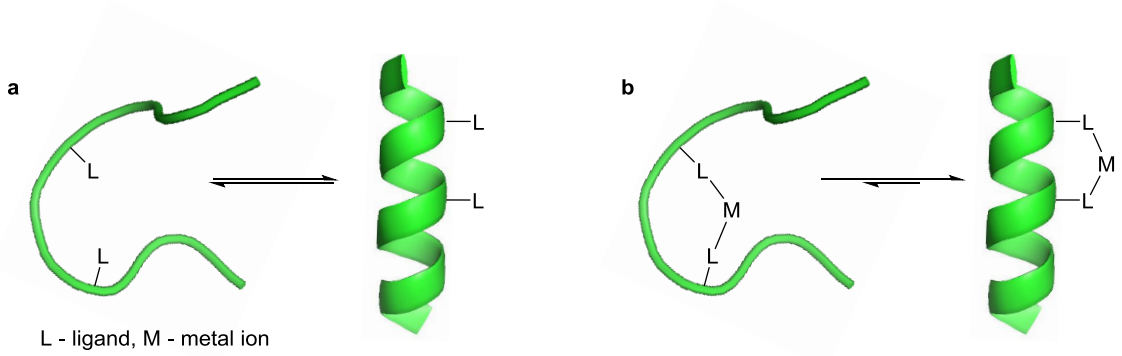


Figure 27: Metal ligation as a means of stabilising an α-helix.

At the same time, Ghadiri and Fernholtz reported the formation of stable α-helical peptides whilst utilising an exchange-inert metal complex.[131] They increased the helical content of a 17-residue peptide from 40% to 80% at 21 ˚C by formation of a macrocyclic *cis*-$[Ru\text{-}(NH_3)_4L_2]^{3+}$ complex, where L_2 are the sides chains of two histidine residues with an *i, i+4* spacing.[131]

Hydrogen bonding surrogates

Hydrogen bond surrogates (HBS) replace one or more of hydrogen bonds in an α-helix with a covalent bond. This locks the conformation of the helix at those positions. Linkages that can mimic the hydrogen bonding motif include hydrazones[145] and alkenes.[132,133] Arora and colleagues utilized the ring closing metathesis reaction in order to create an alkene linkage between an *i* and *i+4* residue in the Bak BH3 peptide (Figure 28).[132] The synthesised peptide consisted of an *N*-allyl dipeptide and a 4-pentenoic acid attached at the *N*-terminal which, following the ring closing metathesis reaction gave the HBS at the *N*-terminus of the peptide.

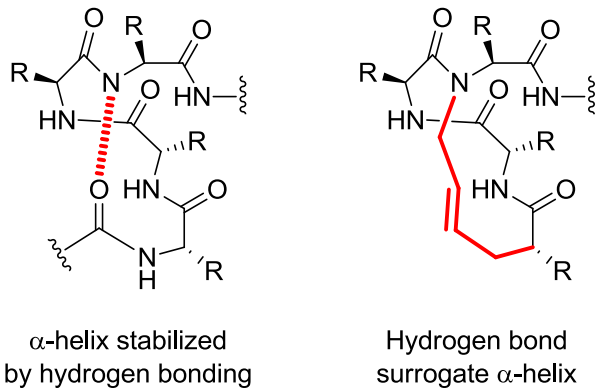


Figure 28: Stabilizing α-helices by hydrogen bonding and by the use of hydrogen bond surrogates.

As a result of the HBS crosslink being placed within the helix system, it therefore does not block any solvent exposed molecular recognition surfaces of the peptide[133] and does not block any side chain groups that are on the peptide. Consequently, the HBS strategy is a useful tool for nucleating an α-helix. Normally α-helices with a maximum of ten amino acid residues are inherently unstable due to low nucleation probabilities.[146,147] It is however possible to constrain much shorter peptides in to a helical conformation through the use of HBS strategy.

Photocontrollable macrocycles

One particular method of constraining an α-helix which allows control of activity, is to use an external stimulus, such as light. In this way, the Alleman group have developed a method in which the helical constraint was induced reversibly through the use of light.[135] Through the use of cysteine residues, an azobenzene crosslinker was introduced into the peptide at the *i* and *i+4*, *i+7* or *i+11*, residues. Upon irradiation with light, the crosslinker adopts a *cis* or *trans* configuration depending on the wavelength of light. Thus, a way of controlling the conformation of the peptide between a random coil and α-helical conformation has been possible.

For peptides with the azobenzene crosslink positioned at *i* and *i+4* residues, the active α-helix requires the azobenzene crosslink to adopt a *cis* configuration

(Figure 29a). However, with the crosslink positioned at *i* and *i+7* residues, the crosslink adopts a *trans* configuration (Figure 29b).[135]

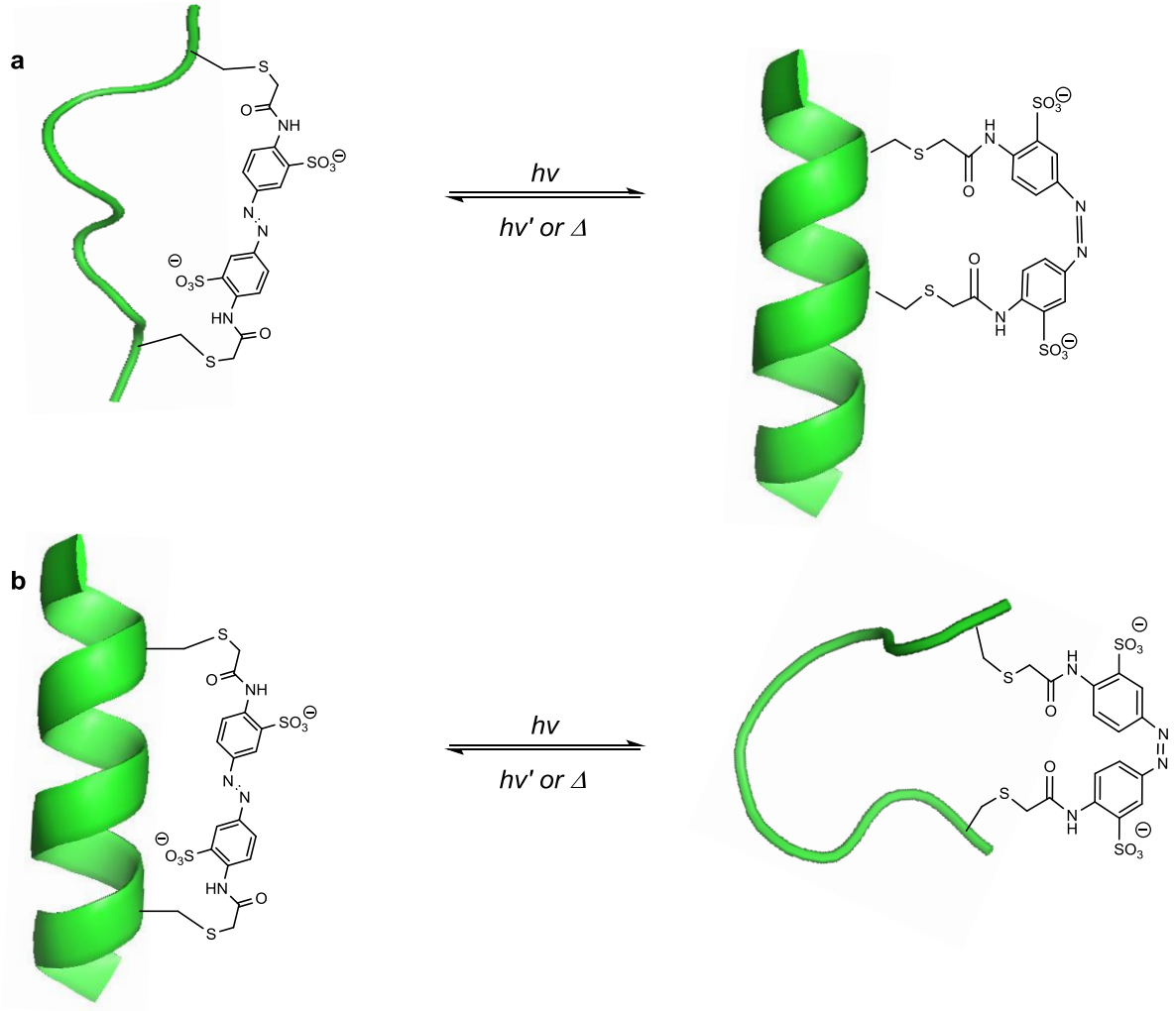


Figure 29: Structural change of the azobenzene crosslinker upon irradiation with light, with a peptide containing a) an *i, i*+7 spacing in which the *cis*-azobenzene stabilizes the α-helix; and b) an *i, i*+4 spacing in which the *trans*-azobenzene stabilizes the α-helix.

However, photocontrollable helix constraints do have a distinct limitation. As a result of the interconversion between the *cis* and *trans*-azobenzene structures, achieving just one of these structures is difficult. Thus, the peptides cannot all be in their helical conformation, and some will still be random coil.

Triazole Staples

Triazole stapling involves a 'Click' reaction between alkenyl and azido side chain residues on a peptide sequence. Compared to many of the other stabilisation techniques, triazole stapling is not as well studied. Consequently, Kawamoto and co-workers followed a systematic study on the effect of the linker length, the position of the triazole within the staple and the use of different stereoisomers used to form the staple (Figure 30).[134]

The group synthesised triazole stapled peptides with varying linker lengths and found a linker length of 8 atoms was the optimum for an *i,i+4* staple.[134] Additionally, they synthesised both a purely L-amino acid triazole staple and a mixed L- and D-amino acid triazole staple, to determine the appropriate stereochemistry of the amino acids required to form the triazole. The placement of the L- and D-amino acids in the mixed triazole staple was also investigated, where the L-amino acid is at the *i* position and the D-amino acid at the *i+4* position and *vice versa*. The group found that the best binding affinity was possible for the first scenario in comparison with the latter or the purely L-amino acid triazole staple.[134]

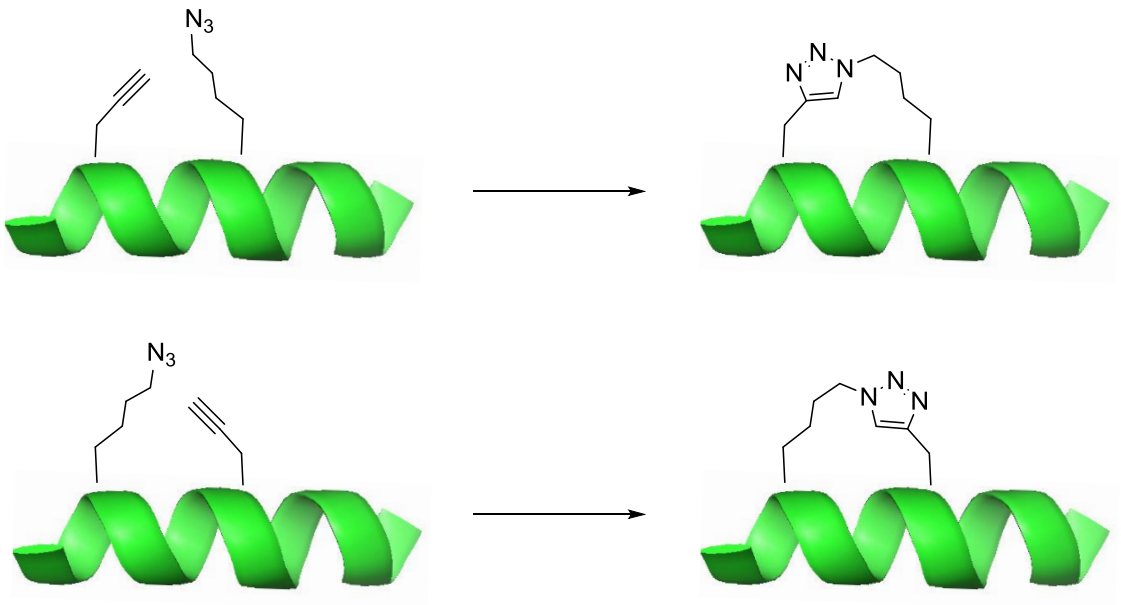


Figure 30: The use of triazole stapling in constraining an α-helical peptide while varying the positions of the alkyne and azide amino acids.

In terms of helicity, the purely D-amino acid triazole staple showed 66% helicity in its stapled form compared to the purely L-amino acid triazole staple, which showed 90% helicity. Both show an increase when compared to the native sequence (44%) and their linear sequences (44% and 45% respectively). The unusually high helicity seen for the purely L-amino acid triazole staple can thus be correlated with its higher binding activity.[134]

α,α-Disubstituted amino acids

The inclusion of α,α-disubstituted amino acids has also been used to help stabilise helical peptides. In fact, the insertion of aminoisobutyric acid (Aib) residues into a peptide has been shown to help stabilise the 3_{10}-helical conformation.[136] This can be explained by the Thorpe-Ingold effect which suggests that the two α-methyl groups restrict the possible conformational space that is sterically allowed as a result of van der Waals clashes.[137,148,149] This moves the other two substituents on the α-carbon closer together. Further on from the use of Aib residues, other α,α-disubstituted amino acids have also been incorporated into helical peptide sequences to help stabilise the helix.

Hydrocarbon staples

As a way of stabilising an α-helical peptide conformation, hydrocarbon stapling is a relatively new technique. However, it is coming of age as a method of stabilising peptides in order to modulate PPIs.

The hydrocarbon stapling technique uses alkene bearing unnatural amino acids, which then undergo a ring closing metathesis reaction in order to form the hydrocarbon bridge between two residues of a peptide (Figure 31). The most common of these is the *i, i+4* staple, but it is possible to introduce the hydrocarbon staple at *i, i+3, i, i+7* and *i, i+11* positions also. Since it's conception, many groups have used this method to stabilize a large variety of peptides including the Grubbs[150,151], Verdine[152–155] and Walensky[156,157] groups to name but a few.

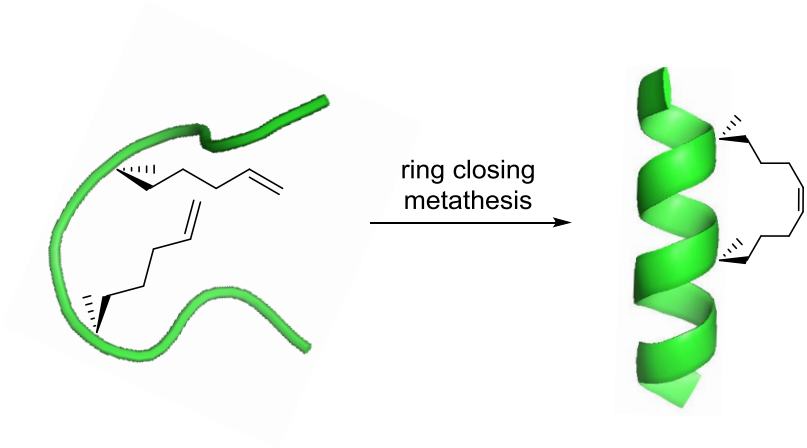


Figure 31: A ring closing metathesis reaction, allows two alkenyl amino acids to be covalently linked to form an all-hydrocarbon staple bridge.

1.4. Stapled peptides

1.4.1 Design of stapled peptides

The field of stapled peptides first came about in the late 1990s from the pioneering work of Miller, Blackwell and Grubbs.[150,151,158] The ring closing metathesis (RCM) of alkenyl functionalised amino acids was first carried out on amino acid monomers, dimers and trimers (Figure 32).[150] Following on from this, RCM was carried out on longer peptide sequences bearing *O*-allyl amino acids.[151,158] However, rather than being α-helical in conformation, these peptides were found to have the structure of a 3_{10}-helix.

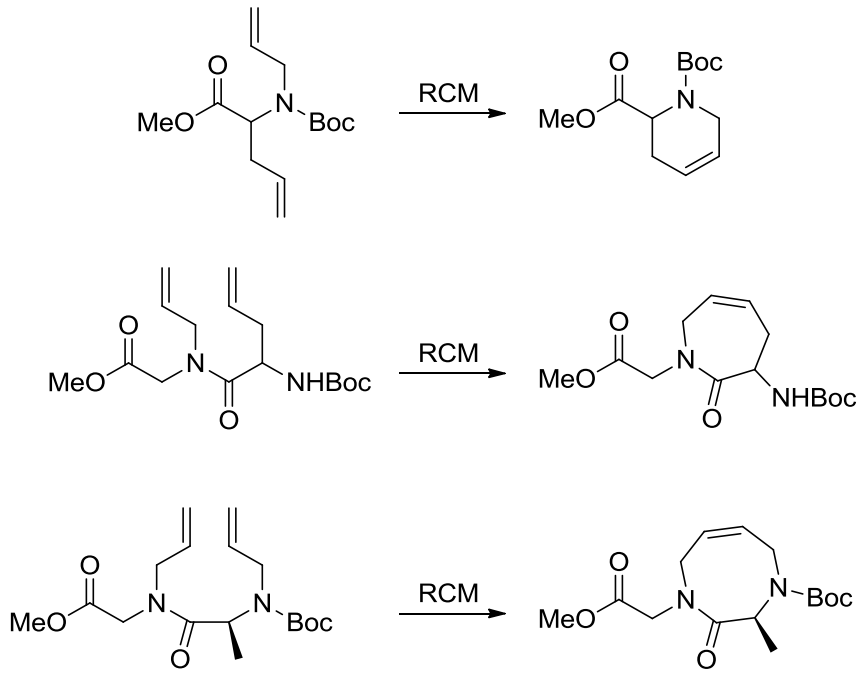


Figure 32: Ring closing metathesis on amino acid monomers, dimers and trimers.[150]

Following the work of Grubbs and colleagues, in 2000, Schafmeister and co-workers combined α,α-disubstitution with alkenyl amino acids.[138] The α,α-disubstitution introduced the Thorpe-Ingold effect into their peptides and to avoid the helix-destabilizing effect of D-amino acids.[138] Korsmeyer, Verdine and Walensky then designed an all-hydrocarbon bridge that would stabilise an α-helix and developed the rules for effective peptide stapling.[138]

The effect of RCM on different peptides hugely depends on the native peptide sequence, its conformation and physicochemical properties; in other words, if it already has the propensity to form a helix upon binding its target.

For RCM to be successful at stabilising peptides, various important considerations need to be taken into account. These include the length of the macrocycle formed as a result of the RCM reaction, the types of heteroatoms, multiple bonds and side chains in the bridge, the configuration of the double bond that is formed, the primary structure of the native peptide, the location and distance apart of the α-carbon bridgehead residues and the stereochemistry of the bridgehead amino acids (Figure 33).[159]

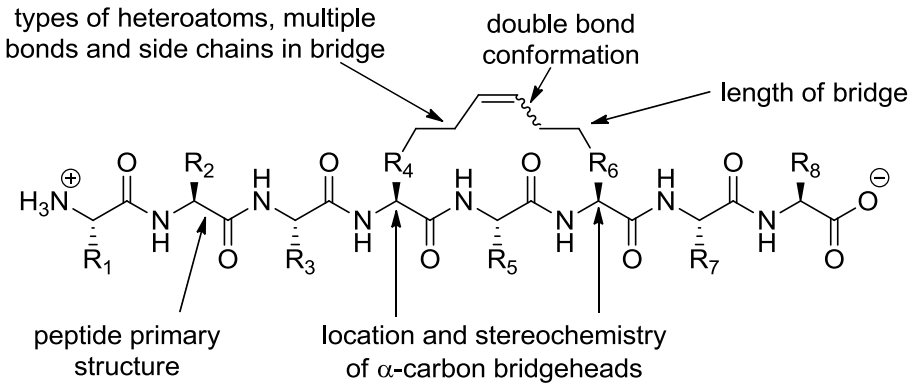


Figure 33: Considerations needed in installing a constraint using ring closing metathesis. (Adapted from Jacobsen *et al.*[159])

A significant feature of hydrocarbon stapling is that the constraint can be incorporated at different positions along the peptide sequence. The most common hydrocarbon staple constraints involve constraining the peptide across either one or two turns of the helix. This involves a side chain-side chain bridge between the *i* and *i+3*[154] (Figure 34a) or *i* and *i+4*[160] (Figure 34b) residues for a single turn constraint and between the *i* and *i+7*[138] (Figure 34c) residues for a two turn helical constraint. For short peptide sequences, one staple is enough to fully constrain the peptide. On the other hand, for longer peptides, two staples can improve the effectiveness of the constraint. The introduction of two staples into a peptide can be achieved through the use of either double stapling[161] (Figure 34d) or by using a stitched staple[162] (Figure 34e).

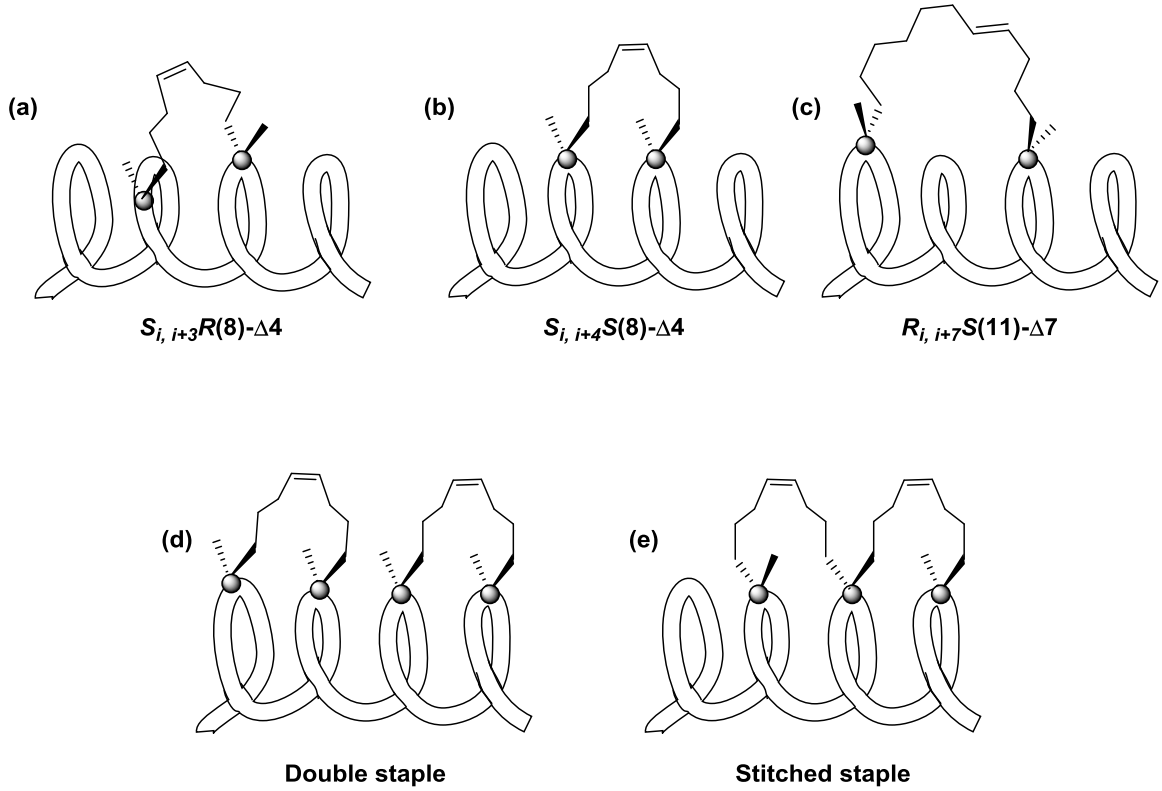


Figure 34: The different possible spacings for stapling α-helices: (a) i,i+3, (b) i,i+4, (c) i,i+7, (d) i,i+4 double staple and (e) i,i+4,i+7 stitched staple.

The optimisation of the *i,i+3*, *i,i+4* and *i,i+7* all hydrocarbon staples include the use of the disubstituted α-methyl, α-alkenyl glycine derived amino acid as the RCM precursor. The optimised staple lengths have been transferrable across different peptides. Nevertheless, it is still necessary too incorporate the hydrocarbon constraint at an appropriate location in the peptide sequence for maximal helix stabilisation.

To achieve these precise lengths and stereogenic centres for an *i,i+3* staple, one R_5 and one S_5 α-methyl, α-alkenyl amino acid are required. The *R* and *S* symbols account for the stereogenic centre at the α-carbon of the backbone amino acid. The subscript number accounts for the chain length relating to the alkenyl moiety of the amino acid.

With the *i,i+4* staple, two S_5 or two R_5 amino acids are required. This fact makes the *i,i+4* staple the easiest one to synthesise because only one type of amino

acid is required. The most common form of the amino acid used of this type is the S_5 amino acid.

The *i,i+7* staple, just like the *i,i+3* staple, requires the two amino acids to have opposite stereogenic centres. In addition, it also requires the two amino acids to have an α-alkenyl group of different carbon chain lengths. In particular, it requires an R_8 and an S_5 amino acid and *vice versa*.

Along with these most common staple lengths encompassing one or two turns of the α-helix, it is also possible to staple across three turns of the α-helix using an *i,i+11* staple.[163]

For each peptide that is stapled, the location where the staple is placed with in the peptide is crucial and can dictate the physicochemical properties of the stapled peptide. Study and understanding of the native structure is generally required to give the best staple positions. These structural studies can only be undertaken when crystal structures of the target peptides in their native environment are known. However, in most cases these crystal structures are not known.

When limited structural information is known regarding a particular PPI, the use of a staple scan can be implemented. A staple scan can help determine the most effective position for the staple constraint within a peptide sequence. This is achieved through the placement of a staple sequentially along the length of the peptide, mutating two residues at a time from the native sequence. In this way, a selective inhibitor for MCL-1, a human cancer resistance factor, was developed.[164] From the staple scan, Stewart and colleagues found that the peptide which had the highest α-helical character also had the highest binding affinity for MCL-1 (Figure 35).[164]

Compound	Sequence
MCL-1 BH3	KALETLRRVGDGVQRNHETAF
MCL-1 $SAHB_A$	KALETLR**X**VGD**X**VQRNHETAF
MCL-1 $SAHB_B$	KAL**X**TLR**X**VGDGVQRNHETAF
MCL-1 $SAHB_C$	KALETLRRV**X**DGV**X**RNHETAF
MCL-1 $SAHB_D$	KALETLRRVGDGV**X**RNH**X**TAF
MCL-1 $SAHB_E$	KALETLRRVGDGVQR**X**HET**X**F

X = S_5

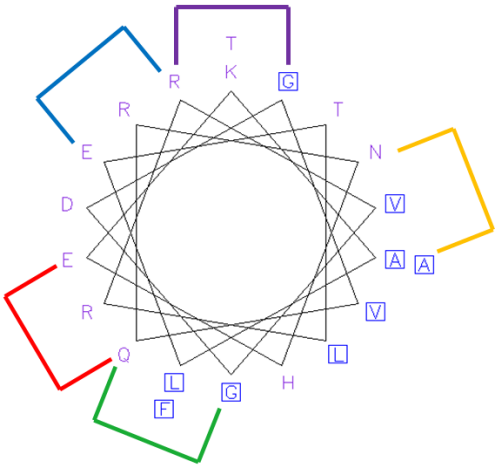


Figure 35: Staple scan and helical wheel diagram of peptides used as inhibitors for MCL-1.

Along with the initial studies carried out by Verdine and co-workers,[138] other groups are also attempting to improve on the all-hydrocarbon stapling technique. Indeed, the placement of the alkene within the *i,i+4* staple macrocycle as well as the optimum macrocycle size has been studied.[165,166] A comparison of the alkene position was carried out using oct-2-enyl, oct-6-enyl and the more usual oct-4-enyl macrocylic constraints.[165] Both the oct-2-enyl (S_3-S_7) (Figure 36c) and oct-6-enyl (S_7-S_3) (Figure 36e) staples failed to undergo RCM to give the corresponding macrocycle at room temperature. However, the oct-2-enyl staple did undergo RCM at an elevated temperature of 60 °C, while the oct-6-enyl did not. Even so, at 60 °C, decomposition of the peptide substrate was observed.[165] Only the RCM on the oct-4-enyl (S_5-S_5) staple (Figure 36d), with the alkene placed at the regular position in the macrocycle, proceeded at room temperature. Although the helicity of the oct-2-enyl hydrocarbon stapled peptide did increase the helicity of the peptide in comparison with the native sequence, it did not match the increase in helicity delivered by the oct-4-enyl stapled peptide.[165]

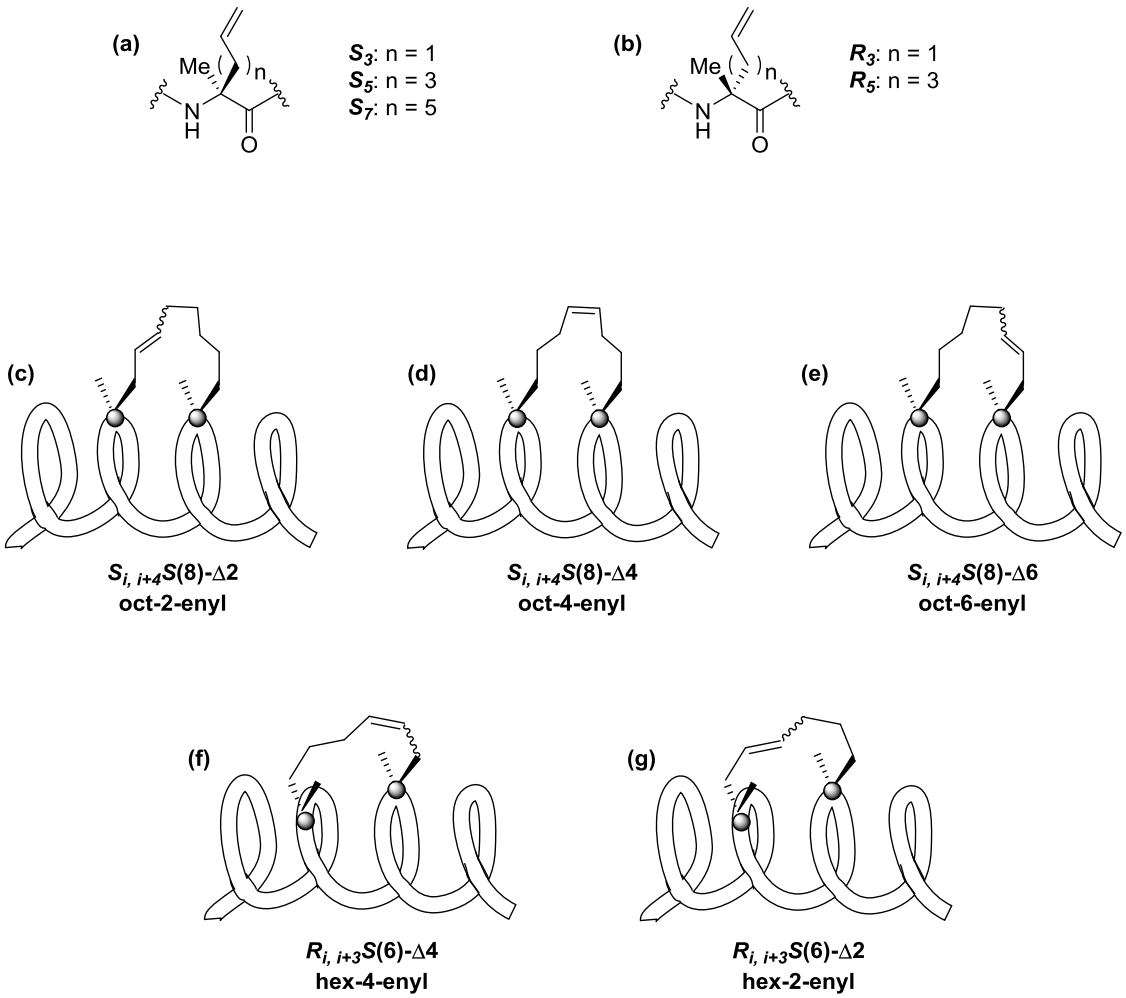


Figure 36: Structures of the (a) *S*- and (b) *R*- α-methyl, α-alkenyl amino acid precursors for stapling and the various positions the alkene moiety in the all-hydrocarbon staple for constraining across one turn of an α-helix for (c-e) an *i,i+4* staple and (f-g) an *i,i+7* staple.

The placement of the alkene within the *i,i+3* staple constraint has also been studied, by contracting the macrocycle ring size.[166] This was achieved by the removal of two methylene units from the macrocycle forming staple.

The removal of the methylene groups also shifts the placement of the alkene moiety within the macrocycle itself, and so the alkene is no longer central for most of the examples. Of course, by removing one methylene group from each side of the alkene, the alkene is still central in the macrocycle.

With the exclusion of two methylene groups, most of the linear peptide sequences failed to undergo RCM to give the corresponding macrocycle. There

were however two constrained peptides which did show the capability to form a macrocycle as a result of the RCM reaction. Although the hex-4-enyl (R_5-S_3) (Figure 36f) and hex-2-enyl (R_3-S_5) (Figure 36g) peptides both formed a macrocycle, the hex-4-enyl was shown to be much more helical (55%) in comparison with 15% for the hex-2-enyl hydrocarbon stapled peptide.[166]

1.4.2 Synthesis of α,α-disubstituted amino acids

The synthesis of a stapled peptide begins with the replacement of two amino acids in a native peptide sequence with two alkenyl bearing amino acid residues. There are some commercially available unnatural (*R*)- and (*S*)-allyl amino acids, as well as a few α,α-disubstituted akenyl amino acids. However, the cost of the α,α-disubstituted akenyl amino acids is expensive relative to standard Fmoc protected amino acids and the availability of more unusual unnatural amino acids for stapling is still limited. Consequently, methods for the efficient synthesis of unnatural alkenyl amino acids are highly desirable.

The asymmetric synthesis of the unnatural α,α-disubstituted alkenyl amino acids required for peptide stapling can be achieved using the method reported by Williams and Im.[167] The method involved the alkylation of a glycine enolate followed by a lithium reduction to obtain the amino acid. However, one particular drawback to this method is the potential for over reduction of the alkene moiety. As a result, the Belokon method, involving the alkylation of a nickel II Schiff base complex has been more widely used.[168,169]

More recently Aillard and co-workers improved on the synthesis further by using a fluorine-modified nickel II Schiff base complex to efficiently synthesise unnatural α,α-disubstituted alkenyl amino acids.[170] Additionally, the group also improved on the diastereoselectivity of the complex to >95:5 dr. A possible reason for the improved diastereoselectivity using the (*S*)-*N*-(2-Benzoylphenyl)-1-(2-fluorobenzyl)pyrrolidine-2-carboxamide (2-FBPB) ligand is due to the fluorine atom creating a partial positive charge on the benzyl group as a result of a dipole that could potentially interact with the amide oxygen, to help stabilise the complex. This is compared with the π-π stacking interaction between the *N*-

benzyl functionality and the proline amide bond which normally helps to stabilise the complex when the BPB ligand is used[170].

1.4.3 Ring closing metathesis

A key step in the synthesis of stapled peptides is the macrocyclisation reaction. Once the unnatural α,α-disubstituted alkenyl amino acids have been incorporated into the peptide sequence, RCM is carried out to form the macrocyclic peptide. This is achieved through the use of ruthenium-based catalysts which have high tolerance for different functional groups. Thus, these catalysts are suitable to use with peptides that can have a plethora of different functionalities.

The first catalyst to be used for RCM with peptides was (diphenylallyldiene)bis(tricyclohexylphosphine)-dichlororuthenium (Figure 37a).[150] However, with the more recent development of bis(tricyclohexylphosphine)benzylidine ruthenium(IV) dichloride, also known as Grubbs' 1st generation catalyst (Figure 37b) and [1,3-bis-(2,4,6-trimethylphenyl)-2-imidazolidinylidene]dichloro(phenylmethylene)(tricyclohexyl-phosphine)ruthenium, commonly referred to as Grubbs' 2nd generation catalyst (Figure 37c), the use of (diphenylallyldiene)bis(tricyclohexylphosphine)-dichlororuthenium has largely been eliminated and the newer catalysts have come to the forefront. Normal catalyst loading for RCM is between 5-40 mol%, with 20 mol% being the most frequently used.[150,155]

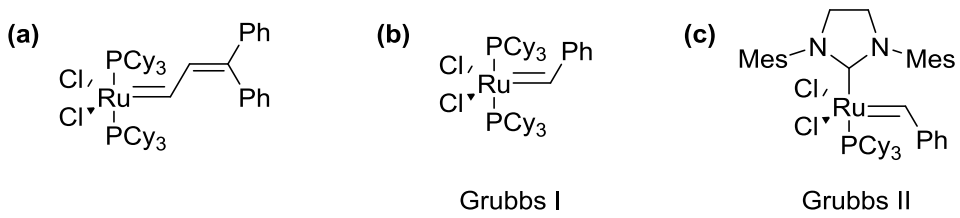


Figure 37: Structures of three different catalysts that have been used for ring closing metathesis on peptide sequences.

The initial RCM reactions on peptides were carried out in solution phase.[150] However, in 1996, Grubbs and colleagues performed the first on-resin RCM of a peptide.[151] This allows for any waste and solvents left over from the reaction to be easily separated from the peptide. Furthermore, on-resin peptide synthesis also allows for a much cleaner overall peptide to be synthesized because of the ability to easily remove by-products. However, there are still some limitations to RCM using ruthenium catalysts. More particularly, the synthesised peptide can still contain some degree of ruthenium contamination and for biological studies this can be an issue. Although the most of the remaining ruthenium can be removed through reverse-phase HPLC during purification, it can then contaminate the chromatography column.

1.4.4 Economic value

Aileron Therapeutics, Inc., recently released details of the first ever successful Phase I clinical trial of a stapled α-helical peptide, ALRN-5281.[171] ALRN-5281 is a long-acting growth-hormone-releasing hormone (GHRH) agonist for the treatment of orphan endocrine diseases, which include, amongst others, adult growth hormone (GH) deficiency and human immunodeficiency virus (HIV) lipodystrophy.[172] Additionally, Aileron Therapeutics, Inc. have also published preclinical data regarding another stapled peptide, ATSP-7041,[173] which is due to start clinical trials in 2014. ATSP-7041 is a first-in-class potent and selective dual inhibitor of MDM2 and MDMX. It is an *i,i+7* hydrocarbon stapled peptide which increases its helicity to 70% with the inclusion of the staple in comparison with only 11% helicity in its linear form at pH 7.0.[173] Its function is to reactivate the p53 tumour suppressor pathway in a mechanism-dependent manner in p53-dependent cancers.[173,174]

1.5. Peptide synthesis

1.5.1 Solid phase peptide synthesis

The field of solid phase peptide synthesis (SPPS) originally came about from the Nobel Prize winning work of Bruce Merrifield in the 1960s.[175] Since that time, it has been of great use to peptide chemists everywhere due to its ease of use, efficiency and relatively low cost. Peptides are synthesized from the *C*- to

N-terminus using the SPPS method; the opposite of ribosomal synthesis in the cell.

The idea behind SPPS involves synthesizing a peptide sequence attached to an insoluble polymer bead. This allows for the easy removal of solvent and unreacted reagents from the reaction vessel (Figure 38). Along with these, impurities can also be removed simply by dissolving them in the appropriate solvent and filtration away from the peptide. The peptide remains securely on the resin bead, and is therefore not washed away.

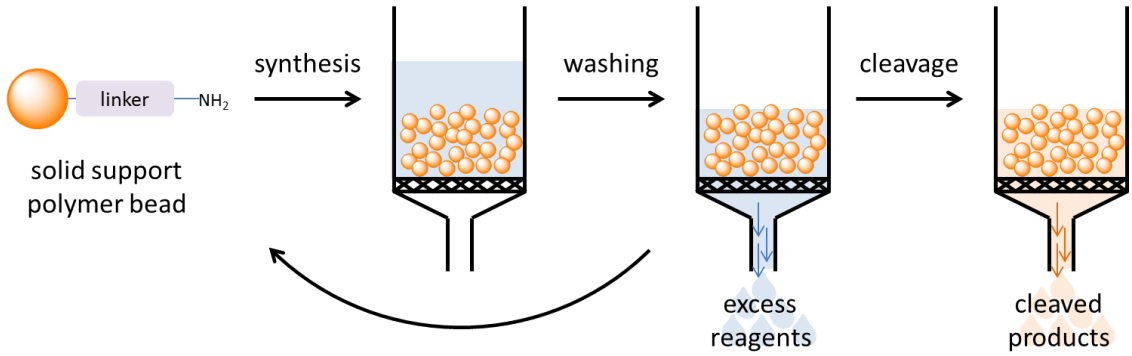


Figure 38: Solid phase peptide synthesis (SPPS) involves the synthesis of the peptide on a solid support polymer bead. Following coupling and deprotection steps the resin is washed to remove excess reagents before the peptide is cleaved from the resin.

For the polymeric solid support itself, one of the main requirements is that it is insoluble in all solvents that would be used in the synthesis of the peptide. In addition, the polymer bead must also be functionalised to allow it to covalently bond to the first protected amino acid residue (the C-terminus). In general, resins used for SPPS also contain a linker between the polystyrene and the peptide. The linker also dictates the functionality that is transferred onto the peptide upon cleavage from the resin such as an acid or an amide to name a few.

Another important consideration is the substitution on the resin, in other words, the number of moles of peptide that can be synthesized per gram of the resin. In general, the longer the peptide required, the lower the resin substitution

There are a number of different resin types available depending on the requirements of the peptide being synthesized.

From the work pioneered by Merrifield came Merrifield resin, the common name for chloromethylated polystyrene crosslinked with 1-2% divinylbenzene (Figure 39a). This remains one of the most popular resins used alongside Boc based peptide synthesis. Attachment of the first protected amino acid onto the resin involves the nucleophilic displacement of the chlorine atom from the polymer. Further elongation of the peptide chain proceeds via the usual peptide coupling steps. Cleavage of the peptide from Merrifield resin is commonly undertaken using hydrofluoric acid (HF), but it can also be done using trifluoromethanesulfonic acid (TFMSA) or trifluoromethanesulfonate (TMSOTf).The need for HF is due to the relative lower stability of the carbocation formed upon cleavage of the peptide from the solid support, when compared to resins cleaved using TFA.

One disadvantage of Merrifield resin is that it is acid labile to some degree, and therefore not particularly suitable to the synthesis of very long peptides. This is because during a Boc deprotection on resin using TFA, some of the peptide is also likely to be cleaved from the resin support. However, the greatest concern when using Boc-based SPPS still remains with the cleavage of the peptide from the solid support, which requires anhydrous liquid HF. This requires very careful handling because HF is extremely toxic, corrosive and volatile. Additionally, specialised all-fluorocarbon apparatus is required as HF reacts with glass.

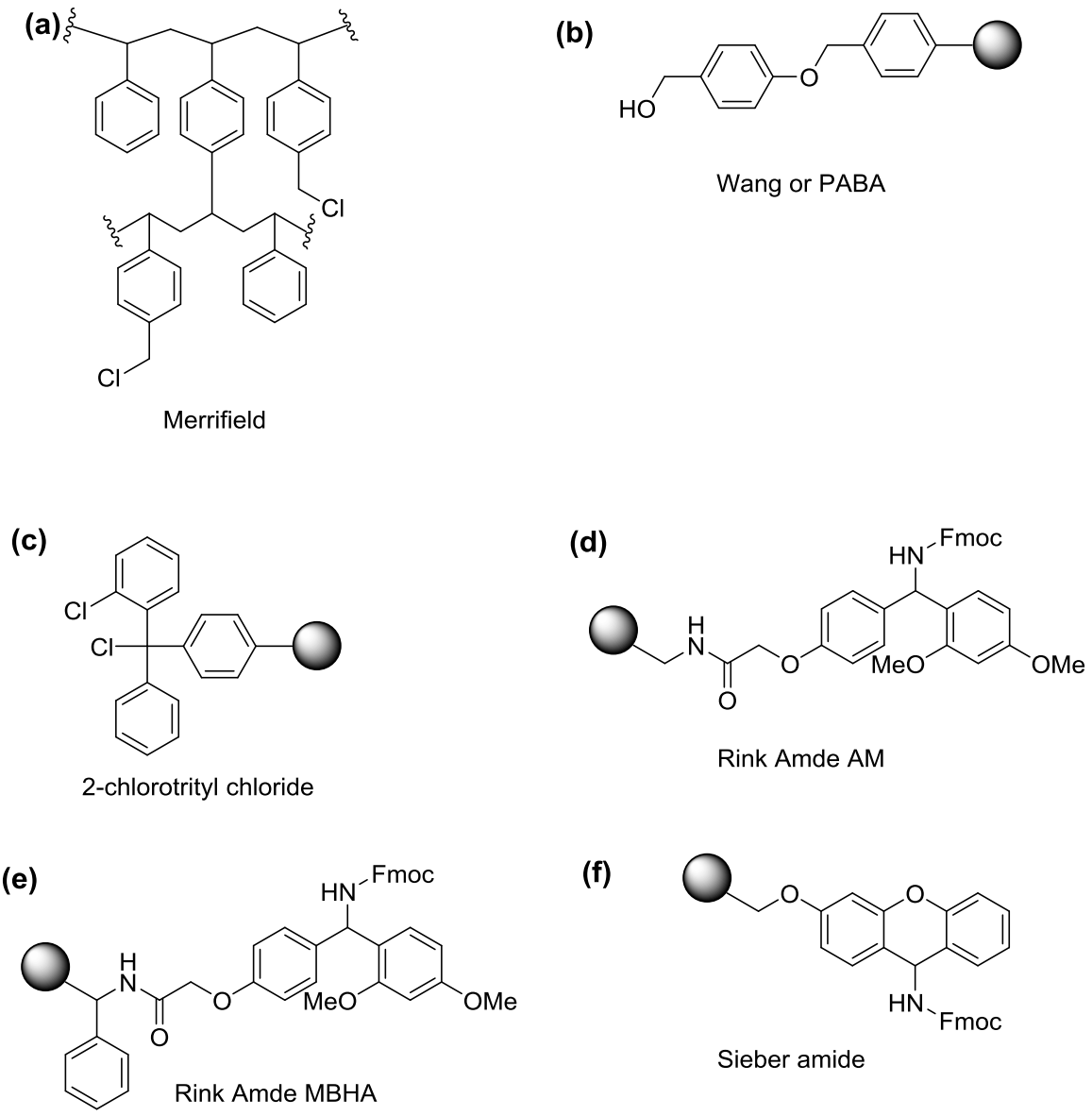


Figure 39: Some of the different types of resins that can be used for SPPS. (a) Merrifield resin, was the original resin used for SPPS and is a chloromethylated polystyrene resin crosslinked with 1-2% divinylbenzene, then modified with linkers to give (b) *p*-alkoxybenzyl alcohol (PABA) and (c) 2-chlorotrityl chloride resins which are used to give *C*-terminal acid peptides and (d) Rink amide amino methyl (AM), (e) Rink amide 4-methylbenzhydrylamin (MBHA) and (f) Sieber amide resins to give *C*-terminal amide containing peptides.

Along with Merrifield resin, other resins use its polymeric backbone, with the addition of various linkers. By this way, many resins with specialised applications have been made. Of particular note, Wang resin, also called PABA (*p*-alkoxybenzyl alcohol) resin (Figure 39b),[176] is the most common resin used

with Fmoc chemistry. As with Merrifield resin, Wang resin is acid labile and after cleavage, leaves a C-terminal acid. However, in comparison, peptides on Wang resin can be cleaved using conditions that are much milder, that is, it uses a cleavage cocktail of 50% v/v TFA/DCM. It is for this reason that Wang resin has risen greatly in popularity. Attachment of the first amino acid onto Wang resin requires an activating agent such as dicyclohexylcarbodiimide (DCC) and a catalytic amount of 4-dimethylamino-pyridine (DMAP). Hydroxybenzotriazole (HOBt) is also required to reduce the amount of epimerization that occurs when the first amino acid is attached to the resin. In addition, racemization at the first amino acid attachment is also very likely when the amino acid in question is cysteine or histidine. After coupling of the first amino acid, acetyl capping is also required to block unreacted sites on the resin. As a consequence of the difficulty of the first amino acid attachment, preloaded resins are commercially available, that is the first Fmoc-protected amino acid is already attached to the Wang resin.

Another acid labile resin that is in common use for Fmoc peptide chemistry is 2-chlorotrityl chloride resin (Figure 39c).[177,178] As with Wang resin, it requires relatively mild cleavage conditions to afford a *C*-terminus acid on the peptide. The mechanism of the first residue attachment to 2-chlorotrityl chloride resins follows an S_N1 route which allows no possibility of racemisation. On the other hand, 2-chlorotrityl chloride resin is moisture sensitive and needs to be stored in a desiccator in a tightly sealed container. Without this precaution, water can hydrolyse the resin to give the alcohol, which is not very reactive. However, the hydrolysed resin can be reactivated back to the 2-chlrotrityl chloride through the reaction with acetyl chloride in DCM in anhydrous conditions. To prevent the likelihood of water hydrolysis 2-chlorotrityl resins are also commercially available with the first amino acid residue already attached.

The need for *C*-terminal amide peptides is also great. For such peptides, Rink amide resin is the most useful. There are two main types of Rink amide resin: Rink amide aminomethyl (AM) resin and Rink amide 4-methylbenzhydrylamine (MBHA) resin. Rink amide AM resin consists of an Fmoc-rink amide linker attached to AM resin (Figure 39d), whereas Rink amide MBHA resin has the Fmoc-Rink amide linker attached to MBHA resin (Figure 39e).[179]

In some cases, there is a need to cleave the peptide from the resin support without also removing the protecting groups. Sieber amide resin (Figure 39f) allows for this because it is highly acid labile and requires only about 1% TFA in the cleavage solution.[180] Thus, this gives a fully protected peptide with a *C*-terminal amide.

More recently, new resin supports have been developed for SPPS which include ChemMatrix® and SpheriTide®. The ChemMatrix® resin (Figure 40a) is a 100% polyethylene glycol polymer which was developed by PCAS BioMatrix Inc. The structure of the resin allows for optimal chemical stability. It was developed to improve the crude peptide purity of long, complex or hydrophobic peptides in comparison to polystyrene based resins.

Additionally, SpheriTide® resin (Figure 40b), from Spheritech Ltd., was also developed to overcome a weakness in traditional polystyrene based resins. In particular that on-resin peptide aggregation is a widely-known problem in SPPS which is often encouraged by a hydrophobic resin environment.[181] As a result of its poly-ε-lysine based cross-link network, SpheriTide® resin gives a more hydrophilic environment, which, like the growing peptide, contains amide bonds. These amide bonds discourage the intra-molecular on-resin peptide aggregation. This in turn allows for easier synthesis and higher purities of long and complex peptide sequences as a result of better efficiencies of the Fmoc-deprotection and amino acid coupling reactions.

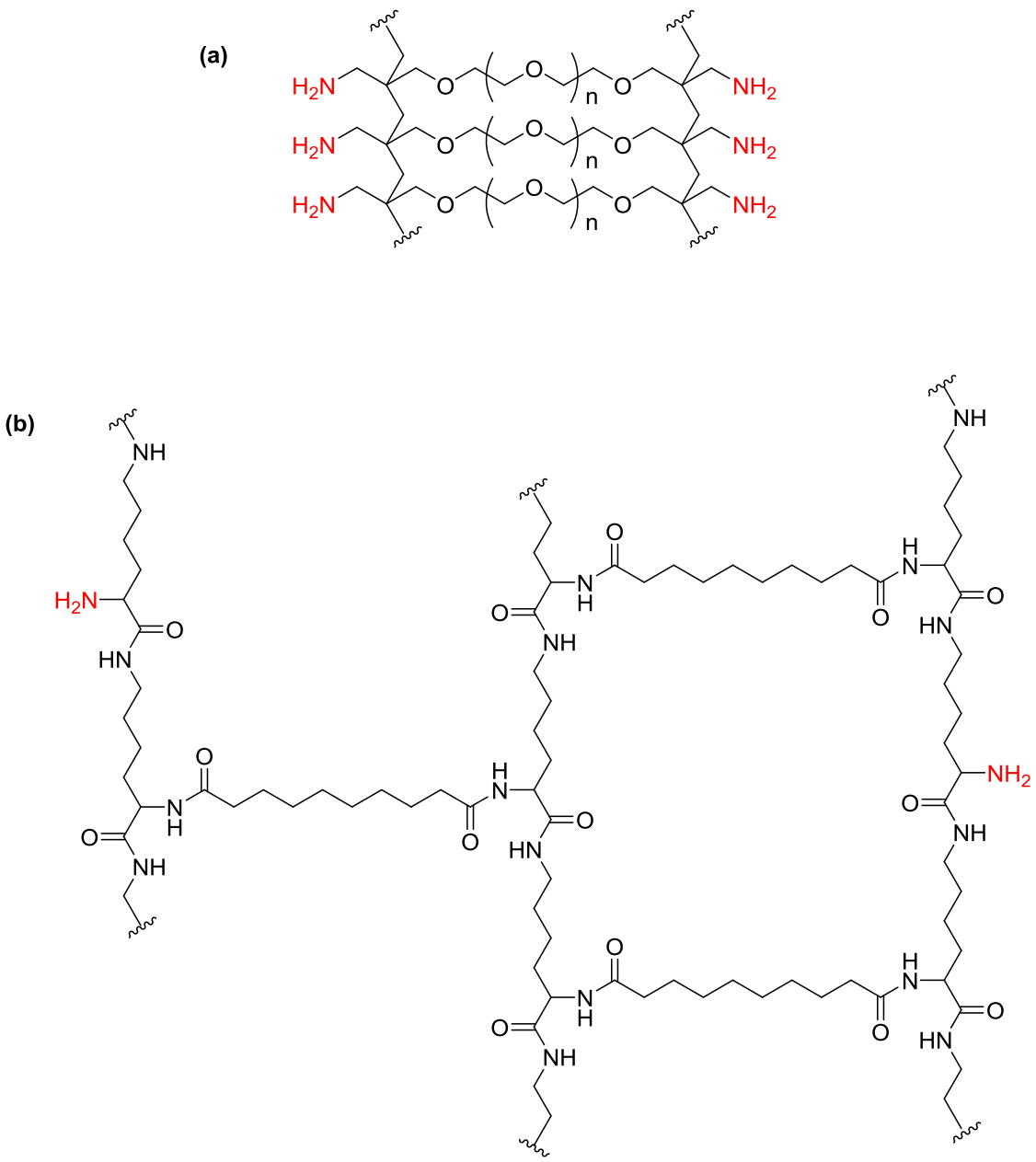


Figure 40: New polymeric resin supports developed for SPPS: (a) ChemMatrix® and (b) SpheriTide® with the amine functionality (red) on the resin highlighted.

1.5.2 Microwave assisted peptide synthesis

Many peptide sequences are a challenge to synthesize even with the development of SPPS. In addition, for peptides of relatively short length, many coupling and deprotection steps are still required to obtain the final peptide sequence and each of these reactions must proceed at very high efficiency to

be able to produce a relatively clean peptide. However this can still be a challenge and many of the coupling and deprotection steps do not go fully to completion and can affect the yield and purity of the peptide. As a result, new methods were investigated to improve the synthetic outcomes of such peptide sequences. Coupled with this need, the growing area of using microwave irradiation for the improvement of organic syntheses, came together to give what is now common place in peptide chemistry, the use of microwave irradiation in the synthesis of peptides.

It has been shown that some sequences can be synthesized in a fast and efficient manner without the aid of microwave power and with retention of product purity.[182] However, there is still the problem of a significant excess in reagents that is needed per coupling step. In addition, the use of unnatural amino acids has become much more popular and these generally require longer reaction times as a result of their associated steric hindrance.[183] Unnatural amino acids are also much more expensive than the natural amino acids and therefore the use of fewer equivalents of reagents in combination with heat is greatly desired. Therefore, the use of microwave irradiation can be looked upon as a way of synthesizing difficult peptide sequences in a fast and efficient manner.

Microwave irradiation has also been proposed to help disrupt on-resin aggregation of peptides as a result of inter- and intra-chain associations by rapid heating of polar and ionic species using microwave energy and so the increase in temperature allows for easier access to the reactive N-terminal amine and therefore more complete reactions.[184,183]

The first reported microwave-assisted peptide synthesis was reported by Yu and colleagues in 1992 in which they used a domestic microwave oven in their synthesis which led to irreproducible conditions.[185] More recently, commercial microwave reactors have been produced that allow for control of power output and so constant reaction temperatures can be achieved and measured in-situ using a fibre optic probe.[186]

There are several reports which compare the use of conventional and microwave heating in the synthesis of peptides.[187–189] In general, the microwave

heating methods gave much higher yields and purities of the synthesized peptides over a much shorter time frame.[188,190,191]

The first reported use of microwave energy for Fmoc deprotection was in 2003[192] and has grown in use owing to the availability of automated and manual microwave-assisted peptide synthesizers.[183] In general, for microwave Fmoc deprotection, an initial deprotection step of ~30 s is first applied followed by a longer second deprotection of ~3 min at T_{max} = 75 ˚C.[183] Recently, the Fmoc-deprotection step was further optimised on the CEM Liberty Blue peptide synthesiser at a temperature of 90 ˚C for just 1 min.[193]

The deprotection steps are usually carried out using 20-30% piperidine in DMF or 5% piperazine in DMF. The deprotection mixes can also include additives such as formic acid, HOBt, 4-nitrophenol and HFIP to reduce base-catalysed side reactions such as aspartimide formation.[194]

As a result of the fulvene-piperidine by-product from Fmoc deprotection, this reaction can be monitored using UV at 301 nm. In fact, in 2010, CEM added UV monitoring capabilities onto their automated Liberty peptide synthesizer system.[195] However, the use of UV monitoring for deprotection is not suitable when additives such as HOBt and OxymaPure are used due to their inherent UV absorbance.

The use of microwave technology for peptide synthesis is certainly an attractive prospect. In fact, intellectual property has been fought over the use of automated microwave-assisted peptide synthesis, specifically between the companies CEM and Biotage. As a result, the peptide synthesisers developed by other companies, in particular Biotage, cannot use microwave-assisted Fmoc deprotection in conjunction with UV monitoring of the fulvene-piperidine by-product.

1.6. Aims of the project

The aims of the project are to dissect the structure activity relationships between HDAC enzymes and their corepressor complexes. The main focus will be the HDAC3/DAD corepressor complex, followed by two of the HDAC1

corepressor complexes: HDAC1/MTA1 from the NuRD complex and HDAC1/CoREST/LSD1 from the CoREST complex.

Isoform specific inhibition of the HDAC3/DAD complex will be achieved using stapled peptide analogues of its corepressor DAD. This will be done by first synthesising unnatural amino acids for incorporation into the peptide sequences. These unnatural amino acids are the precursor residues for peptide stapling as well as active site inhibitor motifs that can be incorporated into substrate peptides. The isoform specificity of the inhibitor is imagined to come from the stapled peptide binding to the HDAC3 surface to displace the natural corepressor, DAD.

For the HDAC1/MTA1 complex, rather than an allosteric inhibitor, an active site inhibitor will be synthesised based on its natural substrate, the histone H4 tail peptide sequence, and a well-known pan-inhibitor. In this example, the isoform specificity of the inhibitor peptide is postulated to come from the sequence of the H3 tail.

In the ternary complex, HDAC1/CoREST/LSD1, rather than finding an inhibitor of the complex, its function will be studied as both a deacetylase and demethylase. This will be achieved through the synthesis of short histone H3 tail peptide sequences, the natural substrate for both HDAC1 and LSD1. These peptides will contain acetylated and/or mono- or dimethylated lysine residues. Following the peptide synthesis, NMR studies will be carried out to monitor the activity of the ternary complex.

Chapter 2: Synthesis of stapled peptide proteomimetics to regulate the HDAC3:SMRT-DAD protein-protein interaction

2.1. Introduction

All current HDAC inhibitors are classed as pan inhibitors because they inhibit all 11 zinc-dependent HDAC isoforms. Thus, inhibitors that are isoform selective are highly desirable. Currently, all the HDAC inhibitors also target the active site of the enzyme. One potential way of achieving an isoform specific HDAC inhibitor is to target an allosteric site of the HDAC enzyme rather than its active site. An allosteric site allows for regulation of the protein without affecting the active site of the enzyme.

As mentioned in Chapter 1, the class I HDAC enzymes need to be recruited to their cognate corepressor complexes in order to become catalytically active. HDAC8 is the exception to the rule because it is active in isolation, without the need for further corepressor recruitment. Of the class I HDACs which do need recruitment to corepressor complexes, HDAC3 is very selective as to the complex that it is recruited to. It is uniquely recruited to the SMRT/NCoR corepressor complex. HDACs 1 and 2 are more promiscuous in their corepressor recruitment being recruited to the Sin3A, NuRD, CoREST and MIDAC complexes.[39]

From this information, the HDAC3:SMRT-DAD corepressor complex was chosen as a suitable candidate with which to achieve isoform selectivity. Once a suitable allosteric inhibitor has been developed, the other complexes can then be investigated.

The allosteric inhibition of the HDAC3 enzyme can be achieved by first studying the available crystal structure of HDAC3 bound to a domain of the corepressor SMRT (Figure 41).[42] This domain is called the deacetylase activation domain (DAD), as it recruits and activates HDAC3. Also in the crystal structure is a molecule of Ins(1,4,5,6)P_4 which is found at the interface of HDAC3:SMART-DAD. From this crystal structure, it can be seen that there is a distinct PPI between the HDAC3 surface and SMRT-DAD, in particular with helices 0 (422-429) and 3 (462-466) of SMRT-DAD. In this study, the focus will be on helix 3 of SMRT-DAD.

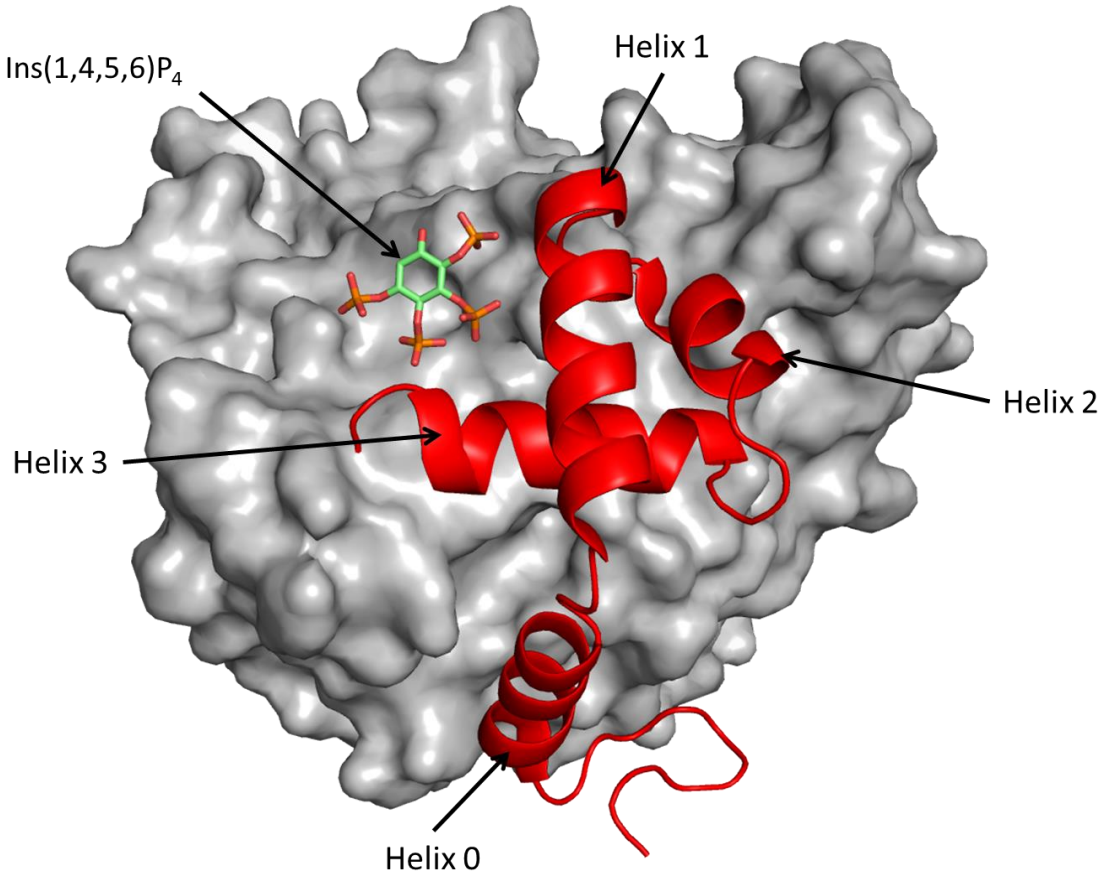


Figure 41: Crystal structure of HDAC3 (grey surface) bound it its corepressor the deacetylase activation domain (DAD, red helices) and a molecule of inositol tetraphosphate (Ins(1,4,5,6)P_4, sticks), which was pulled out from the protein purification step (pdb: 4A69).[42] The four helices of SMRT-DAD that were crystallised are labelled.

Allosteric inhibition can potentially be achieved through the disruption of the PPI between HDAC3 and SMRT-DAD. This can be realised through the design of a conformationally constrained peptide mimic of SMRT-DAD. The constraint allows the peptide mimic to adopt a more helical conformation, without which out with the larger protein there would be no structure.

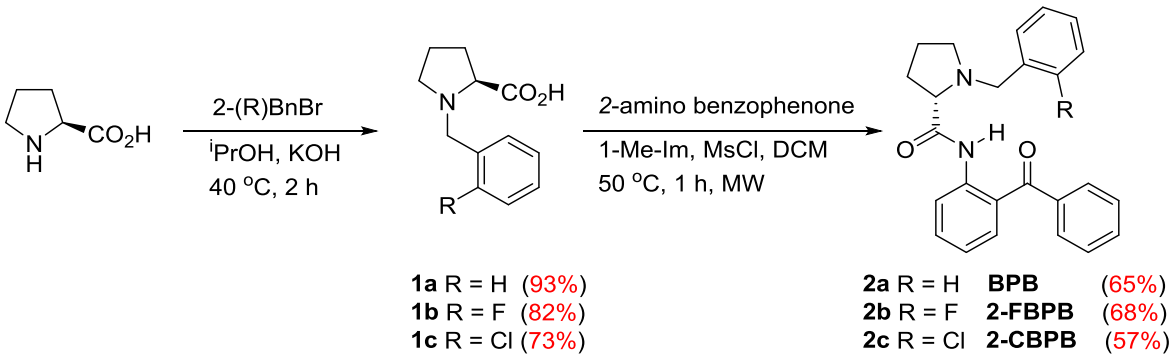
2.2. Aims of this chapter

The aims of this chapter are:

- The development and asymmetric synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-2-methylhept-6-enoic acid an unnatural α,α-disubstituted alkenyl amino acid using Ni (II) Schiff base methodologies.
- The synthesis and conformational analysis of native and stapled SMRT-DAD peptide analogues.
- *In vitro* biological assays using a HDAC fluorescence activity assay.

2.3. Synthesis of (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-2-methylhept-6-enoic acid

The synthesis of substituted benzyl pyrrolidine derivatives (**2.1a-c**) were prepared by using a modified version of Belokon's method from commercially available L-proline and 2-substituted benzyl bromides[168,169] (Scheme 7).



Scheme 7: Synthesis of ligands **2.2a-c** from readily available L-proline.

However, unlike what is suggested in the literature, **1b** and **1c** were not isolatable purely by precipitation and filtration. Instead, after acidification to pH 5 (using a pH meter), the precipitate was filtered out and the filtrate itself was concentrated and washed consecutively with acetone, chloroform and ethyl acetate, before the solvent was evaporated and the product dried to give **1b**

and **1c**, both as waxy solids. **1a** however, was able to be filtered out leaving an off white solid. In our hands, the addition of a halogen (F or Cl) to the aromatic of the benzyl pyrrolidine changed the properties of **1b** and **1c** making it soluble in chloroform.

These benzyl pyrrolidine derivatives were subjected to amide coupling conditions with 2-aminobenzophenone using methanesulfonyl chloride and *N*-methylimadazole (Scheme 7). This gave the corresponding auxiliaries (*S*)-2-[(*N*-benzylprolyl)amino]-benzophenone (BPB, **2a**) and the two 2-substituted auxiliaries (*S*)-*N*-(2-benzoylphenyl)-1-(2-fluorobenzyl)pyrrolidine-2-carboxamide (2-FBPB, **2b**) and (*S*)-*N*-(2-benzoylphenyl)-1-(2-chlorobenzyl)pyrrolidine-2-carboxamide (2-CBPB, **2c**). **2a-c** were recovered in respectable yields, considering the bulky nature of the 2-aminobenzophenone, which being an aniline has relatively low nucleophilicity because the lone pair of electrons on the nitrogen is conjugated into the ring system. A crystal structure of **2b** was also obtained showing hydrogen bonding between the benzophenone carbonyl and the hydrogen from the newly formed amide (Figure 42).

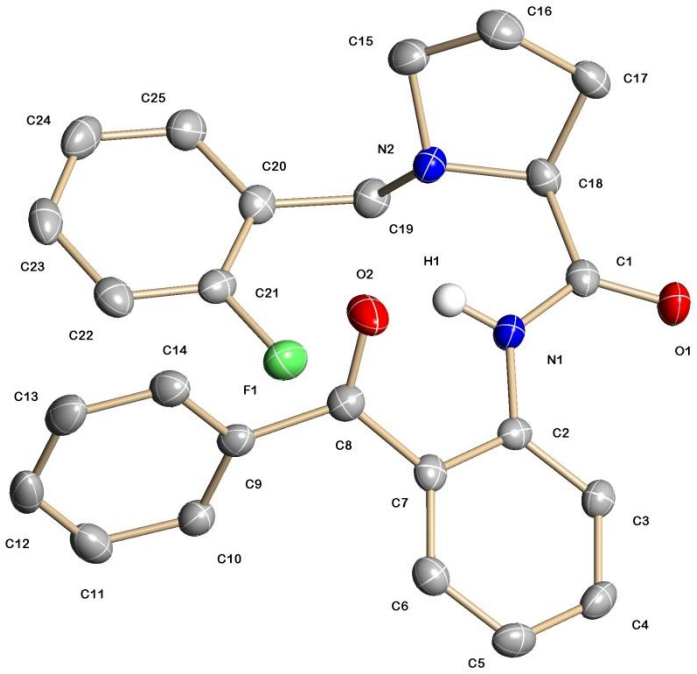
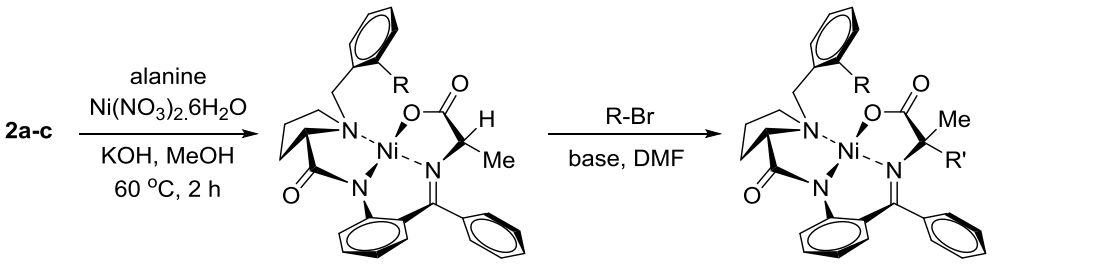


Figure 42: Crystal structure of 2-FBPB (**2b**) showing 50% displacement ellipsoids. The hydrogen atoms (except N-H) have been omitted. (CCDC 102152)

This amide coupling was carried out using a microwave reactor set at 50 °C for 1h after the activation of the acid to a mesylate ester at room temperature. However, due to the use of a microwave reactor, this amide coupling reaction was only suitable for use with a maximum of 3 g scale reactions. For reaction scales larger than 3 g, conventional heating was employed.

2a was isolated as the hydrochloride salt and collected as a precipitate after acidifying with concentrated hydrochloric acid. On the other hand, **2b** and **2c** were isolated in their pure form on extraction with dichloromethane, followed by purification by flash column chromatography.

Complexation of **2a-c** with nickel (II) nitrate hexahydrate and L-alanine under basic conditions gave their corresponding Ni (II) Schiff base complexes with a thermodynamic mixture of diastereomers (Scheme 8). This step to form the initial nickel Schiff base complex also occurred with stoichiometric yields to give red crystals that were stable under air storage. Crystal structures of **3a** and **3b** were also obtained (Figure 43).



Scheme 8: Synthesis and alkylation of nickel (II) Schiff base complexes.

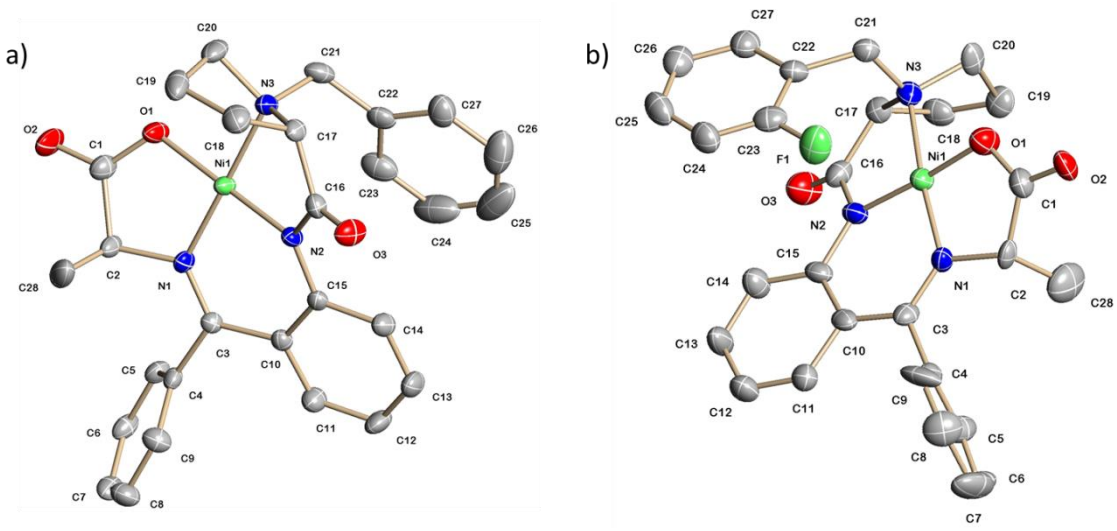


Figure 43: Crystal structures of a) compound **3a** and b) compound **3b** (CCDC 1020154) showing 50% displacement ellipsoids. The hydrogen atoms have been omitted from the diagrams.

3a showed correlation with the literature in that it had an $[\alpha]_D^{20}$ +2499 (0.036, MeOH) compared to the literature value of $[\alpha]_D^{25}$ +2643.[169] As well as this, **3a** showed promise in that it had a reasonable diastereomeric excess of 86%. **3b** and **3c** gave a diastereomeric excess of 98% and 96% respectively.

Alkylation of the Ni (II) Schiff bases was initially carried out with potassium *tert*-butoxide as the base. However, further into the study, it was replaced with sodium hydroxide, which is a much stronger base and makes the deprotonation of the α-proton on the alanine derived Schiff base easier.

For **3a**, the reaction was carried out at room temperature and reached completion within 10 mins as determined by TLC. However, for compounds **3b** and **3c**, the reaction required the formation of the enolate at 0 °C, followed by alkylation at 50 °C to get the same conversion. The resulting alkylated complexes, **4a-c**, were isolated with reasonable yields. For compound **4b** the diastereomeric ratio (>88:12) was determined by ^{19}F NMR on the crude material and reflects the position of the kinetic equilibrium and was isolated as one diastereomer after purification. Absolute stereochemistry of **4a** and **4b** were determined using X-ray crystal structures (Figure 44).

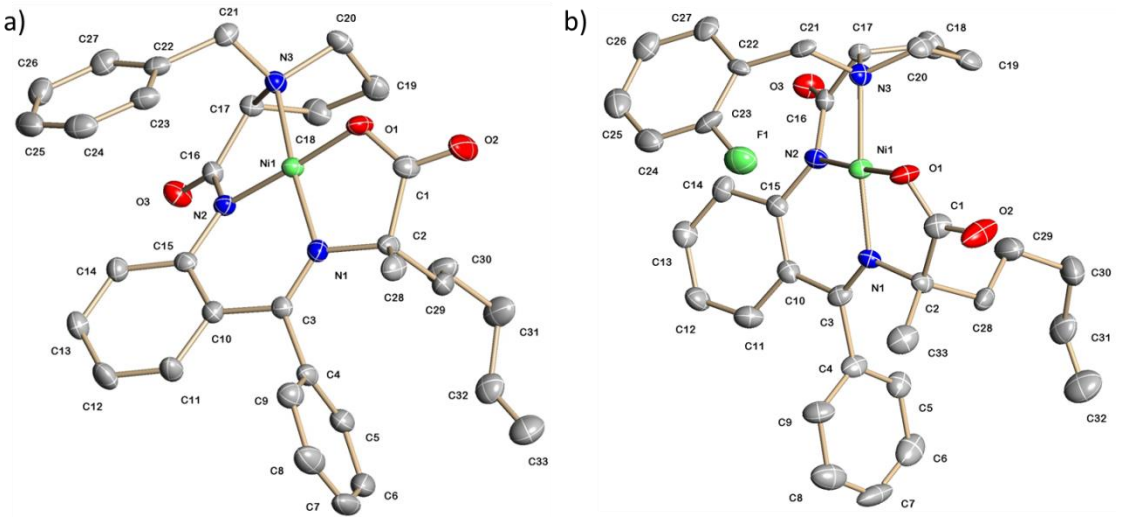
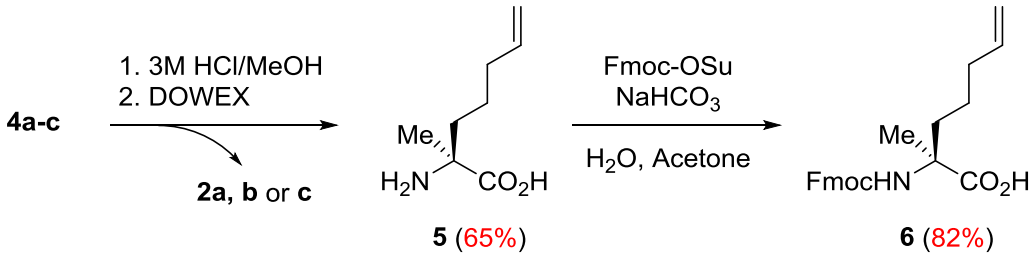


Figure 44: Crystal structures of a) compound **4a** and b) compound **4b** (CCDC 1020157) showing 50% displacement ellipsoids. The hydrogen atoms have been omitted from the diagrams.

The alkylated Ni^{II} Schiff base complexes were purified by column chromatography before decomposition with 3M HCl in methanol. The resulting amino acids were then purified on a Dowex ion-exchange resin column. The chiral auxiliaries **2a-c** were also recovered in good yield on decomposition of the alkylated Schiff base complexes **4a-c** respectively (Scheme 9). Following the decomposition of Ni (II) Schiff base complex, **5** was Fmoc-protected under mildly basic conditions to allow the subsequent use of Fmoc chemistry in the peptide synthesis (Scheme 9). The reaction produced (*S*)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-2-methylhept-6-enoic acid (Fmoc-S_5-OH, **6**) in 82% yield.



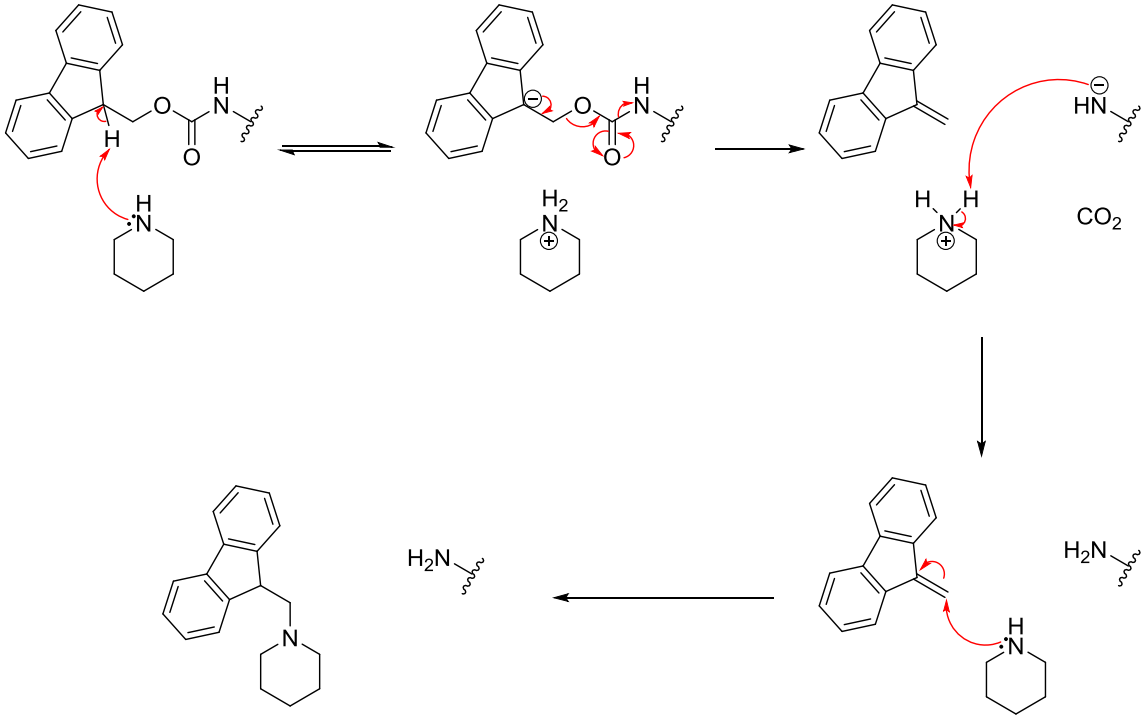
Scheme 9: Decomposition of nickel (II) Schiff base and amino acid purification, followed by Fmoc-protection.

2.4. Solid Phase Peptide synthesis

There are two variations on solid phase peptide synthesis (SPPS), the Boc strategy and the Fmoc strategy. In this body of work the focus is on Fmoc-based SPPS. One of the main advantages of the Fmoc strategy of SPPS is the orthogonality with the side chain protecting groups on the amino acids. Additionally, milder reaction conditions using piperidine are needed for Fmoc deprotection rather than the successive TFA treatment in the Boc strategy which can also result in side chain deprotection. However, an advantage in using the Boc strategy for SPPS is the possibility of synthesising longer peptides with a smaller chance of on resin aggregation occurring.

Fmoc deprotection

The three ring system of the fluorene group of Fmoc is very electron withdrawing thus making the lone proton at the 9-position very acidic (pK_a about 25). This proton is very base labile allowing its abstraction by weak bases such as piperidine, which has a pK_a of 11.2. Abstraction of the proton from the fluorene ring system generates a 14-electron aromatic cyclopentadiene-type intermediate which rapidly undergoes β-elimination (Scheme 10). This abstraction leads to the formation of dibenzofulvene which is a highly reactive compound quickly scavenged by excess piperidine to form the stable fulvene-piperidine adduct. Removal of the dibenzofulvene in turn also generates the free amine by first undergoing a decarboxylation step (Scheme 10). This amine can then be used for the coupling of Fmoc-protected amino acids in order to build up the peptide chain on the resin.



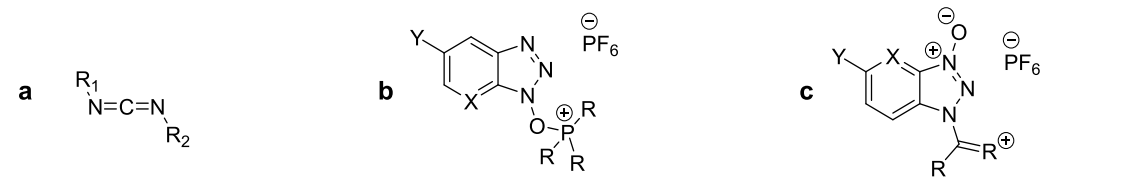
Scheme 10: Mechanism of Fmoc deprotection with piperidine to give the UV active fulvene-piperidine adduct.

Another advantage of using the Fmoc strategy is that the fulvene-piperidine adduct is also highly UV active due to the conjugated fluorene ring system. This allows monitoring of the reaction progression by UV at 301 nm.[181,196]

Amino acid coupling

Following the Fmoc-deprotection step, amino acid coupling can occur. In this manner, an Fmoc-protected amino acid can be coupled onto the deprotected amine of either the resin or another amino acid already coupled onto the resin.

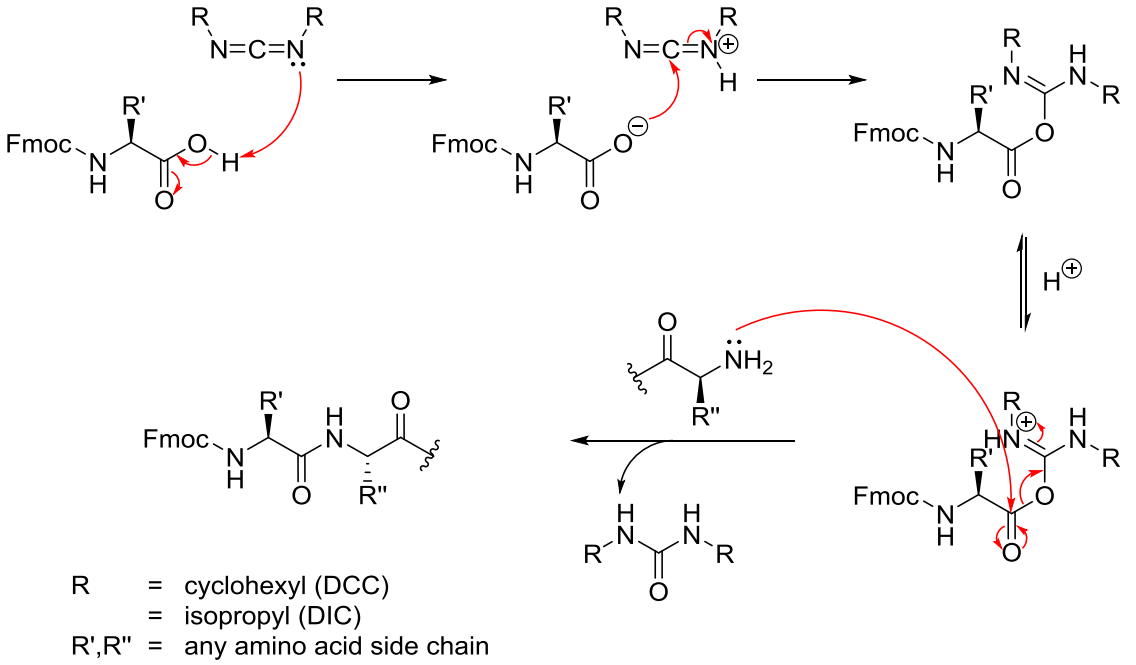
The coupling step involves the activation of the Fmoc-protected amino acid using activators such as carbodiimides (Figure 45a), phosphonium salts (Figure 45b) or uronium salts (Figure 45c).



DIC	$R_1 = R_2$ = isopropyl	**BOP**	R = $N(CH_3)_2$; X = Y = CH	**HATU**	R = $N(CH_3)_2$; X = N; Y = CH
DCC	$R_1 = R_2$ = cyclohexyl	**PyBOP**	R = pyrrolidine; X = Y = CH	**HBTU**	R = $N(CH_3)_2$; X = Y = CH
PIC	R_1 = isopropyl; R_2 = cyclohexyl	**PyAOP**	R = pyrrolidine; X = N; Y = CH	**HCTU**	R = $N(CH_3)_2$; X = CH; Y = Cl
		PyClock	R = pyrrolidine; X = CH, Y = Cl	**HAPyU**	R = pyrrolidine; X = N; Y = CH

Figure 45: Coupling agents that can be used for amino acid coupling: a) carabodiimides, b) phosphonium salts and c) uronium salts.

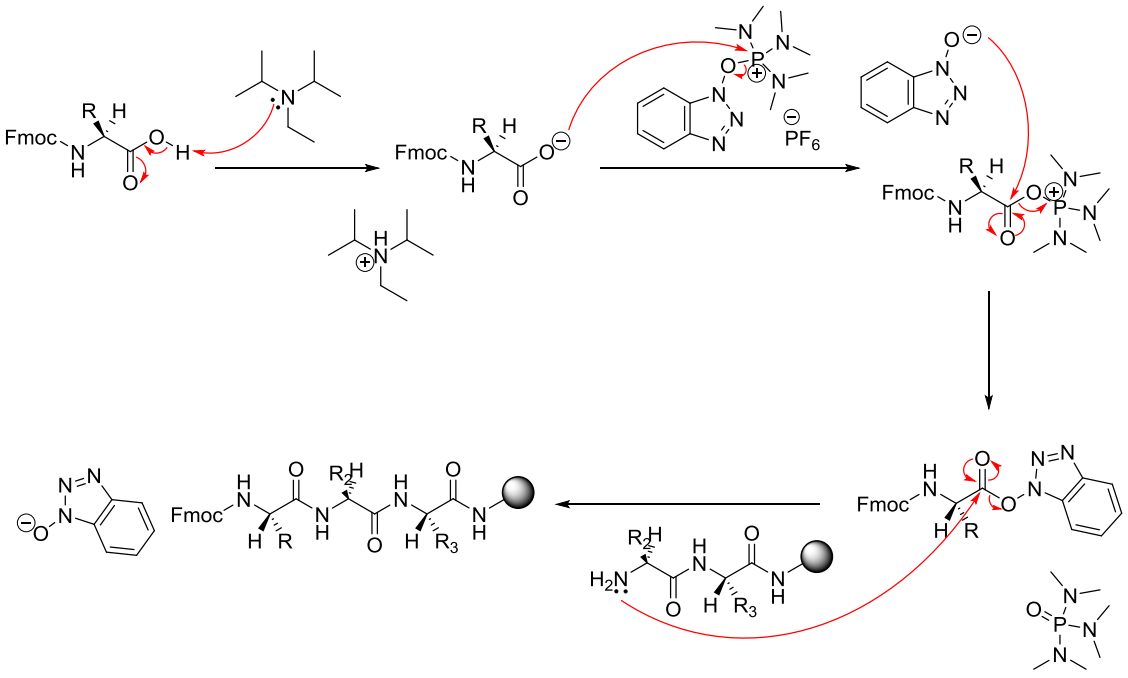
For carbodiimides, the lone pair of electrons on one of the nitrogen atoms abstracts the acidic proton on the Fmoc-protected amino acid (Scheme 11). The negative charge left on the Fmoc-protected amino acid then attacks the carbon of the carbodiimide to form the *O*-acylisourea, which in turn reacts with the deprotected amine of another amino acid to give an acid anhydride as a leaving group and the fully formed amide bond between the two amino acid moieties.



R = cyclohexyl (DCC)
= isopropyl (DIC)
R',R" = any amino acid side chain

Scheme 11: Amino acid coupling mechanism using carbodiimides.

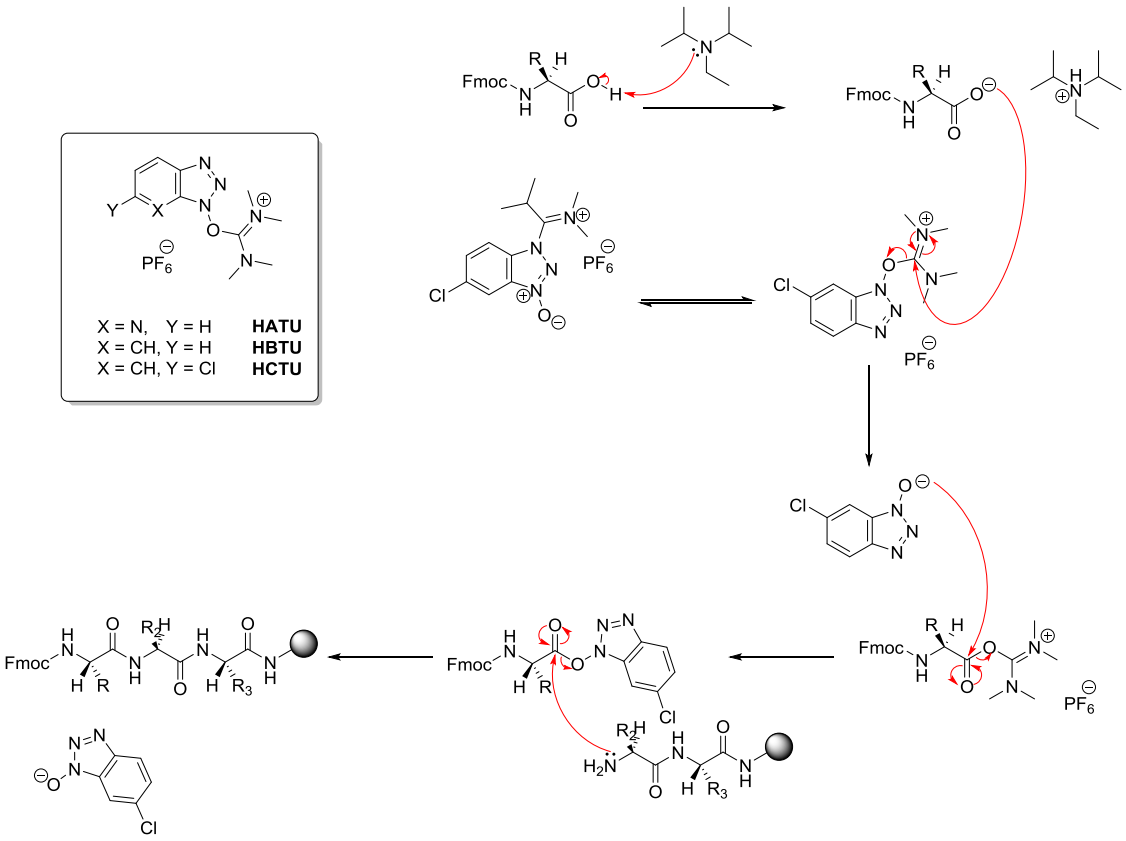
Using phosphonium salts as the coupling agent for amide bond formation, there is also the need to use an appropriate base, such as diisopropylethylamine (DIPEA). DIPEA abstracts the acidic proton of the Fmoc-protected amino acid and the resulting negative charge can attack the positive phosphonium ion (Scheme 12). This displaces the *O*-benzotriazole (OBt) which then attacks the carbonyl of the phosphonium, creating the phosphoramide side product. The lone pair of electrons on the free amine on the peptide chain can attack the carbonyl of the amino acid-benzotriazole complex to finally increase the peptide chain by one amino acid and regenerating the benzotriazole.



Scheme 12: Mechanism of amino acid coupling with the use of a phosphonium salt, BOP.

For the peptides described, we used *N,N,N′,N′*-tetramethyl-*O*-(6-chloro-1H-benzotriazol-1-yl)uronium hexafluorophosphate (HCTU) as the activating agent (Scheme 13). In addition to the activator, an activator base is also required, as with phosphonium salts. In this instance, DIPEA was used to abstract the proton from the Fmoc-protected amino acid to create the carboxylate anion which attacks the iminium ion and creating the free 6-chloro-*O*-benzotriazole

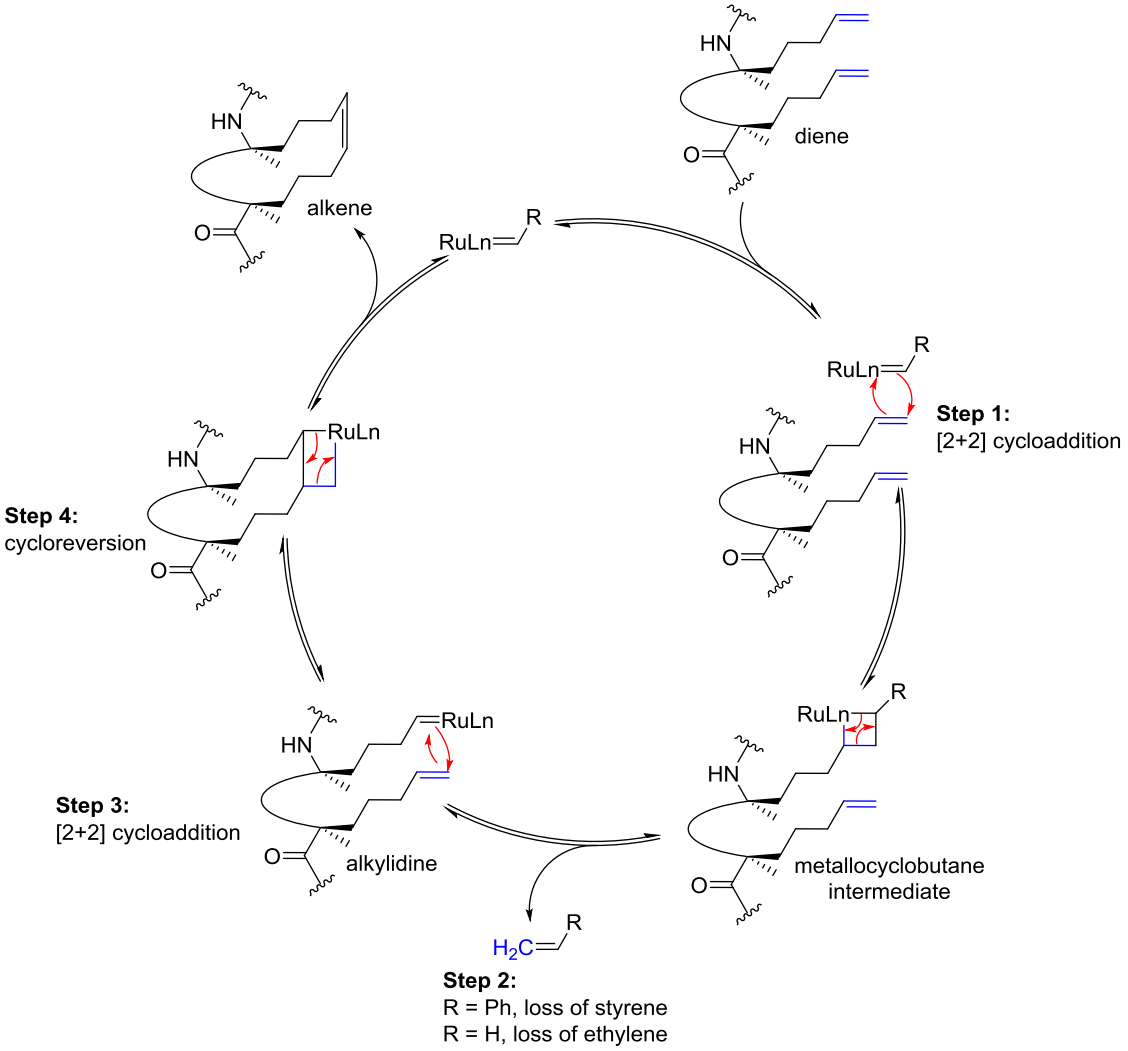
species, which acts in the same way as the benzotriazole from the phosphonium salt to give the extended peptide chain and regenerating the 6-chloro-*O*-benzotriazole species.



Scheme 13: Amino acid coupling mechanism using uronium salts.

Ring closing metathesis

For all the stapled peptides, ring closing metathesis (RCM) is the key cyclisation step of the synthesis. The reaction itself is catalytic, using 20 mol% of Grubbs' 1st generation catalyst, and therefore giving a catalytic cycle (Scheme 14).[197] The catalyst itself is not regenerated in its original form, so could be classed as a precatalyst. Styrene is removed as a by-product after the first catalytic cycle, followed by ethylene for all subsequent cycles. Ethylene and styrene can be easily removed by bubbling nitrogen through the reaction.



Scheme 14: Mechanism of ring closing olefin metathesis.[197]

On the peptides, RCM proceeded without difficulty. The reactions proceed to completion, according to LC-MS, after two sequential reactions of 2 h at room temperature. In fact, the reaction proceeds to near completion after just the first 2 h, but to make sure it reached completion it was necessary to add fresh catalyst to the resin.

Another surprise in this reaction is that only Grubbs' 1st generation catalyst is required for the reaction to reach completion. This is specially the case because the 1st generation catalyst is more difficult to handle and much more air sensitive in comparison to the 2nd generation catalyst, which is also more reactive.

The ring closing metathesis reactions on these peptides proceeded without difficulty. Our hypothesis for this is that the peptides are likely to start forming a helical configuration due to the α-methyl group present on the alkenyl amino acid. This extra α-methyl group gives a gem-dialkyl substitution which induces the Thorpe-Ingold effect bringing the two remaining substituents on the carbon closer together, in the case of peptides, the backbone amine and carbonyl.[148,198]

As a result the two alkene moieties on the peptide come close together ready for the RCM to take place. This leads to a high effective local concentration, and in turn allows the reaction to occur faster and to completion.

2.5. Design of Native SMRT-DAD helix 3 (463-476) peptide

The crystal structure of HDAC3 bound to the deacetylase activation domain (DAD) and a molecule of inositol tetraphosphate (D-myo-inositol-(1,4,5,6)-tetrakisphosphate) [Ins(1,4,5,6)P_4][42] was the most valuable tool in designing the peptides to target HDAC3. The SMRT-DAD contains a conserved SANT (Swi3, Ada2, N-Cor, and TFIIIB) domain (a histone tail binding module)[199] formed from helices 1, 2, and 3 and an extra helix at the *N*-terminus (helix 0) (Figure 46).

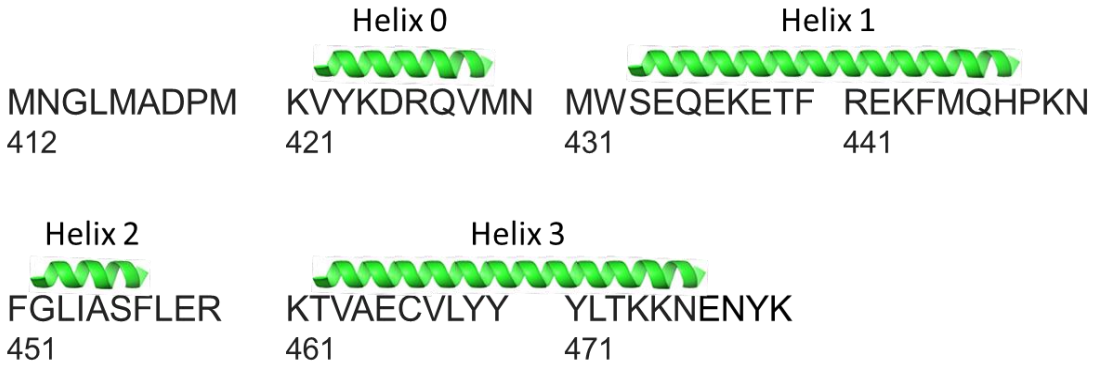


Figure 46: Structure prediction of SMRT-DAD showing the protein sequence and the regions that are helical.

Of the four helices associated with the SMRT-DAD protein, three make direct contact with the HDAC3 surface (helices 0, 2 and 3), while helix 1 does not. Additionally, of the three helices that do make contact with the HDAC3 surface, helix 3 makes the most extensive contacts with HDAC3 and in particular contain residues which play a key role in binding. These include Lys449, Tyr 470 and Tyr471, which when mutated, impairs the interaction of HDAC3:SMRT-DAD in addition to failing to activate HDAC3.[42] To this end, SMRT-DAD helix 3 was chosen as the peptide target.

The native sequence of SMRT-DAD helix 3 contains a cysteine residue at position 466. Cysteine can be thought of as a difficult residue to work with because it is a promiscuous residue. The thiol group of cysteine is very reactive and can easily undergo many transformations. In fact, cysteine can easily form intermolecular disulfide bonds as well as undergo redox reactions, alkylations

and can also non-covalently bind metals. Is it because of its thiol reactivity that cysteines are often incorporated into peptide sequences to aid with biconjugation methods.[200]

Cysteine 466 in the native peptide sequence of SMRT-DAD helix 3 was therefore mutated to a serine residue before the peptide was synthesised (Figure 47). Serine was chosen because of its similarity to cysteine, replacing the thiol functionality with a hydroxyl group, which are bioisosteres of each other. This means that biological systems cannot easily distinguish between the two functional groups.

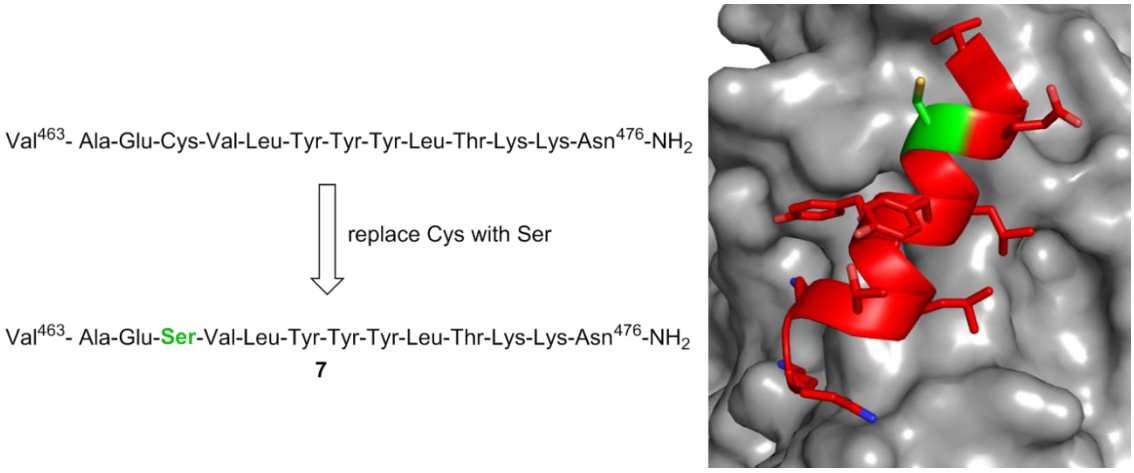


Figure 47: In the synthesised native sequence of SMRT-DAD (**7**), a C466S mutation was introduced to remove the promiscuous cysteine residue for serine which can be classed as a bioisotere of cysteine.

The *N*-terminus of a protein consists of a free amino group, which is charged at physiological pH. However, when taking a short peptide sequence out with the context of its parent protein, the *N*-terminus of the peptide is not charged. Consequently, acetyl capping of the peptide *N*-terminus is carried out in order to give a more physiological relevant model for the protein. In the synthesis of the SMRT-DAD peptides acetyl capping was not introduced as this was found to be detrimental to the solubility of the native peptide in aqueous conditions.

In the crystal structure of HDAC3, helix 3 of SMRT-DAD is located directly at the *C*-terminus.[42] As a result, in the synthesis of the isolated peptide, a free acid should be at the *C*-terminus. However, the SMRT-DAD protein is part of a much larger protein, the silencing mediator of retinoic acid and thyroid hormone

receptor (SMRT) and thus, the *C*-terminus of SMRT-DAD is not the real *C*-terminus of the much longer SMRT protein. From this information, it can be deduced that the *C*-terminus of SMRT-DAD helix 3 should not be charged.

By introducing an amide group, the negative charge of the free acid group, normally at the *C*-terminus, becomes neutral. To achieve a *C*-terminal amide, a resin with amide functionality, namely Rink Amide resin, was used in the synthesis. Upon cleavage of the peptide from the resin, the peptide is left with a *C*-terminal amide.

The 14-residue SMRT-DAD helix 3 native peptide sequence, with the serine mutation at position 466, was synthesised using a state-of-the-art automated, microwave assisted peptide synthesiser. On completion of the synthesis, the resin was washed and the peptide was cleaved from the resin using the cleavage cocktail of TFA/TES/H_2O (95:2.5:2.5) followed by precipitation of the peptide in cold diethyl ether. The peptide was then purified by reverse-phase HPLC and isolated in 39% yield and over 99% purity, before biological testing was carried out.

2.6. Staple scan

In isolation, a short peptide sequence is less likely to adopt its bioactive secondary structure. To overcome this entropic penalty of folding, conformational constraints can be employed. In this work, an all-hydrocarbon staple system has been utilised as the conformational constraint if choice for the SMRT-DAD helix 3 peptide.

The 14-residue SMRT-DAD peptide consists of 3 α-helical turns. This could be conformationally constrained by an *i,i+7* staple. However, in a lot of peptide-protein interactions, some flexibility of the peptide is required for binding to its protein target. Thus, an *i,i+4* staple was instead chosen to constrain the SMRT-DAD peptide across one turn of the helix.

For an effective all-hydrocarbon staple constraint the design of the stapled peptides is crucial. As such, there are a number of suitable positions in which to

install the staple constraint in the SMRT-DAD peptide. To determine the best position for the staple constraint, a staple scan was carried out.

Staple scanning is a useful tool particularly when there is limited structural information available. This strategy has been successfully used to identify a selective inhibitor of MCL-1, a human cancer resistance factor.[201] Additionally, the staple scanning technique echoes alanine scanning in that it can help determine residues on the peptide required for binding.[202]

In our system, the crystal structure of HDAC3:SMRT-DAD is available and details of the binding motif can be studied from the crystal structure. However, the best placement of the staple constraint within the peptide is still unknown. In the design of the SMRT-DAD stapled peptides, the staple constraint was required on the back face of the helix, so as not to interfere with any residues necessary for binding to the HDAC3 surface. As such, working from the crystal structure and the helical wheel projection, three peptides were chosen for the staple scan of SMRT-DAD helix 3 (Figure 48).

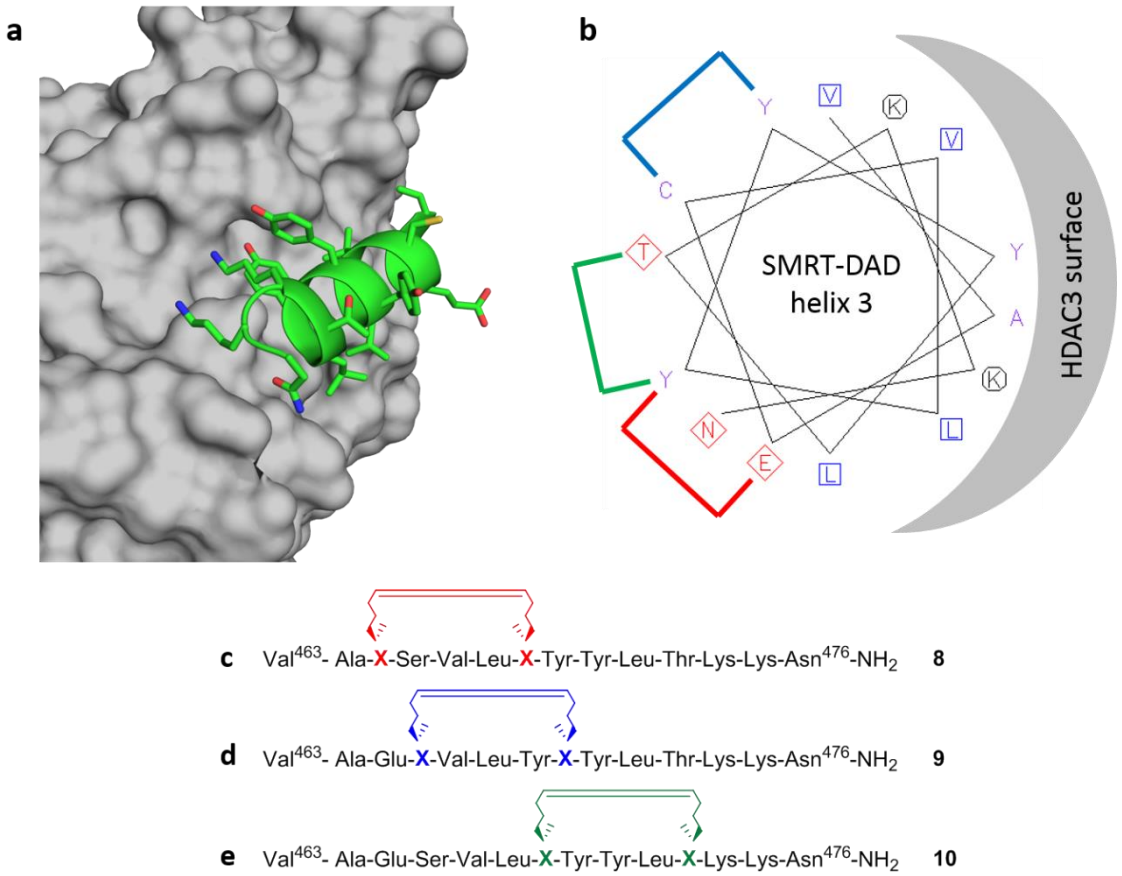


Figure 48: a) Crystal structure of SMRT-DAD helix 3 bound to the surface of HDAC3 and **b)** a helical wheel projection of the staple scan of SMRT-DAD helix 3, where the positions of the staples in compounds **9** (red), **9** (blue) and **10**(green) are shown. Both of these show the binding motif of the SMRT-DAD helix 3 peptide to HDAC3 and the back face of the helix free from any interactions. **c-e)** Sequences of the three peptides synthesised from the staple scan.

Glutamic acid 465 and Tyrosine 469 → S_5 (8)

The first stapled peptide that was designed and synthesised involved mutating an *i* and *i+4* residue on the native peptide, namely Glu465 and Tyr469 to the unnatural α,α-disubtituted S_5 amino acid (Figure 49). These mutated residues are also very close to the N-terminus of the peptide. Studies by Aurora and coworkers suggested that incorporating a constraint close to the *N*-terminus of a peptide aids in the nucleation and formation of the helix.[133]

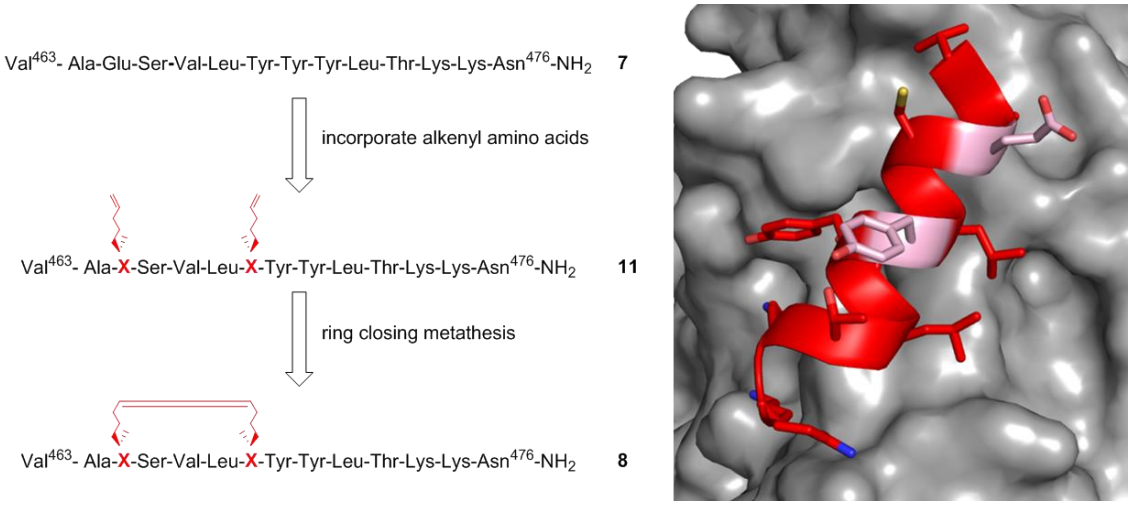


Figure 49: The design of the first stapled peptide (peptide **8**) involves a Cys466Ser mutation and replacement of Glu465 and Tyr469 (pink residues) to the S_5 amino acid. The constraint is formed through a ring closing metathesis reaction.

As with the native peptide sequence, the cysteine residue at position 466 was also mutated to a serine. This mutation allows for the direct comparison between the synthesised native and stapled peptides.

The synthesis of the stapled peptide was carried out using microwave-assisted SPPS. However, for the couplings of the unnatural α,α-disubstituted amino acid required for stapling, manual peptide synthesis was utilised. This was done in order to reduce the number of equivalents of amino acid used for each coupling. The reduction in equivalents can still give highly efficient couplings; and also reduced the cost of each of the couplings. For expensive amino acid residues, such as the unnatural α,α-disubstituted amino acid, this can be an advantage.

The synthesis of opened staple peptide sequence (**11**) was carried out leaving the Fmoc protection on the valine residue at position 463 of the *N*-terminus. This is required because of the sensitivity of the Grubbs 1st generation catalyst to primary amines. Only following the RCM reaction, can the *N*-terminus be deprotected to reveal the free amino group of the *N*-terminus, before the peptide was globally deprotected and cleaved from the resin to give **8**, in a reasonable yield of 14% and 39% crude purity.

Cysteine 466 and Tyrosine 470 → S_5 (9)

The second stapled peptide was designed to incorporate the constraint closer towards the centre of the peptide sequence (Figure 50). Cys466 and Tyr470 were thus mutated to the S_5 α,α-disubstituted unnatural amino acid for stapling. Both of these mutated residues are on the opposite face of the helix to the PPI with HDAC3. Additionally, with the mutation of Cys466, there is no worry of disulfide bond formation, as is the case with the native sequence and other stapled peptides in this series.

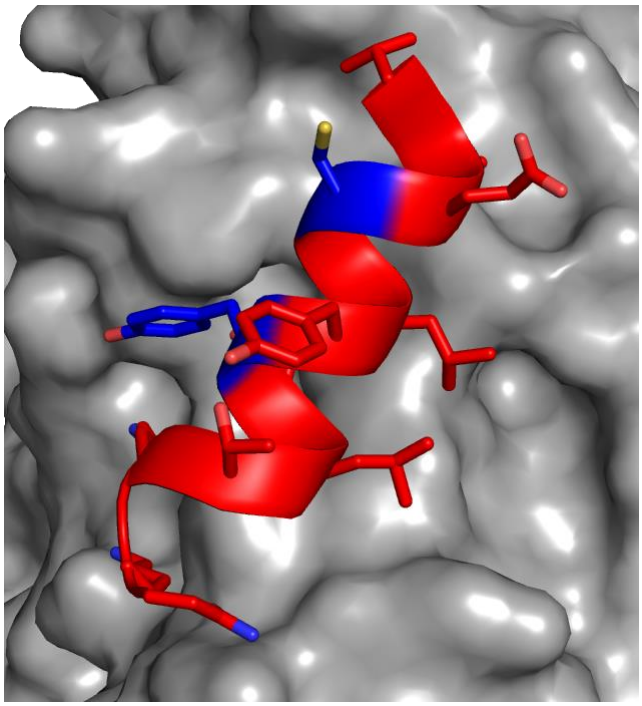


Figure 50: The design of peptide **9**involved the mutation of Cys466 and Tyr470 (blue residues) to the α,α-disubstituted alkenyl amino acid.

Tyrosine 469 and Threonine 473 → S_5 (10)

The final *i,i+4* stapled peptide that was possible in the staple scan was moving the constraint closer to the *C*-terminus. Here, Tyr469 and Thr473 were mutated to the S_5 amino acid (Figure 51). In addition, as with the other peptides in the series, Cys466 was also mutated to a serine residue.

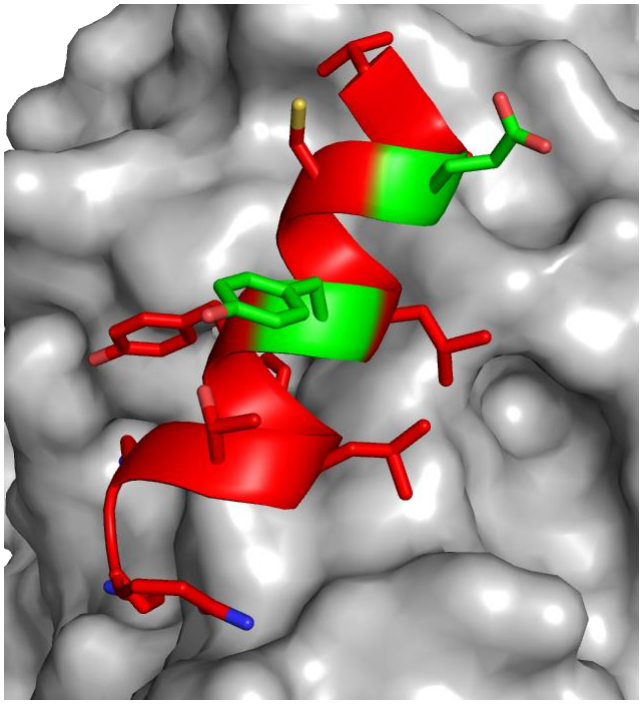


Figure 51: The design of peptide **10** involved the mutation of Tyr469 and Thr473 (green residues) to the α,α-disubstituted alkenyl amino acid. Cys466 is also mutated to Ser.

For this particular peptide, it was impossible to move the staple any closer to the *C*-terminus because the residues that were closer, namely Lys474, Lys475 and Asn476, are all involved in some binding to either the HDAC3 surface or the Ins(1,4,5,6)P_4 molecule. The only other way of incorporating a stapled constraint much closer to the *C*-terminus would be to add an extra residue. However, in this peptide sequence, the crystal structure showed that the *C*-terminus did not present itself greatly to be helical, but more of a random coil. Hence, it was Tyr469 and Thr473 that were chosen to be mutated to the S_5 amino acid, which in turn could form the staple after RCM.

2.6.1. Conformational analysis

Circular dichroism (CD) has been found to be one of the most useful tools to study the secondary structures of peptides and proteins.[203] CD itself can be thought of as differential absorption of two circularly polarised components of plane polarised light: one of which rotates clockwise and the other counter-

clockwise. Both of these circularly polarised components are also of equal magnitude.[204]

Using CD, structural information of a peptide or protein can be accurately determined from a number of spectral regions. For peptides and proteins, the peptide bond is the most telling chromophore, which absorbs in the region below 240 nm.

In the peptide bond region below 240 nm, there are two notable transitions. The first is a weak yet broad transition ($n \rightarrow \pi^*$) which is centred at around 220 nm. The second transition, and more intense of the two, is $\pi \rightarrow \pi^*$ which is around 190 nm.

The different secondary structures can be determined in the spectral range below 240 nm, as each displays a very specific spectrum (Figure 52). A random coil structure usually has a near horizontal CD signal above 210 nm and a minimum at between 190 nm to 200 nm, α-helices display minima at 222 nm and 208 nm leading to a sharp maximum at about 190 nm, and β-sheets have a minimum at about 215 nm and a maximum between 190 nm to 200 nm.

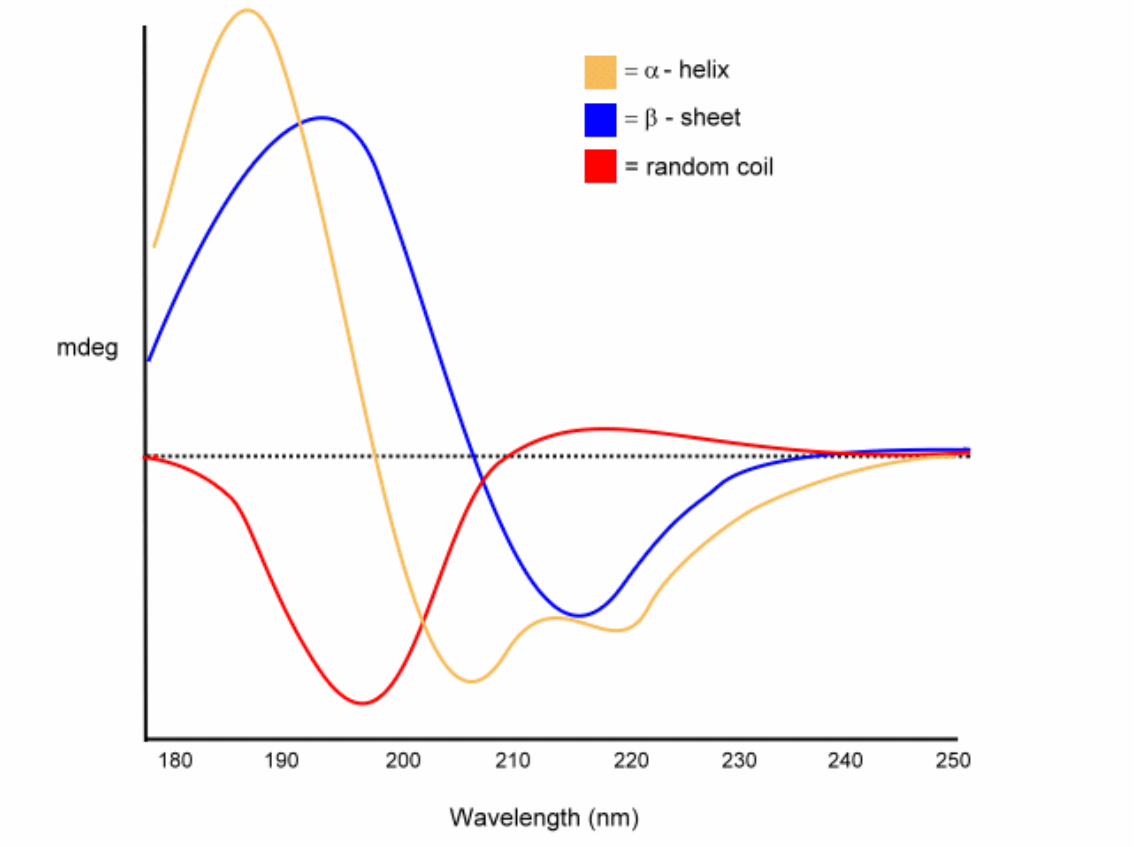


Figure 52: Typical curves for peptide secondary structures: random coil (red line), α-helix (yellow line), β-sheet (blue line).[204]

This method of analysis is widely used for peptides because it requires only a small peptide concentration and allows for the study of the peptide structure at different temperatures.[205]

Mean residue molar ellipticity, [θ], is the usual way in which CD results for proteins and peptides are expressed. It is given in deg.cm^2.dmol^{-1} and calculated as in Equation 1, where θ is the ellipticity in degrees, M_r is the molecular mass, c is the concentration (mg/mL), l is the optical path length (cm) and n is the number of amino acid residues in the peptide.

$$[\theta] = \frac{\boldsymbol{\theta} \, x \mathbf{100} \, x \boldsymbol{M_r}}{\boldsymbol{c} \, x \boldsymbol{l} \, x \boldsymbol{n}} \qquad \textbf{Equation 1}$$

It is also possible to estimate the percentage α-helicity of the peptide from taking the values of the CD signal at 208 nm and 222 nm. This can be achieved using Equation 2,[206] where $[\theta]_{222}$ is the molar ellipticity of the sample at 222 nm, θ_{222}^{∞} is the theoretical molar ellipticity at 222 nm of a completely α-helical peptide (-39,500 deg.cm^2.dmol^{-1}), *i* is the number of helices in the sample, *k* is the wavelength specific constant (2.57 at 222 nm) and *n* is the number of amino acid residues in the peptide.

$$\% \, \alpha - helicity = \frac{[\theta]_{222}}{\theta_{222}^{\infty} \left[1 - \left(\frac{ik}{n}\right)\right]} \, x \, 100 \qquad \textbf{Equation 2}$$

However, this estimate is just that, an estimate. There are many algorithms that are also available both as freeware and commercial software that attempt to estimate percentage α-helicity.

Initially, all the peptides that were synthesised were dissolved in deionised water for the CD measurements. This was done because the natural environment in which the peptides would be found is in the cell and therefore in aqueous conditions.

In addition, each of the peptides were dissolved in a solution of 10% trifluoroethanol (TFE) in deionised water. TFE is known to induce helicity in peptides and proteins. Hence, for each of the peptides, it was expected that due to the stabilising effect of TFE on α-helices, the peptides would show themselves to be even more helical in structure than exactly the same peptides in water only.

As well as this, we could establish the maximal helical character possible for each of the peptides that were synthesised and compare that with the helicity of the peptides purely in deionised water. As a result we can see whether the peptides that were synthesised already had their maximal helical properties.

In water, the native SMRT-DAD helix 3 peptide, **7**, gives a spectrum that is very typical for a peptide that exhibits a random coil structure (Figure 53). In fact, using the equations above, the percentage α-helicity of the native peptide was found to be just 16% in water. This verifies that the native peptide sequence is to be found more as a random coil rather than an α-helix.

The CD spectra for the staple scan show all three of the stapled peptides (peptides **8**, **9** and **10**) exhibit an α-helical conformation (Figure 53). From the traces it can be clearly seen that **8** is the most helical of these peptides (30%, Figure 53**,** red line), closely followed by **10** (25%, Figure 53**,** green line) and the incorporation of the staple in the middle of the peptide as in **9** has not given a very high percentage helicity (19%, Figure 53**,** blue line). It is for these reasons that we chose to incorporate the staple at the position used in **8** for all subsequent peptide analogues.

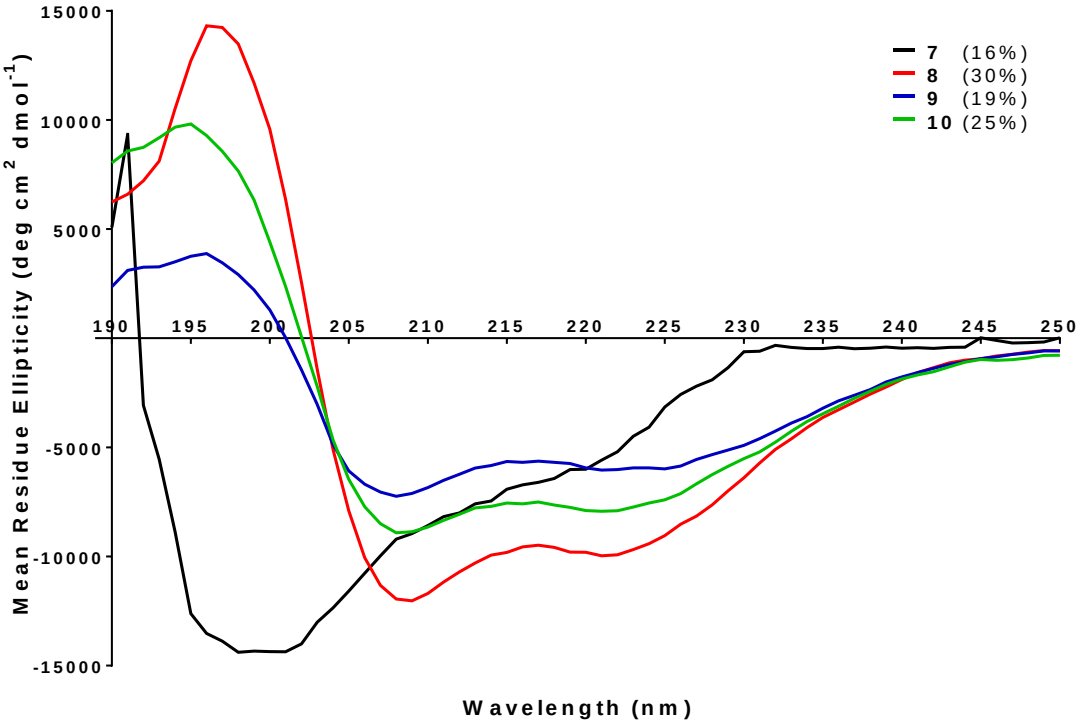


Figure 53: Circular dichroism spectra of the native SMRT-DAD helix 3 peptide (**7**) and the stapled peptides: **8** (red), **9** (blue) and **10** (green), from the staple scan.

On addition of 10% TFE to the samples, each of them showed an increase in α-helical character (Figure 54). In 10% TFE/H_2O, even the native peptide **7** showed helicity (24%, Figure 54**,** black line). However, unlike the samples in the absence of TFE, peptide **9** exhibited the largest helical character (53%, Figure 54**,** blue line), closely followed by **8** (50%, Figure 54**,** red line), which gave the highest helicity in the absence of TFE, and finally, **10** showed the least helicity of the stapled peptides (42%, Figure 54**,** green line).

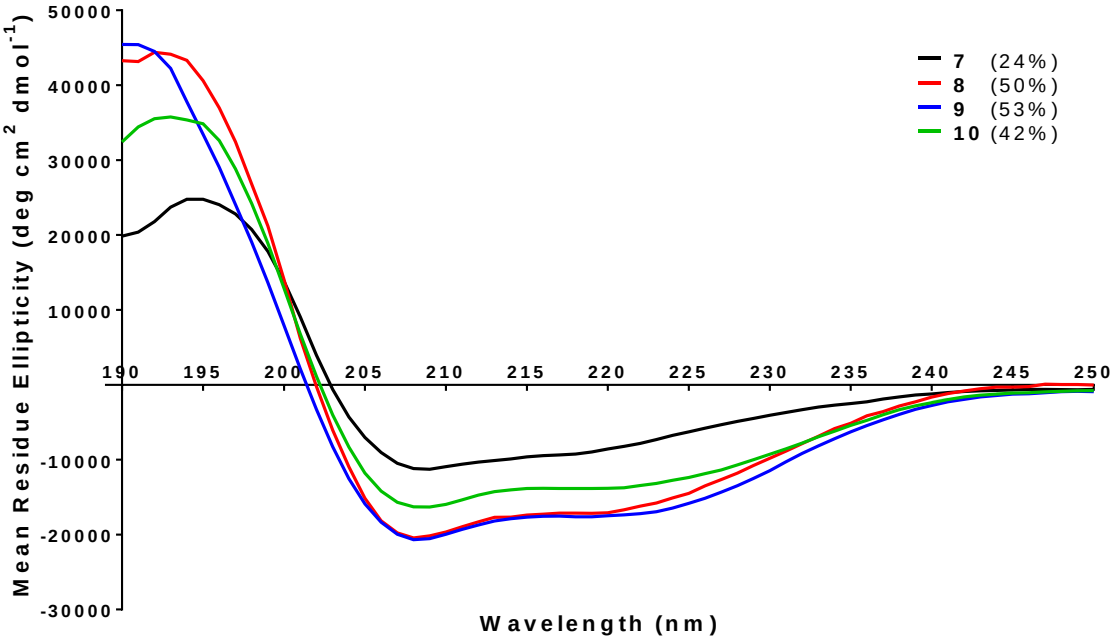
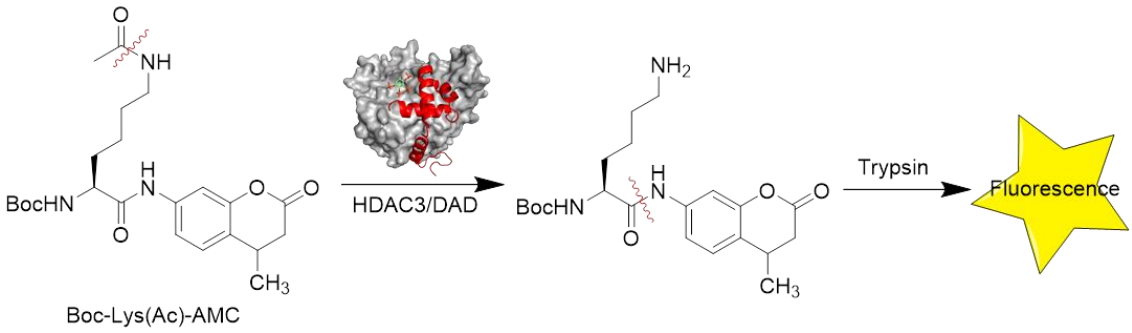


Figure 54: CD spectra of the native peptide **7** and the stapled peptides: **8** (red), **9** (blue) and **10** (green), from the staple scan in 10% TFE/H_2O.

Comparing the two sets of data, in the absence and presence of 10% TFE, it can be seen that there is a fairly big difference between the percentage helicities that can be achieved in the peptides. It shows that the stapled peptides that have been synthesised for the scan are not able to achieve their full potential in terms of α-helical character. This would indicate that even with the staple scan, we have not been able to reach maximum helicity. This is something that could be optimised in the future by using an *i,i+7* staple as a longer constraint. A truncated peptide can also be used so that the *i,i+4* staple only needs to constrain a shorter peptide.

2.6.2. Fluorescence based activity assay

To analyse the biological activities of these peptides, HDAC activity assays were carried out. For these assays, the HDAC3 protein complex was incubated with the native or stapled peptides. Aminomethylcoumarin (AMC) fluorescently labelled, acetylated Boc-protected lysine (Boc-Lys(Ac)-AMC) was then used as a substrate for the HDAC3 enzyme complex. With the active HDAC3, the acetyl group is cleaved from the Boc-Lys(Ac)-AMC. Trichostatin A (TSA), an HDAC inhibitor was then added to quench the activity of the HDAC enzyme, before trypsin was added. Trypsin is used to cleave the amide bond between the deacetylated lysine residue and the AMC fluorophore (Scheme 15), but will only do so if the acetyl group on lysine has been removed. As a result, a quantifiable fluorescence readout can be obtained based on the amount of AMC released. This gives a measurement of the deacetylase activity of the HDAC3 complex.



Scheme 15: The HDAC activity assay involves Boc-Lys(Ac)-AMC being deacetylated by the HDAC3 complex, then trypsin cleaves the amide bond liberating the AMC fluorophore.

The need for comparison between different assays and different batches of proteins is important. To be able to do this another assay was designed which first takes into account the activity of HDAC3 in its native form. Next it takes into account the HDAC activity once the Ins(1,4,5,6)P_4 that was co-purified with the complex, has been removed by high ionic strength buffer (“salt stripping”).

Finally, the HDAC3 activity once 100 µM of Ins(1,4,5,6)P_4 has been added back to the protein is also assessed. 100 µM of Ins(1,4,5,6)P_4 was chosen because it is this concentration of Ins(1,4,5,6)P_4 which gives the maximal HDAC activity. Consequently, during each assay run the HDAC activity without added peptides was also carried out as a direct comparison and to normalise the data.

As expected, after purification the HDAC3:SMRT-DAD complex gives a good signal for HDAC activity. Once salt stripped, and therefore the Ins(1,4,5,6)P_4 molecule removed, the HDAC activity is reduced to about a third of the natural activity. Upon addition of 100 µM Ins(1,4,5,6)P_4, the HDAC activity then increases to about a third higher than the natural activity (Figure 55). It has also been shown that with different histone tail substrates for the HDAC3:SMRT-DAD complex, there is variation in the Km, kcat and enzymatic activity (kcat/Km).[44] More particularly, for the preferred substrate for HDAC3:SMRT-DAD (H3 23-29(K27Ac)) a 3-fold greater kcat was observed in the absence of inositol phosphate.[44] However, inositol phosphates do not show a real effect on Km but has been shown to increase kcat. As a result, addition of Ins(1,4,5,6)P_4 allosterically activates the HDAC3:SMART-DAD comple by increasing the catalytic efficiency. Overall, the binding of Ins(1,4,5,6)P_4 can therefore be decoupled from the binding of the substrate.

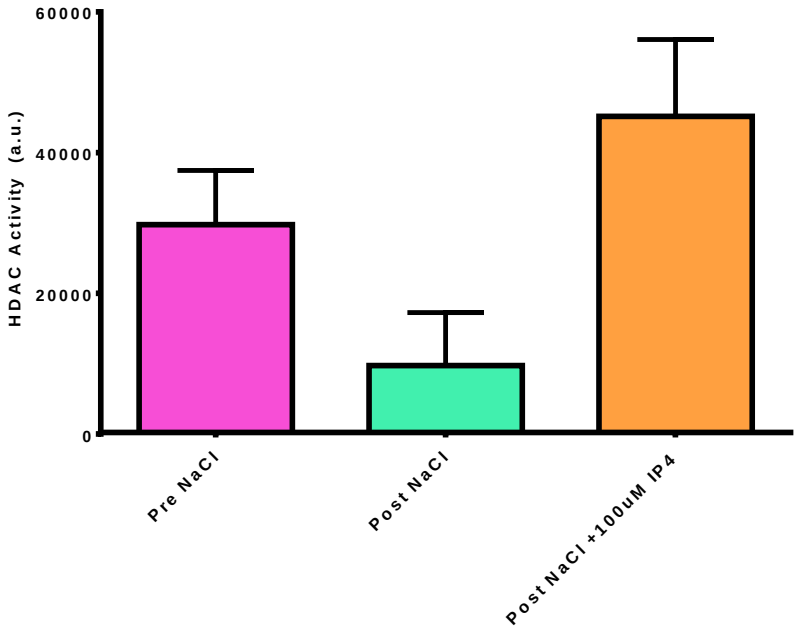


Figure 55: HDAC activity comparison of the protein before and after removal of Ins(1,4,5,6)P_4 and after removal, but in the presence of 100 µM Ins(1,4,5,6)P_4. Error bars represent ± s.e.m. (n = 3).

In order to work on salt-stripped HDAC3/SMRT-DAD complex, an extended SMRT-DAD construct (Xt-DAD) was utilised. This was required because when the Ins(1,4,5,6,)P_4 molecule is removed from the HDAC3:SMRT-DAD:IP4 complex, the complex fall apart. In other words, the Ins(1,4,5,6)P_4 is needed to help stabilise the complex.The extended SMRT-DAD protein allows the HDAC3 complex to remain intact when the Ins(1,4,5,6)P_4 is removed. This is a result of the extended part of the SMRT-DAD protein wrapping itself around the HDAC3 enzyme and acts to have an anchoring effect holding the HDAC3/Xt-DAD complex together. This echoes the findings with HDAC1 and its corepressor MTA1.[45]

Working with the salt stripped HDAC3/Xt-DAD protein, the assays were carried out once again with the peptides of interest. The native peptide **7** was compared with peptides **8**, **9** and **10** from the staple scan at a protein concentration of 25 nM. The data showed that the native SMRT-DAD peptide **7** showed activation of HDAC3, as did peptides **8**, **9** and **10** (Figure 56).

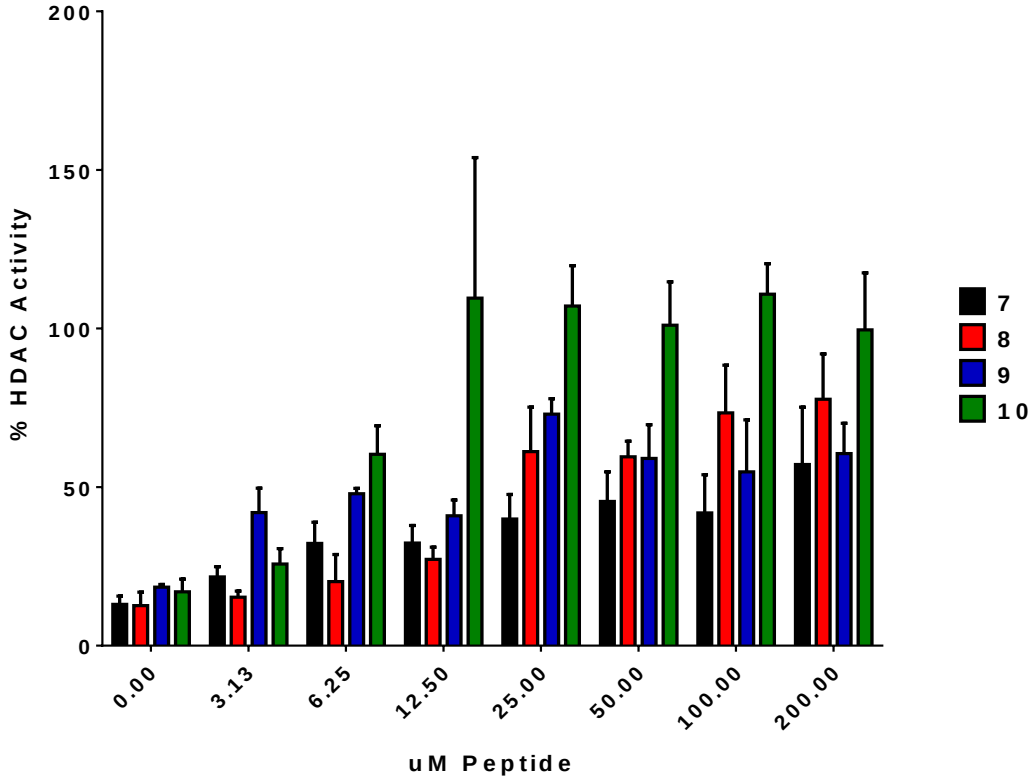


Figure 56: HDAC activity of the native SMRT-DAD peptide **7** versus the peptides in the staple scan: **8**, **9** and **10**. Error bars represent ± s.e.m. (n = 3).

The higher activity of HDAC3 on addition of **10** could be directly due to the positioning of the hydrocarbon staple constraint. In **10**, the staple is positioned close to the *C*-terminal. In the crystal structure of HDAC3 bound to SMRT-DAD and Ins(1,4,5,6)P_4, two of the residues situated near the *C*-terminus (Tyr470 and Tyr471) point toward the Ins(1,4,5,6)P_4 molecule forming hydrogen bonding motifs. These two tyrosine residues are directly in the middle of the staple constraint in **10**, on the opposite face of the helix to where the staple constraint is placed. This position, allows Tyr470 and Tyr471 to be held robustly in the correct configuration for the interaction with the Ins(1,4,5,6)P_4 molecule. Therefore, this could explain the greater percentage HDAC activity seen with **10** in comparison with **8** and **9**.

As a comparison, the four peptides were also tested in the presence of 10 μM Ins(1,4,5,6)P_4 and normalised using the maximal HDAC activity seen of the protein with 100 μM Ins(1,4,5,6)P_4. It can be seen that with 10 μM Ins(1,4,5,6)P_4, all four peptides activate HDAC3 much more than the maximal activity seen with just 100 μM Ins(1,4,5,6)P_4 (Figure 57).

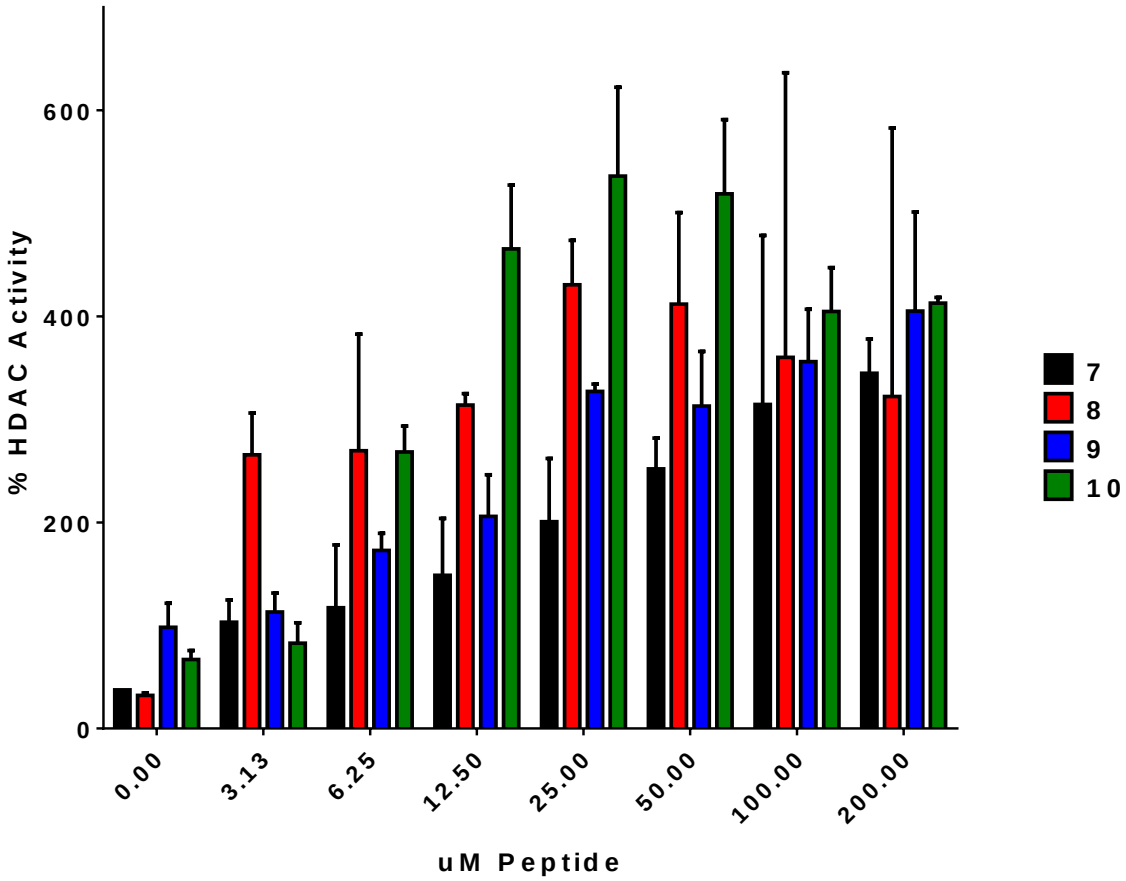


Figure 57: HDAC activity of the native SMRT-DAD peptide **7** versus the peptides in the staple scan: **8**, **9** and **10** in the presence of 10 μM Ins(1,4,5,6)P_4. Error bars represent ± s.e.m. (n = 3).

On addition of 10 μM Ins(1,4,5,6)P_4 to the HDAC activity assay, it can be seen that at 200 μM peptide concentration, the native peptide **7** induces about 300% HDAC3 activity (Figure 57). However, the stapled peptides surpass this activity. In fact, peptide **8** and **10** induce their maximal percentage HDAC3 activity of 430% and 530% respectively at 25 μM peptide concentration. Peptide **9** induces its maximal percentage HDAC3 activity of about 400% at 200 μM peptide concentration.

Directly comparing the two assays, upon addition of Ins,(1,4,5,6)P_4, the increase in percentage HDAC activity overshadows any small increase in activity without Ins(1,4,5,6)P_4. What this data shows is that Ins(1,4,5,6)P_4 plays

a key role in the activation of HDAC3. Ins(1,4,5,6)P_4 seems to act as a molecular glue between the HDAC3 surface and the SMRT-DAD protein.

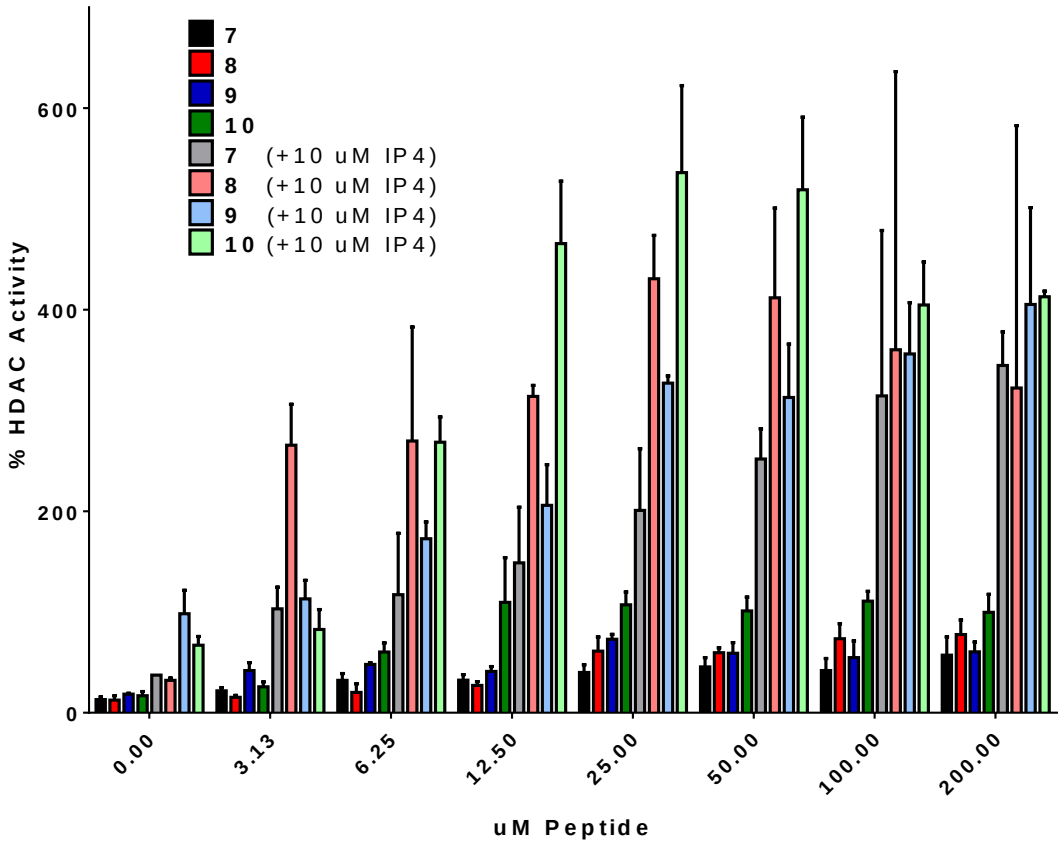


Figure 58: Difference in HDAC activities of the native SMRT-DAD peptide **7** versus the peptides in the staple scan: **8**, **9** and **10**upon addition of 10 µM Ins(1,4,5,6)P4. Error bars represent ± s.e.m. (n = 3).

2.7. Rational design of phosphotyrosine stapled peptide analogues of SMRT-DAD helix 3

From the crystal structure of HDAC3:SMRT-DAD:Ins(1,4,5,6)P_4 complex, it can be clearly seen that Tyr470 and Tyr471 both point towards the 5-position phosphate on Ins(1,4,5,6)P_4 (Figure 59). In fact, from the hydroxyl on each of the tyrosine residues, to the phosphorus at the 5-position of Ins(1,4,5,6)P_4, the distance is exactly the same at 3.7 Å.

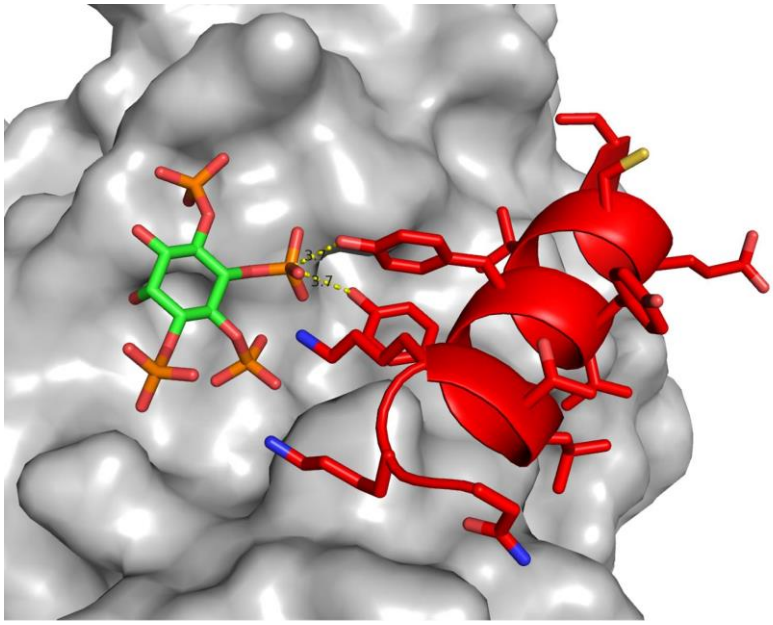


Figure 59: The crystal structure of HDAC3-DAD-Ins(1,4,5,6)P_4 shows both Tyr470 and Tyr471 point directly toward the phosphate group at the 5-position on Ins(1,4,5,6)P_4.

Using this information, it was thought a stapled peptide that could be used to mimic both the SMRT-DAD and Ins(1,4,5,6)P_4. A possible way of creating a dual mimic is to covalently attach Ins(1,4,5,6)P_4 onto the stapled peptide. However, another route could be to mimic only part of Ins(1,4,5,6)P_4, such as one or more of the phosphate groups.

This can be achieved through incorporation of a phosphate group on the peptide, which is much more straightforward than covalently attaching Ins(1,4,5,6)P_4 to the stapled peptide. From the crystal structure, the best position for a phosphate group would be at the tyrosine which points directly at Ins(1,4,5,6,)P_4. As such, the second generation stapled peptides were synthesised containing a phosphotyrosine residue at positions 470 and 471 (Figure 60).

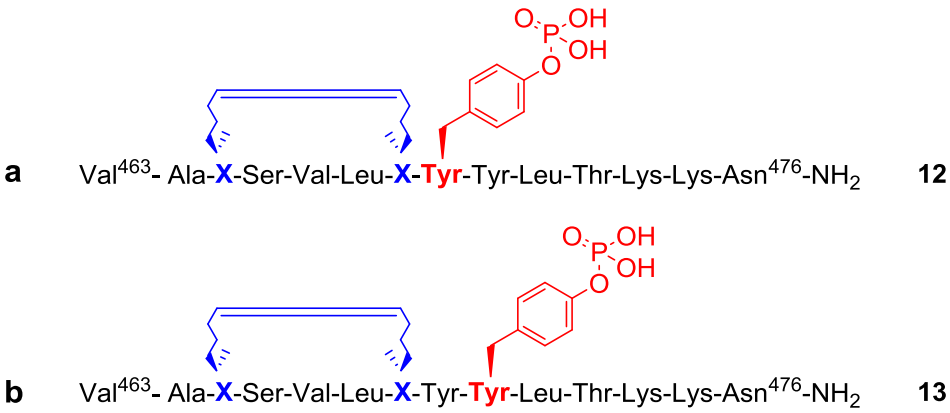


Figure 60: The two phosphotyrosine containing stapled peptides based on **8** with a) the phosphotyrosine at position 470 (**12**) and b) with the phosphotyrosine residue at position 471 (**13**). The C466S mutation is also present in both analogues.

Tyr470pTyr phosphotyrosine stapled peptide analogue (12)

First phosphotyrosine stapled peptide analogue synthesised involved the mutation of Tyr470 to a phosphotyrosine residue (Figure 60a). The mutation of Tyr470 to phosphotyrosine was achieved through the use of an Fmoc-protected phosphotyrosine amino acid, namely Fmoc-Tyr(PO(OBzl)OH)-OH, which is the most commonly used phosphotyrosine derivative.[207] The synthesis of the sequence was achieved in the usual fashion without much difficulty until it came to removing the benzyl protecting group from the phosphotyrosine residue itself. Previous literature indicated that the removal of the benzyl protecting group followed the normal procedure for cleavage from the resin.[207] However, it was found that this did not occur in this manner even after a 3h cleavage in the usual cleavage cocktail of TFA/TES/H_2O (95:2.5:2.5).

Due to this, it was thought that prolonged exposure to the cleavage cocktail of the benzyl-protected phosphotyrosine containing peptide might result in the desired deprotected product. However, all this procedure returned was a very messy LC-MS with the main peak being shown to be the stapled peptide still containing the benzyl-protected phosphotyrosine residue.

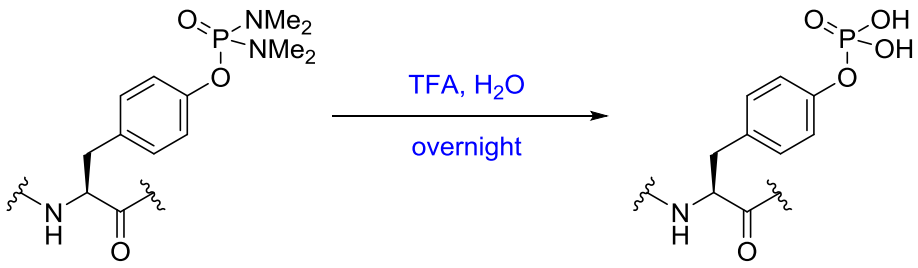
It was then thought that after the 3h cleavage, it might then be possible to deprotect the phosphotyrosine residue by hydrolysing the benzyl group with an excess amount of water. To achieve this, the cleaved peptide was subjected to water overnight and LC-MS carried out on it once again. However, this proved

to be futile and the only mass that was returned was that of the benzyl protected phosphotyrosine peptide.

Another problem with the use of Fmoc-Tyr(PO(OBzl)OH)-OH is the fact that there is still an acidic hydroxyl group on the phosphotyrosine that can cause problems during subsequent coupling reactions.[207]

Consequently, instead of pursuing this particular analogue of a phosphotyrosine mutated stapled peptide, a different Fmoc-protected phosphotyrosine amino acid was chosen.

The new Fmoc-protected phosphotyrosine amino acid that was chosen was Fmoc-Tyr(PO(NMe_2)$_2$)-OH. This was chosen because the side chain phosphate on the tyrosine is fully protected and it is compatible with the coupling methods employed in the synthesis of the peptide. In addition, the phosphodiamidate is easily converted to the phosphate by acid catalysed hydrolysis using the normal TFA cleavage cocktail followed by addition of 10% water (v/v TFA) to the cleavage cocktail after 3h, which was then left overnight to allow for the complete hydrolysis of the phosphate side chain (Scheme 16).[207,208]



Scheme 16: Deprotection of the Tyr(PO(NMe_2)$_2$) residue within a peptide proceeds under normal TFA cleavage conditions followed by addition of 10% water and incubation at room temperature overnight.

Tyr471pTyr phosphotyrosine stapled peptide analogue (13)

As can be evidently seen in the crystal structure of HDAC3, Tyr471 also points directly at the 5-position phosphate on Ins(1,4,5,6)P_4 (Figure 59). Additionally, Tyr471 of SMRT-DAD helix 3 is shown to point into a pocket very close to where Ins(1,4,5,6)P_4 is bound, hence the desire to synthesise this particular phosphotyrosine stapled peptide analogue. Thus, the second phosphotyrosine containing peptide synthesised involved the mutation of Tyr471 to a

phosphotyrosine residue (Figure 60b). This stapled peptide was also synthesised using the same Fmoc-Tyr(PO(NMe$_2$)$_2$)-OH protected amino acid.

As with the Tyr470pTyr stapled peptide analogue, the Tyr471pTyr analogue was hypothesised to bind to the HDAC3 surface without the need for Ins(1,4,5,6)P$_4$ to act as a molecular glue.

Conformational analysis of phosphotyrosine stapled peptides

The phosphotyrosine based peptides (**12** and **13**) also showed very promising percentage helicities: 32% for **12**(Figure 61, orange line) and 29% for **13**(Figure 61, pink line). In fact, they were comparable to the percentage helicity of **8** in H$_2$O. This would suggest that the addition of a phosphotyrosine residue onto the peptide allows it to retain the helicity provided by the staple.

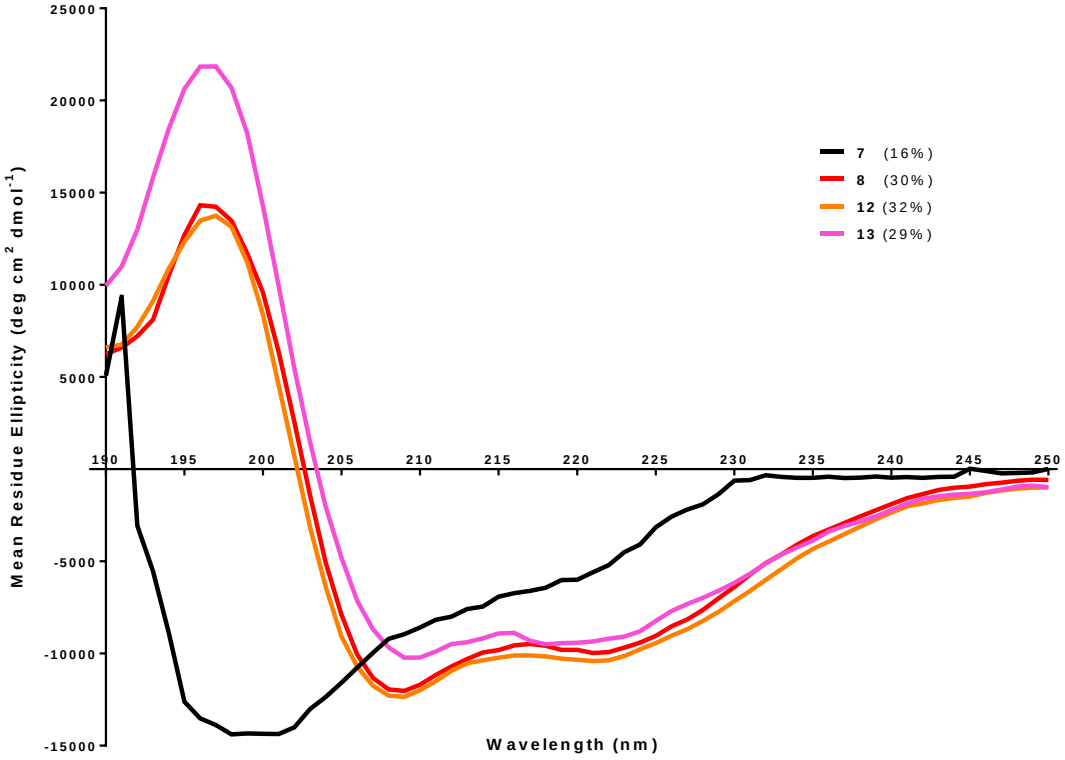


Figure 61: Circular dichroism spectra of the native SMRT-DAD helix 3 peptide **7** (black) and stapled peptide **8** (red) compared to the two phosphotyrosine containing stapled peptides **12** (orange) and **13** (pink).

Fluorescence based activity assay with stapled peptides 12 and 13

Both of the phosphotyrosine stapled peptides activated HDAC3 to a much higher degree than either of the native or original stapled peptides (Figure 62). In fact, the activity for both the phosphotyrosine stapled peptides is more comparable to the percentage HDAC activity seen on addition of 10 μM Ins(1,4,5,6)P_4 along with the native or the original stapled peptide.

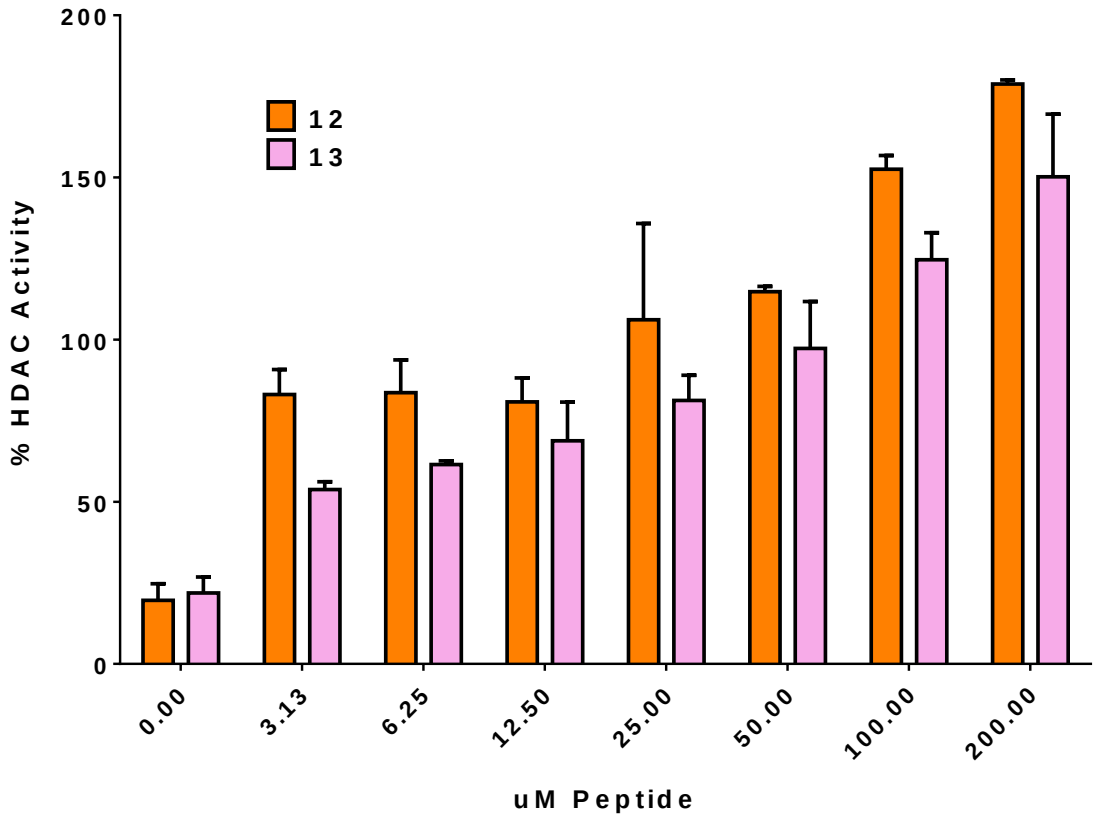


Figure 62: HDAC activities of the phosphotyrosine based peptides, **12** (Y470pY) and **13** (Y471pY). Error bars represent ± s.e.m. (n = 3).

Both stapled peptides **12** and **13** exhibit similar activities, with **12** exhibiting a slightly higher activity than **13** (20-30% more). In fact, with “salt stripped” complex, the HDAC3 activity observed for the phosphotyrosine containing peptides is higher than any of the other stapled peptide analogues.

This data suggests that rather than the phosphotyrosine stapled peptides inhibiting HDAC3 due to their phosphate groups competing with the Ins(1,4,5,6)P_4 molecule binding, these peptides actually activate HDAC3. A possible reason would be that the peptides are mimicking not only SMRT-DAD Helix 3, but also mimicking the Ins(1,4,5,6)P_4 molecule. In other words, these phosphotyrosine stapled peptides are dual mimics.

One way to test if this is the case, is by carrying out the assay with the addition of 10 μM Ins(1,4,5,6)P_4. Due to the phosphotyrosine mutation on these peptides, in comparison with the native sequence and the original staple, on addition of 10 μM Ins(1,4,5,6)P_4, there should be very little change in percentage HDAC activity. This is because the phosphate groups on the peptides occupy part of the binding site of the Ins(1,4,5,6)P_4 molecule and therefore does not require the Ins(1,4,5,6)P_4 molecule to increase the percentage HDAC activity.

However, in the presence of 10 μM Ins(1,4,5,6)P_4, both of the phosphotyrosine peptides still showed a significant increase in HDAC activity, contrary to what was thought. In fact, at a peptide concentration of 200 μM, the HDAC activity with 2.13 is shown to be over 1100% on the addition of 10 μM Ins(1,4,5,6)P4 (Figure 63), compared to only about 200% HDAC activity without Ins(1,4,5,6)P4). 2.14 shows a similar profile inducing a maximal HDAC activity of 1020% at 200 μM peptide and 10 μM Ins(1,4,5,6)P4 (Figure 63).

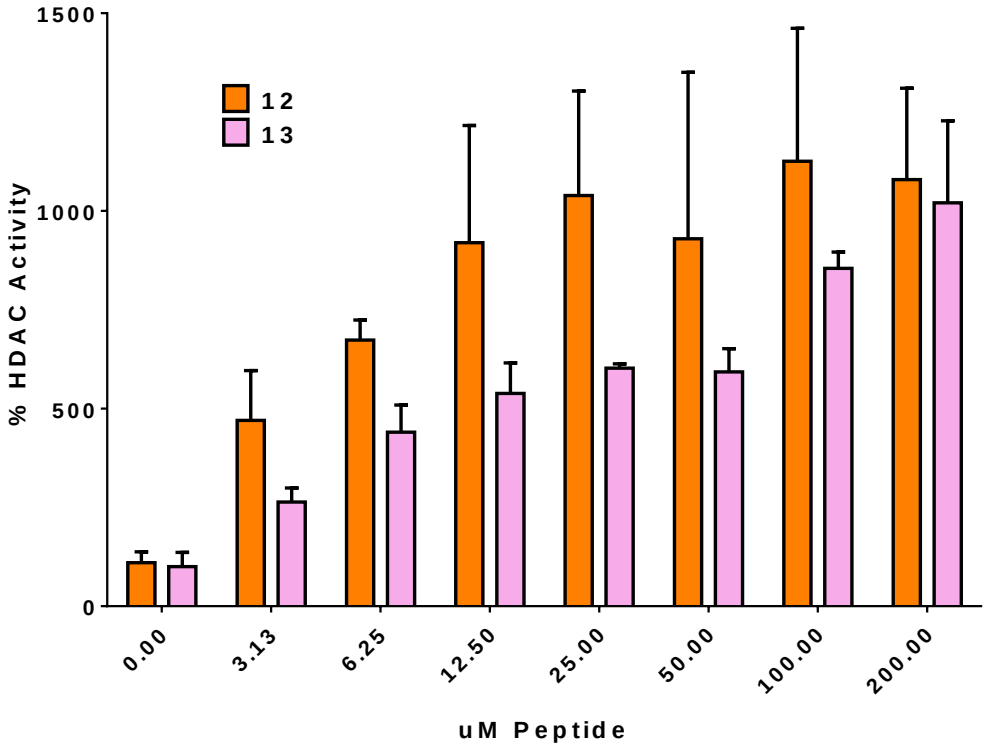
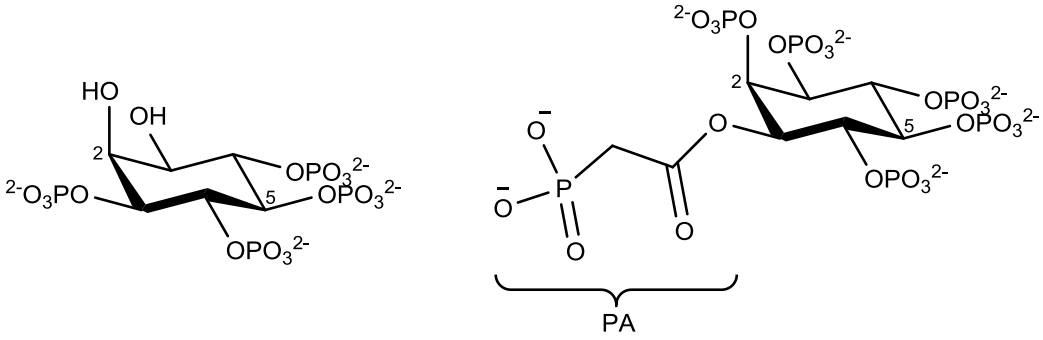


Figure 63: HDAC activities of the phosphotyrosine based peptides, **12** (Y470pY) and **13** (Y471pY), in the presence of 10 µM Ins(1,4,5,6)P_4. Error bars represent ± s.e.m. (n = 3).

In the presence of 10 µM Ins(1,4,5,6)P_4 the percentage HDAC activity still increases, which was not expected and is certainly of great interest. This could indicate that the peptide may not be binding to the site that they are thought to bind. In fact, they could be binding to a completely different surface of HDAC3, and therefore not able to displace the SMRT-DAD of HDAC3. Also this would mean that the mode of activation of HDAC3 is still very much an unknown.

A distinct possibility is that there could be another binding site close to the current SMRT-DAD binding site, which would allow another SANT domain to bind and also interact with the Ins(1,4,5,6)P_4 molecule bound to the HDAC3 surface. To test this hypothesis, an experiment was designed which would test if the peptides are actually bound on the opposite side of the Ins(1,4,5,6)P_4 to where the SMRT-DAD binds. This was achieved using 10 µM 1-PA-InsP_5 (Figure 64) instead of Ins(1,4,5,6)P_4. 1-PA-InsP_5 is able to mimic Ins(1,4,5,6)P_4 and activate HDAC3. The difference between these two inositol phosphates is the positioning of the phosphate groups. Ins(1,4,5,6)P_4 does not bear

phosphate groups at positions 2 and 3 on the cyclohexane ring and instead these are hydroxyls. In 1-PA-InsP$_5$, a phosphonoacetate (PA) is used to mimic a pyrophosphate at position 1 of the cyclohexane ring and positions 2 and 3 bear phosphates groups.



Ins(1,4,5,6)P$_4$ (**14**) 1-PA-InsP$_5$ (**15**)

Figure 64: Structures of Ins(1,4,5,6)P$_4$ and 1-PA-InsP$_5$.

The theory behind using 1-PA-InsP$_5$ is that if the peptides bind to the side of Ins(1,4,5,6)P$_4$ which does not have phosphate groups, then maybe the Tyr470 and Tyr471 of the native peptide sequence point into this phosphate free area and when these residues are mutated to Tyr470*p*Tyr and Tyr471*p*Tyr, the phosphate groups on the phosphotyrosine residues also point toward the Ins(1,4,5,6)P$_4$. However, when 1-PA-InsP$_5$ is used, the phosphate free sites of Ins(1,4,5,6)P$_4$ now contain phosphate groups also. Therefore, if the peptides did bind at this position then they should repel the 1-PA-InsP$_5$ and in turn reduce the percentage HDAC activity.

Nevertheless, upon carrying out the experiment it was clear that changing the inositol phosphate molecule from Ins(1,4,5,6)P$_4$ to 1-PA-InsP$_5$ did not make much of a difference in the percentage HDAC activity caused by each of the different peptides (Figure 65). This data therefore suggests that the binding site of the stapled peptide analogues may not be where it was originally thought. The only real way of determining without doubt the binding site of the peptides would be to co-crystallize them with HDAC3, SMRT-DAD and Ins(1,4,5,6)P$_4$.

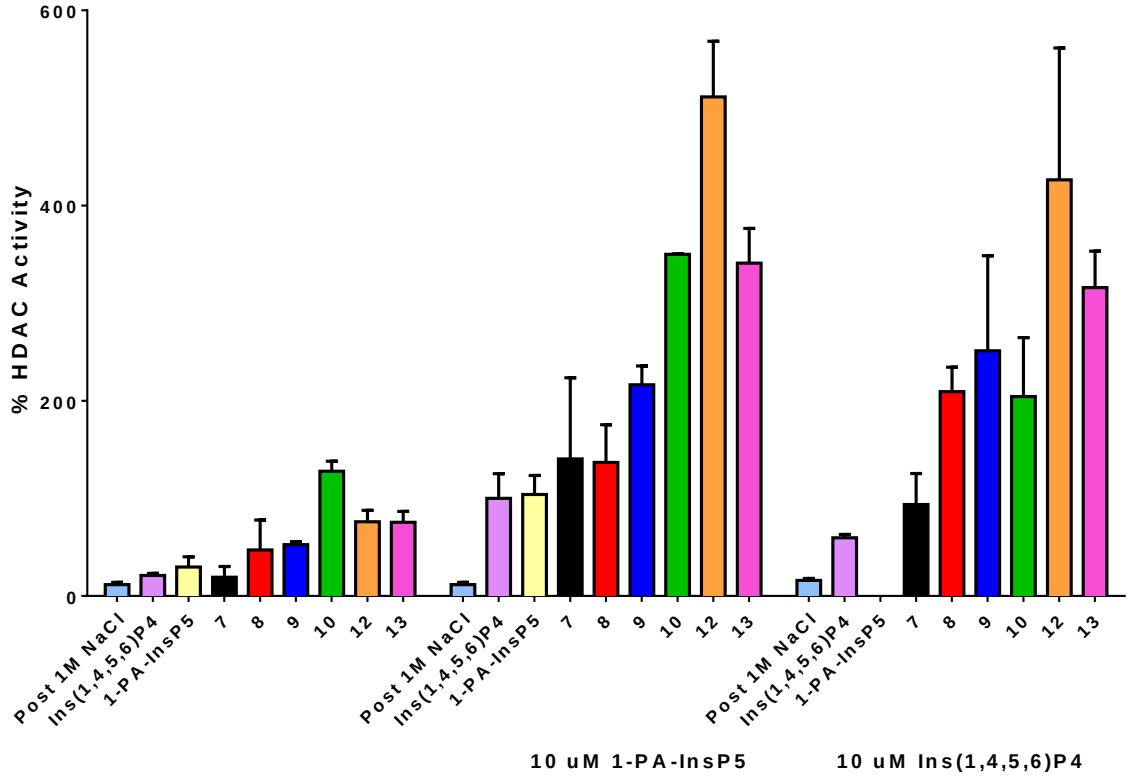


Figure 65: Effect of the different stapled peptide analogues at 50 μM concentration on the HDAC activity of the HDAC3-SMRT-DAD complex with and without the addition of 10 μM 1-PA-InsP_5 or Ins(1,4,5,6)P_4. Error bars represent ± s.e.m. (n = 3).

2.8. Effect of α,α-disubstitution on stapled peptide conformation

As a direct comparison with the initial stapled peptide **8**, a peptide was synthesised which incorporated the two unnatural α,α-disubstituted alkenyl amino acids in place of Glu465 and Tyr469. However, this peptide was not subjected to the RCM reaction to form the macrocyclic constraint, thus the staple is open (Figure 66).

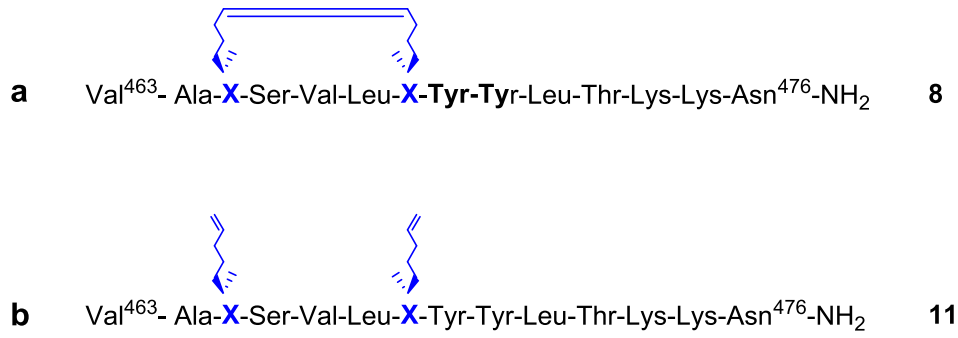


Figure 66: Structures of a) peptide **8** with the closed macrocycle and b) unmetathesised peptide **11**, where the macrocycle has been left open. The C466S mutation is also present in the open stapled peptide just like in the closed stapled peptide.

The difference between the open staple **11** and closed stapled peptide **8** was studied by circular dichroism spectroscopy, along with the native peptide **7**. Surprisingly, **11** showed a level of helicity comparable with that of the **8** (Figure 67). Although it was expected for that **11** would be a little helical, it was not expected to be as helical as **8**.

The only difference between these two peptides is the stapling. Thus, from this data, we could hypothesise that the stapling of the peptide seems to have little effect on its percentage helicity. Another possibility is that the percentage peptide helicity could stem from other factors such as the placement of the unnatural amino acid residues into the peptide. This is particularly true because in comparison to the native peptide **7**, peptides **8** and **11** only differ in the mutations of Glu465 and Tyr469 to the α,α-disubstituted unnatural amino acid, S_5. Another factor could be the inclusion of the α-methyl group into the unnatural amino acid, which forces the nucleation of helicity.

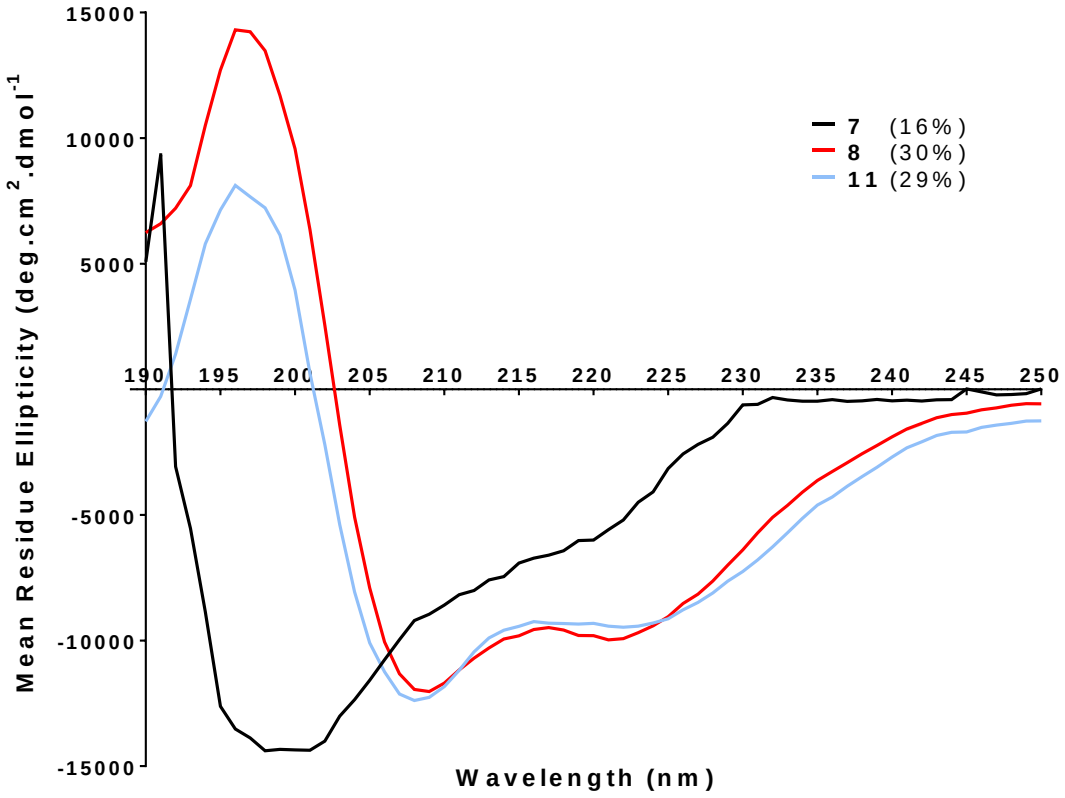
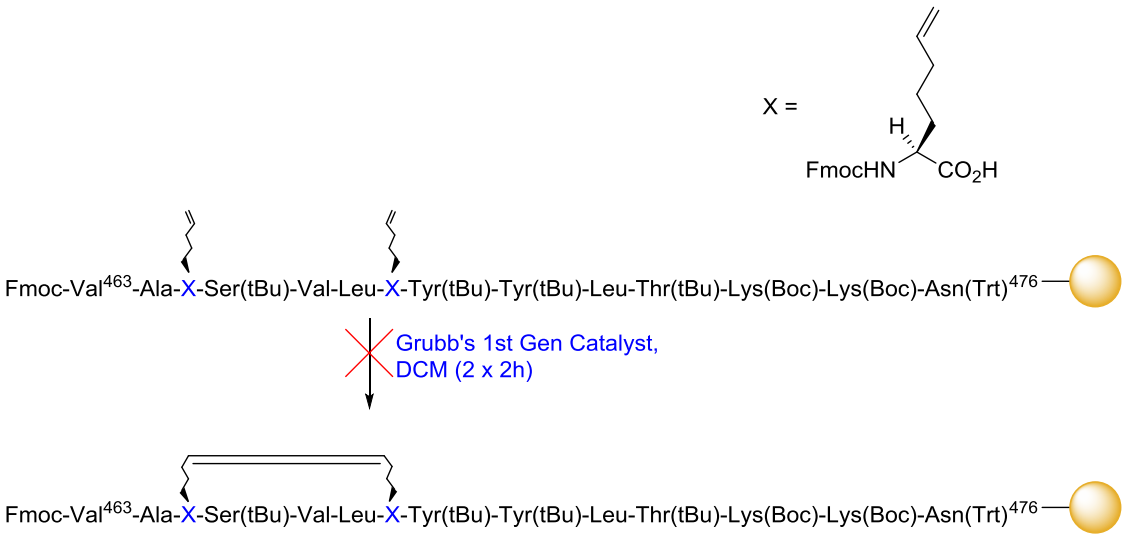


Figure 67: Circular dichroism spectra of the native SMRT-DAD helix 3 peptide (**7**, black) and the closed (**8**, red) and opened (**11**, blue) stapled peptides.

In addition, the extra α-methyl group on the amino acid was investigated further. This was achieved through the incorporation of an S_5 amino acid analogue which does not have the extra α-methyl group.

The mono-substituted unnatural amino acids were incorporated into the SMRT-DAD helix 3 peptide sequence in the same manner as for the α,α-disubstituted amino acids and in the same positions (Scheme 17).



Scheme 17: RCM of the SMRT-DAD helix 3 stapled peptide which does not contain α,α-disubstitution at the alkenyl amino acid residues. This synthesis was unsuccessful.

However, unlike the α,α-disubstituted amino acids stapled peptide derivatives, the peptide containing the mono-substituted amino acids did not undergo RCM. This was true even after different conditions were utilised for the RCM reaction including increasing the catalyst loading, changing the solvent and prolonging the time of the reaction (Table 1). This was even more surprising because there is literature precedent to suggest that the RCM on mono-substituted stapled peptides does in fact proceed to completion after just two iterations of 2 h each and 20 mol% Grubbs' 1st generation catalyst.[209]

Table 1: Attempted RCM on the mono-substituted, alkenyl amino acid containing peptides.

Catalyst loading	Solvent	Time	Conversion
20 mol%	DCM	2 x 2 h	n/a
100 mol%	DCM	2 x 2 h	n/a
20 mol%	DCE	2 x 2 h	n/a
100 mol%	DCE	2 x 2 h	n/a
20 mol%	DCE	2 x 24 h	n/a

From this data, it can be postulated that the α,α-disubstitution is particularly important to help the helical constraint form in this SMRT-DAD helix 3

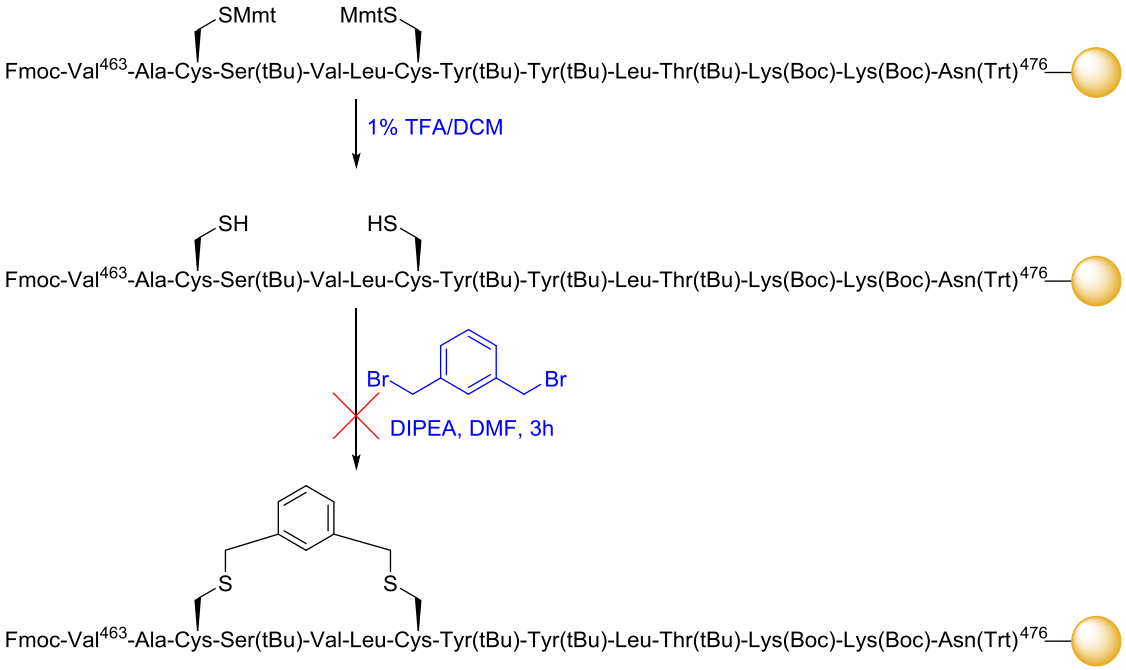
sequence. Although other groups have had success without the α,α-disubstitution,[209–211] this did not work in our system. Thus, different peptide sequences must have different requirements when it comes to incorporation of helical constraints. If a native peptide sequence already has a high propensity to form a helix, it may not be necessary to introduce α,α-disubstitution into the peptide sequence. However, in our peptide this does not seem to be the case.

2.9. Xylene-based peptide constraint

In addition to using all hydrocarbon stapled peptide constraints, we wanted to see if there were other ways in which we could constrain a peptide into an α-helical structure.

One way of doing this was to introduce a xylene moiety into the peptide that bridges two cysteine residues with an *i,i+4* spacing. The resulting bridge would be formed across one turn of the α-helix. For the bridging part, we chose to use *α,α'*-dibromo-*m*-xylene because, from a screening study of different thiol reactive crosslinkers, it was found to be the best at stabilizing the α-helix structure using a model peptide sequence, which in its native form showed a low to moderate level of helicity.[212]

For our xylene bridge containing peptide, the native residues mutated to cysteine were Glu465 and Tyr469. Moreover, the Fmoc-protected cysteine residues used were also side-chain protected by a monomethoxytrityl (Mmt) group, which is able to be cleaved selectively using a 1% solution of TFA in DCM for 10 min and was repeated until no more of the yellow colour indicative of the Mmt protecting group was seen (Scheme 18).



Scheme 18: Attempted macrocycle formation using *α,α'*-dibromo-*m*-xylene onto two cysteine residues at *i* and *i+4* positions. This synthesis was unsuccessful.

Crosslinking of the *α,α'*-dibromo-*m*-xylene with the deprotected cysteine residue was carried out using 2 equivalents of the alkylating agent and 4 equivalents of Hunig's base. After shaking for 3 h at room temperature in an SPPS tube, a cleavage test was taken to monitor the reaction. But, none of the desired cross-linked peptide was seen upon LC-MS analysis. It did however return unreacted starting material containing the deprotected cysteine residues.

Taking half of this unreacted peptide (~25 mg), which was still bound to the resin, the reaction was repeated, but instead of the reaction being carried out at room temperature for 3 h, it was carried out at 75 ˚C for 1 h in a microwave reactor. On removal from the microwave, the solvent was drained and the resin was washed in the usual manner. However, even after the crosslinking reaction was carried out at elevated temperatures, only the starting material was observed by LC-MS after a cleavage test was taken.

With the other half of the unreacted peptide on resin, a different approach was taken from that above. The peptide on resin was instead swollen in NH_4HCO_3 buffer (pH 8, 0.1 M) for 1 h, followed by the addition of *α,α'*-dibromo-*m*-xylene (5 equiv.) dissolved in the minimum volume of DMF and the resulting suspension

shaken in an SPPS tube for 3 h, as detailed in the literature.[212] After washings with DMF, MeOH and DCM, a cleavage test was undertaken using the usual cleavage cocktail. Once again, LC-MS showed there to be none of the desired product and as before it only showed the presence of the starting material. As a result for our 14-residue peptide system, this method of helical constraint seems to be of no use.

Once again the data obtained from this crosslinking method reiterates the fact that in our SMRT-DAD helix 3 peptide sequence, it seems that α,α-disubstitution is once again needed to help constrain the peptide into its helical form. To test this further, the same crosslinking reaction using *α,α'*-dibromo-*m*-xylene could be tried again, but instead using α,α-disubstituted cysteine residues in the peptide in place of Glu465 and Tyr469. If our hypothesis is correct, the α,α-disubstitution of the cysteine residues should start to induce a helical conformation in the peptide which should allow the crosslinking reaction with *α,α'*-dibromo-*m*-xylene to proceed with greater ease.

2.10. Crystal trials of SMRT-DAD stapled peptide 9 with HDAC3

The most conclusive way to determine where the SMRT-DAD staple peptides binds to HDAC3, is by obtaining a crystal structure. As a result, crystallisation experiments were set up. This was carried out using a Cartesian Honey Bee set up to deliver 100 nL drops of both the peptide and a buffer. As a control a parallel drop was used with 100 nL peptide and 100 nL water. Three different 96 condition crystal screens were used namely JCSG+, ProPlex and PACT.

Unfortunately, from each of these screens no usable crystals were seen. However, what were found were star-like structures with very thin needle projections (Figure 68), none of which were suitably large enough to obtain an x-ray diffraction pattern. These structures were found in drop F10 of the JCSG+ screen and drops C7 and F7 from the ProPlex screen. Optimisation of the drop conditions which contained star-like needle structures was carried out by varying salt conditions, pH and additives. However, these changes did not result in any crystals.

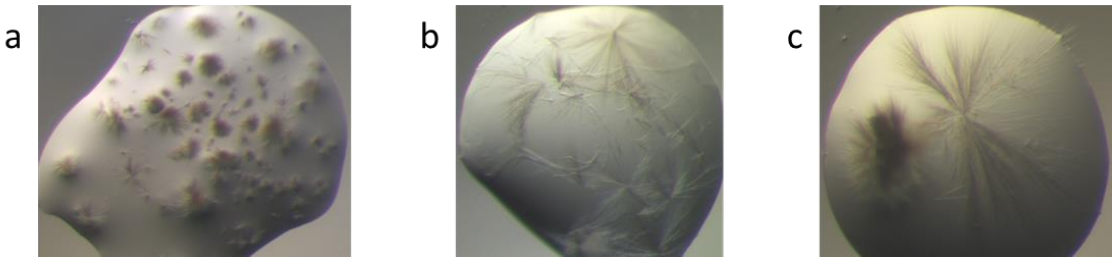


Figure 68: Photographs of crystal trial drops a) JCSG+ F10, b) ProPlex C7 and c) ProPlex F7 showing some of the varying structures which were found.

2.11. Conclusions and future work

In this project, a series of nickel (II) Schiff base complexes were synthesised. These underwent a decomplexation to provide the α,α-disubstituted alkenyl amino acids required for stapling. The amino acids were Fmoc-protected before incorporation into rationally designed peptide sequences. The key step in the synthesis of the stapled peptides was a ring closing metathesis reaction to form the macrocycle.

When compared to the native peptide, the stapled peptides showed higher percentage helicities, however, the helicities were still not optimal. Even with the suboptimal helicities, the biological activities of the peptides were still studied. However, none of the synthesised peptides turned out to be inhibitors of HDAC3. In fact, they were the opposite and activated HDAC3 to a very high degree.

Although not what was expected, these results lead to peptides that are able to regulate the activity of HDAC3. In fact, it is generally much more difficult to activate an enzyme than it is to inhibit it.

Through the synthesis of stapled analogues of SMRT-DAD helix 3, a series of peptide mimetic agonists were instead synthesised. These mimetics do not just mimic a short peptide sequence but instead mimic a whole protein. Thus, they can therefore also be termed proteomimetics. To be able to mimic a full protein with a 14-residue peptide sequence is quite an achievement.

Even though HDAC activity was observed upon addition of the stapled peptide analogues, unfortunately, no direct binding of the stapled peptides to the HDAC3 enzyme was able to be directly determined. This is still an important factor in determining the mechanism of action of HDAC enzymes.

Conformationally constraining peptides into their bioactive conformation can be achieved by using stapled peptides. In the system used for the stapling of SMRT-DAD helix 3 only an *i,i+4* staple was utilised. In the crystal structure of HDAC3 bound to SMRT-DAD, it can be seen that helix 3 of SMRT-DAD is made up of two helical turns and not just one. Thus, it would also be possible to staple SMRT-DAD helix 3 across two turns of the helix and so an *i,i+7* stapled peptide is also possible on this sequence (Figure 69). In addition, two different stitched stapled peptides are also possible on the SMRT-DAD helix 3 sequence. These involve mutating three residues on the peptide to three alkenyl amino acids, including a bis-alkenyl amino acid in the middle (Figure 70). For a peptide with an *i,i+4, i,i+8*, stitched staple (Figure 69b), the residue at the *i* position would be an S_5 amino acid, a B_5 amino acid at the *i+4* position, while at the *i+8* position an R_5 amino acid would be incorporated. The other possible stitched peptide for the SMRT-DAD helix 3 sequence would be an *i,i+3, i+7* stitched peptide (Figure 69c). This involves the mutation of the residues at the *i* and *i+7* positions to the S_5 amino acid and the *i+3* residue to the B_5 amino acid.

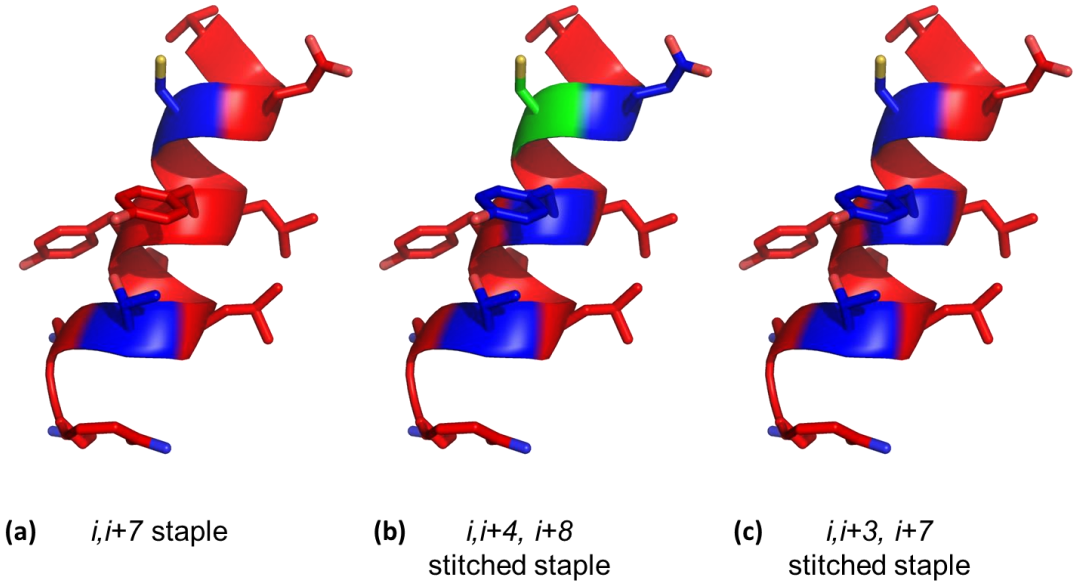


Figure 69: Possible positions for stapling on the SMRT-DAD helix 3 peptide (blue) to give (a) an *i,i+7*, (b) an *i,i+4, i+8* stitched and (c) an *i,i+3, i+7* stitched stapled peptides.

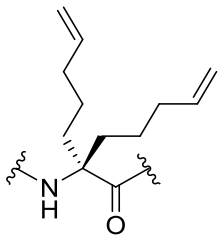


Figure 70: The B_5 amino acid used as the middle amino acid in a stitched peptide, which then undergoes RCM.

Going back to the original aims of the project, an isoform selective, allosteric peptide inhibitor of HDAC3 is still desirable. From the crystal structure of HDAC bound to SMRT-DAD, another helix of SMRT-DAD, helix 0, has direct contact with the HDAC3 surface. Thus, the helix 0 sequence could be constrained in the same manner to see whether this sequence would have a different effect from the helix 3 stapled peptide analogues.

If an allosteric staple peptide inhibitor is not possible, then going back to the active site of the enzyme may be the only option. In this regard, rather than using a small molecule inhibitor, like all the current HDAC inhibitors, a peptide

inhibitor would still be explored. The hypothesis is by using a peptide, the sequence may help to give specificity to one isoform over the others.

Chapter 3: Hydroxamic acid peptide inhibitor of HDAC1

3.1. Introduction

The natural substrates for many of the HDAC enzymes are acetylated lysine residues. As a posttranslational modification, lysine acetylation has been found in approximately 3600 different sites, across 1750 proteins, which include 61 acetylated histone sites.[27] On the histone proteins themselves, the histone tails undergo a highly dynamic process of posttranslational modifications including acetylation and deacetylation. On lysine residues it is the ε-amino group which is acetylated. Deacetylation of lysine residues in histone tails occurs through their interaction with the active site of HDAC enzymes.

There are four drug families that can be used as inhibitors of HDAC enzymes. In fact, four small molecule inhibitors have been approved by the FDA for the treatment of T-cell lymphomas and multiple myeloma, as discussed in Chapter 1.[62–64,79,83,213] Of these four FDA approved HDAC inhibitors that are out in the clinic, three contain a hydroxamic acid motif.

All four of these HDAC inhibitors are classed as *pan inhibitors*. They are small molecule inhibitors that are not isoform selective. Therefore these drugs have acute side effects. As a result, new strategies of inhibiting HDAC enzymes are needed which give selectivity to one isoform over the others.

In this study, we aimed to obtain a crystal structure of the natural peptide substrate bound to HDAC1. This is important to investigate how the histone 4 tail peptide makes interactions with the HDAC1 surface. From the data acquired a logical approach can be taken to design an isoform specific inhibitor peptide.

In order to obtain a crystal structure the native acetylated histone 4 tail peptide cannot be used. This is because the peptide would be turned over by the HDAC enzyme and then the deacetylated peptide released. To trap the peptide in the active site an inhibitor functionality is required on the peptide.

By using the histone 4 tail peptide sequence, residues 12-18, in the design of the inhibitor, it was proposed that HDAC isoform selectivity would be realised. It was hypothesised that the different HDAC isoforms may preferentially deacetylate certain lysine residues in histone tail sequences over others. Thus,

by incorporating an inhibitor functionality into a specific tail sequence, the possibility of finding an isoform specific sequence could become a reality.

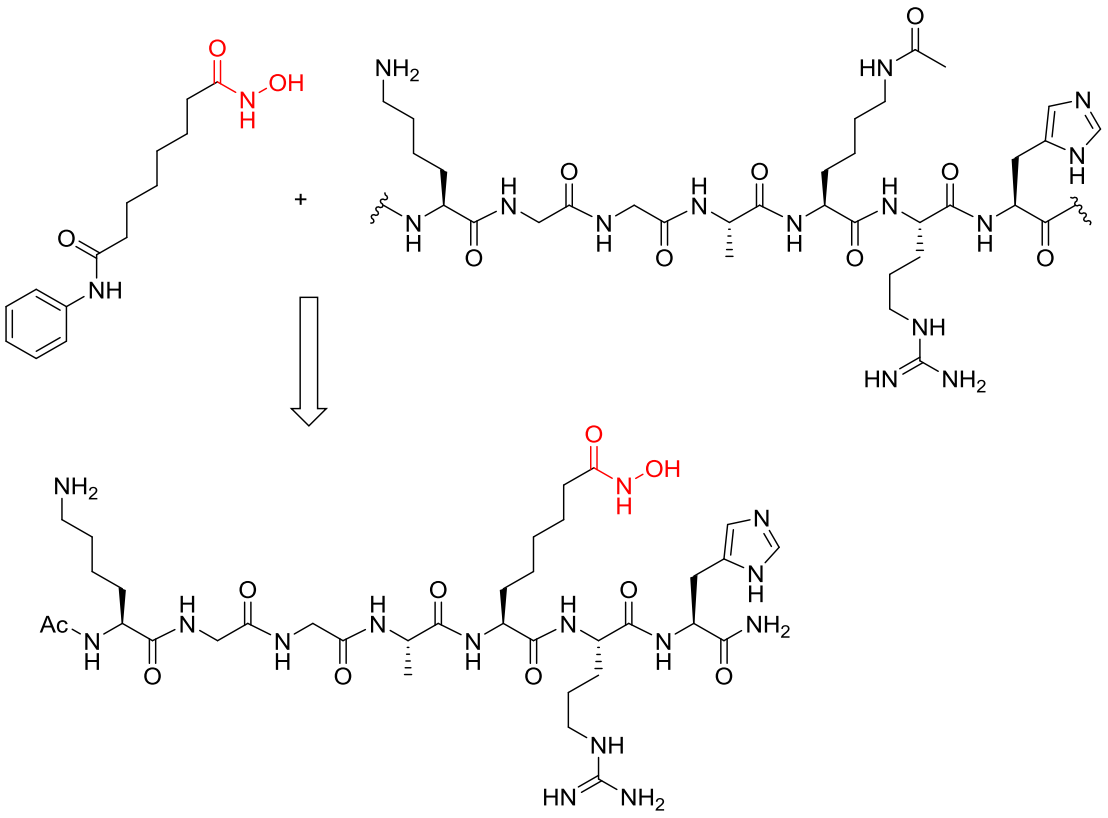
3.2. Aims of this chapter

The aims of this chapter are:

- To synthesise an unnatural amino acid with a zinc binding group at the side chain.
- To incorporate the unnatural inhibitor amino acid into the histone 4 (12-18) tail peptide sequence in place of the acetyl lysine residue.
- To test the inhibitor containing peptide against the HDAC1:MTA1 complex to assess it as an inhibitor
- To co-crystallise the inhibitor containing peptide with HDAC1:MTA1.

3.3. Design of hydroxamic acid containing histone tail peptide

The design of the hydroxamic acid containing histone tail peptide comes from both the native histone 4 tail sequence and the known HDACi SAHA. The length of the aliphatic chain for the hydroxamic acid containing amino acid is based on the length between the α-carbon and the side chain carbonyl on the acetyl lysine residue. Thus, the same distance from the α-carbon and the side chain carbonyl has been designed into the hydroxamic acid containing amino acid.



Scheme 19: Design of histone tail inhibitor H4 (12-18) K16Hxa based on SAHA and H4 (12-18) K16Ac.

3.4. Synthesis of hydroxyurea amino acid

Hydroxamic acids have been characterised as tight binding HDAC inhibitors. However, they possess poor pharmacokinetics and are known to be toxic in cells. This is because hydroxamic acids coordinate zinc ions readily and do not discriminate between different HDAC isoforms.[214] Although hydroxamic acids have been shown to be successful in the treatment of haematological malignancies, in solid tumours, their outcomes have been poor.[215,216]

Thus, a new zinc binding functionality to use as an HDAC inhibitor was required. Many different groups have tackled this problem and a wide set of bidentate zinc binding groups have been suggested (Figure 71).[217,218] The bidentate zinc chelator is required to displace the zinc-bound water molecule in the active site to inactivate the enzyme.[219]

The hydroxyurea functionality (Figure 71b) is very similar to a hydroxamic acid (Figure 71c) and has also been reported as a bidentate zinc-binding

motif.[218,220,221] In a computational study the hydroxyurea functionality (Figure 71b) was found to be a stronger zinc binder than the hydroxamic acid.[218,222] The hydroxyurea functionality also fits well into the HDAC pocket and forms good interactions with neighbouring residues in the active site pocket of the HDAC.[218] In addition, hydroxyureas are hydrolytically stable, unlike hydroxamic acids.

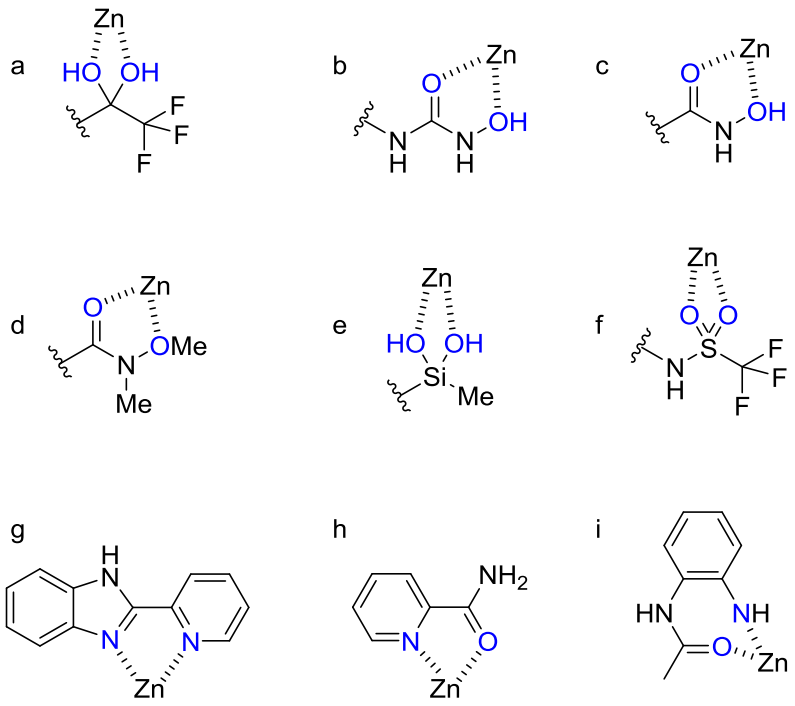
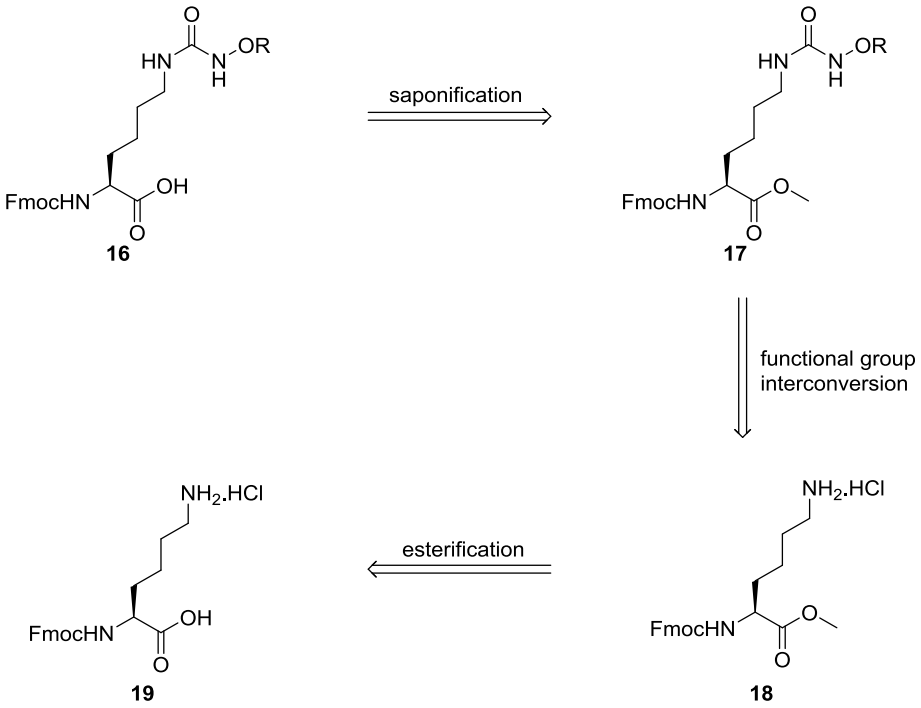


Figure 71: A selection of zinc binding motifs which include a) trifluoromethyl diol, b) hydroxyurea, c)hydroxamic acid, d) *N*-methyl-*N*-methoxy amide, e) silanediol, f) trifluorosulfonamide, g) 2-pyridylbenzimidazole, h) 2-pyridylcarboxamide and i) 2-aminoacetalide. The atoms which coordinate the zinc are highlighted in blue.[217,218]

For the easy incorporation of the hydroxyurea functionality into a peptide sequence, an Fmoc-protected unnatural amino acid containing a side-chain hydroxyurea functionality was required.

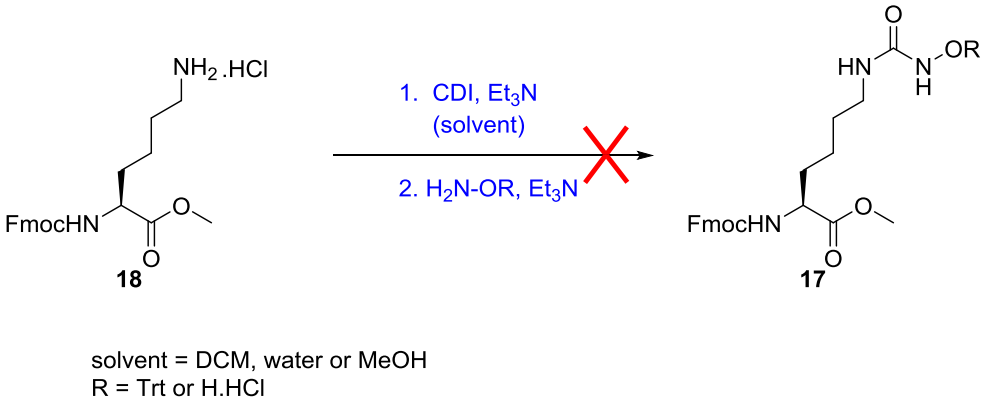
From the retrosynthetic analysis (Scheme 20), the synthesis of the hydroxyurea amino acid can be made in three steps starting from commercially available Fmoc-Lys-OH hydrochloride **19**.



Scheme 20: Retrosynthetic analysis for Fmoc-hydroxyurea-OR, where R = H or protecting group.

Fmoc-Lys-OH hydrochloride **19** was used as the starting material because it already has an Fmoc-protected α-amino group which is useful for Fmoc-based SPPS. As a result, no other base labile protecting groups can be used in the synthesis of a hydroxyurea containing amino acid. Additionally, with Fmoc being base labile, in the synthetic scheme, any steps requiring basic conditions must also be closely monitored.

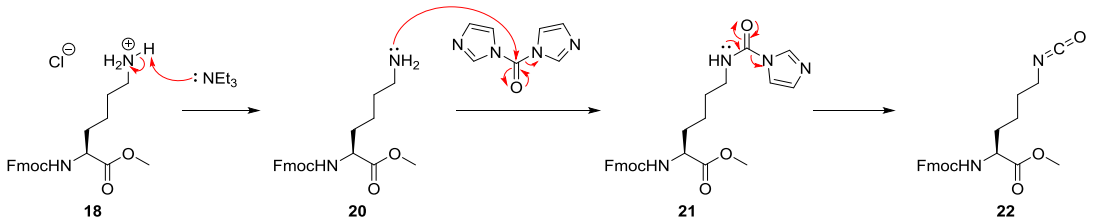
In the forward synthesis, the esterification of the acid was carried out using thionyl chloride and methanol and proceeded with near quantitative yields. With the Fmoc-Lys-OMe hydrochloride **18** in hand, the second step, introducing the hydroxyurea functionality, was carried out. However, this reaction did not yield the expected product **17** (Scheme 21). Systematic analysis of each step was then carried out to determine the reaction pathway.



solvent = DCM, water or MeOH
R = Trt or H.HCl

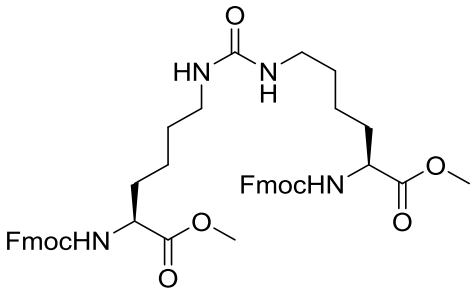
Scheme 21: Attempted synthesis of Fmoc-hydroyurea-OMe **17**.

Carbonyl diimidazole (CDI) was used in an attempt to form the isocyanate intermediate **22** in the route to synthesise the protected hydroxyurea (Scheme 22). However, this very reactive isothiocyanate intermediate was not seen and instead an imidazole urea **21**was recovered (Scheme 22).



Scheme 22: Mechanism to form the reactive isocyanate **22** using carbonyl diimidazole.

In addition to **21**, a bis-urea lysine by-product **23** (Figure 72) was also isolated. This is as a result of another molecule of the starting material reacting with an activated imidazole urea intermediate, resulting in the imidazole acting as a leaving group.



23

Figure 72: Unexpected bis-urea lysine by-product isolated from the reaction to form the isothiocyanate.

3.5. Synthesis of hydroxamic acid containing histone tail peptide

The natural substrates for HDAC enzymes are acetylated lysine residues on histone tails. To target the HDAC1:MTA1 corepressor complex, it was thought that mimicking a native substrate of the enzyme would help in achieving an isoform specific HDAC inhibitor.

The histone H4(12-18)K16ac peptide sequence has been shown to be a physiological substrate for HDAC1.[223] As a result this sequence was chosen and the acetylated lysine was replaced for a hydroxamic acid amino acid to give H4(12-18)K16Hxa as the target peptide.

3.5.1. Hydroxamic acid amino acid precursor synthesis

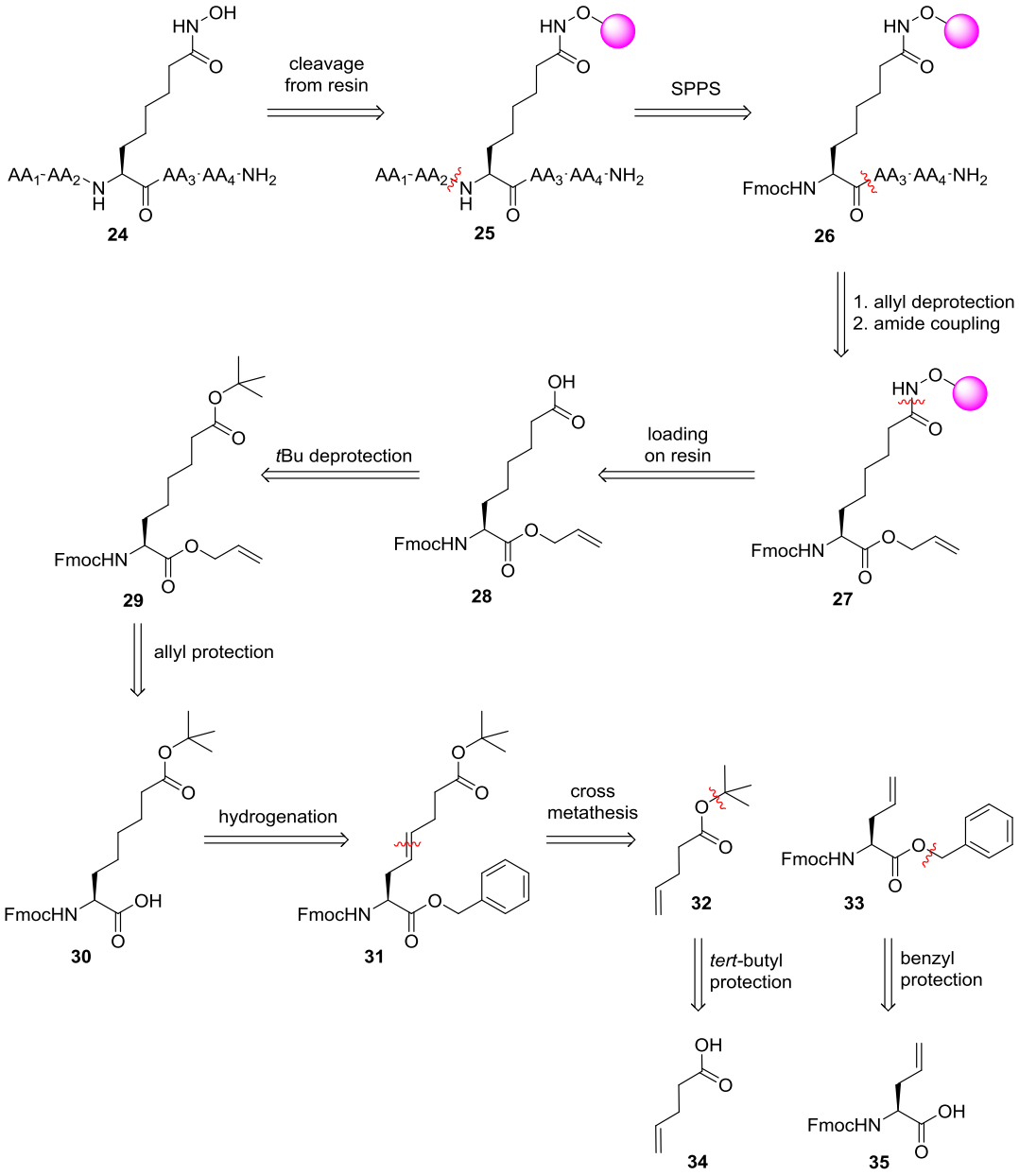
The complex synthesis of the hydroxyurea amino acid led to a rethink in the synthetic scheme. As a consequence, a hydroxamic acid moiety was once again desirable as a target compound, even with the associated toxicities.[214] For ease of incorporation into a histone tail peptide sequence, the hydroxamic acid moiety needs to be positioned at the side chain of an Fmoc-protected amino acid. Fairlie and colleagues have previously described the synthesis of a hydroxamic acid containing amino acid.[224] Based on this method, we have developed a route from commercially available Fmoc-L-allylglycine **34**.

In the retrosynthetic analysis (Scheme 23), peptide **24** can be cleaved from the resin following SPPS at the *N*-terminus of **26**. The *C*-terminus of **26** can be synthesised through an amide coupling following an allyl deprotection of **27**.

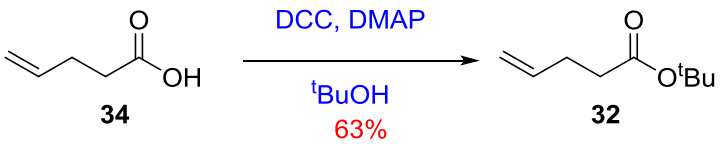
The protected hydroxamic acid functionality of **27** can be installed through an amide coupling of **28** at the side chain with 2-chlorotrityl hydroxylamine resin. The acid side chain of **28** is revealed through a *tert*-butyl deprotection of **29**, while the α-carboxylic acid of **30** is installed through an allyl protection.

Full saturation of the side chain and benzyl-deprotection of the α-acid to give **30** can be achieved through the hydrogenation of **31**. This in turn can be synthesised from the cross metathesis of **32** and **33**. The cross metathesis components, **32** and **33**, can both be synthesised through the esterification of commercially available starting materials, **34** and **35** as a *tert*-butyl and benzyl ester respectively.

The forward synthesis of the precursor began with a tertiary butyl protection of pentenoic acid **32**to give **34** in 63% yield as a colourless oil (Scheme 24).

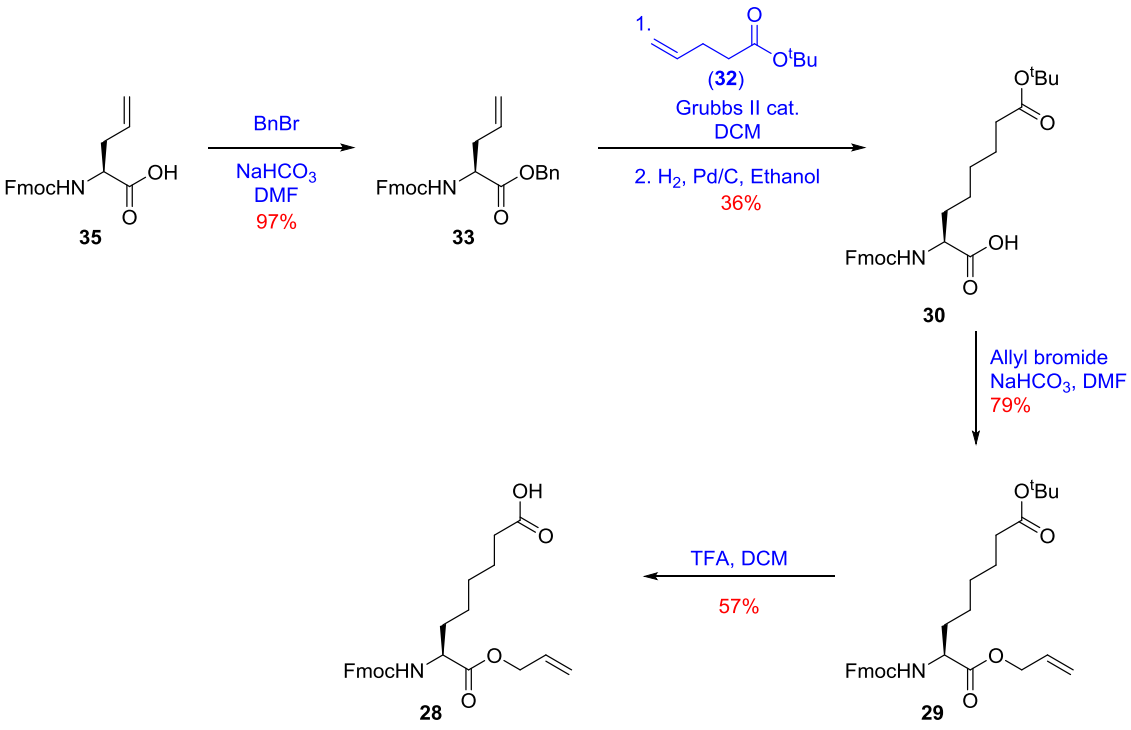


Scheme 23: Retrosynthetic analysis for the synthesis of a hydroxamic acid-containing peptide **24** from commercially available Fmoc-L-allylglycine **35**, based on work by Fairlie and colleagues.[224]



Scheme 24: Synthesis of *t*-butylpent-4-enoate **32**.

In parallel, the benzyl ester protection of commercially available Fmoc-L-allyl glycine **35** was also carried out with near quantitative yields (Scheme 25). Once protected, the allyl side chain of **33** can then undergo a cross metathesis reaction with **32** using Grubbs' 2nd generation catalyst. This cross metathesis reaction gives the unsaturated heterodimer **31**. Grubbs' 1st generation catalyst is not suitable for this cross metathesis reaction because both the alkene moieties concerned are so called Type 1 alkenes. Type 1 alkenes are renowned for undergoing self-condensation reactions with Grubbs' 1st generation catalyst to give the homodimers of the starting materials in addition to the heterodimer.[225]



Scheme 25: Synthetic route towards the precursor amino acid, Fmoc-oxooctanoic acid-OAllyl **28,** used in the synthesis of the hydroxamic acid containing amino acid.

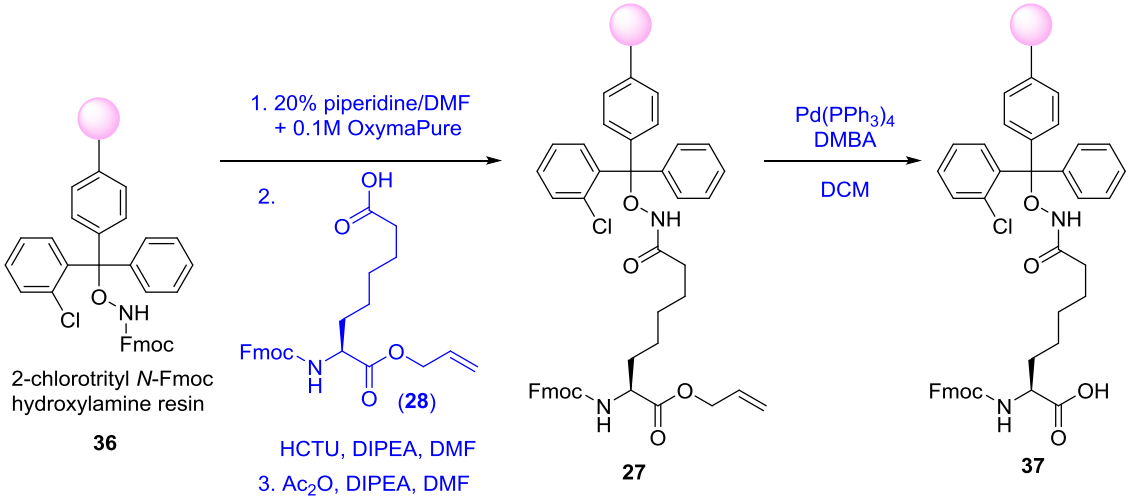
The cross metathesis product was then subjected to a hydrogenation reaction. In addition to giving a saturated hydrocarbon chain through the reduction of the alkene, hydrogenation also removes the benzyl ester protecting group on the amino acid to give **30**. As a result of the benzyl deprotection, the acid is then

further protected by an allyl ester to give **30**, because an allyl ester is orthogonal to Fmoc based SPPS. This protection-deprotection-protection strategy is necessary due to the lack of complete orthogonality between the other protection schemes and the reactions that the amino acid undergoes.

Deprotection of the tertiary-butyl ester on the amino acid side chain was then carried out under acidic conditions, to give Fmoc-oxooctanoic acid-OAll **28**. The deprotected side chain then becomes the precursor for the hydroxamic acid functionality on the amino acid.

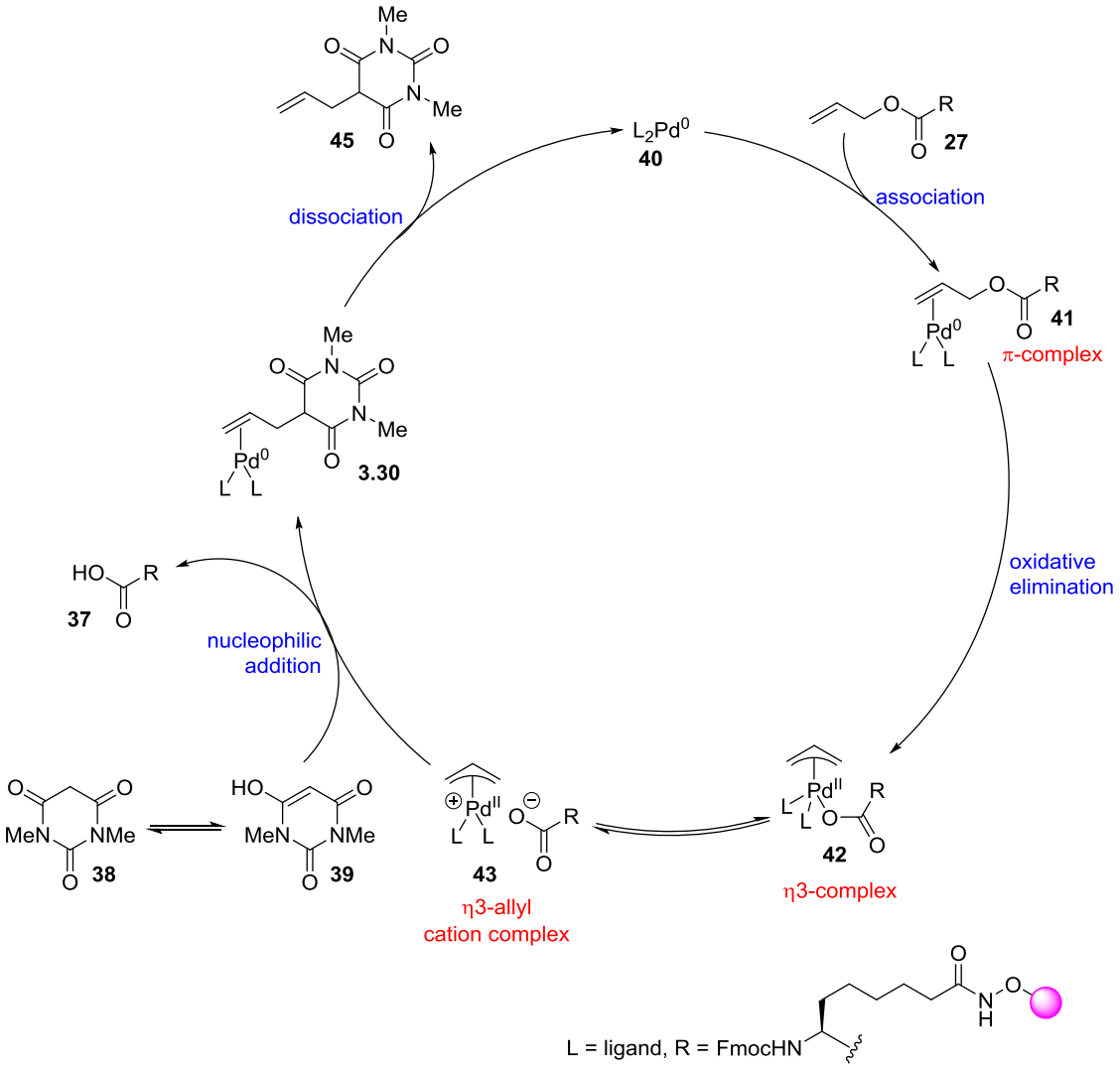
3.5.2. Amino acid loading on to 2-chlorotrityl resin preloaded with Fmoc-hydroxylamine

During SPPS, all reactive amino acid side chains are protected to prevent unwanted side reactions. The same needs to be true for a hydroxamic acid side chain. As with the Fairlie synthesis, the protected hydroxamic acid moiety is installed onto the amino acid side chain using 2-chlorotrityl resin preloaded with Fmoc hydroxylamine (**36**).[224] In this manner **2** was prepared from **28** after an initial Fmoc deprotection (Scheme 26). Capping of the resin was also performed to block any unreacted amine sites on the resin.



Scheme 26: Loading of Fmoc-oxooctanoic acid-OAll (**28**) onto 2-chlorotrityl *N*-Fmoc hydroxylamine resin (**36**), followed by a deallylation to give **37**.

The allyl ester on the α-acid of **27** was then removed using tetrakis(triphenylphosphine)palladium(0) and 1,3-dimethylbarbituric acid (DMBA, **38**) to give the free acid ready for subsequent coupling reactions. In fact, the acid of **37** is the side product of the allylation of DMBA (Scheme 27).



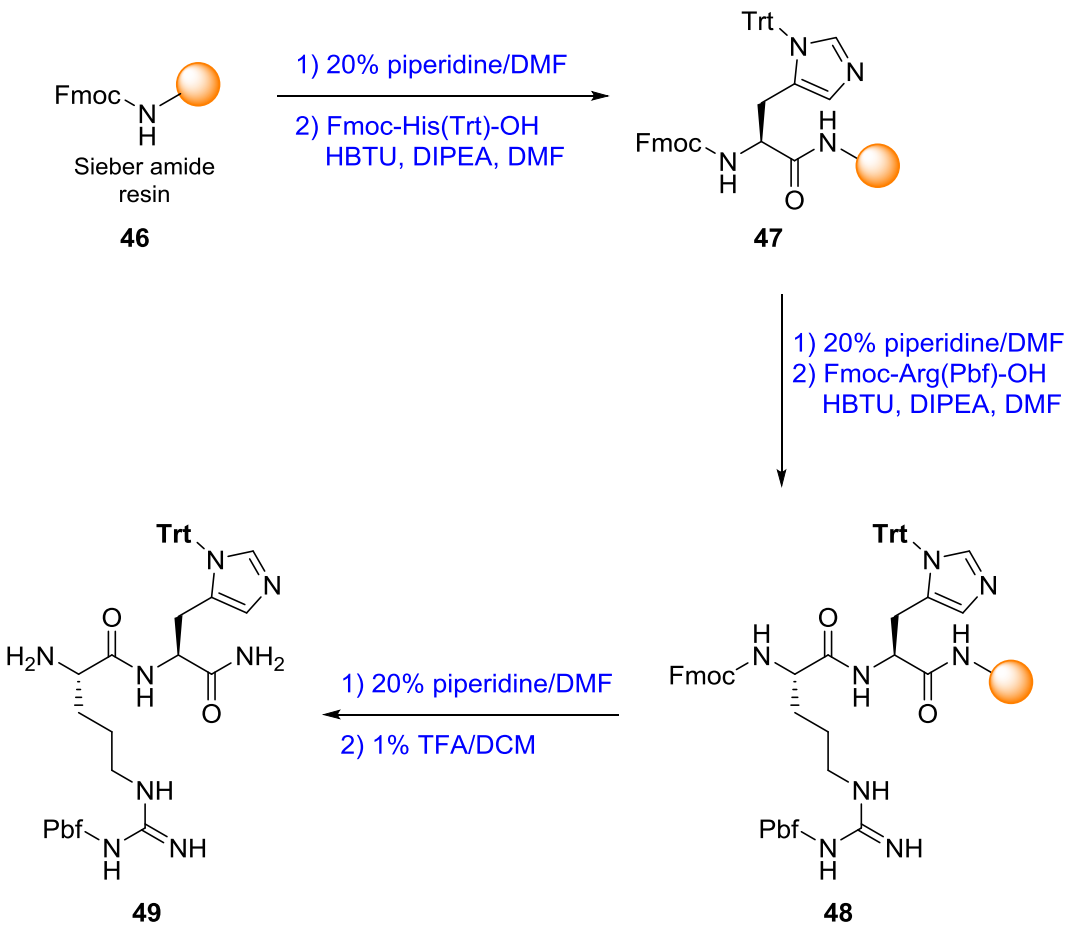
Scheme 27: Palladium catalysed allylation of 1,3-dimethylbarbituric acid (DMBA) also produces the free acid of **37** as a by-product of the nucleophilic addition step.

3.5.3. Synthesis of histone 4 tail (12-18) K16Hx

In H4 (12-18) K16ac, the acetylated lysine residue is situated towards the middle of the peptide and not at one end. This acetylated lysine is the residue mimicked by the hydroxamic acid. As a result, the dipeptide *C*-terminus of the

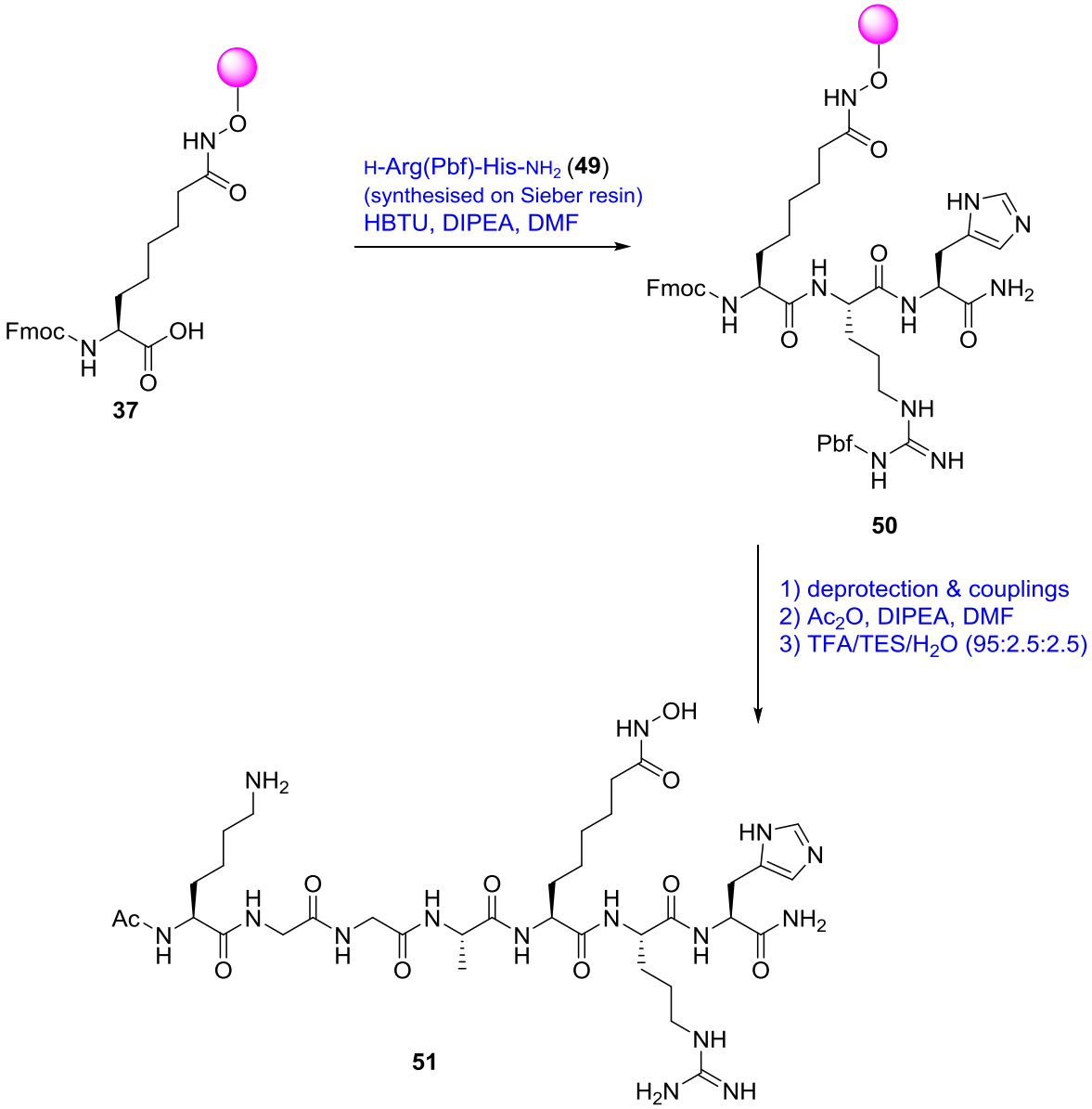
histone tail peptide needed to be synthesised first. Following this, the dipeptide can then be coupled onto the hydroxamic acid containing amino acid bound to the 2-chlorotrityl resin. Thus, the *C*-terminus of the histone tail peptide was synthesised on Sieber amide resin. This was done in order to have a *C*-terminal amide and hence, no negative charge on the *C*-terminus. Additionally, cleavage of the peptide from the resin can be achieved using relatively mild cleavage conditions, whilst leaving all the side chain protecting groups on the dipeptide intact.

The dipeptide was therefore synthesised in the usual manner on Sieber amide resin **46**. Upon a final Fmoc deprotection, the dipeptide was cleaved from the resin with 1% TFA in dichloromethane, while leaving both the Pbf and Trt protecting groups on the arginine and histidine respectively to give **49** (Scheme 28). **49**was then coupled onto the deprotected acid of the hydroxamic acid amino acid bound to the 2-chlorotrityl resin through the side chain **3.19** (Scheme 29).



Scheme 28: Synthesis of the H-Arg(Pbf)-His(Trt)-NH_2 **3.23** dipeptide on Sieber amide resin.

Next, the rest of the histone tail peptide sequence, towards the *N*-terminus, was synthesised through sequential deprotection and coupling steps using automated, microwave-assisted SPPS. The full peptide then underwent a final Fmoc deprotection and acetyl capping. This was followed by cleavage from the resin and at the same time global deprotection of the side chain protecting groups to give the hydroxamic acid containing histone tail peptide **51**with 1% recovered yield (Scheme 29). Following the isolation and purification of the hydroxamic acid containing histone tail peptide **51** it was used in biological studies and crystal trials.



Scheme 29: Synthesis of H4 (12-18) K16Hxa **51**.

3.6. Inhibition assay of hydroxamic acid containing peptide with HDAC1:MTA1

An inhibition assay was performed with the hydroxamic acid containing histone tail peptide by Chris Millard (Department of Molecular and Cell Biology, University of Leicester). This showed that H4 (12-18) K16Hxa inhibits the HDAC1:MTA1 complex with an IC_{50} of 336 nM (Figure 73). This is similar to the

values reported for other hydroxamic acid based inhibitors of class I HDAC enzymes.[226]

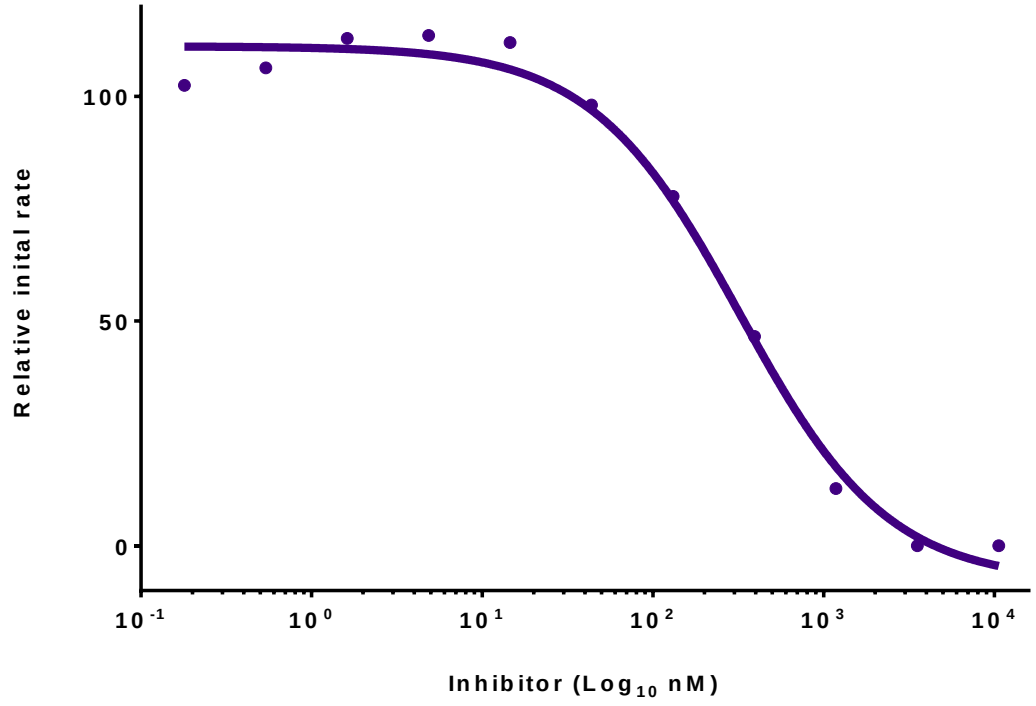


Figure 73: Inhibition curve of HDAC1:MTA1 by the histone H4 (12-18)K16Hx inhibitor peptide.

3.7. Crystal structure of hydroxamic acid containing peptide bound to HDAC1:MTA1:Ins(1,2,3,4,5,6)P_6

The hydroxamic acid containing histone tail peptide sequence, H4 (12-18) K16Hxa, was co-crystallised with the HDAC1:MTA1 corepressor complex before soaking with Ins(1,2,3,4,5,6)P_6 . The crystal structure of was solved by Chris Millard (Department of Molecular and Cell Biology, University of Leicester) (Figure 74).

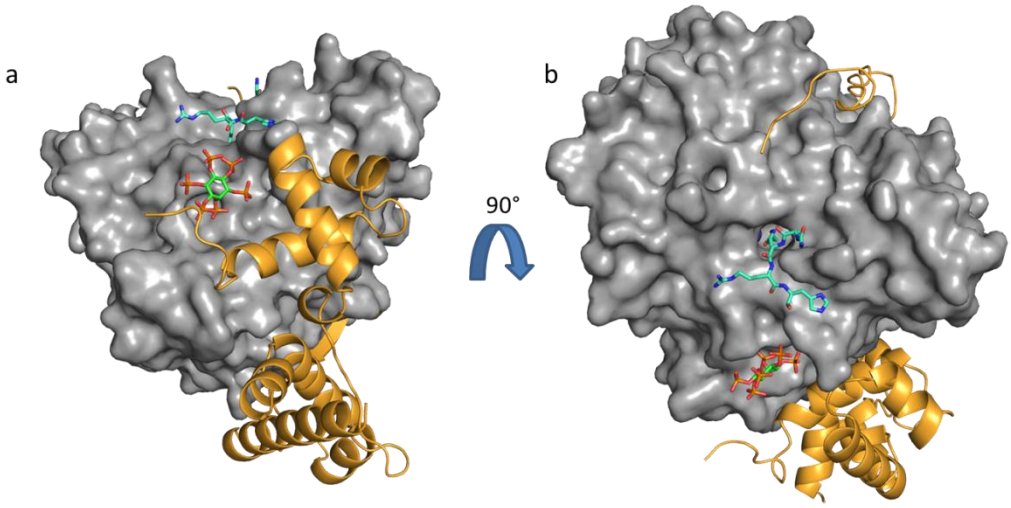


Figure 74: a) Crystal structure of HDAC1/MTA1/IP6/H4K16Hxa from the front face and b) rotated 90° to show the complex from the top face with the H4K16Hxa peptide bound in the active site. (pdb: 5ICN)

From the crystal structure data, it was observed that the hydroxamic acid side chain occupies a narrow hydrophobic channel in the HDAC1 active site (Figure 75). Furthermore, the hydroxamic acid functionality itself was coordinated to the zinc ion, buried deep within the active site. The coordination of the hydroxamic acid to the zinc ion in a bidentate fashion (Figure 75) echoes the binding of other class 1 HDAC inhibitors like SAHA and trichostatin A (TSA).[41]

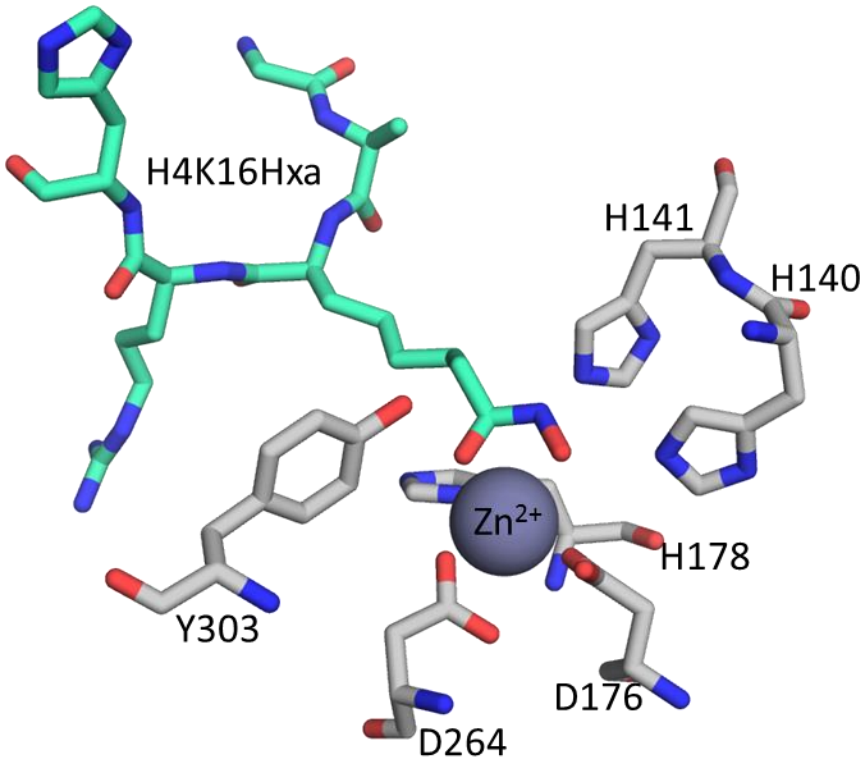


Figure 75: The positions of the active site residues (grey) involved in the coordination of zinc and the binding of H4K16Hxa (green). (pdb: 5ICN)

From the crystal structure, it can be seen that with the H4(12-18)K16Hxa peptide bound, the peptide backbone wraps around the side chain of Asp99 from HDAC1 (Figure 76a). Interestingly, Asp99 is a conserved residue in all class I and class II HDAC isoforms. Even more significant is the fact that in this bound state, the position of Asp99 and its neighbour residues have moved significantly in comparison to the HDAC1 complex without the bound peptide (Figure 76b). The movement in Asp99 mediates several hydrogen bonds with the H4K16Hxa peptide backbone, which can help stabilise the peptide inhibitor in the active site.

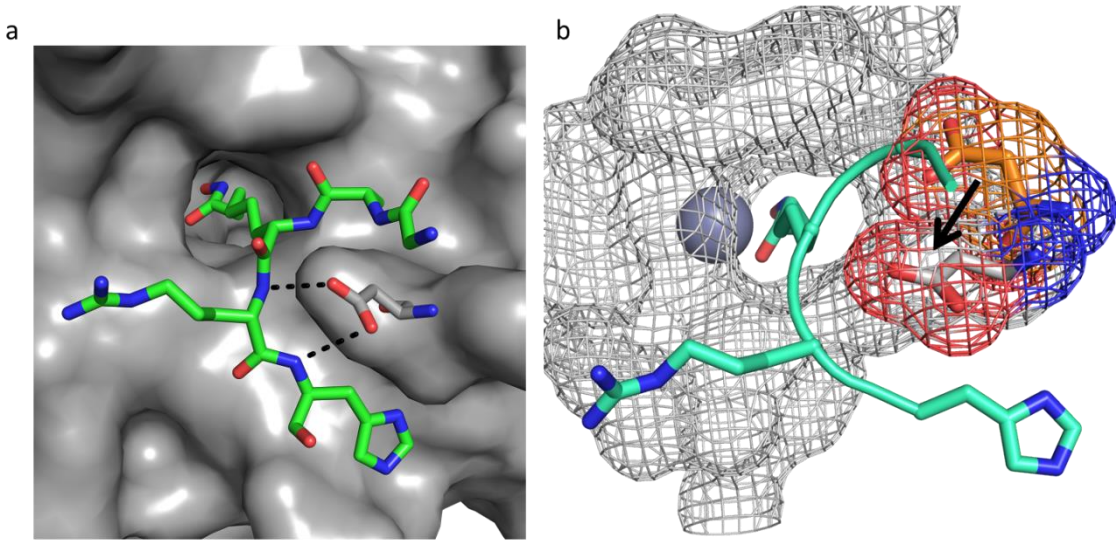


Figure 76: The H4 (12-18) K16Hxa peptide in the narrow active site channel and showing the backbone hydrogen bonding interactions of the peptide with the side chain of Asp99 on HDAC1 and b) the movement of Asp99 from the ligand free HDAC1 (orange mesh and sticks) to the HDAC1 with H3K16Hxa bound (grey mesh and sticks). The arrow shows the movement of Asp99 from its unbound to bound state. The H4K16Hxa (green) and zinc (grey sphere) are also shown. (pdb: 5ICN)

It has also been shown that Asp99 is an important residue implicated in deacetylase activity.[50,44] It was found that upon an alanine mutation of Asp99 (D99A), HDAC1 loses all its deacetylase activity. Furthermore, on addition of Ins(1,2,3,4,5,6)P_6 to the D99A HDAC1 mutant, deacetylase activity is not recovered, unlike with wild type HDAC1, which is maximally activated.[44]

With the knowledge of the importance of Asp99 to HDAC1 activity, this fact can be used in the design of second generation peptide inhibitors. More particularly, Asp99 can be used as an anchoring residue for the peptide inhibitors.

As a result of the peptide inhibitor H4K16Hxa wrapping itself around Asp99 in the crystal structure, a cyclic peptide can be hypothesised to be a good substrate for HDAC1. The cyclised peptide would wrap itself around the Asp99 residue, while still allowing the hydroxamic acid functionality to enter into the narrow active site channel of the HDAC1 enzyme. In addition, cyclic peptides are also more resistant to peptidases in comparison with linear peptide sequences. Another benefit of the cyclisation of the peptide inhibitor is that cyclic peptides are more likely to be cell penetrative. Indeed, Romidepsin, one

of the FDA approved HDAC inhibitors is a cyclic peptide, in fact it is a bicyclic peptide.

3.8. Improved synthesis of the hydroxamic acid amino acid via the Ni (II) Schiff base complex

The initial synthesis of H4(12-18)K16Hxa is not amenable to library peptide synthesis. This is as a result of the hydroxamic acid functionality coming from the resin. Therefore, to be able to synthesise a library of hydroxamic acid containing peptides, a different synthetic route was required.

For ease of incorporation of the hydroxamic acid containing amino acid into a peptide by standard Fmoc SPPS, the amino acid requires an Fmoc-protected α-amino group. An orthogonal protection strategy is also needed for the other protecting groups on the amino acid. In particular, protection of the α-acid is required during the synthesis of the amino acid and amino acid side chain during SPPS.

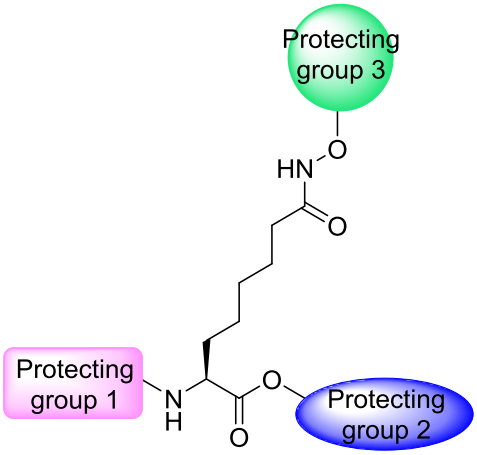
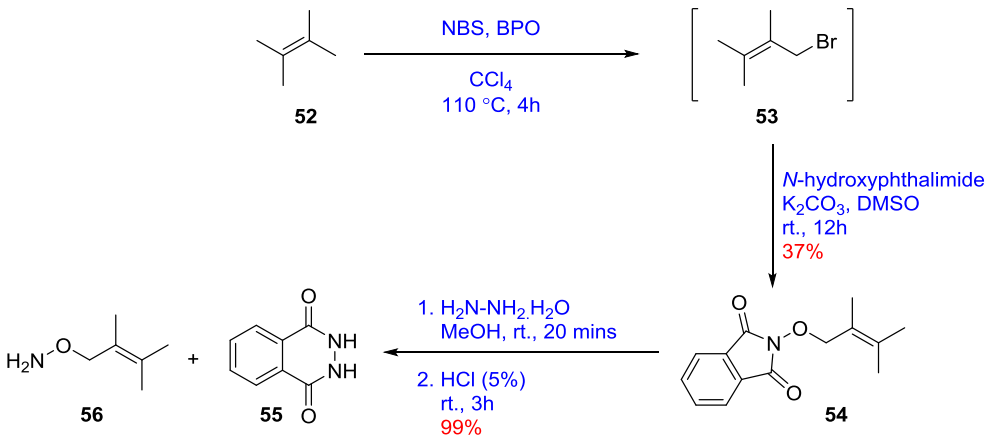


Figure 77: Protecting group strategy for orthogonal protection of the hydroxamic acid-containing amino acid.

The orthogonality of the protecting groups is such that the Fmoc group is removed in basic conditions and the side chain protecting group removed in acidic conditions. These allow for the use of the amino acid in solid phase

peptide synthesis. However, the protection of the acid group during the synthesis therefore requires a group which is neither acid nor base labile. In this respect, an allyl group was chosen, allowing cleavage of the protecting group with palladium tetrakis.

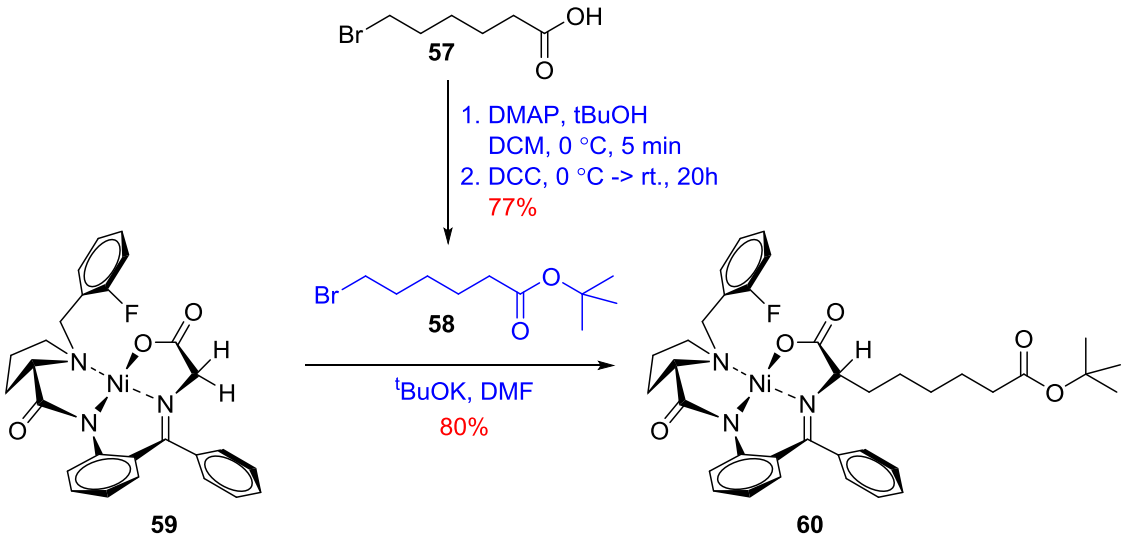
Protection of the side chain hydroxamic acid functionality was chosen to be an *O*-2-methylprenyl group based on the work by Nikitjuka and Jirgensons.[227] Starting from tetramethylethylene **52** and *N*-bromosuccinimide, and using benzoyl peroxide as a radical initiator, 1-bromo-2,3-dimethyl-2-butene **53** was synthesised. As a result of its volatility, **53** was not isolated, instead it was used in the next step without further purification to give *N*-(2-methylprenyl)oxy phthalimide **54** (Scheme 30). Using the Gabriel synthesis, **54** was cleaved with hydrazine monohydrate to give a stable cyclic by-product, 2,3-dihydrophthalazine-1,4-dione **55**, and the deprotected *O*-2-methylprenyl hydroxylamine **56**. A careful work up and evaporation of the ethereal solvent was required as a result of the volatility of **56**.



Scheme 30: Synthesis of O-2-methylprenyl hydroxylamine **56**.

The synthesis of the prerequisite amino acid for the hydroxamic acid-containing amino acid was achieved through the use of a Ni Schiff base complex pioneered by Belokon and coworkers[168,169] and improved upon by Aillard *et al.*[170]

The glycine Ni Schiff base complex **59** was used and alkylated with *tert*-butyl 6-bromohexanoate **58**, which was synthesised from 6-bromohexanoic acid **57** and *tert*-butanol following a previously described method.[228] The alkylation of the Ni Schiff base complex was carried out at room temperature for 15 mins to give the alkylated product **60** in 80% yield (Scheme 31).

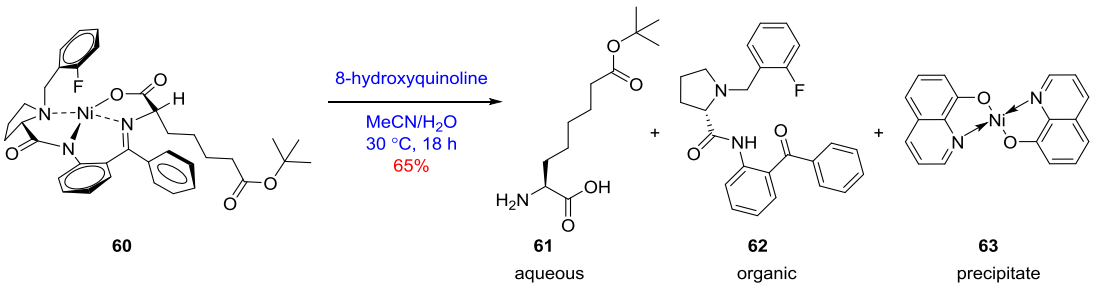


Scheme 31: Alkylation of Gly-Ni-2FBPB **59** with *tert*-butyl 6-bromohexanoate **58**.

Decomplexation of the alkylated Ni Schiff base complex is usually accomplished by heating an acidic solution of methanol to yield the free amino acid. However, the *tert*-butyl group on the acid side chain of the amino acid is also acid labile, therefore this method of decomplexation is not suitable.

In order to keep the acid labile protecting group intact, another method was required for the decomplexation. Nickel is well known at chelating different ligands. In fact this is one of the reasons that nickel is used for this method of unnatural amino acid synthesis. However, because of this, a ligand was needed that is able to out-compete the 2FBPB ligand. Patent literature highlighted a number of different possible ligands including citric acid, tartaric acid, oxalic acid, EGTA, EDTA, acetylacetone and 8-hydroxyquinoline to name a few.[229] From this patent, it was found that 8-hydroxyquinoline (Hqol) was the most effective Ni^{II} complexing agent and so Hqol was used in this synthesis (Scheme 32). Hqol was used at twice the excess of the Ni Schiff base complex to out-

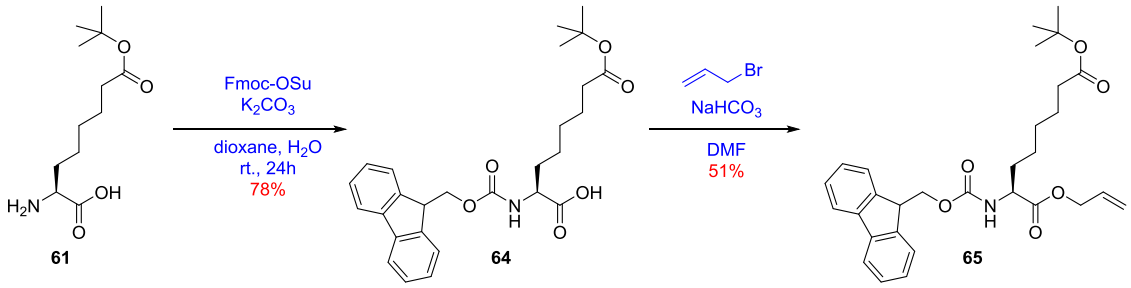
compete the 2-FBPB ligand **60**. The bis-hydroxyquinoline Ni (II) complex, [Ni(qol)$_2$] **61**, has previously been reported.[230,231]



Scheme 32: Decomplexation of the alkylated nickel (II) Schiff base complex **60** using 8-hydroxyquinoline.

The decomplexation of the Ni (II) Schiff base by bis-chelation with 8-hydroxyquinoline is an improvement on the previous acidic decomplexation reaction. Unlike the acidic decomplexation reaction carried out previously, the decomplexation using 8-hydroxyquinoline was carried out at 30 °C and over 18 hours. Although an increase in reaction time was required, the purification of the amino acid became simpler. The workup for the chelation does not require the purification of the amino acid through a Dowex ion exchange column. Instead, the workup requires only the extraction of the amino acid into water and lyophilisation of this solution to give the amino acid as a white powdery solid. In the decomplexation reaction with Hqol, **63** itself is filtered out as a yellow-green precipitate (Scheme 32). Any remaining complex and excess 8-hydroxyquinoline is extracted out from the reaction mixture with dichloromethane. The 2FBPB ligand **62** can also be recovered from the organic portion of the reaction workup following chromatographic purification.

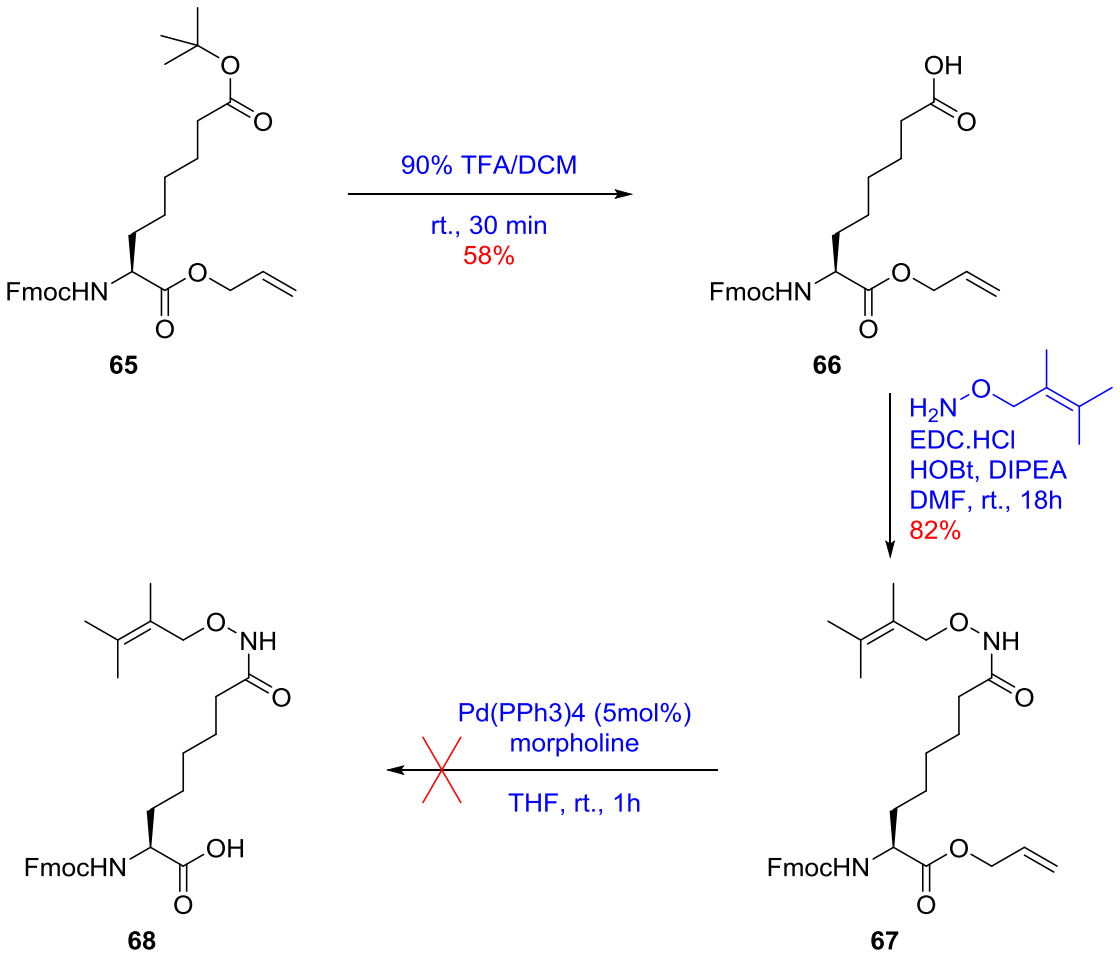
Subsequent to the decomplexation, an Fmoc-protection of the α-amine was carried out, followed by an allyl protection of the acid to give **65** (Scheme 33). An allyl protecting group was chosen because it is cleavable with the use of palladium tetrakis which is orthogonal to both the Fmoc-protection and the *tert*-butyl protection.



Scheme 33: Two step synthesis towards Fmoc-octanoate(*t*Bu)-OH **65** from aminosuberic acid 8-*tert*-butyl ester **61**.

However, the *O*-2-methylprenyl protecting group on the hydroxamic acid is very similar to an allyl group. Thus, it was thought that the *O*-2-methylprenyl group could also be removed upon deprotection of the allyl ester. When the *O*-2-methylprenyl protecting group was initially reported for hydroxamic acids, it was shown that the *O*-2-methylprenyl protecting group had high tolerance for many reaction conditions, and is even stable to 5 mol% Pd catalysts.[227] Thus, in addition to being orthogonal to both Fmoc and *tert*-butyl protecting groups, the allyl ester was also proposed to be orthogonal to the *O*-2-methylprenyl protecting group.

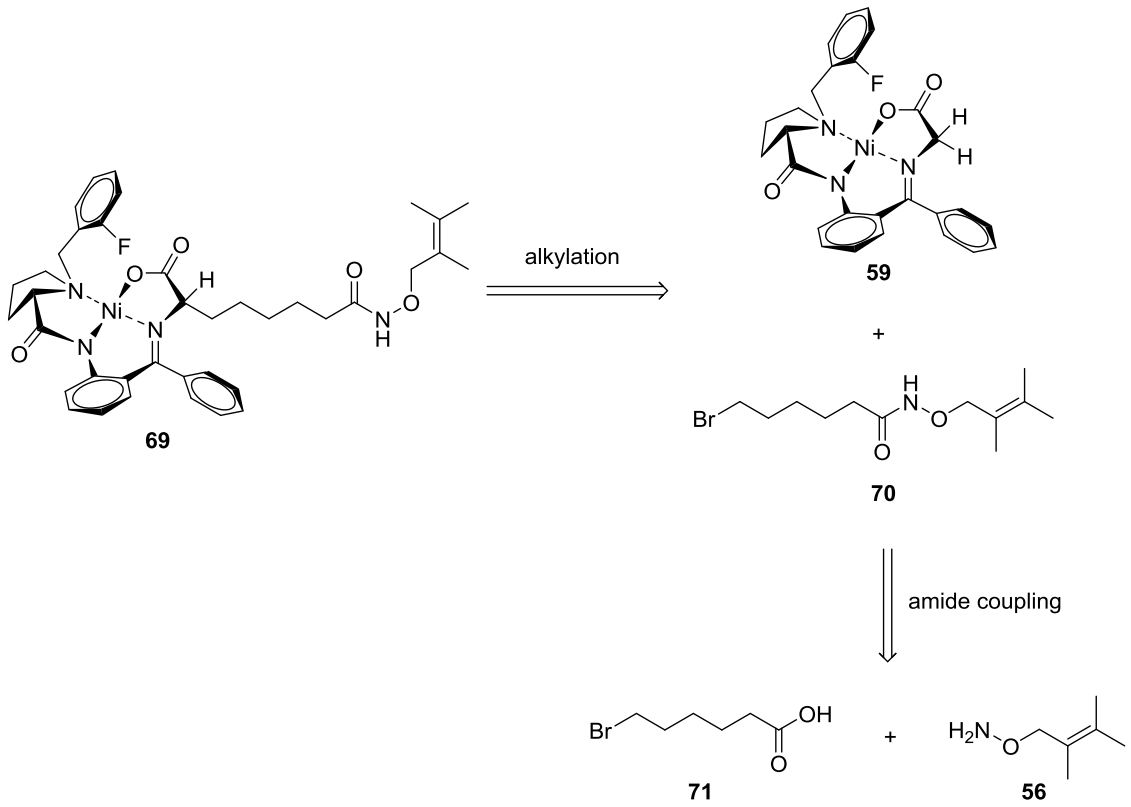
To introduce the hydroxamic acid moiety into the amino acid side chain, the *tert*-butyl protecting group was removed with a 90% solution of trifluoroacetic acid in dichloromethane. This was followed by coupling of *O*-2-methylprenyl hydroxylamine **56** to give fully protected **67** (Scheme 34).



Scheme 34: The final step in the synthesis of Fmoc-amino-8-(((2-methylprenyl)oxy)amino)-8-oxooctanoic acid **68** was unsuccessful.

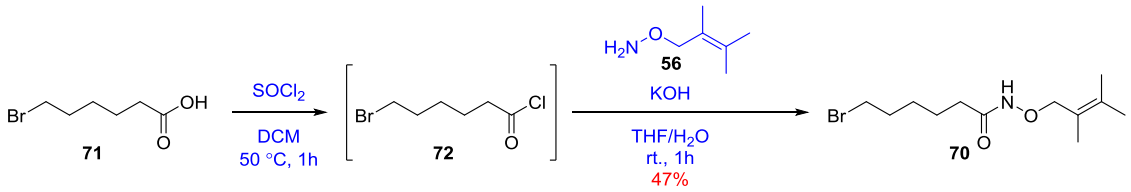
The final step in the synthesis of the hydroxamic acid containing amino acid was the deprotection of the allyl ester with tetrakis(triphenylphosphine) palladium(0) to give **65** ready for incorporation into a peptide sequence (Scheme 34). However, this final step was unsuccessful and none of the product was isolated. The difficulty in this synthesis was that the crucial step of removing the allyl protection on the α-acid is the final step in the synthetic scheme.

Another method for synthesising the hydroxamic acid containing amino acid also uses the nickel Schiff base as a starting point. The alkylation of the Schiff base is carried out using an alkyl halide electrophile **70**, which already has a hydroxamic acid moiety protected with a methylprenyl group (Scheme 36).



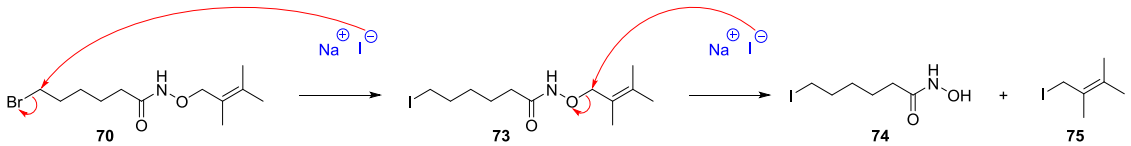
Scheme 35: Retrosynthetic analysis towards an alkylated Ni Schiff base containing an *O*-2-methylprenyl protected hydroxylamine.

The synthesis of the protected alkyl halide started from commercially available 6-bromohexanoic acid **71**. It was converted to the acid chloride using thionyl chloride (Scheme 36). Thionyl chloride is used in excess in comparison to 6-bromohexanoic acid and on completion of the reaction is removed *in vacuo* at the same time as the solvent. This leaves only the product of the reaction, which is 6-bromohexanoyl chloride **72** and was used without further purification. Coupling of **56** can then be carried out with potassium hydroxide acting as the base to give the protected hydroxamic acid alkyl bromide, 6-bromo-*N*-((2-methylprenyl)oxy)hexanamide **70** (Scheme 36).



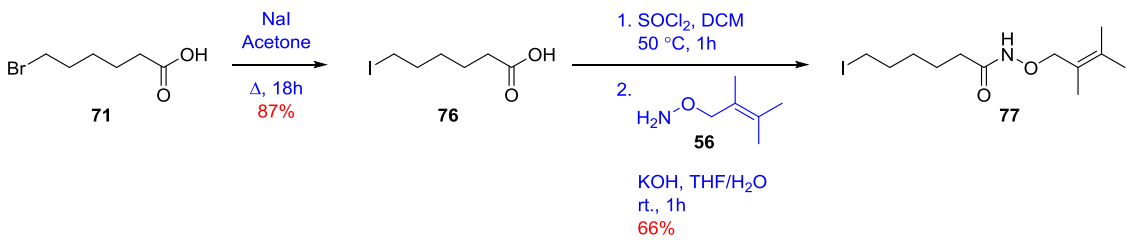
Scheme 36: Synthesis of 6-bromo-*N*-((2-methylprenyl)oxy)hexanamide **70** from commercially available 6-bromohexanoic acid **71**.

Alkyl bromides are well known nucleophiles and is thus a suitable electrophile for the alkylation of the nickel Schiff base complex.[170] However, in this instance, the alkylation using the **70** was unsuccessful; only staring material was recovered. Thus, a Finkelstein reaction was carried out to give a more reactive alkyl iodide electrophile. Although the Finkelstein reaction seemed to progress, including the precipitation of the sodium bromide salt, upon work up of the reaction, the product was not seen. After further investigation, the deprotected iodo-hydroxamic acid **74** was seen instead. This could be a result of a second S_N2 reaction taking place on the methylprenyl protecting group either before or after the S_N2 reaction on the alkyl halide moiety of the electrophile (Scheme 37).



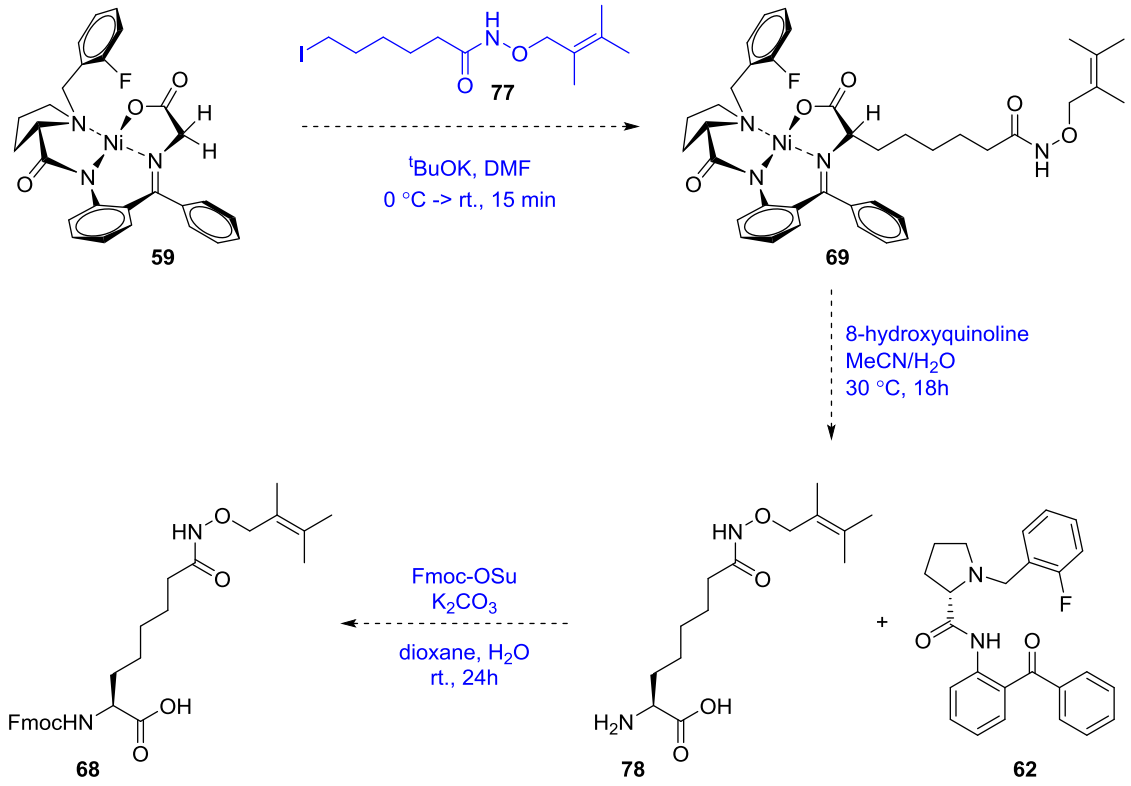
Scheme 37: Potential mechanism for the formation of *N*-hydroxy-6-iodohexanamide in the Finkelstein reaction.

To overcome this phenomenon, the Finkelstein reaction was instead carried out on **71**, which was followed by the coupling of **56** to give **77** as the more reactive electrophile (Scheme 38).



Scheme 38: Synthesis of 6-iodo-N-((2-methylprenyl)oxy)hexanamide **77**.

Through the synthesis of **61** and thus its use in the alkylation of the nickel Schiff base complex, extra steps in the previous synthesis of the hydroxamic acid containing amino acid are eliminated. These include the extra protection-deprotection steps of the *tert*-butyl acid along with the allyl protection and deprotection of the acid of the amino acid. By eliminating these steps, the length of the synthesis of the amino acid is significantly reduced from seven to three steps starting from the glycine Ni Schiff base complex **59**(Scheme 39). The completion of this proposed scheme was not possible due to time constraints.



Scheme 39: Completion of the synthesis of Fmoc-amino-8-(((2-methylprenyl)oxy)amino)-8-oxooctanoic acid **68** requires a three step synthesis from the Ni Schiff base complex **59**.

3.9. Conclusion and future work

In this project, Fmoc-oxooctanoic acid-OAll was successfully synthesised as a precursor for the hydroxamic acid functionality. This amino acid was coupled onto 2-chlorotrityl resin preloaded with hydroxylamine. Automated, microwave-

assisted SPPS was then used to synthesise the rest of the peptide. The hydroxamic acid containing peptide, H4(12-18)K16Hxa, was successfully synthesised. H4(12-18)K16Hxa was found to be an inhibitor of the HDAC1:MTA1 corepressor complex with a similar IC_{50} as TSA.[226]

The crystal structure of H4(12-18)K16Hxa bound in the active site of HDAC1 has also been solved. This is the first known inhibitor mimetic based on the native peptide substrate. The HDAC1:MTA1:Ins(1,2,3,4,5,6)P_4:H4(12-18)K16Hxa crystal structure has provided significant information regarding the interactions of peptide substrates to the HDAC1 surface. Like with HDAC8,[50] the importance of Asp99 in HDAC1 has also been highlighted.

A novel synthetic strategy towards a side chain protected, hydroxamic acid containing amino acid has also been designed and developed. Although the synthesis has not yet been fully achieved, this work is ongoing. Successful synthesis of this amino acid will lead to a library of hydroxamic acid containing histone tail peptides. This library would include variety in the lengths of the histone tail peptide sequence as well as the positioning of the hydroxamic acid containing amino acid for each of the different histone tail peptides. Through the synthesis of a series of short histone tail peptide inhibitors it could be possible to determine the specificity of each of the tail sequences for HDAC1. Additionally, each of the histone tail inhibitor peptide sequences need to be tested against the different class I HDAC isoforms, not just HDAC1. This would allow the potential determination of specificity of the different histone tail inhibitor peptides for each of the class I HDAC isoforms. Additionally, the different hydroxamic acid containing peptides need to be tested against the different corepressor complexes for the class I HDAC enzymes. This is important because except for HDAC8, all other class I HDAC enzymes need to be recruited into their corepressor complexes in order to be active.

To validate the strategy of using hydroxamic acid containing histone tail peptide inhibitors, there are two tail sequences that have been shown by our group to be particularly good at targeting either HDAC1 or HDAC3. These are H4K16Ac and H3K27ac. By confirming the specificity of each peptide with each HDAC

enzyme, it can be reasoned that there is potential of making isoform specific class I HDAC inhibitors based on histone tail peptide sequences.

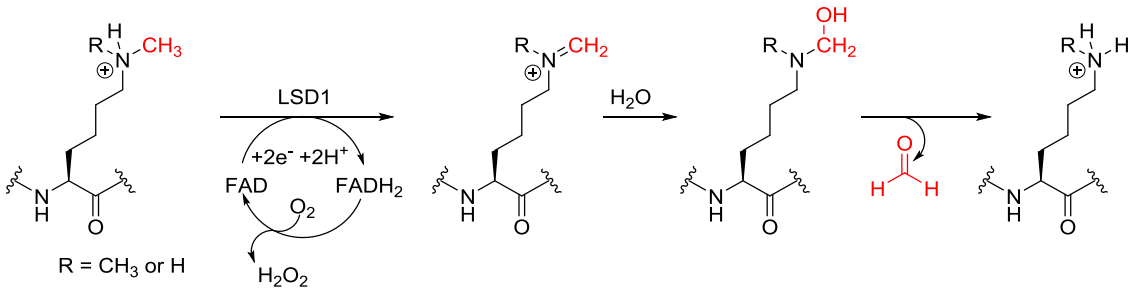
Along with the differing histone tail peptide sequences, the zinc binding group functionality within the peptide could also be altered. This will be achieved by the incorporation of other binding motifs, such as those reported by other groups.[220,218] Using a selection of different zinc binding motifs, it may be possible to find a histone tail peptide inhibitor that has a lower associated toxicity than those which contain hydroxamic acids.

Studying the crystal structure obtained, H4(12-18)K16Hxa wraps around the Asp99 shelf on HDAC1. This fact can also be used in future studies to design and synthesise cyclic peptides based on H4(12-18)K16Hxa, that would completely wrap around the Asp99 shelf. This could help lock the peptide in position and improve its inhibitor capabilities towards the HDAC1:MTA1 corepressor complex.

Chapter 4: Synthetic analogues of histone 3 (1-21) tail sequences for NMR studies

4.1. Introduction

In order to be active, class I HDAC enzymes need to be recruited into much larger corepressor complexes. This chapter describes experiments on the CoREST complex. The CoREST complex consists of HDAC1, LSD1 and CoREST.[39] LSD1 (lysine specific demethylase 1) was discovered to be a demethylase enzyme in 2004 by Shi *et al.*[232] It is a nuclear homolog of the flavin-dependant amine oxidases and demethylates by catalysing an oxidative process.[233,234] The side chain amino group of a methylated lysine residue on the histone tail peptide is oxidised to the corresponding imine, which in turn is then hydrolysed generating formaldehylde (Scheme 40).[233] In the oxidative process of demethylation, the FAD (flavin adenine dinucleotide) cofactor is reduced to $FADH_2$ which is then reoxidised by oxygen to produce hydrogen peroxide.[235]



Scheme 40: Demethylation reaction catalysed by LSD1 with molecular oxygen acting as an electron acceptor while the methylated lysine amine is oxidised to give the corresponding imine, which is then hydrolysed to give the demethylated lysine residue and forming a molecule of formaldehyde in the process.

LSD1 is a specific enzyme that will only demethylate mono- or dimethylated H3K4 substrates and does not have a strong preference for one or the other.[232–234,236] Trimethylated H3K4 however, is not a substrate for LSD1. This is likely due to the inherent chemistry of the flavin-containing amino oxidases, which requires a protonated nitrogen in the substrates.[232] As a result, only mono and dimethylated lysine residues are suitable substrates for LSD1.

As well distinguishing between different methylation states of H3K4, LSD1 can also discriminate between methylation at different lysine residues and different PTMs on the histone tail peptide.[234] If methylation at H3K9 is present, although

it is not demethylated, it does not affect the activity of LSD1. On the other hand, acetylation at H3K9 reduces the activity of LSD1. The most detrimental histone tail PTM is the phosphorylation of H3S10 which stops the peptide binding to LSD1 completely.[234]

This suggests that the charge on the peptide substrate is important for the recognition and binding of the peptide to LSD1. More particularly, the addition of PTMs that alter the charge on the histone peptide side chains, such as acetylation and phosphorylation, is detrimental to the activity of LSD1.

The binding of the H3 tail peptide into the LSD1 pocket is facilitated by a number of different intermolecular interactions including Arg2, Gln5, Arg8 and Ser10 (Figure 78).[236] Arg2 and Arg8 in particular are crucial in the binding. The Arg2 side chain forms H-bonds with the carbonyl backbones of both Gly11 and Gly12 as well as with the side chain of Ser10. It also forms salt bridges with the carboxylate groups of Asp553 and Asp556. Arg8 also forms interactions with the carbonyl of Cys360, and the carboxylates of Asp375 and Glu379.

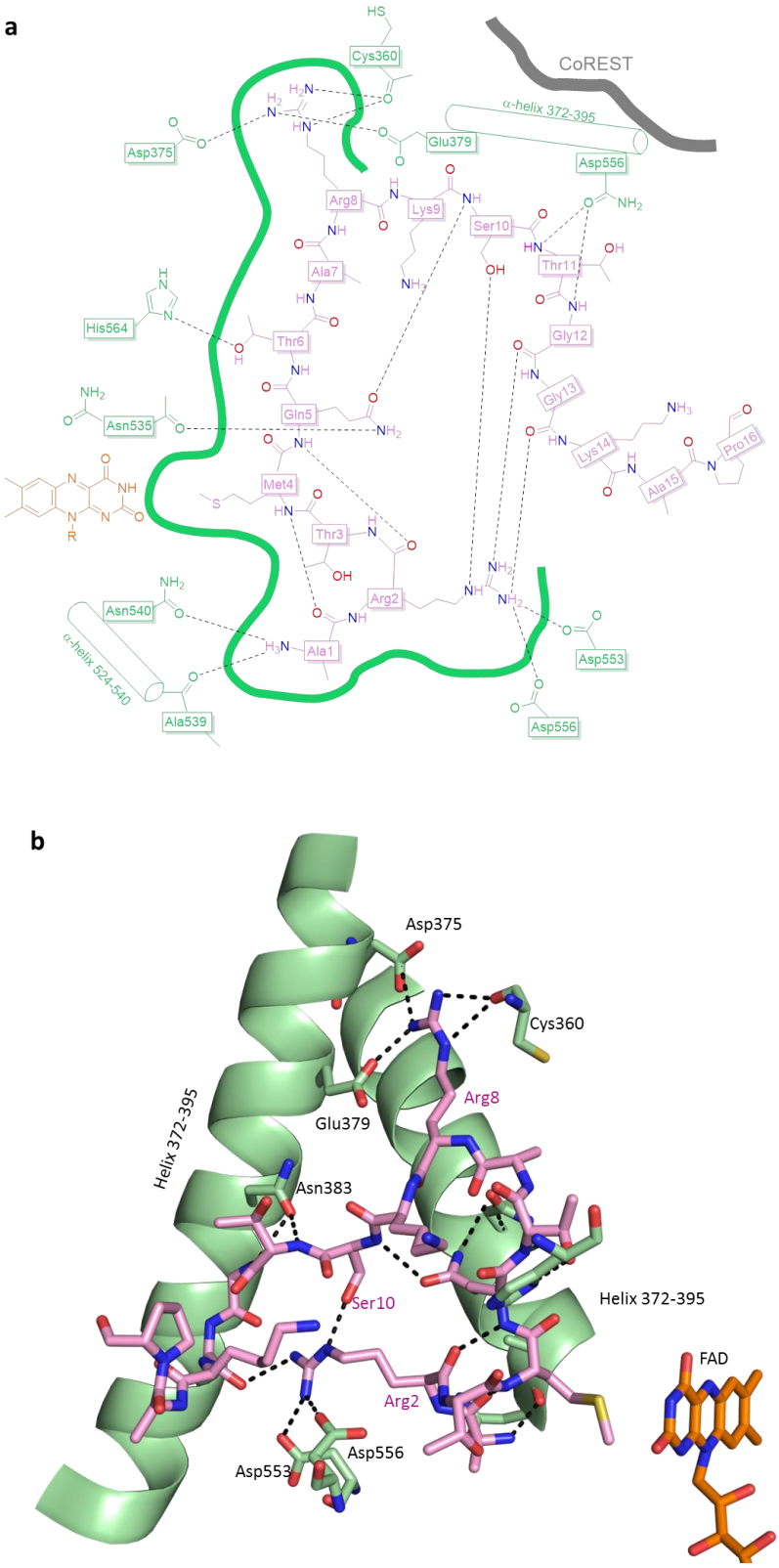


Figure 78: a) Schematic representation of interactions present between the H3(1-16) peptide and LSD1; and b) the three-dimentional structure showing how the peptide binds. These show the H3(1-16) peptide (pink) bound to LSD1 (green) through a number of different H-bonds (black dotted lines). The cofactor FAD (orange) points directly at Met4 of the H3 peptide, which was mutated from Lys4 to produce an inhibitor in order to obtain a crystal structure of the peptide bound. This mutation also gave a 30-fold increase in the binding of the H3 peptide to LSD1. The CoREST surface (black) is also depicted. (Adapted from Forneris *et. Al.*, pdb: 2V1D)[236]

The crystal structure of the H3 peptide bound to LSD1 shows only 16 residues, while the peptide used in the crystal trials were 21 residues in length. Residues 17-21 were found to be disordered and the flexibility was thought to allow the tail to project from the nucleosomal particle into the binding pocket.[236] In fact, previous studies had shown that H3 tail peptides of 16 residues had a barely detectable LSD1 activity, and peptides of even shorter length induced no LSD1 activity.[234] However, longer length peptides, 21 and 30 residues, have been shown to be good substrates for LSD1.[234] This suggests that the length of the histone H3 substrate is important in the demethylase activity of LSD1.

Additionally, the crystal structure of LSD1 bound to the H3 peptide and the cofactor FAD, shows that LSD1 binds CoREST through its tower domain (Figure 79).[236] Thus, LSD1 is also thought to help stabilise the CoREST-HDAC1/2 complexes.[234]

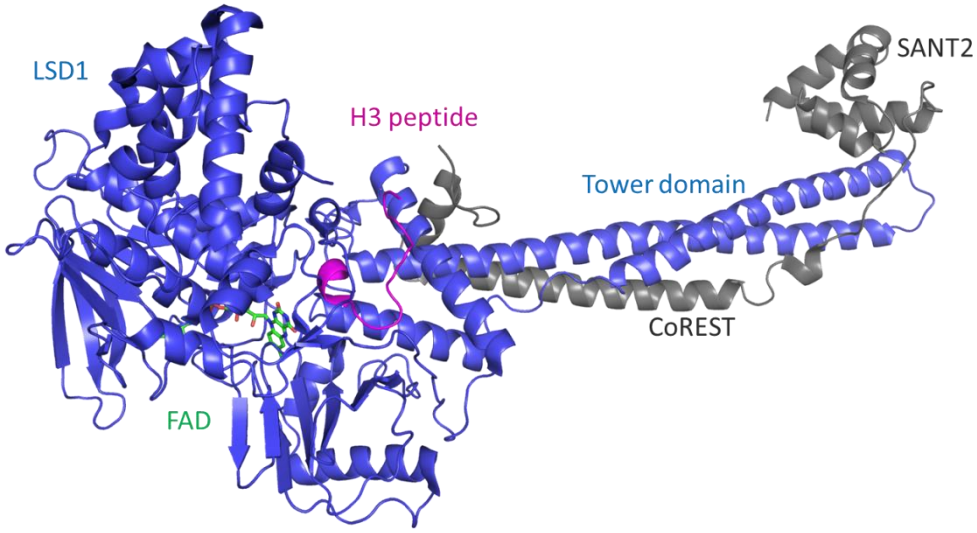


Figure 79: Crystal structure of LSD1 (blue) bound to an H3 tail peptide (magenta) and a molecule of FAD (green sticks). CoREST (grey), and its SANT2 domain at the *C*-terminal, is also bound through the LSD1 tower domain (pdb: 2V1D).[236]

4.2. Aims of this chapter

The aims of this chapter are:

- To synthesise a series of histone H3 tail (1-21) sequences (Figure 80) including mono- or di-methylation at Lys4 and on some of the analogues acetylation at Lys9 or Lys14.
- To dissect the mechanism of action of the multidomain complex of LSD1-CoREST-HDAC1 provided by Yun Song.

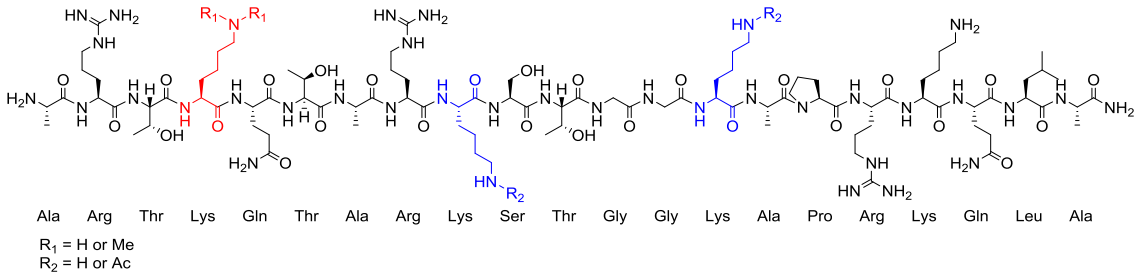
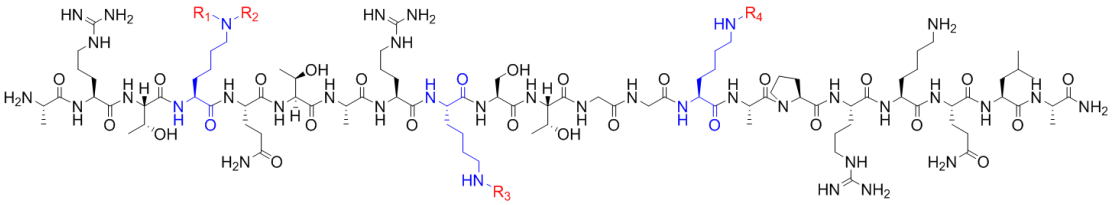


Figure 80: Structure of the histone 3 (1-21) tail sequence used to synthesise the analogues containing the mono- or dimethylated Lys4 and acetylated Lys9 or Lys14.

4.3. Synthesis of H3 (1-21) sequences

Five histone H3 (1-21) tail sequences were synthesised in order to determine the activity of the HDAC1-CoREST-LSD1 complex. These peptides were H3K4me1, H3K4me2, H3K4Me1K9ac, H3K4me2K9ac and H3K4me2K14ac (Figure 81).



		R_1	R_2	R_3	R_4
79	H3K4me1	Me	H	H	H
80	H3K4me2	Me	Me	H	H
81	H3K4me1K9ac	Me	H	Ac	H
82	H3K4me2K9ac	Me	Me	Ac	H
83	H3K4me2K14ac	Me	Me	H	Ac

Figure 81: Structures of the five peptide sequences synthesised based on the histone 3 (1-21) tail sequence.

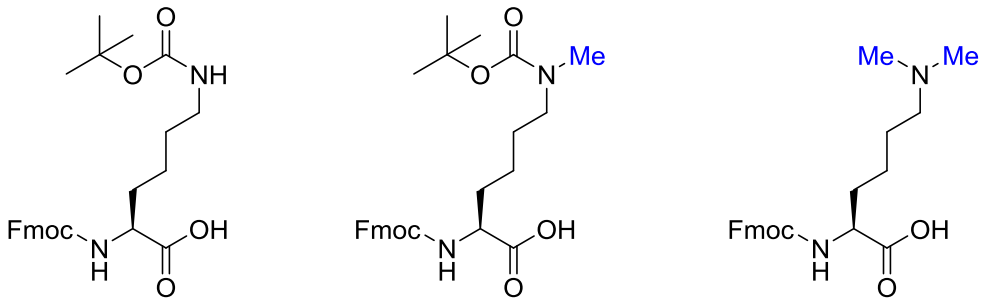
Acetylation of Lys9 and Lys14 is postulated to reduce the LSD1 activity on H3K4me1K9ac, HeK4me2K9ac and H3K4me2K14ac until the acetyl groups are removed by HDAC1.[234]

4.3.1. Synthesis of H3 (1-21) peptide analogues using standard microwave-assisted SPPS

Initial synthesis of the histone tail analogues used a standard microwave peptide synthesis strategy performed using a 4-fold excess Fmoc-AA-OH with 1:1:2 AA/HCTU/DIPEA for 5 minutes at 75 °C for all residues except Fmoc-Arg(Pbf)-OH. Fmoc-Arg(Pbf)-OH was coupled for 45 minutes at room temperature, followed by 5 minutes at 75 °C. Deprotection was performed using standard conditions of 30 seconds followed by 3 minutes at 75 °C using 20% piperidine with 0.1 M OxymaPure in DMF. However, using these conventional methods, none of the desired histone H3 peptides were found by LCMS. Thus, optimisation of the synthesis was required.

4.3.2. Room temperature couplings and deprotections of mono- and dimethylated lysine residues

One of the most prevalent impurities found during the synthesis of the H3 (1-21) sequences involved a molecular mass greater than the desired product. In fact, it was an extra 157 Da which equates to an additional arginine residue. This impurity was found in the H3 analogues containing a dimethylated lysine at position 4. The Fmoc-Lys(Me_2)-OH (Figure 82) residue used in the synthesis of the dimethylated H3 (1-21) analogues does not have any side chain protection, unlike for Fmoc-Lys(Boc)-OH or the monomethylated version, Fmoc-Lys(Me,Boc)-OH (Figure 82).



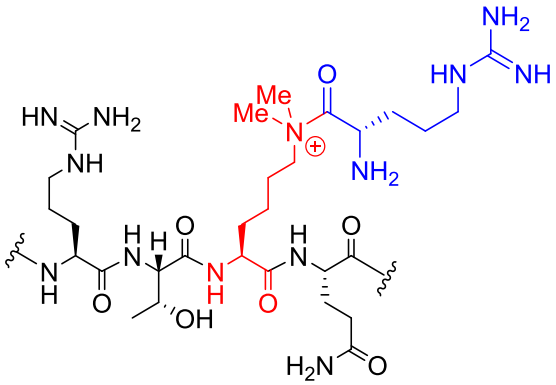
Fmoc-Lys(Boc)-OH **84**

Fmoc-Lys(Me,Boc)-OH **85**

Fmoc-Lys(Me_2)-OH **86**

Figure 82: Structures of the three different Fmoc-Lysine residues showing the side-chain Boc protection of both lysine and mono-methyl lysine residues, whereas the dimethyl lysine residue is not side-chain protected.

As a result of this knowledge, it was hypothesised that the extra arginine residue was caused by branching from the side chain of the dimethylated lysine residue (Figure 83). Thus, the protection of the lysine was important. However, since side chain protection of the Fmoc-Lys(Me_2)-OH is not possible another method was required.



87

Figure 83: Branching of an arginine residue (blue) from the dimethyl lysine residue (red) on the histone 3 tail peptide sequence.

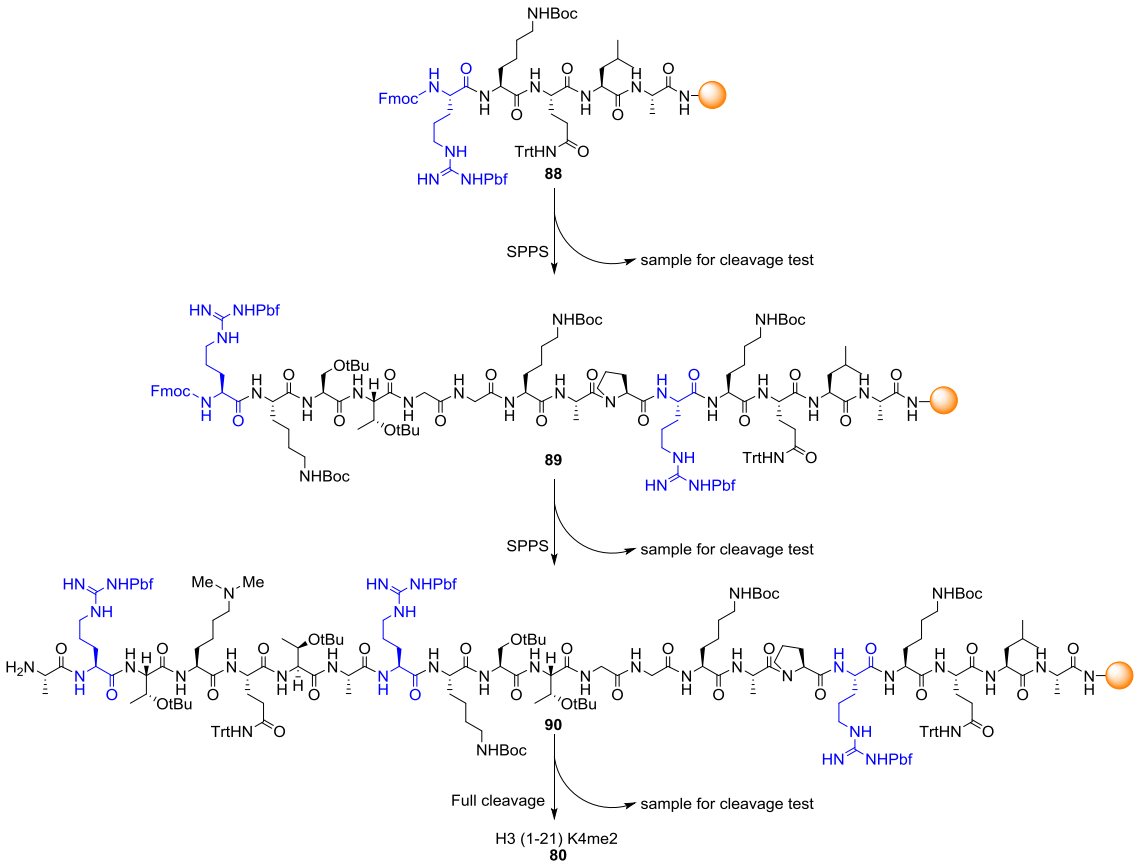
The side chain tertiary amino group of the dimethyl lysine residue is nucleophilic, thus allowing branching at this position. This is especially true when utilising elevated temperatures during the coupling reactions. The elevated temperature can encourage branching through overcoming the activation barrier of the reaction.

Consequently, the temperature of the reactions for both couplings and deprotections were investigated. Since the branching was only an issue after the addition of the dimethyl lysine residue at position 4 of H3, the peptide was synthesised using standard microwave-assisted SPPS until the dimethyl lysine residue. For the coupling of the dimethyl lysine residue to the rest of the peptide sequence, heating was omitted and the coupling was carried out at room temperature for an extended time period of 60 mins. For the Fmoc-deprotection reactions, 0.1M OxymaPure was also added to the 20% piperidine in DMF deprotection mix. This was done to make the deprotection solution slightly more acidic in order to protonate the side chain amine on dimethylated lysine residue and thus unavailable to react. The Fmoc-deprotection of the dimethyl lysine residue was also carried out at room temperature with two incubations at 15 minutes each. The subsequent couplings and deprotections of the remaining residues were also performed at room temperature. The result of this optimisation was the removal of any branching at the side chain of the dimethyl lysine residue.

4.3.3. Synthesis of H3 (1-21) peptide analogues using pseudoproline dipeptides

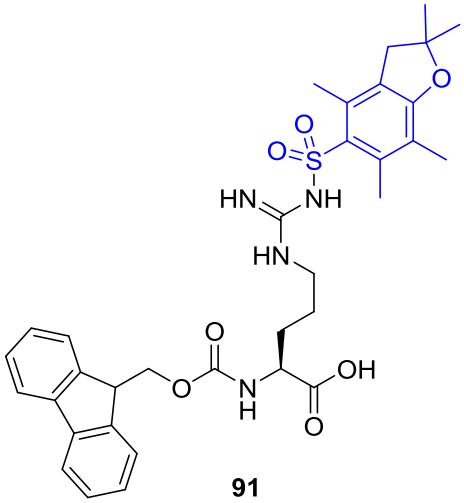
In addition to branching, other impurities were also found in the initial peptide synthesis. Many of the compounds seen by LCMS showed truncated peptide sequences. This could be due to a number of different reasons. Firstly, deletion sequence peptides could be as a result of incomplete couplings of different Fmoc amino acid residues. Secondly, truncations could occur as a result of the free amino group on the *N*-terminus of the peptide being unavailable for coupling.

The most substantial deletion sequence involved an arginine deletion. However, in the H3 (1-21) peptide sequence, there are three different arginine residues Arg2, Arg8 and Arg17and the deletion could have been at any of these residues. Consequently, each arginine residue was investigated by synthesising the peptide in a stepwise manner (Scheme 41). After the incorporation of each arginine residue incorporated into the sequence, a sample for a cleavage test was taken to determine whether the coupling was successful. The synthesis was then restarted until all cleavage tests had been taken. For the last arginine residue, an Fmoc deprotection and coupling of the final alanine in the sequence was also carried out, followed by a final Fmoc deprotection before a final cleavage test was taken.



Scheme 41: The stepwise synthesis of H3 (1-21) K4me2, while taking cleavage tests after each of the arginine couplings to determine which arginine residue does not couple correctly to leave an arginine deletion sequence.

For the first two arginine residues incorporated into the H3 sequence, no evidence of an arginine deletion was seen from the cleavage tests. However, upon the completion of the peptide synthesis, an arginine deletion was observed. This led to the realisation that it was the final arginine residue, Arg2, which was the cause of the deletion sequence problem. Fmoc-Arg(Pbf)-OH (Figure 84) is the largest residue when compared to all other 19 native Fmoc amino acid residues. As a result of this, it can be a challenge to couple arginine residues onto a lengthening peptide chain. One way to overcome this is to heat the coupling reaction. However, since heating caused branching of Lys4, this was not an option for the H3 peptide.



91

Figure 84: Structure of Fmoc-Arg(Pbf)-OH showing the large 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl (Pbf) protecting group (blue).

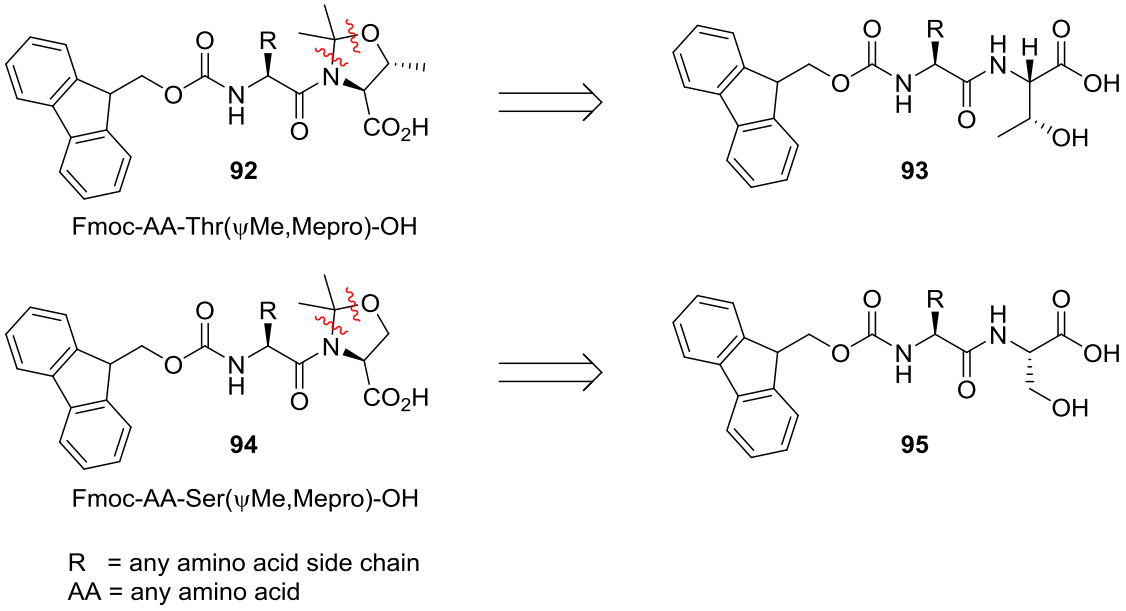
Another reason for incomplete couplings of the arginine residue may be because the free amine at the *N*-terminus of the peptide on resin is not actually available to react with the acid of the Fmoc-Arg(Pbf)-OH residue. This could be because some form of hydrophobic collapse on the resin and peptide aggregation has occurred, shielding the reactive *N*-terminus amine away from the reactive activated acid group. Long peptide sequences in particular are prone to hydrophobic collapse and aggregation, as well as forming on resin secondary structures. As a result, it is generally much harder to synthesise longer peptide sequences.[237]

To overcome this problem of hydrophobic collapse, and so also overcome the problem of the arginine deletion sequence, another optimisation technique was utilised. This optimisation method used pseudoproline dipeptide residues.

Pseudoproline dipeptide residues have been used to improve the efficiency of the synthesis of long peptide seuences.[238,239] These pseudoproline residues break secondary structures that form on resin and allows for much cleaner and predictable peptide couplings.

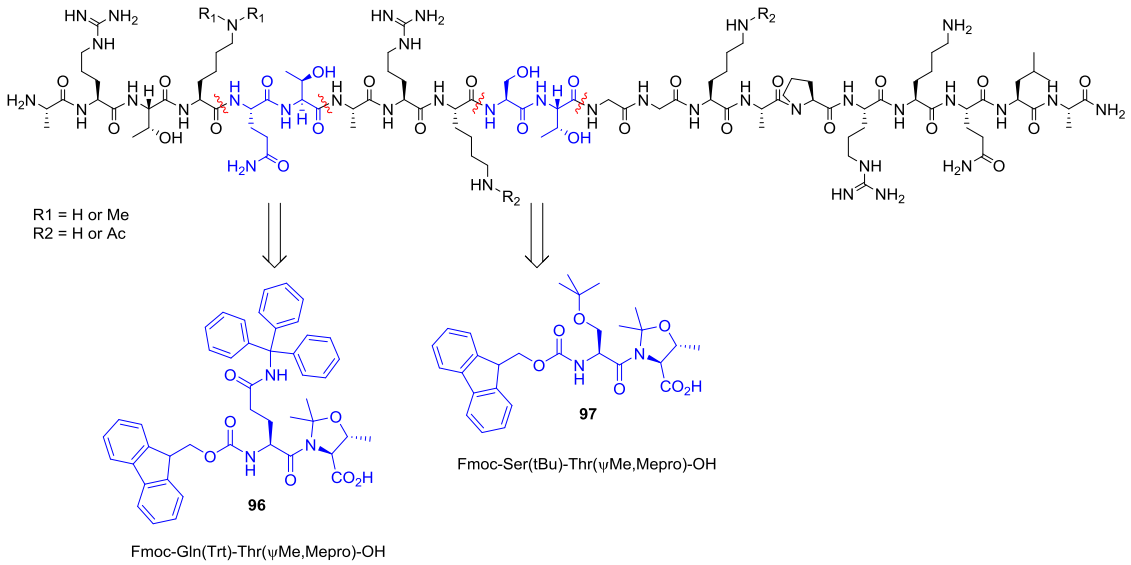
For pseudoproline dipeptide residues it is important for one of the residues to be either a serine or a threonine at the *C*-terminus of the dipeptide. Thus, not all peptides can include the incorporation of pseudoproline dipeptide residues. The

threonine and serine are required to help form the proline-like structure (Scheme 42), which is known to prevent on resin secondary structure formation by forming kinks in the peptide sequence.



Scheme 42: Retrosynthetic analysis of pseudoproline dipeptide residues containing threonine or serine, showing how the proline-like structure is achieved.

Incorporation of the pseudoproline dipeptide residues is straightforward. The pseudoproline dipeptide residues are synthesised so that they have an Fmoc-protected *N*-terminus and a free *C*-terminus acid ready for coupling using an Fmoc-based peptide synthesis. To incorporate more than one pseudoproline dipeptide into a longer sequence, the minimum separation between two pseudoproline residues is two residues. However, a separation of five to six amino acid residues is optimal for two pseudoprolines residues.[240,241] In the H3 (1-21) tail peptide sequence, there are four residues that can be substituted for two pseudoproline dipeptide residues (Scheme 43). These are Gln5 and Thr6 as well as Ser10 and Thr11 which gives a separation of five residues between the two ψPro residues.



Scheme 43: Retrosynthetic analysis starting at an H3 (1-21) tail sequence showing the starting Fmoc-protected pseudoproline dipeptide residues. For Gln5 and Thr6, Fmoc-Gln(Trt)-Thr(ψMe,Mepro)-OH is utilised, while for Ser11 and Thr12, Fmoc-Ser(tBu)-Thr(ψMe,Mepro)-OH is used.

The pseudoproline dipeptide residues only affect the peptide only whilst it is still on the resin. Upon cleavage of the peptide from the resin and total global deprotection of the peptide, the pseudoproline protection is also cleaved to give the original dipeptide residues.

The addition of pseudoproline dipeptide residues into the H3 (1-21) tail peptide sequences did in fact help in successfully synthesising the full H3 (1-21) sequence. However, the resulting isolated peptides were very low yielding and because these peptides were for KDM and HDAC enzyme assays, a much larger quantity was required.

4.3.4. Use of Rink Amide SpheriTide resin

Originally, the H3 (1-21) tail analogue peptides were synthesised on a PEG grafted resin, Rink Amide ChemMatrix resin (Biotage), with a substitution of 0.45 mmol/g. Thus, when trying to synthesise larger quantities of peptides the amount of resin used can become a hindrance. This is a particular problem since the amount of solvent required to swell the resin, is more than the volume of the reaction vessel.

As a consequence, the resin used in the synthesis was also investigated. A *C*-terminal amide was required for the H3 (1-21) peptide sequence because residue 21 is not the true *C*-terminus and thus needs to be uncharged. Therefore Rink Amide SpheriTide resin (CEM) was used instead.

The SpheriTide resin is a much higher loading resin (1.05 mmol/g substitution). Intuitively, a higher loading resin would seem to encourage on resin peptide aggregation. However, due to the poly lysine crosslinked polymer structure of the SpheriTide resin, higher loading is possible without the associated aggregation. Using high loading resin allows a smaller amount of resin to be used for the same amount of peptide when compared to the Rink Amide ChemMatrix resin. The peptides that were isolated and purified from Rink Amide SpheriTide resin had higher crude purities and greater isolated yields than those from the Rink Amide ChemMatrix resin.

4.4. NMR concentration determination of H3 (1-21) peptide analogues

Following the synthesis, purification and characterisation of all the synthesised H3 (1-21) tail analogue peptides, the peptide concentration also needed to be determined. However, in all the H3 (1-21) tail peptide sequences, there are no aromatic residues that can be used to determine UV concentrations.

Determination of the peptide concentration is also difficult due to the formation of TFA salts of the peptides. TFA salts add a significant molecular mass to the synthesised peptides and so the mass of peptide could be significantly different to the actual amount of peptide in the sample.

A new NMR based methodology has been developed to calculate the concentration and peptide content of non-chromophore containing peptides by Yana Rennie (Department of Chemistry, University of Leicester). In this new method, *p*-nitrophenol (*p*NP) is used as an NMR standard. This is possible because of the two distinct doublet signals for the aromatic protons of *p*NP. Since the peptides have no aromatic residues, there are therefore no other aromatic signals that should be seen in the same region of the NMR spectrum.

This new methodology also requires a distinct signal in the peptide NMR, such as that produced by an *N*-terminal acetyl group. It is from this signal that the concentrations can be calculated using a known standard. In the H3 (1-21) peptide tail sequences, there is no need for *N*-terminal acetylation. Nevertheless, there are other groups on the peptides that give a distinct signal in the NMR. The PTMs on the peptides all contain distinct methyl groups, more particularly in the mono- or dimethylated lysine residues and in the acetylated lysine residues. By NMR analysis, the most distinct of these peaks are those from the methyl groups on the mono- or dimethylated lysine residues.

To determine the peptide content of each of the H3 (1-21) peptides, first an accurate stock concentration of *p*NP was assessed by UV-Vis spectroscopy. This was achieved by taking 20 µL of the *p*NP stock solution in 980 µL H_2O. The absorbance was taken and the accurate concentration calculated using a modified Beer-Lambert equation (Equation 3), where c is the concentration of the stock solution, A_{320} is the absorbance at 320 nm, d is the dilution factor, ε is the extinction coefficient for *p*NP (10.172 $M^{-1}cm^{-1}$), and l is the path length.

$$pNP_{conc}^{stock} = \frac{A_{320}d}{\varepsilon l} \qquad \textbf{Equation 3}$$

Next, a weighed amount of lyophilised peptide was dissolved in D_2O to give an approximate peptide stock concentration of 5 mg/mL. This relatively high stock concentration is required in order to observe the peptide peaks in the NMR spectrum relative to the *p*NP peaks. 200 µL of the peptide stock was then mixed with 300 µL of *p*NP stock to give a final NMR volume of 500 µL in the NMR tube. The concentration of *p*NP in the NMR tube (pNP_{conc}^{NMR}) was then calculated using Equation 4, where pNP_{conc}^{stock} is the stock concentration of *p*NP, v_{pNP} is the volume added to the NMR tube of the stock and v_{tot} is the total volume of solvent in the NMR tube.

$$pNP_{conc}^{NMR} = \frac{pNP_{conc}^{stock}v_{pNP}}{v_{tot}} \qquad \textbf{Equation 4}$$

The spectra of the H3 (1-21) peptide tails were analysed by comparing the integrations of the most downfield *p*NP doublet, with the methyl singlet of the peptide taking into account the number of protons each peak represents. By setting the *p*NP doublet integration as 1 for reference, the integration for the peptide singlet can then be assessed and the ratio between the *p*NP protons and the protons in the singlet peak of the peptide (H_{ratio}) can then be calculated with Equation 5. In Equation 5, P_{int} is the integration of the singlet NMR peak for the peptide, pNP_H is the number of protons in the *p*NP doublet peak (ie. 2) and P_H is the number of protons in the singlet peptide peak (ie. 3). From Equation 6, the concentration of the peptide in the NMR tube (P_{conc}^{NMR}) can then be calculated, where H_{ratio} is taken from Equation 5 and pNP_{conc}^{stock} is the concentration of the *p*NP stock solution.

$$H_{ratio} = \frac{P_{int} \,.\, pNP_H}{P_H} \qquad \textbf{Equation 5}$$

$$P_{conc}^{NMR} = H_{ratio} \,.\, pNP_{conc}^{stock} \qquad \textbf{Equation 6}$$

Finally, the concentration of the peptide stock solution (P_{conc}^{stock}), was calculated using Equation 7 where P_{conc}^{NMR} is taken from Equation 6, v_{tot} is the total volume of solvent in the NMR tube and v_P is the volume of peptide stock in the NMR tube.

$$P_{conc}^{stock} = \frac{P_{conc}^{NMR} \,.\, v_{tot}}{v_P} \qquad \textbf{Equation 7}$$

To obtain the percentage peptide content, the calculated concentration of the peptide stock, based on a weighed amount of peptide, is used along with the experimental value for the peptide concentration, as in Equation 8.

$$\%\, peptide\ content = \frac{Calculated\ conc}{Experimental\ conc} \text{ x } 100 \qquad \textbf{Equation 8}$$

The NMR spectra obtained for H3K4me1 and H3K4me2 clearly show the presence of the two *p*NP peaks in the aromatic region used as the reference. Additionally, the appearance of singlet peaks at 2.72 ppm and 2.89 ppm for the H3K4me1 and H3K4me2 respectively can also be easily identified (Figure 85).

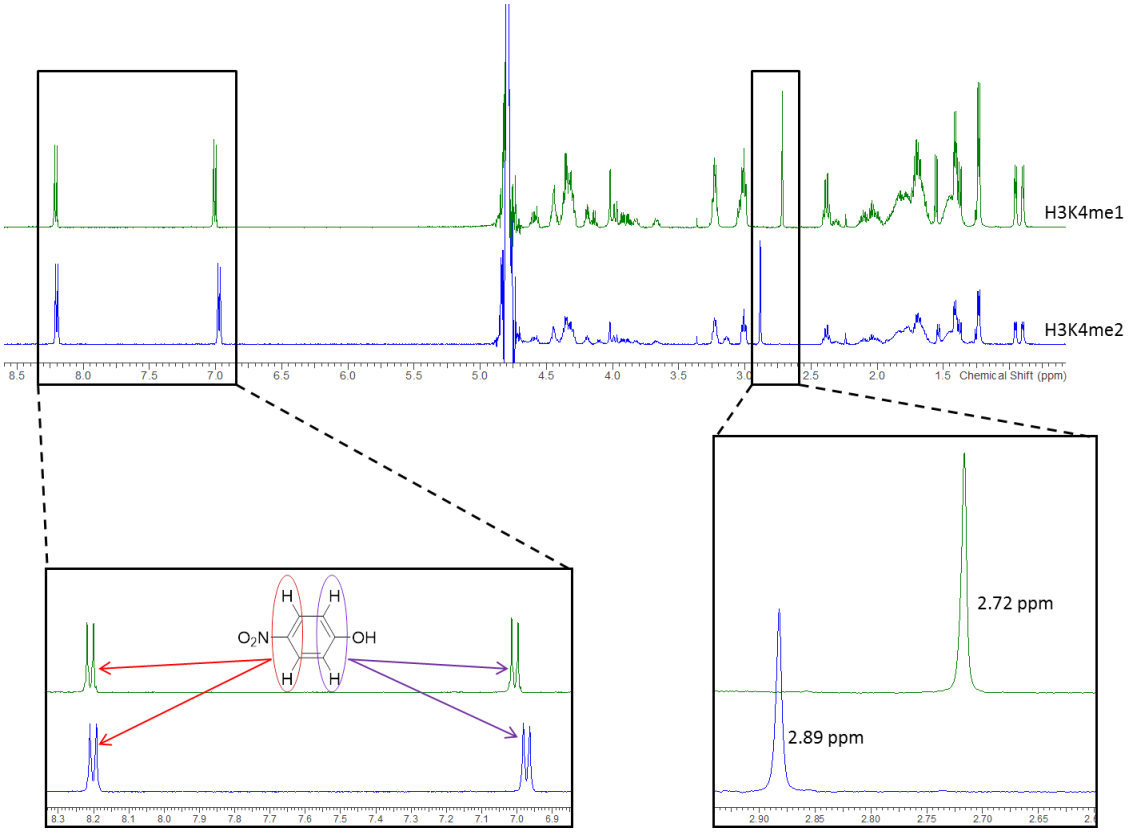


Figure 85: NMR spectra of H3K4me1 (green) and H3K4me2 (blue), and showing an expanded aromatic region where the protons of *p*-nitrophenol are present (bottom left) and an expanded region where the distinct methyl peaks from the mono- and dimethylated lysine residues (bottom right). The monomethylated lysine has a distinct peak at 2.72 ppm, while the dimethylated lysine peak lies at 2.89 ppm.

The peptides that were synthesised also contain an acetylated lysine at either Lys9 or Lys14 and the singlet peak indicative of the acetyl group at 1.98 ppm can also been distinctly seen on the NMR spectrum (Figure 86).

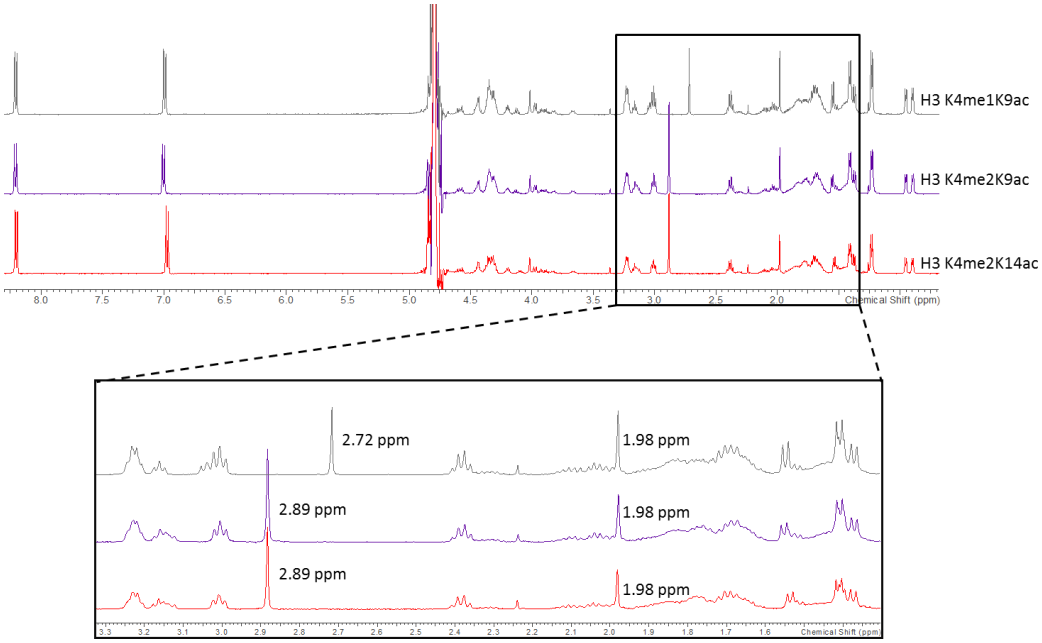


Figure 86: NMR spectra of H3K4me1K9ac (grey), H3K4me2K9ac (purple) and H3K4me2K14ac (red). The expansion shows the region with the distinct singlet methyl peaks for the mono- (2.72 ppm) and dimethylated (2.89 ppm) lysine residues and the singlet for the acetyl group (1.98 ppm).

As can been seen in the NMR spectra above, changing the methylation state of the lysine residue from mono- to dimethylated also changes the position of the singlet methyl peak. However, changing the residue position on which the acetyl group sits does not alter the position of the singlet acetyl peak in the NMR spectrum.

From the NMR concentrations it was deduced that the synthesised H3(1-21) tail peptide analogues had a % peptide content of between 81-93% (Table 2). The large difference of up to about 20% in peptide content based on concentration shows the importance of calculating peptide content. This is especially true when accurate concentrations of peptides are required for biological assays.

Table 2: Peptide content determination for the five H3 (1-21) histone tail peptide analogues calculated using Equation 6.

	Calculated conc (mM)	Experimental conc (mM)	% peptide content
H3K4me1	5.82	5.39	93
H3K4me2	2.19	1.77	81
H3K4me1K9ac	3.42	2.86	84
H3K4me2K9ac	2.99	2.61	87
H3K4me2K14ac	2.15	1.81	84

The rest of the weight of the peptide (up to 20%) can be attributed to TFA salts and water coordinating to the peptide. TFA salts in particular have great affinity for the peptides due to their high positive charge and coordinate to the peptide through both the cleavage and purification conditions used.

4.5. NMR substrate studies on H3 (1-21) peptide analogues

The experiments with the H3 (1-21) peptides were performed in collaboration with Yun Song (Department of Molecular and Cell Biology, University of Leicester), Micha Kunze (Institute of Structural & Molecular Biology, University College London), Lisbeth Andersen (Institute of Structural & Molecular Biology, University College London) and Flemming Hansen (Institute of Structural & Molecular Biology, University College London).

4.5.1. Arginine deletion peptides

For the initial NMR substrate studies the H3 (1-21) tail peptide analogues that were used contained a minor impurity of an arginine deletion. Initially, it was thought that this would not be a problem for the NMR studies.

The H3 (1-21) tail peptide analogues were added to the ternary HDAC1-CoREST-LSD1 complex to determine the relative rates of demethylation and deacetylation of the peptides. Taking H3K4me2K9ac as the initial substrate peptide, the first NMR spectrum was obtained after 2 mins and shows a peak

for the dimethyl lysine at 2.76 ppm and a peak for the acetylated lysine at 1.87 ppm (Figure 87).

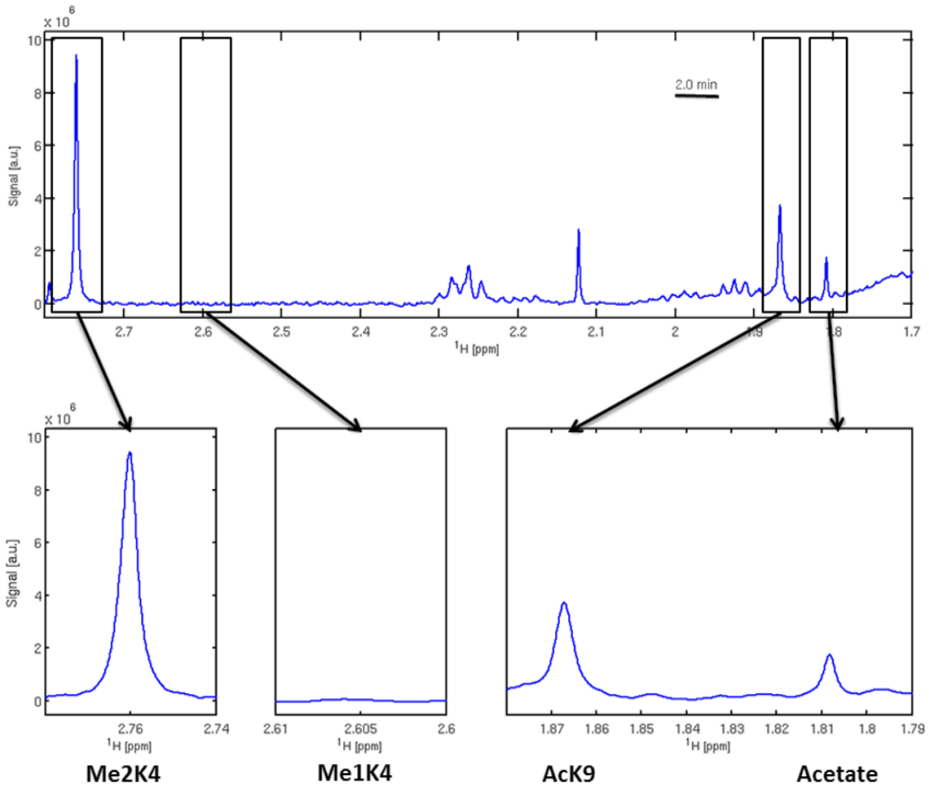


Figure 87: NMR spectrum of the H3K4me2K9ac peptide at 2 mins which shows a dimethyl lysine singlet at about 2.76 ppm, an acetylated lysine singlet at about 1.87 ppm and an acetate singlet at about 1.81 ppm.

Subsequent spectra of the reaction were taken to monitor the demethylation and deacetylation. After 30 mins, the acetylated lysine residue at position 9 had been completely turned over which is shown by the disappearance of the singlet methyl peak at about 1.87 ppm (Figure 88). Additionally, the appearance and increase in size of singlet peak at around 1.81 ppm signifies the production of the deacetylation product, which is an acetate molecule (Figure 88).

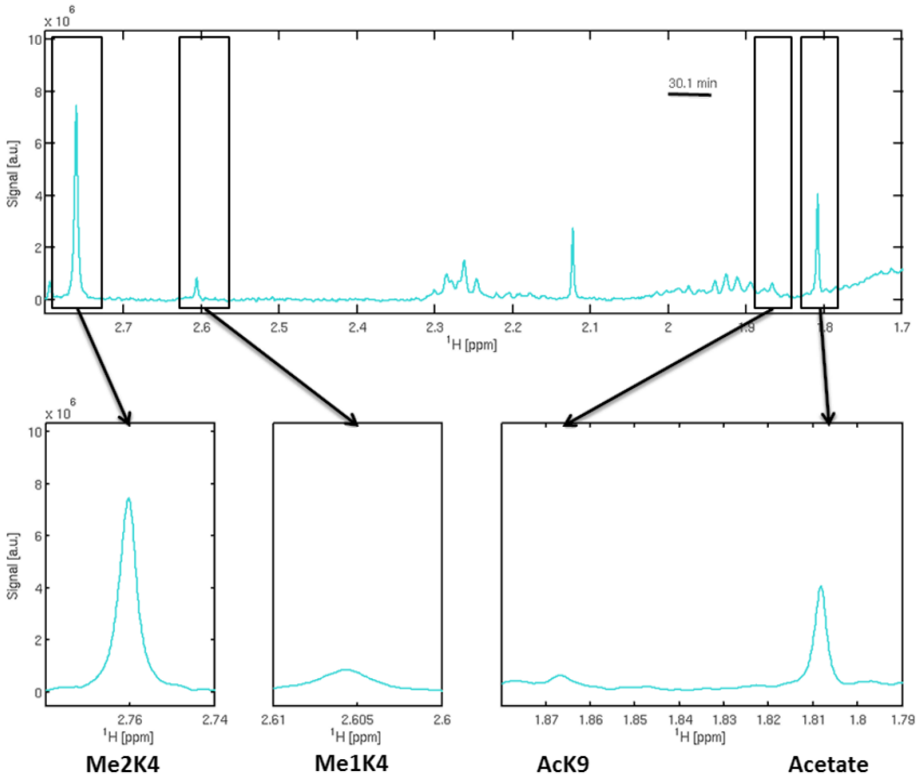


Figure 88: NMR spectrum of the H3K4me2K9ac peptide after 30 mins showing the disappearance of the acetyl group and the appearance and increase in the acetate peak at about 1.81 ppm. The dimethyl peak at about 2.76 ppm has also started to decrease in size while a monomethyl singlet peak appears at about 2.61 ppm.

Thus, from this data, it can be observed that the deacetylase activity of the HDAC1:CoREST:LSD1 complex is much faster than its demethylase activity. This is illustrated by the dimethyl lysine peak slowly decreasing in size and while the appearance of the monomethylated lysine peak at 2.61 ppm is also slow (Figure 88). The appearance of the monomethylated lysine peak indicates the demethylation of one of the methyl groups on the dimethyl lysine.

At about 150 mins, the monomethyl peak reaches its highest point, before it then also starts decreasing in size (Figure 89a) and finally disappears completely at about 600 mins (Figure 89b). However, with the dimethylated lysine, its associated singlet peak never decreases more than two thirds of its original size, even after nearly 950 mins.

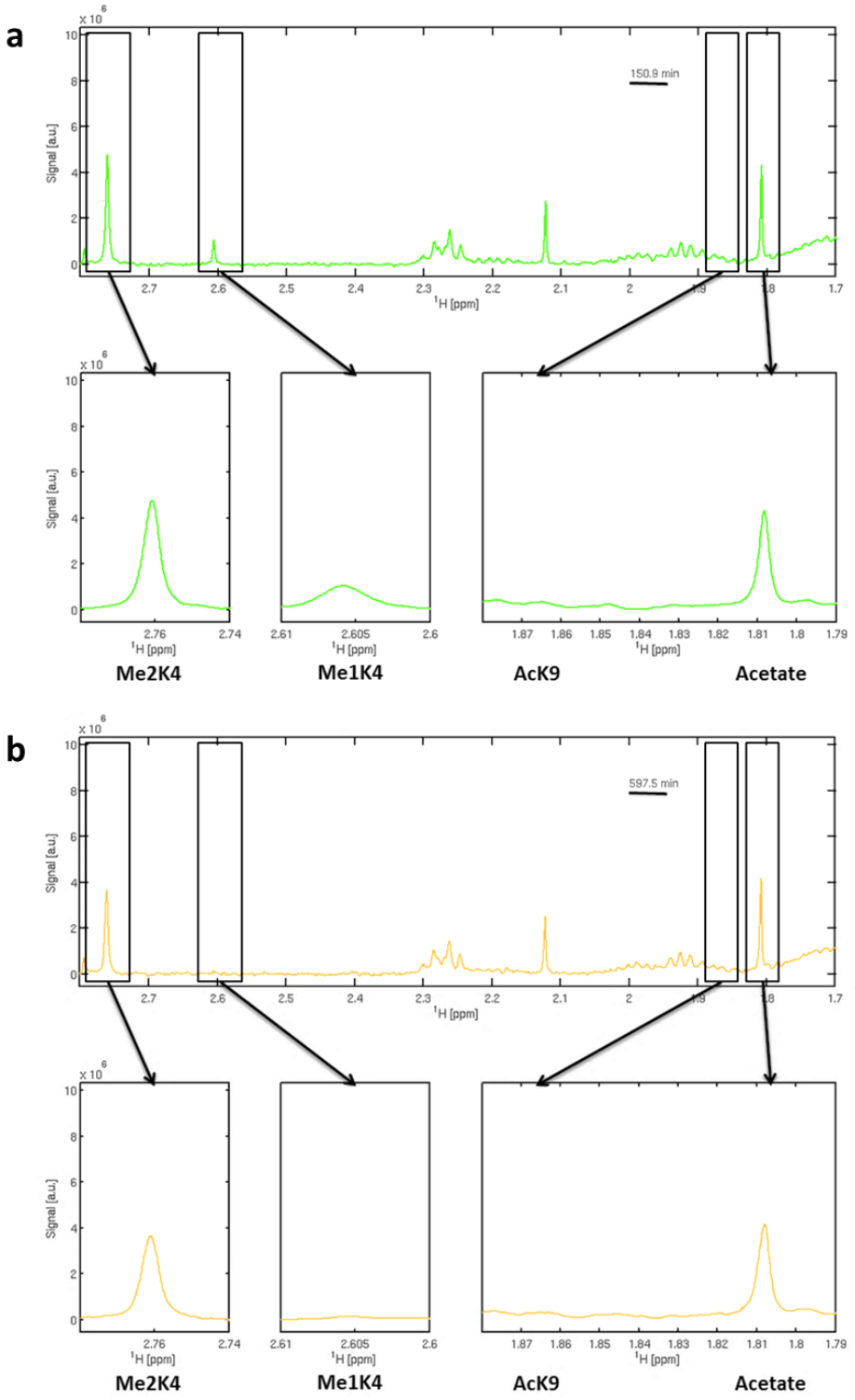


Figure 89: NMR spectra of H3K4me2K9ac peptide after (a) 150 mins, (b) 600 mins and (c) 950 mins. Between them, these show the disappearance of the monomethylated lysine peak and the incomplete disappearance of the dimethylated lysine peak.

As a result of incomplete disappearance of the dimethylated lysine peak, it was hypothesised that the product of the demethylation inhibits the LSD1 enzyme. However, it was determined that product inhibition was not the case for

incomplete demethylation. This was confirmed through varying substrate concentration, which still gave incomplete demethylation. This indicated that demethylation is not dependent on substrate concentration and thus, incomplete demethylation is not due to product inhibition.

It was therefore hypothesised that the arginine deletion at position 3 in the initial NMR substrate studies was actually a vital residue in the recognition of the substrate peptide to be demethylated. Thus, the peptides were resynthesised and the synthesis optimised to remove the arginine deletion.

4.5.2. Optimised H3 (1-21) peptide analogues

The newly synthesised H3 (1-21) peptide analogues were found not to contain arginine deletions. However, although every attempt was made to remove branching of residues, by MALDI-TOF analysis it was found that H3K4me2K9ac and H3K4me2K14ac did still contain some arginine branching (Figure 90).This data indicates further optimisation is still required in order to synthesise purer histone tail peptides. Even with branching present, NMR assays were carried out once again, this time on the newly synthesised H3 (1-21) peptide analogues.

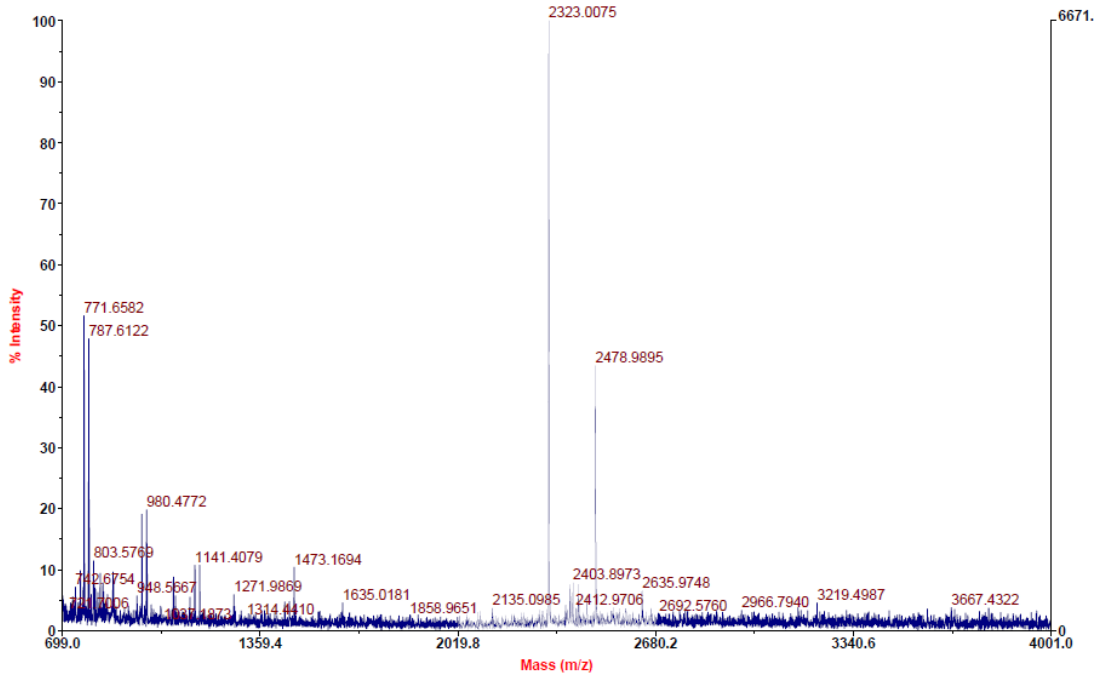


Figure 90: MALDI-TOF spectrum of H3K4me1K9ac showing the molecular $[M+H]^+$ 2323.0075 and the arginine branched peptide $[M+Arg+H]^+$ 2478.9895.

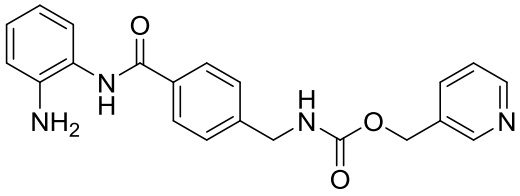
In this section, the focus will be on the monomethylated H3 (1-21) peptide analogues. Demethylation of H3K4me1 was studied and showed an initial rate of 8.40 nM/sec (Table 3, Entry 1). However, the demethylation did not reach completion. In fact only about two thirds of the peptide was demethylated. For H3K4me1K9ac, the initial demethylation rate was 7.42 nM/sec (Table 3, Entry 2). This result agrees with previous studies which suggest slower demethylation rates when Lys9 is acetylated. The deacetylation rate of H3K4me1K9ac is markedly fast at 377.74 nM/sec (Table 3, Entry 2). This shows that the demethylation activity of LSD1 is inherently slower than the deacetylase activity of HDAC1.

Table 3: Initial rates of demethylation and deacetylation of histon H3 (1-21) petide analogues.

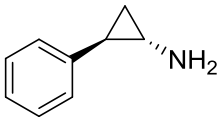
	Peptide	Inhibitor	V_0 demethylation (nM/sec)	V_0 deacetylation (nM/sec)
1	H3K4me1		8.4	n/a
2	H3K4me1K9ac		7.4	337.7
3	H3K4me1	**98**	9.5	n/a
4	H3K4me1K9ac	**98**	8.3	13.8
5	H3K4me1K9ac	**99**	1.6	427.8
6	H3K4me1K9ac + IP_6		12.0*	984.7*
7	H3K4me1K9ac + IP_6	**98**	12.1*	29.3*

* fitted to straight line (f(x)=ax+b) for the first 30 minutes.

When an HDAC inhibitor, entinostat **98** (Figure 91), was added to the experiment, the initial rate of demethylation of the H3K4me1 increased slightly to 9.5 nM/sec (Table 3, Entry 3). This suggests that there is competition for the peptide substrate from both LSD1 and HDAC1 in the CoREST complex. By inhibiting HDAC1, it cannot compete for the substrate so the rate of demethylation increases. A similar increase in the initial rate of demethylation was also observed for H3K4me1K9ac, which increased to 8.3 nM/sec (Table 3, Entry 4). Another effect of the inhibition of HDAC1 with **98** resulted in an expected reduction in the deacetylase activity of HDAC1 on H3K4me1K9ac to 13.8 nM/sec (Table 3, Entry 4).



entinostat
98



2-PCPA
99

Figure 91: Structures of the HDAC inhibitor entinostat **98** and LSD1 inhibitor trans-2-phenylcyclopropylamine, 2-PCPA **99**.

Inhibition of LSD1 in the CoREST complex was also studied using trans-2-phenylcyclopropylamine (2-PCPA, **99**) (Figure 91), a known demethylase inhibitor. The initial rate of demethylation of H3K4me1K9ac was observed to be 1.6 nM/sec (Table 3, Entry 5), which is greater than a 5-fold difference in the initial demethylation rate. On the other hand, the initial rate of deacetylation of H3K4me1K9ac increases to 427.8 nM/sec (Table 3, Entry 5) upon inhibition of the CoREST complex with **99**.

Inositol phosphates have been shown to aid in the activation of HDACs. In the NMR experiments on the CoREST complex, this phenomenon was also studied using Ins(1,2,3,4,5,6)P_6. With Ins(1,2,3,4,5,6)P_6, the initial rate of deacetylation of H3K4me1K9ac increases to 3-fold to 984.7 nM/sec (Table 3, Entry 6). The initial demethylation rate of LSD1 also increases slightly to 12.0 nM/sec (Table 3, Entry 6). This shows crosstalk between the HDAC1 and LSD1 in the CoREST complex when Ins(1,2,3,4,5,6,)P_6.

Recovery of a small amount of HDAC1 deacetylase activity was possible when Ins(1,2,3,4,5,6,)P_6 is added alongside **98**. This gave an initial deacetylation rate of 29.3 nM/sec (Table 3, Entry 7), a 2-fold increase in comparison to addition of **98** alone. Initial rate of demethylation also increases to 12.1 nM/sec (Table 3, Entry 7), when compared to addition of **98** alone. However, this initial rate is indistinguishable from the initial rate of demethylation when Ins(1,2,3,4,5,6)P_6 is added without any inhibitor (Table 3, Entry 6).

From this data, Ins(1,2,3,4,5,6)P_6 has been shown to increase the deacetylase activity of HDAC1 both in the absence and presence of an HDAC inhibitor. Additionally, through crosstalk between HDAC1 and LSD1 in the LSD1:CoREST:HDAC1 ternary complex, Ins(1,2,3,4,5,6)P_6 has moderately increased the demethylation activity of LSD1.

4.6. Synthesis of dual inhibitors of LSD1 and HDAC1 for crystal trials

The crystal structure of the LSD1-CoREST complex bound to a SNAG domain peptide of the transcription factor SNAIL1 was solved by Baron *et al*.[242] The

SNAIL1 peptide mimics the histone H3 peptide tail and binds to the LSD1 active site pocket. In fact, the interaction of the SNAIL1 peptide with LSD1 is much tighter than the H3 peptide and thus can act as an inhibitor of LSD1.[242]

From this crystal structure it is hypothesised that a set of peptides could be designed which would bind both the LSD1 and HDAC1 enzymes and act as a dual inhibitor. This might restrict the movement of LSD1 with respect to HDAC1 and increase the chances of obtaining a crystal structure of the ternary LSD1-CoREST-HDAC1 complex. The peptides were designed based on histone H3 and the distance between Lys4, Lys14 and Lys18. Additionally, the peptides incorporate the SNAIL1 peptide at the *N*-terminus. The design incorporated an inhibitor functionality for HDAC1 at the *C*-terminus based on the well-known inhibitor SAHA (Figure 92).

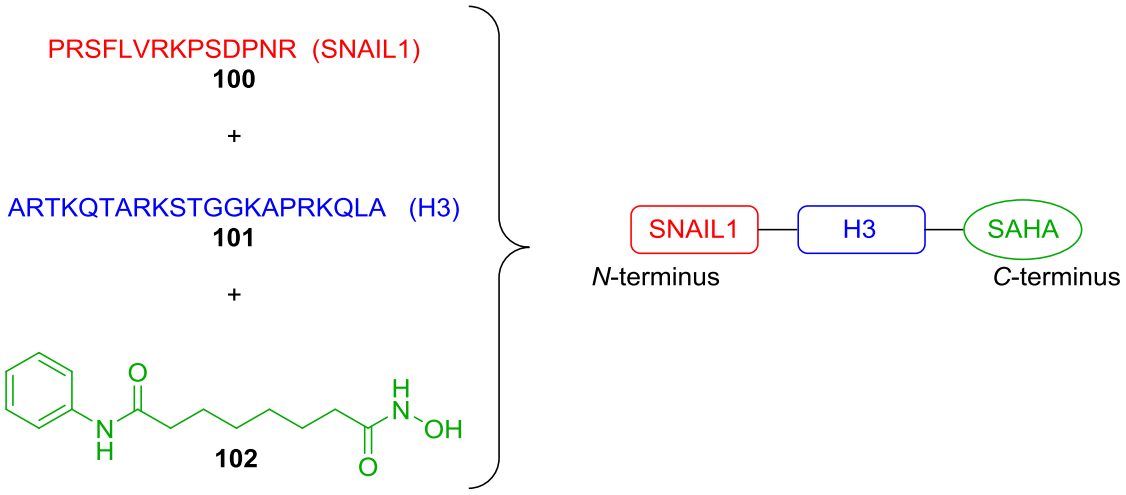


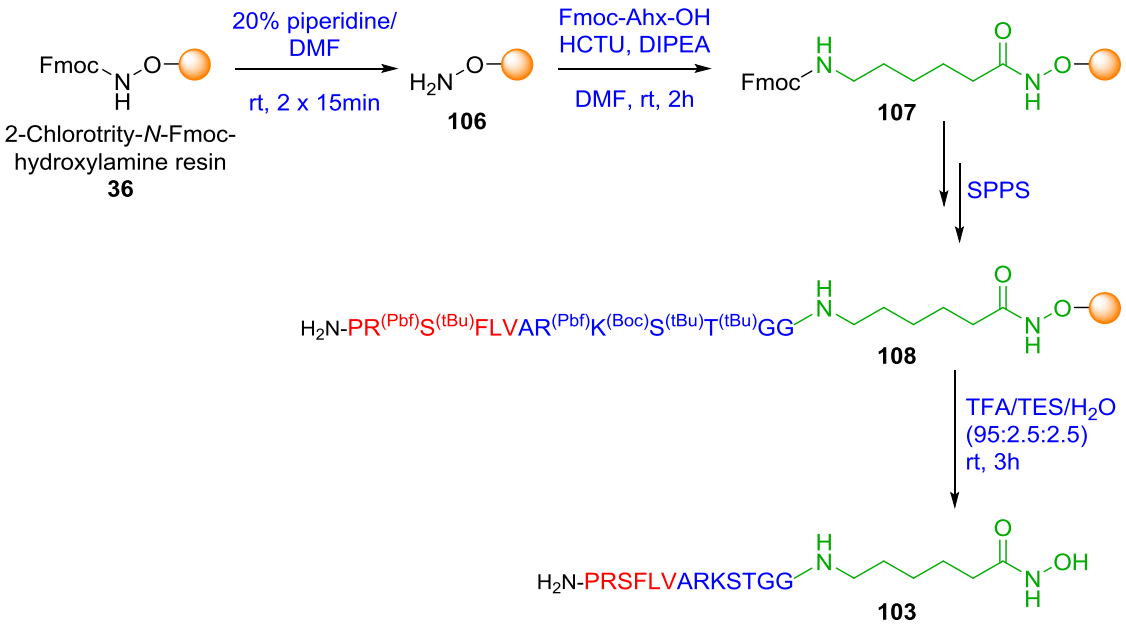
Figure 92: The design of a dual inhibitor of both HDAC1 and LSD1 based on the SNAIL1 (red) and histone 3 tail (H3, blue) peptides and the HDAC inhibitor SAHA (green).

The idea is to vary the lengths of the SNAIL1 and H3 peptides in order to obtain an inhibitor that binds to both the LSD1 and HDAC1 enzymes in the ternary complex simultaneously. Three different peptides are proposed as the tool compounds for this investigation (Figure 93). For the SAHA or HDAC inhibitor part of the peptide, the synthesis is designed to so that inhibitor is synthesised as part of the SPPS itself, based on SAHA, but not actually be SAHA.

H_2N-PRSFLVARKSTGG-inhibitor	(9aa)	**103**
H_2N-PRSFLVRKPSDPNR-inhibitor	(10aa)	**104**
H_2N-PRSFLVARKSTGGKAPR-inhibitor	(13aa)	**105**

Figure 93: The three peptide sequences based on the SNAIL1 (red) and H3 (blue) sequences where a SAHA-based inhibitor (green) at the *C*-terminus. The distance between the *N*-terminus SNAIL1 LSD1-binding residues and the inhibitor functionality (green) is based on SAHA.

In fact the synthesis of the full peptides was achieved with the use of 2-chlorotrityl-*N*-Fmoc-hydroxylamine resin **36** (Scheme 44), which was deprotected to give the free amine **106**. Fmoc-aminohexanoic acid (Fmoc-Ahx-OH) was then coupled to **106** to create **107** with a protected hydroxamic acid functionality at the *C*-terminus. Automated microwave-assisted SPPS was then used to synthesise the rest of the peptide leaving a free amine at the *N*-terminal to give **108.** The final step involved TFA cleavage to allow global deprotection of all protecting groups and removal of the peptide from the resin leaving a hydroxamic acid functionality at the *C*-terminus to give **103** (1% yield) following peptide precipitation and purification. Peptides **104** and **105** were prepared in the same manner giving 7% and 6% yields respectively. This work is still ongoing and with the peptides in hand, crystallisation experiments can be carried out.



Scheme 44: The synthesis of the SNAIL-H3 dual inhibitor peptides starting from 2-chlorotrityl-*N*-Fmoc-hydroxylamine resin to finally give a hydroxamic acid functionality at the *C*-terminus when the peptide is cleaved from the resin.

4.7. Conclusions and future work

The successful synthesis of five histone H3 (1-21) peptide tail analogues has been achieved following multiple optimisation methods. Arginine deletion sequences of the histone tail peptides have been eliminated to give purer peptides. However, some arginine branching has still been observed. The optimisation of the synthesis involved the replacement of Ser-Thr and Gln-Thr dipeptides to their pseudoproline derivatives. This has also led to the more efficient synthesis of the histone tail peptide analogues by eliminating on-resin peptide aggregation.

From the NMR studies carried out on the histone H3 (1-21) peptide analogues, it is clear that the inherent deacetylase activity of HDAC1 is faster than the demethylase activity of LSD1. All of the K9ac containing peptides were completely deacetylated. However, only about two thirds of the K4me2 containing peptides were demethylated. If this is a result of branching at Lys4, then incomplete demethylation could be because the quaternary ammonium centre at the lysine side chain can no longer be a substrate for LSD1. This is

because the chemistry of the flavin-containing amino oxidases, requires a protonated nitrogen in the substrates. This however, does not explain the incomplete demethylation of the K4me1containing peptides.

The crosstalk between LSD1 and HDAC1 in the CoREST complex is clear from the separate inhibition experiments with entinostat and 2-PCPA. Inhibition of HDAC1 with entinostat increases the activity of LSD1, while the HDAC1 activity is increased when LSD1 is inhibited with 2-PCPA. Ins(1,2,3,4,5,6)P_6 has also been shown to activate HDAC1 even in the presence of entinostat. However, Ins(1,2,3,4,5,6)P_6 cannot completely recover the activity of HDAC1 when entinostat is present. The activity of LSD1 in the presence of Ins(1,2,3,4,5,6)P_6 is also affected, with a slight increase in activity.

In order to fully complete this work, further studies are required. In particular the synthesis of H3K9ac is necessary as a control sequence. This peptide would undergo experiments with both the entinostat and Ins(1,2,3,4,5,6)P_6 alone and in combination. Additionally, 2-PCPA can also be used to investigate the effect of inhibiting LSD1 on the deacetylase activity of HDAC1.

Furthermore, the synthesis of a native histone H3 (1-21) sequence with no PTMs is also required. This would be used as a control sequence to help detect products formed from the deacetylase and demethylase experiments. Once the NMR experiments have been run, the products can be analysed by LCMS. Through this method of analysis the peptides which have not been demethylated could be determined and the incomplete demethylation of the peptides resolved.

In the second part of this project, the synthesis of three peptide based dual inhibitors of LSD1 and HDAC1 has been successfully achieved. However, this is still ongoing work and the crystallisation experiments need to be carried out. Additionally, activity assays need to be carried out to determine the IC_{50} values of the peptides for LSD1, HDAC1 and the ternary complex. These dual inhibitor peptides can also be useful in targeting cancers such as AML, where combination therapies have previously be used.[243] In fact, a dual inhibitor for LSD1 and HDAC, 4SC-202 (Figure 94) has been evaluated in phase I clinical trials in haematological cancers and demonstrated favourable safety and

tolerability profiles[244,245] 4SC-202 is a tosylate salt of a benzamide HDAC inhibitor containing a *N*-sulfonylpyrrol scaffold (Figure 94).

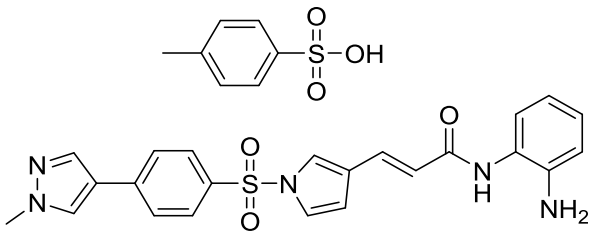


Figure 94: Structure of 4SC-202, a combined LSD1 and HDAC inhibitor.

However, the length of 4SC-202 does not preclude to binding both LSD1 and HDAC1 simultaneously. This is the reason behind the different lengths of the peptide dual inhibitors. Additionally, with the LSD1 inhibitor and the HDAC1 inhibitor at different ends of the peptide, it can be hypothesised that the dual peptide inhibitor can bind both LSD1 and HDAC1 simultaneously.

To establish this dual binding, melting curves can be studied using CD. When the dual inhibitor is bound, the complex should be stabilised better which would be shown by an increase in the melting temperature. Through these experiments, the distance between the LSD1 and HDAC1 in the CoREST complex can be determined. Furthermore, the possibility of binding both LSD1 and HDAC1 simultaneously can be assessed. The IC_{50} of the peptide based dual inhibitors can then be evaluated and compared to known HDAC inhibitiors, LSD1 inhibitors and dual LSD1/HDAC inhibitors.

Chapter 5: Experimental

5.1. General information

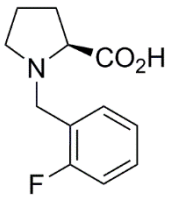
All experiments were carried out under an atmosphere of nitrogen, using anhydrous solvents unless otherwise stated. All chemicals were purchased from Sigma-Aldrich and Fmoc-protected amino acids were purchased from Novabiochem or Pepceuticals, and used as received. Thin layer chromatography was carried out on Merck Kieselgel 60 F_{254} plates and visualisation achieved with ultraviolet light. Flash column chromatography was performed using Kieselgel 60 (230-400) mesh). NMR spectra were recorded on Bruker DPX 300 (^{1}H, 300 MHz, ^{13}C, 75 MHz, ^{19}F, 282 MHz), Bruker DPX 400 (^{1}H, 400 MHz, ^{13}C, 100 MHz, ^{19}F, 376 MHz) or Bruker Avance III 500 (^{1}H, 500 MHz, ^{13}C, 125 MHz) spectrometers as indicated. NMR shifts (δ) are quoted in ppm relative to the residual non-dueterated solvent peak of chloroform (^{1}H δ 7.26, ^{13}C δ 77.0) and coupling constants (*J*) are quoted to the nearest 0.1 Hz.

LC-MS were run using a Xevo QTof mass spectrometer (Waters) coupled to an Acquity LC system (Waters) using an Acquity UPLC BEH C18 column (2.1 x 50 mm, Waters). The flow rate was 0.6 ml min-1 and the gradient was as follows: 95% Solvent A (0.1% formic acid in water) with 5% solvent B (0.1% formic acid in acetonitrile) was held constant for 0.5 min, followed by a linear gradient to 100% B over the next 2.1 min. After 1 min at 100% solvent B, the gradient was returned to 95% solvent A and 5% solvent B over 0.2 min. The ESI capillary voltage was 3 kV, cone voltage 30 V and collision energy 4 eV. The MS acquisition rate was 10 spectra per second and m/z data ranging from 50 to 2000 Da was collected. Mass accuracy was achieved using a reference lock mass scan, once every 10 seconds.

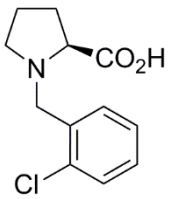
Small molecule X-ray data collection was carried out on a Brucker Apex 2000 CCD diffractometer using graphite-monochromated Mo-Kα radiation (λ = 0.71073 (ollection was carried out on a Brucker Apex 2000 CCD diffractometer using graphite-monochromated Mo-K18 column (2.1 x solution by direct method and structure refinement on F2 employed SHEXTL version 6.10. Hydrogen atoms were included in calculated positions (C‑H = 0.95-1.00 luded in calculated positions (Cme isotropic displacement parameters set 1.5 Ueq for methyl H atoms and 1.2 Ueq for all other H atoms. All non-H atoms were

refined with anisotropic displacement parameters. Figues were generated using the same software. Protein crystal structure figures were generated using the PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC.

5.2. Chemical syntheses

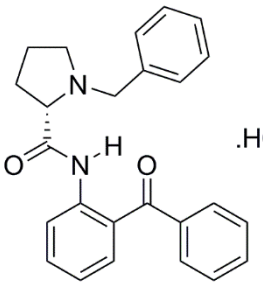


(*S*)-*N*-(2-fluoro)benzyl proline (1b) (*S*)-proline (5.0 g, 43.5 mmol, 1.0 eq.) and potassium hydroxide (7.3 g, 130.4 mmol, 3 eq.) were taken up in isopropyl alcohol (44 mL, 1 M) in a 250 mL round bottom flask at 40 °C. 2-Fluorobenzyl bromide (5 mL, 41.5 mmol, 0.95 eq.) was added dropwise and the resulting solution stirred at room temperature for 12 h. The solution was then acidified to pH 5 with concentrated hydrochloric acid with the aid of a pH probe before the addition of chloroform (100 mL) and the precipitate removed by filtration. The filtrate was concentrated to give an orange residue to which acetone (100 mL) was added and the resulting white precipitate removed by filtration. The filtrate was concentrated in vacuo to give (*S*)-*N*-(2-fluoro)benzyl proline (7.96 g, 82%) as an orange waxy solid. ^{19}F NMR (282 MHz; CDCl3) –116.3 ^{1}H-NMR (300 MHz, $CDCl_3$) δ: 9.95 (1H, s, O*H*), 7.52 (1H, *ddd*, J = 7.5, 7.5, 1.6 Hz, Ar-C*H*), 7.349-7.29 (1H, *m*, Ar-C*H*), 7.13 (1H, *ddd*, Ar-C*H*), 7.12-7.03 (1H, *m*, Ar-C*H*), 4.37 (1H, *d*, *J* = 13.2 Hz, N-C*H*H), 4.20 (1H, *d*, *J* = 13.2 Hz, N-CH*H*), 3.74 (1H, *dd, J* = 9.2, 6.2 Hz, α-C*H*), 3.63-3.52 (1H, *m*, δ-C*H*), 2.91-2.77 (1H, *m*, δ'-H), 2.39-2.11 (2H, *m*, β-CH_2), 2.05-1.82 (2H, *m*, γ-CH_2). ^{13}C-NMR (75 MHz, $CDCl_3$) δ: 171.8 (*C*=O), 163.0 (q*C*), 159.7 (q*C*), 132.9 (Ar-*C*H), 131.5 (Ar-*C*H), 124.8 (Ar-*C*H), 119.3 (q*C*), 116.0 (Ar-*C*H), 67.3 (α-C), 53.6 (δ-C), 50.8 (N*C*H_2Ar), 29.0 (β-C), 23.0 (γ-C). IR (ν_{max}/cm^{-1}, neat): 3458, 3013, 2970, 1736, 1618, 1443, 1369, 1229, 1109, 898, 758. HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{12}H_{14}FNO_2$: 224.1087. Found: 224.1089 (Δ 0.89 ppm); m.p. 78-80 °C (EtOH); $[\alpha]_D^{20}$ -25.0 (c 1.0, EtOH).



(*S*)-*N*-(2-chloro)benzyl proline (1c) (*S*)-proline (1.97 g, 17.1 mmol, 1 eq.) and potassium hydroxide (2.88 g, 51.3 mmol, 3 eq.) were taken up in isopropyl alcohol (17 mL, 1 M) and heated to 40

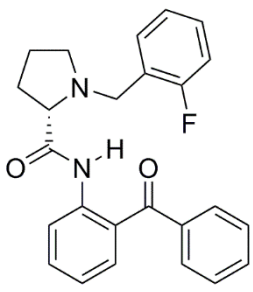
°C. After the solution had become transparent, 2-chlorobenzyl bromide (2.67 mL, 20.5 mmol, 1.2 eq.) was added dropwise and stirred at 40 °C for 2h. The course of the reaction was monitored by TLC (20% ethyl acetate/petroleum ether). The reaction was acidified to pH 5.32 with concentrated hydrochloric acid with the aid of a pH probe, before the addition of chloroform (17 mL). The solution left to stir for 1 h before a white precipitate was filtered. The filtrate was concentrated to a yellow oil before the addition of acetone. The solvent was evaporated and the residue washed with chloroform (3 x 25 mL), followed by ethyl acetate (3 x 25 mL) and chloroform (3 x 25 mL), before being dried in vacuo to give (*S*)-*N*-(2-chloro)benzyl proline (2.99 g, 73%) as a waxy yellow solid. ^{1}H-NMR (300 MHz, $CDCl_3$) δ: 7.59 (1H, *dd*, *J*=7.2 Hz, *J*= 2.1 Hz, Ar-C*H*), 7.40 (1H, *dd*, J=7.7, 1.6 Hz, Ar-C*H*), 7.32-7.20 (2H, *m*, Ar-C*H,* Ar-C*H*), 4.38 (1H, *d*, *J*=13.5 Hz, C*H*HAr), 4.24 (1H, *d*, *J*=13.2 Hz, CH*HAr*), 3.74 (1H, *dd*, *J*=9.3 Hz, *J*=5.4 Hz, α-C*H*), 3.49 (1H, *dt*, *J*=10.8 Hz, *J*=5.4 Hz, δ-C*H*), 2.84 (1H, *dt*, *J*=10.5 Hz, *J*=8.6 Hz, δ'-C*H*), 2.39-2.12 (2H, *m*, β-CH_2), 1.98-1.95 (2H, *m*, γ-CH_2). ^{13}C-NMR (75 MHz, $CDCl_3$) δ: 172.4 (*C*=O), 134.7 (Ar-*C*), 132.5 (Ar-*C*H), 130.7 (Ar-*C*H), 130.5 (Ar-*C*), 127.5 (Ar-*C*H), 67.7 (α-*C*), 55.1 (*C*H_2Ph), 53.9 (δ-*C*H_2), 29.2 (β-*C*H_2), 23.4 (γ-*C*H_2). HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{12}H_{15}NO_2Cl$: 240.0791, found: 240.0794 (Δ 1.23 ppm); $[M+Na]^+$ calcd for $C_{12}H_{15}NO_2ClNa$: 262.0611, found: 262.0615 (Δ 1.53 ppm); $[\alpha]_D^{20}$ -22.1° (c 1.0, MeOH), Lit. -21.2° [246]



.HCl

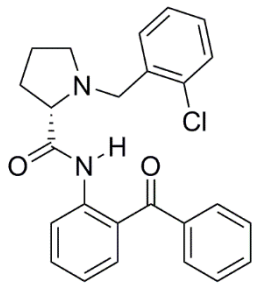
***(S)-N*-(2-benzoylphenyl)-2-benzylpyrrolidine-1-carboxamide hydrochloride (BPB.HCl, 2a)** (*S*)-*N*-benzylproline (2.0 g, 9.76 mmol, 1 eq.) was taken up in dichloromethane (20 mL, 0.5 M) in an oven dried 100 mL RBF. 1-Methylimidazole (1.7 mL, 8.76 mmol, 2.2 eq.) was added and the solution stirred until it became transparent, before being cooled to 0 °C in an ice bath. Methane sulfonyl chloride (0.76 mL, 9.76 mmol, 1 eq.) was added dropwise resulting in an orange to red colour change. The solution was allowed to warm to room temperature before 2-aminobenzophenone (1.74 g, 8.76 mmol, 0.9 eq.). The RBF was then placed in a microwave reactor, fitted with a condenser, set at 50 °C for 1 h. The solution

was then cooled to room temperature and quenched with saturated sodium hydrogen carbonate solution (20 mL). The two layers were separated and the aqueous layer was extracted with dichloromethane (3 x 15 mL). The combined organic layers were washed with brine (20 mL) and dried over $MgSO_4$, before being concentrated to a brown oil. Acetone (50 mL) was added and the solution acidified with concentrated hydrochloric acid (1.68 mL, 19.52 mmol, 2 eq.). Precipitation was allowed to occur over 2 h in the fridge. The precipitate was collected by filtration, washed with acetone (50 mL) and dried in vacuo to give BPB.HCl (2.39 g, 65%) as an off white solid. ^{1}H-NMR (400 MHz, DMSO-d_6) δ: 11.24 (1H, broad *s*, NH), 9.75 (1H, broad *s*, Ar-C*H*), 7.74-7.64 (3H, *m*, Ar-C*H*), 7.64-7.55 (2H, *m*, Ar-C*H*), 7.50 (2H, *t*, *J* = 7.6 Hz, Ar-C*H*), 7.47-7.31 (6H, *m*, Ar-C*H*), 7.26 (1H, *d*, *J* = 7.8 Hz, Ar-C*H*), 4.41-4.15 (3H, *m*, α-C*H*, C*H*$_2$Ph), 3.49-3.36 (1H, *m*, δ-C*H*), 3.30-3.18 (1H, *m*, δ'-C*H*), 2.29-2.14 (1H, *m*, β-C*H*), 2.02-2.88 (1H, *m*, γ-C*H*), 1.73-1.56 (1H, *m*, γ'-C*H*), 1.37-1.21 (1H, *m*, β'-C*H*). ^{13}C-NMR (101 MHz, DMSO-d_6) δ: 194.2 (C=O), 165.7 (C=O), 136.6 (q*C*), 134.1 (Ar-*C*H), 133.0 (Ar-*C*H), 131.4 (q*C*), 131.0 (Ar-*C*H), 129.6 (Ar-*C*H), 129.5 (Ar-*C*H), 129.4 (q*C*), 128.6 (Ar-*C*H), 128.4 (Ar-*C*H), 125.5 (Ar-*C*H), 124.4 (Ar-*C*H), 123.1 (q*C*), 119.5 (Ar-*C*H), 56.7 (α-*C*), 54.1 (*C*H$_2$Ph), 35.4 (δ-*C*), 27.8 (β-*C*), 22.1 (γ-*C*). IR (ν_{max}/cm^{-1}, neat): 3276, 1682, 1647, 1577, 1512, 1444, 1286, 1266, 1102, 923, 768, 697; HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{25}H_{24}N_2O_2F$ 403.1822, found 403.1825 (Δ 0.7 ppm); $[\alpha]_D^{20}$ -37.5 (c 1.0, MeOH) Lit: -40.2 for HCl salt.[247]



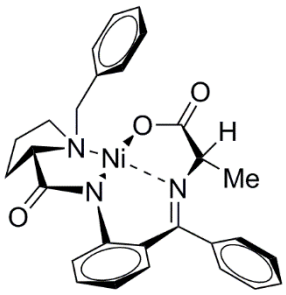
(*S*)-*N*-(2-Benzoylphenyl)-1-(2-fluorobenzyl)pyrrolidine-2-carboxamide (2FBPB, 2b) Methane sulfonyl chloride (2.99 mL, 38.6 mmol, 1.0 equiv.) was added to a solution of 2-fluorobenzyl proline (10.0 g, 38.6 mmol, 1.0 equiv.) and 1-methyl imidazole (9.85 mL, 124 mmol, 3.2 equiv.) in dichloromethane (80 mL, 0.5 M) at 0 °C. After 5 mins, 2-aminobenzophenone (6.85 g, 34.7 mmol, 0.9 equiv.) was added. The reaction mixture was heated at 50 °C for 18 h, cooled, then saturated sodium hydrogen carbonate solution (80 mL) was added. The two layers were separated and the aqueous layer extracted with dichloromethane (3 x 80 mL). The combined organic layers were washed with brine (3 x 250 mL), dried ($MgSO_4$), filtered and

concentrated *in vacuo*. Acetone (100 mL) was added to the brown oil and acidified with concentrated hydrochloric acid (37%). The product was allowed to precipitate in the fridge overnight. The precipitate was filtered and washed with acetone to give (*S*)-*N*-(2-benzoylphenyl)-1-(2-fluorobenzyl)pyrrolidine-2-carboxamide hydrochloride (11.5 g, 68%) as an off-white solid. ^{19}F NMR (376 MHz, $CDCl_3$) δ: –117.5 ^{1}H-NMR (300 MHz, $CDCl_3$) δ: 11.42 (1H, *s*, N*H*), 8.55 (1H, *dd*, *J* = 8.4 Hz, *J* = 0.8 Hz, Ar-C*H*), 7.79-7.72 (2H, *m*, Ar-C*H*), 7. 64-7.56 (1H, *m*, Ar-C*H*), 7.56-7.41 (5H, *m*, Ar-C*H*), 7.16-7.04 (2H, *m*, Ar-C*H*), 6.92 (1H, *ddd*, *J* = 7.5 Hz, *J* = 7.5 H, *J* = 1.1 Hz, Ar-C*H*), 6.83-6.74 (1H, *m*, Ar-C*H*), 3.90 (1H, *d*, *J* = 13.3 Hz, C*H*HAr), 3.73 (1H, *d*, *J* = 13.2 Hz, CH*H*Ar), 3.35 (1h, *dd*, *J* = 10.1, *J* = 4.5 Hz, α-C*H*), 3.23 (1H, *m*, δ-C*H*), 2.46 (1H, *m*, δ'-C*H*), 2.24 (1H, *m*, β-C*H*), 1.94 (1H, *m*, β'-C*H*), 1.80 (2H, *m*, γ-CH_2). ^{13}C-NMR (75 MHz, $CDCl_3$) δ: 197.8 (Ph*C*=O), 174.4 (NH*C*=O), 139.0 (q*C*), 138.6 (q*C*), 133.3 (Ar-*C*H), 132.5 (Ar-*C*H), 131.7 (Ar-*C*H), 130.1 (Ar-*C*H), 128.9 (Ar-*C*H), 128.8 (Ar-*C*H), 128.2 (Ar-*C*H), 125.6 (q*C*), 125.0 (q*C*), 124.8 (q*C*), 123.9 (Ar-*C*H), 122. 3 (Ar-*C*H), 121.5 (Ar-*C*H), 115.3 (Ar-*C*H), 115.0 (Ar-*C*H), 67.92 (α-*C*), 53.7 (δ-*C*), 52.0 (*C*H_2Ar), 31.1 (β-*C*), 24.3 (γ-*C*). HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{25}H_{23}FN_2O_2$: 403.1822, found: 403.1817 (Δ -1.24 ppm). HPLC (OD-H column, hexane (5%)/iPrOH isocratic): 15.2 min. IR (ν_{max}/cm^{-1}, neat): 3276, 1682, 1647, 1577, 1512, 1444, 1286, 1266, 1102, 923, 768, 697; m.p: 88-90 °C; $[\alpha]_D^{20}$ -125 (c 1.0, MeOH).



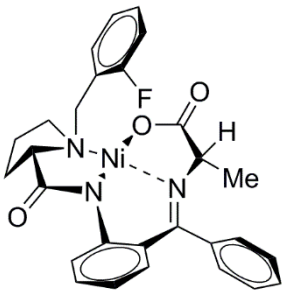
(*S*)-*N*-(2-Benzoylphenyl)-1-(2-chlorobenzyl)pyrrolidine-2-carboxamide (2CBPB, 2c) (*S*)-*N*-(2-chloro)benzyl proline (2.46 g, 10.3 mmol, 1.5 eq.) was dissolved in dry DCM (14 mL, 0.5 M) in an oven dried 100 mL RBF under a nitrogen atmosphere. 1-Methylimidazole (1.80 mL, 22.6 mmol, 3.3 eq.) was added before being cooled to 0 °C in an ice bath. Methane sulfonyl chloride (0.79 mL, 10.3 mmol, 1.5 eq.) was added dropwise over 5 mins. The resulting solution was allowed to reach room temperature before 2-aminobenzophenone (1.35 g, 6.84 mmol, 1 eq.) was added and the RBF placed in a microwave reactor equipped with a condenser and under a nitrogen atmosphere, set at 50 °C for 1 h (the temperature reached a maximum

of 51 °C). The reaction was monitored by TLC (20% ethyl acetate/petroleum ether). The reaction was quenched with saturated ammonium chloride solution (40 mL) and extracted with dichloromethane (3 x 40 mL). The combine organic layers were washed with brine (50 mL), dried over $MgSO_4$ and concentrated to give (*S*)-*N*-(2-benzoylphenyl)-1-(2-chlorobenzyl)pyrrolidine-2-carboxamide (2.45 g, 57%) as a yellow oil. ^{1}H-NMR (300 MHz, $CDCl_3$) δ: 11.29 (1H, *s*, N*H*), 8.42 (1H, *dd*, *J* = 8.4 Hz, *J* = 0.8 Hz, Ar-C*H*), 7.69-7.67 (1H, *m*, Ar-C*H*), 7.66-7.64 (1H, *m*, Ar-C*H*), 7.60-7.56 (1H, *m*, Ar-C*H*), 7.56-7.46 (1H, *m*, Ar-C*H*), 7.46-7.37 (4H, *m*, Ar-C*H*), 7.13-7.08 (1H, *m*, Ar-C*H*), 7.04-6.94 (3H, *m*, Ar-C*H*), 3.83 (2H, *s*, CH_2Ar), 3.34 (1H, *dd*, *J* = 9.9 Hz, *J* = 4.5 Hz, α-C*H*), 3.28-3.19 (1H, *m*, δ-C*H*), 2.40 (1H, *dt*, *J* = 9.1 Hz, *J* = 7.0 Hz, δ'-C*H*), 2.29-2.14 (1H, *m*, β-C*H*), 1.95-1.82 (1H, *m*, β'-C*H*), 1.82-1.68 (2H, *m*, γ–CH_2). ^{13}C-NMR (75 MHz, $CDCl_3$) δ: 197.9 (*C*=O), 174.4 (*C*=O), 139.0 (q*C*), 138.6 (q*C*), 135.8 (Ar-*C*H), 133.8 (Ar-*C*H), 133.3 (Ar-*C*H), 132.5 (Ar-*C*H), 132.4 (Ar-*C*H), 131.1 (Ar-*C*H), 130.0 (Ar-*C*H), 129.2 (Ar-*C*H), 128.2 (Ar-*C*H), 126.7 (q*C*), 125.5 (q*C*), 122.3 (Ar-*C*H), 121.6 (Ar-*C*H), 68.6 (α-*C*), 56.5 (δ-*C*), 54.1 (*C*H_2Ph), 31.0 (β-*C*), 24.3 (γ-*C*). HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{25}H_{24}N_2O_2Cl$: 419.1526, found: 419.1523 (Δ -0.72 ppm). $[\alpha]_D^{20}$ -549 (c1.0, MeOH), Lit: -40.2 (c 1.0, MeOH) for HCl salt.[247]



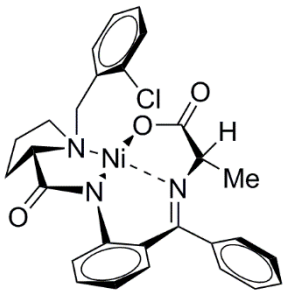
(*S*)-Ala-Ni-BPB (3a) BPB.HCl (3.0 g, 7.13 mmol, 1 eq.) was dissolved in methanol (75 mL, 0.1 M). Nickel (II) nitrate hexahydrate (4.14 g, 14.26 mmol, 2 eq.) was added, followed by alanine (1.27 g, 14.26 mmol, 2 eq.) and heated to 50 °C (alanine dissolves on heating). Potassium hydroxide (3.19 g, 57.0 mmol, 8 eq.) dissolved in methanol (40 mL) was then added and stirred at 60 °C for 2 h. The solution was neutralised with acetic acid (3.27 mL, 57.0 mmol, 8 eq.) and diluted with water (15 mL). Methanol was evaporated and the aqueous extracted with dichloromethane (3 x 20 mL). The combined organic layers were washed with brine (25 mL), dried over $MgSO_4$ and the solvent evaporated to give (*S*)-Ala-Ni-BPB (3.27 g, 90%) as a red powder. ^{1}H-NMR (400 MHz, $CDCl_3$) δ: 8.03-7.99 (3H, *m*, Ar-C*H*), 7.47-7.35 (3H, *m*, Ar-C*H*), 7.29 (2H, *t*, *J* = 7.59, Ar-C*H*), 7.18-7.03 (3H, *m*, Ar-C*H*), 6.89-6.87 (1H, *m*, Ar-C*H*), 6.61-6.53 (2H, *m*, Ar-C*H*), 4.42

(1H, *d*, *J* = 12.6 Hz, C*H*HAr), 3.84 (1H, *q*, *J* = 7.0 Hz, α-C*H*), 3.65 (1H, *m*, CH*H*Ar), 3.74-3.58 (1H, *m*, γ-C*H*), 3.49-3.42 (1H, *m*, δ-C*H*), 3.41 (1H, *dd*, *J* = 11.2 Hz, *J* = 5.8 Hz, α-C*H* pyrrolidine) 2.71-2.39 (2H, *m*, β-CH_2), 2.21-2.12 (1H, *m*, γ'-C*H*) 2.05-1.95 (1H, *m*, δ'-C*H*), 1.52 (3H, *d*, *J* = 7.0 Hz, CH_3). ^{13}C-NMR (101 MHz, $CDCl_3$) δ: 180.5 (*C*=N), 180.4 (*C*=O), 170.3 (*C*=O), 142.1 (q*C*), 133.5 (q*C*), 133.3 (q*C*), 133.1 (Ar-*C*H), 132.1 (Ar-*C*H), 131.5 (Ar-*C*H), 129.7 (Ar-*C*H), 128.9 (Ar-*C*H), 128.9 (Ar-*C*H), 127.5 (Ar-*C*H), 127.2 (Ar-*C*H), 126.5 (q*C*), 123.9 (Ar-*C*H), 120.8 (Ar-*C*H), 70.3 (α-*C* pyrrolidine), 66.6 (α-*C*), 63.1 (*C*H_2Ph), 57.3 (γ-*C*), 30.8 (β-*C*), 24.1 (δ-*C*), 21.8 (*C*H_3). HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{28}H_{28}N_3NiO_3$: 514.1439, found 514.1462 (Δ 4.5 ppm); $[\alpha]_D^{20}$ +2499 (c 0.036, MeOH), Lit. +2643;[169] dr 88:12.



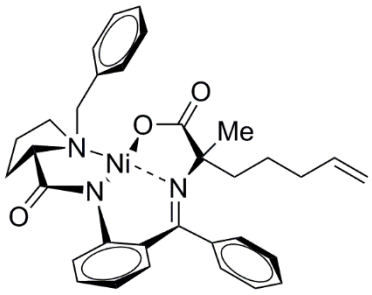
(*S*)-{({2-[1-(2-Fluorobenzyl)pyrrolidine-2-carboxamide]phenyl} phenylmethylene)-glycinato-*N*,*N*',*N*'',*O*}-nickel(II) ((*S*)-Ala-Ni-2FBPB, 3b) (*S*)-*N*-(2-benzoylphenyl)-1-(2-fluorobenzyl)pyrrolidine-2-carboxamide hydrochloride (10.0 g, 22.8 mmol, 1.0 equiv.), nickel nitrate hexahydrate (13.27 g, 45.6 mmol, 2.0 equiv.) and glycine (3.43 g, 45.6 mmol, 2.0 equiv.) were dissolved in methanol (230 mL, 0.1 M) at 50 °C. Potassium hydroxide (10.24 g, 182.5 mmol, 8.0 equiv.) was added and the mixture heated to 70 °C for 1 h. The reaction mixture was cooled and concentrated. Saturated sodium hydrogen carbonate solution (500 mL) was added to the residue and extracted with dichloromethane (3 x 500 mL). The combined organic layers were washed with brine (3 x 1.0 L), dried ($MgSO_4$) and the solvent removed *in vacuo*. Purification was achieved by flash column chromatography (0-40% acetone in dichloromethane). Fractions containing the desired product were combined and the solvent evaporated to afford (*S*)-Ala-Ni-CBPB (11.9 g, 99%) as a red solid. ^{19}F NMR (376 MHz, $CDCl_3$) δ –113.9 ^{1}H NMR (400 MHz, $CDCl_3$) δ: 8.33 (1H, *td*, *J* = 7.3 Hz, Ar-C*H*), 8.13 (1H, *dd*, *J* = 8.6 Hz, Ar-C*H*), 7.53-7.49 (2H *m*, Ar-C*H*), 7.46 (1H, *m*, Ar-C*H*), 7.27-7.20 (2H, *m*, Ar-C*H*), 7.20-7.13 (2H, *m*, Ar-C*H*), 7.06 (1H, *ddd*, *J* = 9.8 Hz, *J* = 8.2 Hz, Ar-C*H*), 6.96 (1H, *m*, Ar-C*H*), 6.71-6.63 (2H, *m*, Ar-C*H*), 4.41 (1H, *dd*, *J* = 12.9 Hz, *J* = 1.2 Hz, N-C*H*H), 3.91 (1H, *q*, *J* = 7.0 Hz, α-C*H*), 3.83

(1H, *dd*, *J* = 12.9 Hz, *J* = 0.8 Hz, N-CH*H*), 3.70 (1H, *m*, β-C*H*), 3.55-3.46 (2H, *m*, α-C*H* pyrrolidine, γ-C*H*), 2.38 (1H, *m*, δ-C*H*), 2.57 (1H, *m*, δ'-C*H*), 2.22 (1H, *m*, γ'-C*H*), 2.07 (1H, *m*, β'-C*H*), 1.58 (3H, *d*, *J* = 7.0 Hz, C*H*$_3$). ^{13}C NMR (100 MHz, $CDCl_3$) δ 180.4 (*C*=N), 180.1 (*C*=O), 170.3 (*C*=O), 142.1 (q*C*), 134.2 (Ar*C*H), 133.5 (q*C*), 133.2 (Ar*C*H), 132.1 (Ar*C*H), 129.7 (Ar*C*H), 128.9 (Ar*C*H), 127.5 (Ar*C*H), 127.2 (Ar*C*H), 126.6 (q*C*), 124.5 (Ar*C*H), 123.9 (Ar*C*H), 120.8 (Ar*C*H), 120.5 (q*C*), 119.8 (q*C*), 117.9 (ArCH), 116.2 (Ar*C*H), 116.0 (Ar*C*H), 70.3 (α-*C* pyrrolidine), 66.6 (α-*C*H), 57.1 (γ-*C*H_2), 55.6 (N-*C*H_2), 30.7 (β-*C*H_2), 24.1 (δ-*C*H_2), 21.8 (*C*H_3). IR (ν_{max}/cm^{-1}, neat) 1680, 1626, 1593, 1440, 1359, 1330, 1302, 1266, 1231, 1168, 1064, 7621, 748, 710, 682. HRMS (m/z) $[M+H]^+$ calcd for $C_{28}H_{27}FN_3O_3Ni$: 530.1390, found 530.1397 (Δ 1.3 ppm), $[M+H]^+$ calcd for $C_{28}H_{27}N_3O_3F^{60}Ni$: 532.1344, found 532.1381 (Δ 7.0 ppm); m.p. 284-286 °C; $[\alpha]_D^{20}$ +3664.5 (c 0.05, $CHCl_3$); dr 99:1.



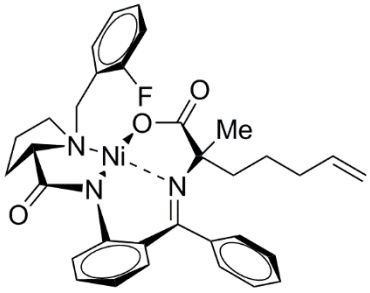
(*S*)-{({2-[1-(2-Chlorobenzyl)pyrrolidine-2-carboxamide]phenyl} phenylmethylene)-glycinato-*N*,*N*',*N*'',*O*}-nickel(II) ((*S*)-Ala-Ni-2CBPB, 3c) (*S*)-*N*-(2-benzoylphenyl)-1-(2-chlorobenzyl)pyrrolidine-2-carboxamide (2.0 g, 4.77 mmol, 1 eq.) was dissolved in methanol (37 mL, 0.1 M). Nickel (II) nitrate hexahydrate (1.98 g, 7.32 mmol, 1.5 eq.) was added, followed by alanine (625 mg, 7.32 mmol, 1.5 eq.) and the resulting green solution heated to 50 °C (alanine dissolves on heating). Potassium hydroxide (1.64 g, 29.3 mmol, 6 eq.) dissolved in methanol (21 mL) was then added and stirred at 60 °C for 2 h. The solution was neutralised with acetic acid (1.68 mL, 29.3 mmol, 6 eq.) and diluted with water (8 mL). Methanol was evaporated and the aqueous extracted with dichloromethane (3 x 10 mL). The combined organic layers were washed with brine (10 mL), dried over $MgSO_4$ and the solvent evaporated to give (S)-Ala-Ni-CBPB (2.56 g, 98%) as a red powder. ^{1}H-NMR (300 MHz, $CDCl_3$) δ: 8.11 (1H, *dd*, *J* = 7.6 Hz, *J* = 1.2 Hz, Ar-C*H*), 7.88 (1H, *d*, *J* = 8.5 Hz, Ar-C*H*), 7.40 (4H, *m*, Ar-C*H*), 7.22 (3H, *m*, Ar-C*H*), 7.08 (2H, *m*, Ar-C*H*), 6.86 (1H, *d*, *J* = 6.9 Hz, Ar-C*H*), 6.57 (2H, *m*, Ar-C*H*), 4.39 (1H, *d*, *J* = 12.9 Hz, C*H*HAr), 3.79 (2H, *m*, α-C*H*, CH*H*Ar), 3.65 (1H, *m*, δ-C*H*), 3.48 (1H, *dd*, *J* = 11.1 Hz, *J* = 6.1

Hz, α-C*H* pyrrolidine), 3.40 (1H, *dd*, *J* = 10.1 Hz, *J* = 6.5 Hz, γ-C*H*), 2.85 (1H, *m*, β-C*H*), 2.55 (1H, *m*, β'-C*H*), 2.16 (1H, *m*, δ'-C*H*), 2.01 (1H, *m*, γ'-C*H*), 1.49 (3H, *d*, *J* = 7.0 Hz, CH_3). ^{13}C-NMR (75 MHz, $CDCl_3$) δ: 180.3 (*C*=N), 179.6 (*C*=O), 170.2 (*C*=O), 141.9 (q*C*), 135.7 (q*C*), 134.1 (Ar-*C*H), 133.4 (q*C*), 133.3 (Ar-*C*H), 133.2 (Ar-*C*H), 131.2 (q*C*), 130.5 (Ar-*C*H), 130.3 (Ar-*C*H), 129.7 (Ar-*C*H), 128.9 (Ar-*C*H), 129.0 (Ar-*C*H), 128.2 (Ar-*C*H), 127.4 (Ar-*C*H), 127.2 (Ar-*C*H), 127.1 (Ar-*C*H), 126.7 (q*C*), 123.8 (Ar-*C*H), 120.8 (Ar-*C*H), 70.6 (α-*C* pyrrolidine), 66.6 (α-*C*), 59.6 (*C*H_2Ar), 57.4 (γ-*C*), 30.6 (β-*C*), 24.1 (δ-*C*), 21.8 (α-*C*H_3). HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{28}H_{26}ClN_3NiO_3$: 546.1094. Found: 546.1102 (Δ 1.65 ppm); 98:2.



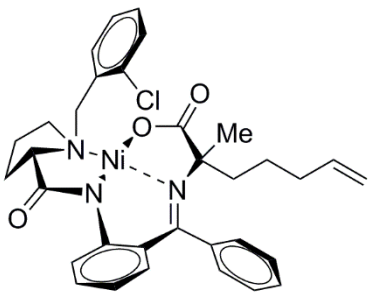
(*S*)-{({2-[1-(benzyl)pyrrolidine-2-carboxamide]phenyl}phenylmethylene)-(*S*)-6-heptenoato *N*,*N*',*N*'',*O*}-nickel(II) (4a) Potassium tert-butoxide (329 mg, 2.9 mmol, 3 eq.) was charged into an oven dried 10 mL round bottom flask and heated with a heat gun, under a nitrogen atmosphere. Dry DMF (4 mL, 0.25 M) was added and the solution heated to 50 °C before the addition of (*S*)-Ala-Ni-BPB (500 mg, 0.98 mmol, 1 eq.). 5-bromopentene (350 µL, 2.9 mmol, 3 eq.) was added dropwise over 5 mins and the reaction allowed to stir at 50 °C for 1 h. The reaction was cooled to room temperature, concentrated and the residue diluted with water (15 mL), then extracted with dichloromethane (3 x 15 mL). The combine organic layers were washed with 5% lithium chloride solution (3 x 15 mL), brine (15 mL) and dried over $MgSO_4$ before the solvent evaporated. Purification was carried out by flash column chromatography (20% acetone/dichloromethane). The fractions containing product were combined and concentrated to give the product (244 mg, 43%) as a red powder. ^{1}H-NMR (300 MHz, $CDCl_3$) δ: 8.10-8.03 (2H, *m*, Ar-C*H*), 8.0 (1H, apparent *d*, *J* = 8.24 Hz, Ar-C*H*), 7.50-7.45 (2H, *m*, Ar-C*H*), 7.44-7.34 (4H, *m*, Ar-C*H*), 7.34-7.29 (1H, *m*, Ar-C*H*), 7.12 (1H, *ddd*, *J* = 8.5 Hz, *J* = 5.8 Hz, *J* = 2.7 Hz, Ar-C*H*), 7.00-6.94 (1H, *m*, Ar-C*H*), 6.68-6.58 (2H, *m*, Ar-C*H*), 5.59-5.78 (1H, *m*, C*H*=CH_2), 5.16-4.97 (2H, *m*, CH=CH_2), 4.50 (1H, *d*, *J* = 12.67 Hz, C*H*HAr), 3.70 (1H, *d*, *J* = 12.67 Hz, CH*H*Ar), 3.67-3.59 (1H, *m*, δ-

C*H*), 3.48-3.41 (1H, *m*, α-C*H* pyrrolidine), 3.37-3.14 (1H, *m*, δ'-C*H*), 2.79-2.60 (2H, *m*, β-CH_2 pyrrolidine), 2.58-2.31 (2H, *m*, γ-CH_2), 2.22-1.93 (4H, *m*, γ-CH_2 pyrrolidine, δ-CH_2), 1.78-1.63 (2H, *m*, β-CH_2), 1.2 (3H, *s*, CH_3). ^{13}C-NMR (75 MHz, $CDCl_3$) δ: 182.4 (C=N), 180.5 (C=O), 172.4 (C=O), 141.5 (q*C*), 137.8 (*C*H=CH_2), 136.5 (q*C*), 133.3 (Ar-*C*H), 131.7 (Ar-*C*H), 131.5 (Ar-*C*H), 130.3 (Ar-*C*H), 129.4 (Ar-*C*H), 129.0 (q*C*), 128.9 (Ar-*C*H), 128.7 (Ar-*C*H), 127.9 (q*C*), 127.3 (Ar-*C*H), 127.0 (Ar-*C*H), 124.0 (Ar-*C*H), 120.7 (*C*H=CH_2), 115.44 (CH=*C*H_2), 78.0 (q*C*), 70.0 (α-*C* pyrrolidine), 63.4 (*C*H_2Ar), 57.0 (δ-*C*), 39.8 (β-*C*), 33.7 (β-*C* pyrrolidine), 30.6 (δ-*C* pyrrolidine), 29.6 (*C*H_3), 25.3 (γ-*C* pyrrolidine), 23.3 (γ-*C*). $[\alpha]_D^{20}$ +2729 (c 0.05, $CHCl_3$).



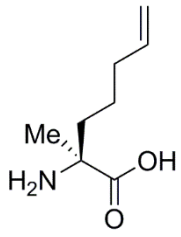
(*S*)-{({2-[1-(2-Fluorobenzyl)pyrrolidine-2-carboxamide]phenyl} phenylmethylene)-(S)-6-heptenoatoN,N',N'',O}-nickel(II) (4b) Sodium hydroxide (34 mg, 0.85 mmol, 4.5 eq.) was placed into an oven dried 10 mL round bottom flask and placed back in the oven. (*S*)-Ala-Ni-2FBPB (100 mg, 0.19 mmol, 1 eq.) was added and the flask purged with copious amounts of nitrogen. Dry DMF (500 µL) was added and the reaction heated at 50 °C for 5 mins to activate the enolate. 5-bromopentene (70 µL, 0.57 mmol, 3 eq.) was dissolve in dry DMF (260 µL) then added to the reaction flask. The reaction was allowed to stir at 50 °C for 2 h and the reaction followed by TLC (20% acetone/dichloromethane) The reaction was quenched with 5% acetic acid solution (1.5 mL) and concentrated. Water was added to the residue and extracted with dichloromethane (3 x 8 mL). The combined organic layers were washed with 5% lithium chloride solution (3 x 10 mL) and brine (3 x 10 mL), then dried over $MgSO_4$ before the solvent evaporated. Purification was carried out by flash column chromatography (20% acetone/dichloromethane). The fractions containing the product were combined and he solvent evaporated to give the product (71 mg, 62%) as a red powder. ^{19}F NMR (376 MHz, $CDCl_3$) δ: –113.7 ^{1}H-NMR (500 MHz, $CDCl_3$) δ 8.29 (1H, *td*, *J* = 7.4 Hz, *J* = 1.0 Hz, Ar-C*H*), 8.03 (1H, *d*, *J* = 8.5 Hz, Ar-C*H*), 7.51-7.44 (2H, *m*, Ar-C*H*), 7.38 (1H, *m*, Ar-C*H*), 7.33 (1H, *m*, Ar-C*H*), 7.29 (1H, *m*, Ar-C*H*), 7.20 (1H, *t*, *J* = 7.4 Hz, Ar-

C*H*), 7.16 (1H, *ddd*, *J* = 8.4 Hz, *J* = 6.2 Hz, *J* = 2.2 Hz, Ar-C*H*), 7.06 (1H, *t*, *J* = 9.1 Hz, Ar-C*H*), 6.97 (1H, *d*, *J* = 7.6 Hz, Ar-C*H*), 6.68-6.61 (2H, *m*, Ar-C*H*), 5.86 (1H, *ddt*, *J* = 17.0 Hz, *J* = 10.3 Hz, *J* = 6.5 Hz, C*H*=CH_2), 5.08 (1H, *dd*, *J* = 17.0 Hz, *J* = 1.0 Hz, CH=CH_{cis}H), 5.02 (1H, *d*, *J* = 10.3 Hz, CH=CHH_{trans}), 4.52 (1H, *d*, *J* = 13.1 Hz, C*H*HAr), 3.95 (1H, *d*, *J* = 13.1 Hz, CH*H*Ar), 3.60 (1H, *dd*, *J* = 9.9 Hz, *J* = 6.5 Hz, α-C*H* pyrrolidine), 3.41 (1H, *dd*, *J* = 10.7, 6.4 Hz, δ-C*H* pyrrolidine), 3.26 (1H, *m*, β-C*H* pyrrolidine), 2.78 (1H, *m*, γ-C*H* pyrrolidine), 2.52 (1 H, m, γ'-C*H* pyrrolidine), 2.40 (1H, *m*, γ-C*H*), 2.17-1.98 (5H, *m*, δ-CH_2, γ'-CH, δ-C*H* pyrrolidine, β'-C*H* pyrrolidine), 1.75-1.62 (2H, *m*, β-CH_2), 1.23 (3H, *s*, CH_3). ^{13}C NMR (126 MHz, $CDCl_3$) δ 182.3 (*C*=N), 180.1 (*C*=O), 172.4 (*C*=O), 141.5 (q*C*), 137.8 (*C*H=CH_2), 136.5 (q*C*), 134.2 (Ar-*C*H), 133.4 (Ar-*C*H), 131.6 (Ar-*C*H), 131.3 (Ar-*C*H), 130.3 (Ar-*C*H), 129.4 (Ar-*C*H), 128.7 (qC), 127.9 (Ar-*C*H), 127.3 (q*C*), 126.9 (Ar-*C*H), 124.5 (Ar-*C*H), 124.0 (Ar-*C*H), 120.8 (Ar-*C*H), 120.3 (q*C*), 116.2 (Ar-*C*H), 116.0 (Ar-*C*H), 115.4 (CH=*C*H_2), 78.1 (α-*C*), 70.1 (α-*C* pyrrolidine), 56.7 (*C*H_2Ar), 55.9 (δ-*C*), 39.8 (β-*C*), 33.7 (β-*C* pyrrolidine), 30.5 (δ-*C* pyrrolidine), 29.6 (*C*H_3), 25.2 (γ-*C* pyrrolidine), 23.2 (γ-*C*). IR (ν_{max}/cm^{-1}, neat) 2927, 2854, 1662, 1635, 1435, 1354, 1334, 1254, 1234, 1163, 1111, 1064, 765, 745, 708, 698, 673. HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{33}H_{35}N_3O_3F^{58}Ni$: 598.2016, found 598.2004 (Δ = -2.0 ppm), $[M+H]^+$ calcd for $C_{33}H_{35}N_3O_3F^{60}Ni$: 600.1970, found 600.1999 (Δ = 4.8 ppm); m.p. 190-192 °C; $[\alpha]_D^{20}$ +2271 (c 0.05 in $CHCl_3$); dr >88:12.



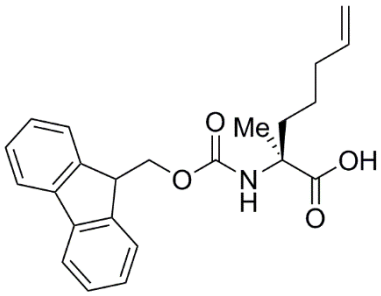
(*S*)-{({2-[1-(2-Chlorobenzyl)pyrrolidine-2-carboxamide]phenyl} phenylmethylene)-(*S*)-6-heptenoato*N*,*N*',*N*'',*O*}-nickel(II) (4c) Potassium tert-butoxide (92 mg, 0.82 mmol, 4.5 eq.) was charged into an oven dried 10 mL two-necked flask and dried with a heat gun. Dry DMF (1.8 mL, 0.1 M) was added followed by (*S*)-Ala-Ni-2CBPB (98 mg, 0.18 mmol, 1 eq.) and heated at 50 °C for 5 mins to form the enolate. 5-bromopentene (65 µL, 0.55 mmol, 3 eq.) was then added dropwise to the reaction and stirred at 50 °C for 1 h. The solvent was removed and the residue diluted with water (5 mL) and extracted with dichloromethane (3 x 5 mL). The combined organic layers were

washed with 5% lithium chloride solution (3 x 10 mL) and brine (10 mL), dried over $MgSO_4$ and the solvent evaporated. Purification was carried out by flash column chromatography (20% acetone/dichloromethane) and the fractions containing the product were combined and the solvent evaporated to give the product (28 mg, 25%) as a red powder. ^{1}H-NMR (300 MHz, $CDCl_3$) δ: 8.13 (1H, *dd*, *J* = 7.6 Hz, *J* = 1.8 Hz, Ar-C*H*), 7.88 (1H, apparent *d*, *J* = 8.3 Hz, Ar-C*H*), 7.52-7.43 (3H, *m*, Ar-C*H*), 7.42-7.31 (5H, *m*, Ar-C*H*), 7.30-7.21 (1H, *m*, Ar-C*H*), 7.15-7.01 (1H, *m*, Ar-C*H*), 6.68-6.58 (1H, *m*, Ar-C*H*), 5.86 (1H, *ddt*, *J* = 17.0 Hz, *J* = 10.2 Hz, *J* = 6.57 Hz, C*H*=CH_2), 5.14-4.99 (2H, *m*, *C*H=C*H*$_2$), 4.60 (1H, *d*, *J* = 12.9 Hz, C*H*HAr), 3.95 (1H, *d*, *J* = 12.9 Hz, CH*H*Ar), 3.64-3.53 (1H, *m*, δ-C*H*), 3.47 (1H, *dd*, *J* = 10.5 Hz, 6.6 Hz, α-C*H* pyrrolidine), 3.39-3.18 (1H, *m*, δ'-C*H*), 2.94-2.79 (2H, *m*, β-*CH*$_2$ pyrrolidine), 2.67-2.48 (1H, *m*, γ-C*H*), 2.20-1.94 (5H, *m*, γ'-C*H*, δ-*CH*$_2$ pyrrolidine, γ-*CH*$_2$ pyrrolidine), 1.79-1.62 (2H, *m*, β-*CH*$_2$), 1.56 (3H, *s*, CH_3). ^{13}C-NMR (75 MHz, $CDCl_3$) δ: 182.3 (C=N), 179.7 (C=O), 172.2 (C=O), 141.4 (q*C*), 137.8 (Ar-*C*H), 136.6 (q*C*), 135.9 (q*C*), 134.1 (Ar-*C*H), 133.4 (Ar-*C*H), 131.5 (Ar-*C*H), 131.2 (Ar-*C*H), 130.6 q*C*), 130.4 (Ar-*C*H), 129.4 (Ar-*C*H),128.9 (Ar-*C*H),128.0 (Ar-*C*H), 127.3 (q*C*), 127.1 (Ar-*C*H), 126.9 (Ar-*C*H), 123.9 (Ar-*C*H), 120.8 (*C*H=CH_2), 115.4 (CH=*C*H$_2$), 78.1 (α-*C*), 70.4 (α-*C* pyrrolidine), 60.0 (*C*H$_2$Ar), 57.0 (δ-*C*), 39.7 (β-*C*), 33.7 (β-*C* pyrrolidine), 30.5 (δ-*C* pyrrolidine), 29.7 (*C*H$_3$), 25.2 (γ-*C* pyrrolidine), 23.2 (γ-*C*). $[\alpha]_D^{20}$ +2730 (c 0.05, $CHCl_3$).



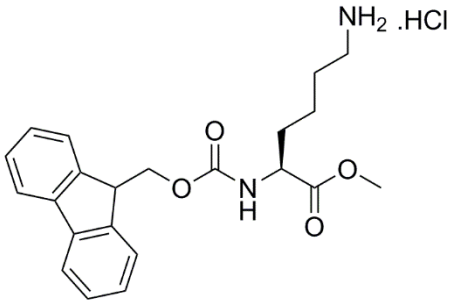
(*S*)-2-amino-2-methylhept-6-enoic acid (5) (*S*)-{({2-[1-(Benzyl)pyrrolidine-2-carboxamide]phenyl}phenylmethylene)-(*S*)-6-heptenoato*N*,*N*',*N*'',*O*}-nickel(II) (4.2 g, 7.23 mmol, 1 eq.) was dissolved in methanol (35 mL, 0.2 M) and added dropwise into 3 M hydrochloric acid (1:1, v/v, 35, 0.1 M) at 70 °C over 10 mins and stirred for 30 mins. A colour change from red to yellow/green was observed over the course of the reaction. The solvent was evaporated and the residue diluted with water (30 mL) and extracted with dichloromethane (3 x 30 mL). The combined organic layers were washed with brine (30 mL), dried over $MgSO_4$ and the solvent evaporated to give BPB (2.65 mg, 96%). Meanwhile the aqueous was loaded onto a Dowex column preconditioned with EtOH/H_2O (1:1)

and washed through with more EtOH/H_2O (1:1). The product was eluted with 10% NH_4OH/EtOH solution. The fractions containing the product were concentrated and lyophilized to give (*S*)-2-amino-2-methylhept-6-enoic acid (715 mg, 65%) as an off white powder. ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.84 (1H, *ddt*, *J* = 17.0 Hz, *J* = 10.2 Hz, *J* = 6.7 Hz, C*H*=CH_2), 5.11-4.96 (2H, *m*, CH=C*H*$_2$), 2.07 (2H, *q*, *J* = 7.0 Hz, δ-C*H*$_2$), 1.86 (1H, *m*, β-C*H*), 1.71 (1H, *m*, β'-C*H*), 1.54 -1.40 (4H, *m*, C*H*$_3$, γ-C*H*), 1.31 (1H, *m*, γ'-C*H*). ^{13}C NMR (100 MHz, $CDCl_3$) δ 177.0 (*C*=O), 138.7 (*C*H=CH_2), 115.1 (CH=*C*H_2), 61.6 (α-*C*), 36.7 (β-*C*), 32.8 (δ-*C*), 22.5 (γ-*C*), 22.5 (*C*H_3). IR (ν_{max}/cm^{-1}, neat) 3124, 3033, 2982, 1738, 1595, 1456, 1397, 1367, 1316, 1232, 912. HRMS (m/z) $[M+H]^+$ calcd for $C_8H_{16}NO_2$: 158.1181, found 158.1175 (Δ = –3.8 ppm); m.p. 250-252 °C; $[\alpha]_D^{20}$ +8.1 (c 0.05 in MeOH), Lit. +10.8 (c 0.83, MeOH).[248]

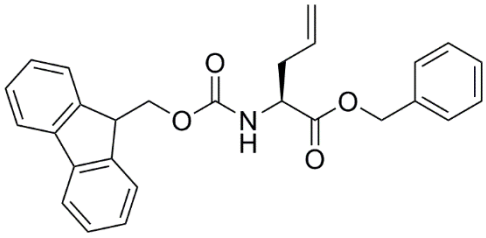


(*S*)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-2-methylhept-6-enoic acid (Fmoc-S$_5$-OH, 6) (*S*)-2-amino-2-methylhept-6-enoic acid (90 mg, 0.57 mmol, 1 eq.) was dissolved in water (2.85 mL, 0.2 M) and sodium hydrogen carbonate (96 mg, 1.14 mmol, 2 eq.) added. The solution was cooled to 0 °C and Fmoc-OSu (212 mg, 0.63 mmol, 1.1 eq) dissolved in acetone (2.85 mL, 0.2 M) was added dropwise to the amino acid solution. The reaction was left to stir overnight then concentrated. Water (10 mL) was added and extracted with diethyl ether (3 x 15 mL). The aqueous layer was acidified and extracted with ethyl acetate (3 x 15 mL). The combined organic layers were dried over $MgSO_4$ and the solvent evaporated to give (S)-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-2-methylhept-6-enoic acid (177 mg, 82%) as an off white powdery gum. ^{1}H-NMR (300 MHz; CDCl3) δ 7.78 (2H, *d*, *J* = 7.5 Hz, Ar-C*H*), 7.60 (2H, *d*, *J* = 7.2 Hz, Ar-C*H*), 7.41 (2H, *t*, *J* = 7.4 Hz, Ar-C*H*), 7.33 (2H, *t*, *J* = 7.5 Hz, Ar-C*H*), 5.76 (1H, *m*, C*H*=CH_2), 5.50 (1H, *s*, N*H*), 5.09–4.89 (2H, *m*, CH=C*H*$_2$), 4.53–4.33 (2H, *m*, C(O)OC*H*$_2$CH), 4.23 (1H, *t*, *J* = 6.6 Hz, C(O)OCH$_2$C*H*), 2.22–1.95 (2H, *m*, δ-C*H*$_2$), 1.95–1.78 (1H, *m*, γ-C*H*), 1.60 (3H, *s*, C*H*$_3$), 1.50–1.26 (2H, *t*, *J* = 7.1 Hz, β-C*H*$_2$). ^{13}C NMR (75 MHz; CDCl3) δ 178.3 (*C*=O), 154.8 (q*C*), 143.8 (q*C*), 141.4 (q*C*), 138.0 (Ar-

*C*H), 127.7 (Ar-*C*H), 127.1 (Ar-*C*H), 125.0 (Ar-*C*H), 120.0 (Ar-*C*H), 115.1 (*C*H=CH_2), 66.5 (CH=*C*H_2), 59.7 (α-*C*), 47.2 (C(O)O*C*H_2CH), 36.2 (β-*C*), 33.4 (δ-*C*), 30.9 (γ-*C*), 23.3 (*C*H_3). HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{23}H_{26}NO_4$: 380.1862, found 380.1860 (Δ –0.5 ppm), $[M+Na]^+$ calcd for $C_{23}H_{25}NO_4Na$: 402.1681, found 402.1692 (Δ 2.7 ppm) $[\alpha]_D^{20}$ +3.5 (c 1.0, MeOH).

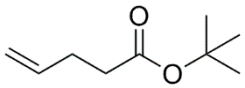


(*S*)-methyl-2-((((9H-fluoren-9-yl)methoxy)carbonyl)amino)-6-aminohexanoate hydrochloride (Fmoc-Lys-OMe.HCl, 18) Fmoc-Lys-OH hydrochloride (100 mg, 0.25 mmol, 1 equiv.) was dissolved in methanol (2.5 mL, 0.1 M) at 0 °C and under a nitrogen atmosphere. Thionyl chloride (36 µL, 0.49 mmol, 2 equiv.) was added dropwise and the reaction stirred at 0 °C for 20 mins., then heated to reflux for 45 mins. The reaction was cooled and the solvent evaporated *in vacuo* to give Fmoc-Lys-OMe hydrochloride (103 mg, 99%) as a white powder. ^{1}H NMR (300MHz) δ: (exchangeable protons not seen) 7.80 (2H, *d*, *J* = 7.6 Hz, Ar-C*H*), 7.67 (2H, *t*, *J* = 6.6 Hz, Ar-C*H*), 7.40 (2H, *t*, *J* = 7.6 Hz, Ar-C*H*), 7.31 (2H, *td*, *J* = 7.3 Hz, *J* = 1.2 Hz, Ar-C*H*), 4.47-4.31 (2H, *m*, C(O)OC*H*$_2$), 4.26-4.14 (2H, *m*, C(O)OCH$_2$C*H*, α-C*H*), 3.72 (3H, *s*, C*H*$_3$), 2.91 (2H, *t*, *J* = 7.6 Hz, ε-C*H*$_2$), 1.87 (1H, *m*, β-C*H*), 1.79-1.61 (3H, *m*, β'-C*H*, δ-C*H*$_2$), 1.52-1.39 (2H, *m*, γ-C*H*). ^{13}C NMR (75 MHz) δ: (quaternary carbon signals not seen) 128.8 (Ar-*C*H), 128.2 (Ar-*C*H), 126.2 (Ar-*C*H), 121.0 (Ar-*C*H), 67.9 (C(O)O*C*H$_2$), 55.1 (C(O)OCH$_2$*C*H), 52.8 (*C*H$_3$), 48.7 (α-*C*), 40.5 (ε-*C*), 32.0 (β-*C*), 28.0 (δ-*C*), 23.9 (γ-*C*). IR (ν_{max}/cm^{-1}, neat): 2948, 1710, 1611, 1520, 1449, 1324, 1213, 1169, 1119, 1045, 993, 759, 738. HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{22}H_{27}N_2O_4$: 383.1971, found 383.1962 (Δ -2.3 ppm); mp: 56-58 °C; $[\alpha]_D^{20}$ -1.2 (*c* 1.0, MeOH).



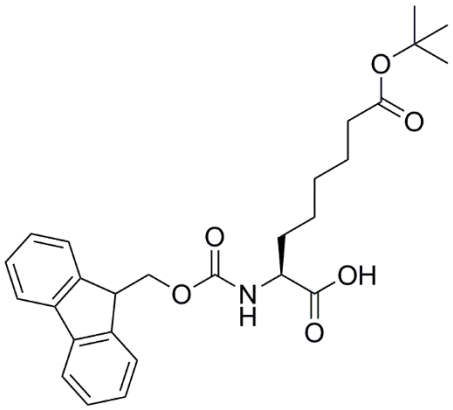
Benzyl 2-{[(9*H*-fluoren-9-ylmethoxy)carbonyl]amino}pent-4-

enoate (Fmoc-allylglycine-OBn, 33) Fmoc-allylglycine-OH (100 mg, 0.30 mmol, 1 equiv.) and sodium hydrogen carbonate (124 mg, 1.48 mmol, 5 equiv.) were taken up in *N,N*-dimethylformamide (0.6 mL, 0.5 M). Benzyl bromide (71 µL, 0.59 mmol, 2 equiv.) was added and the reaction allowed to stir at room temperature for 20 h. Water (10 mL) was added to crash out the product. The resulting solution was extracted with dichloromethane (3 x 10 mL). The combined organic layers were washed with 5% aqueous lithium chloride solution (3 x 30 mL) and brine (3 x 30 mL), dried ($MgSO_4$) and concentrated to give a colourless oil. The oil was dissolved in chloroform and left to slowly crystallise, yielding Fmoc-allylglycine-OBn as white crystals (125 mg, 97%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.78 (2H, *d*, *J* = 7.6 Hz, Ar-C*H*), 7.61 (2H, *d*, *J* = 7.3 Hz, Ar-C*H*), 7.45-7.29 (9H m, Ar-C*H*), 5.75-5.61 (1H, m, C*H*=CH_2), 5.37 (1H, *d*, *J* = 7.8 Hz, N*H*), 5.27-5.06 (4H, *m*, CH_2Ar, β-C*H*, CH=CH_2), 4.59-4.50 (1H, *m*, C(O)OCH_2C*H*), 4.41 (2H, *d*, *J* = 7.4 Hz, C(O)OCH_2CH), 4.24 (1H, *t*, *J* = 6.8 Hz, α-C*H*), 2.70-2.49 (2H, *m*, C(O)OCH_2Ph); ^{13}C NMR (100MHz, $CDCl_3$) δ 171.6 (*C*=O), 155.7 (*C*=O), 143.8 (q*C*), 141.3 (q*C*), 135.2 (*C*H=CH_2), 131.9 (Ar-*C*H), 128.6 (Ar-*C*H), 128.5 (Ar-*C*H), 128.4 (Ar-*C*H), 127.7 (Ar-*C*H), 127.0 (Ar-*C*H), 125.1 (Ar-*C*H), 120.0 (Ar-*C*H), 119.5 (CH=*C*H_2), 67.3 (*C*H_2Ar), 53.3 (C(O)O*C*H_2CH), 47.1 (α-*C*), 36.7 (C(O)OCH_2*C*H); IR (ν_{max}/cm^{-1}, neat): 3328, 3036, 2952, 1728, 1691, 1536, 1450, 1259, 1182; HRMS-ESI (*m/z*) $[M+Na]^+$ calcd for $C_{27}H_{25}NO_4Na$, 450.1681; found 450.1677 (Δ -0.89 ppm); m.p. 76-78 °C ($CHCl_3$); $[\alpha]_D^{20}$ -10.2 (c = 0.5 in $CHCl_3$).

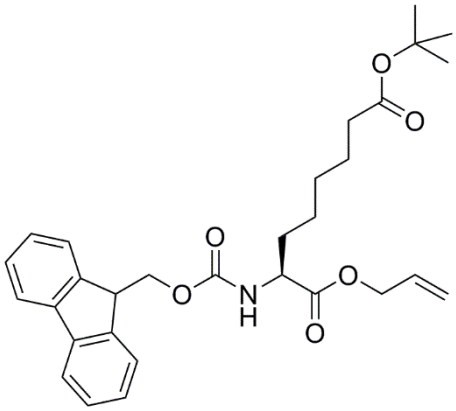


***Tert*-Butyl pent-4-enoate (32)** To a solution of 4-pentenoic acid (22 g, 220 mmol, 1 equiv.) 4-(dimethylamino)pyridine (436 mg, 3.6 mmol, 0.02 equiv.) and *t*-butanol (85 mL, 892 mmol, 4 equiv.), was added dicyclohexylcarbodiimide (62 g, 300 mmol, 1.5 equiv.) at 0 °C. The reaction was stirred at 0 °C for 25 mins, then left to stir at room temperature for 18 h. The dicyclohexylurea precipitate was removed by filtration through celite and the filtrate concentrated *in vacuo*. Purification was carried out by flash column chromatography (0-20% ethyl acetate/hexane) to give *t*-butyl pent-4-enoate (23.4 g, 63%) as a colourless oil; ^{1}H NMR (300 MHz, $CDCl_3$) δ: 5.69-5.88 (1H, *m*, C*H*=CH_2), 4.91-5.07 (2H, *m*, CH=CH_2), 2.22-2.39 (4H, *m*,

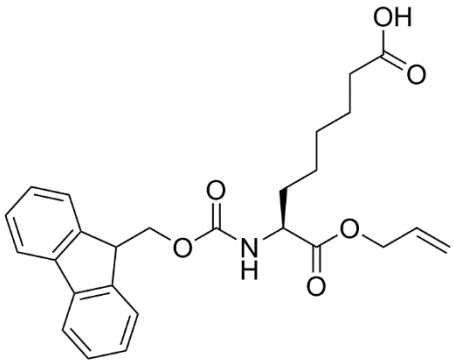
C*H*$_2$C*H*$_2$), 1.42 (9H, *s*, C*H*$_3$); ^{13}C NMR (75 MHz, $CDCl_3$) δ: 172.6 (*C*=O), 137.3 (*C*H=CH$_2$), 115.5 (CH=*C*H$_2$), 80.4 (q*C*), 34.8 (*C*H$_2$), 29.2 (*C*H$_2$), 28.2 (*C*H$_3$); IR (ν_{max}/cm^{-1}, neat): 2979, 2931, 1730, 1367, 1256, 1148. HRMS (*m/z*) $[M+H]^+$ calcd for $C_9H_{17}O_2$, 157.1229; found 157.1230 (Δ 0.64 ppm).



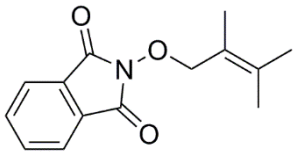
8-*tert*-Butoxy-2-{[(9*H*-fluoren-9-ylmethoxy)carbonyl]amino}-8-oxooctanoic acid (Fmoc-octanoate(*t*Bu)-OH, 30) To a solution of Fmoc-allylglycine-OBn (1.0 g, 4.66 mmol, 1 equiv.) and ^{t}Bu pent-4-enoate, (1.99 g, 23.3 mmol, 5 equiv.) in dichloromethane (24 mL, 0.1 M) was added Grubbs 2nd generation catalyst (100 mg, 0.23 mmol, 5 mol%). The solution was heated to reflux for 2 h, cooled to room temperature and the solvent evaporated *in vacuo*. The residue was taken up in ethanol (24 mL, 0.1 M) followed by the addition of Pd/C (20 mg, 10% loading). The flask was flushed with hydrogen then left under a positive hydrogen pressure for 72 hours at room temperature. The solution was filtered through celite and the solution concentrated *in vacuo*. Purification was carried out using flash column chromatography (0-5% methanol/dichloromethane) to give Fmoc-octanoate(*t*Bu)-OH as an off white powder (774 mg, 36%). ^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.77 (2H, *d*, *J* = 7.6 Hz, Ar-C*H*), 7.59 (2H, *d*, *J* = 7.3 Hz, Ar-C*H*), 7.39 (2H, *t*, *J* = 7.3 Hz, Ar-C*H*), 7.33 (2H, *t*, *J* = 7.0 Hz, Ar-C*H*), 4.42 (2H, *d*, *J* = 6.7 Hz, C(O)OC*H*$_2$CH), 4.23 (1H, *t*, *J* = 6.7 Hz, α-C*H*), 2.22 (2H, *td*, *J* = 7.0 Hz, *J* = 2.3 Hz, ζ-C*H*$_2$), 1.82-2.00 (1H, *m*, β-C*H*), 1.67-1.81 (1H, *m*, β'-C*H*), 1.50-1.67 (2H, *m*, ε-C*H*$_2$,), 1.44-1.47 (9H, *m*, C*H*$_3$), 1.28-1.42 (4H, *m*, γ-C*H*$_2$, δ-C*H*$_2$); ^{13}C NMR (500 MHz, $CDCl_3$) δ: 176.5 (*C*=O), 173.3 (*C*=O), 156.1 (*C*=O), 143.7 (q*C*), 141.3 (q*C*), 127.7 (Ar-*C*H), 127.0 (Ar-*C*H), 125.1 (Ar-*C*H), 120.0 (Ar-*C*H), 80.2 (q*C*), 67.1 (C(O)O*C*H$_2$CH), 53.7 (α-*C*), 47.2 (C(O)OCH$_2$*C*H), 35.4 (ζ-*C*), 33.6 (β-*C*), 28.6 (δ-*C*), 28.1 (*C*H$_3$), 24.9 (γ-*C*), 24.7 (ε-*C*); IR (ν_{max}/cm^{-1}, neat) cm^{-1}: 3331, 2932, 2858, 1715, 1526, 1449, 1211, 1151; HRMS (*m/z*) $[M+H]^+$ calcd for $C_{27}H_{34}NO_6$, 468.2386; found 468.2383 (Δ -0.64 ppm). $[\alpha]_D^{20}$ -9.6 (c = 0.5 in $CHCl_3$).



8-*tert*-butyl 1-prop-2-en-1-yl 2-{[(9*H*-fluoren-9-ylmethoxy)carbonyl]amino} octanedioate (Fmoc-octanoate(*t*Bu)-OAllyl, 29) To a suspension of Fmoc-octanoate(*t*Bu)-OH (100 mg, 0.21 mmol, 1 equiv.) and sodium hydrogen carbonate (81 mg, 0.96 mmol, 4.5 equiv.) in *N,N*-dimethylformamide (3.6 mL, 0.06 M) was added allyl bromide (23 µL, 0.27 mmol, 1.25 equiv.). The reaction was allowed to stir at room temperature for 18h, and then concentrated *in vacuo.* The residue was taken up in ethyl acetate (5 mL) and washed with 5% aqueous lithium chloride solution (3 x 5 mL) followed by brine (3 x 5 mL). The organics were combined, dried ($MgSO_4$) and concentrated *in vacuo*. Purification was carried out by flash column chromatography (20% ethyl acetate/hexane) to give Fmoc-octanoate(*t*Bu)-OAllyl (85 mg, 79%) as a colourless gum; ^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.77 (2H, *d*, *J* = 7.3 Hz, Ar-C*H*), 7.60 (2H, *d*, *J* = 7.6 Hz, Ar-C*H*), 7.37-7.44 (2H, *m*, Ar-C*H*), 7.33 (2H, *td*, *J* = 7.6 Hz, *J* = 1.2 Hz, Ar-C*H*), 5.92 (1H, *ddt*, *J* = 5.6 Hz, *J* = 11.4 Hz, *J* = 16.4 Hz, C*H*=CH_2), 5.24-5.40 (2H, *m*, CH=C*H*$_2$), 4.66, (2H, *d*, *J* = 5.6 Hz, C*H*$_2$CH=CH_2), 4.36-4.47 (3H, *m*, C(O)OC*H*$_2$CH, α-C*H*), 4.24 (1H, *t*, *J* = 6.7 Hz, C(O)OCH$_2$C*H*), 2.21 (2H, *t*, *J* = 7.5 Hz, ζ-C*H*$_2$), 1.80-1.95 (1H, *m*, β-C*H*$_2$), 1.53-1.77 (3H, *m*, β'-C*H*$_2$, ε-C*H*$_2$), 1.46 (9H, *s*, C*H*$_3$), 1.30-1.41 (4H, *m*, γ-C*H*$_2$, δ-C*H*$_2$); ^{13}C NMR (75 MHz, $CDCl_3$) δ: 173.0 (*C*=O), 172.3 (*C*=O), 155.9 (*C*=O), 143.8 (q*C*), 141.3 (q*C*), 131.2 (*C*H=CH_2), 127.7 (Ar-*C*H), 127.0 (Ar-*C*H), 125.1 (Ar-*C*H), 120.0 (Ar-*C*H), 118.9 (CH=*C*H_2), 80.0 (q*C*), 67.0 (C(O)O*C*H$_2$CH), 65.9 (*C*H$_2$CH=CH_2), 53.9 (α-*C*), 47.2 (C(O)OCH$_2$*C*H), 35.4 (ζ-*C*), 32.5 (β-*C*), 28.6 (δ-*C*), 28.1 (*C*H$_3$), 24.9 (γ-*C*), 24.8 (ε-*C*); IR (ν_{max}/cm^{-1}, neat) cm^{-1}: 3336, 2935, 1722, 1525, 1450, 1366, 1248, 1151; HRMS-ESI (m/z) $[M+Na]^+$ calcd for $C_{30}H_{37}NO_6Na$: 530.2519, found: 530.2540 (Δ 4.0 ppm). $\alpha_{[D]}^{20}$: +11.23 (c = 0.05, $CHCl_3$).

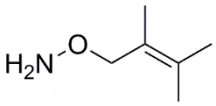


7-{[(9*H*-fluoren-9-ylmethoxy)carbonyl]amino}-8-oxo-8-(prop-2-en-1-yloxy) octanoic acid (Fmoc-oxooctanoic acid-OAllyl, 28) Fmoc-octanoate(*t*Bu)-OAllyl (75 mg, 0.15 mmol, 1 equiv.) was stirred in 90% trifluoroacetic acid/dichloromethane (1 mL) for 30 minutes then concentrated. Purification was carried out by flash column chromatography (0-3% methanol/dichloromethane) to give the product (38 mg, 57%) as a white solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.77 (2H, *d*, *J* = 7.3 Hz, Ar-C*H*), 7.61 (2H, *d*, *J* = 7.0 Hz, Ar-C*H*), 7.41 (2H, *t*, *J* = 7.0 Hz, Ar-C*H*), 7.32 (2H, *t*, *J* = 7.3 Hz, Ar-C*H*), 5.93 (1H, *ddt*, *J* = 16.4 Hz, *J* = 11.1 Hz, *J* = 5.8 Hz, C*H*=CH_2), 5.22-5.38 (2H, *m*, CH=CH_2), 4.66 (2H, *d*, *J* = 5.3 Hz, CH_2CH=CH_2), 4.42 - 4.39 (3H, *m*, C(O)OCH_2CH, α-C*H*), 4.24 (1H, *t*, *J* = 6.7 Hz, C(O)OCH_2C*H*), 2.35 (2H, *t*, *J* = 7.5 Hz, ζ-CH_2), 1.80-2.00 (1H, *m*, β-CH_2), 1.54-1.79 (3H, *m*, β'-CH_2, ε-CH_2), 1.29-1.48 (4H, *m*, γ-CH_2, δ-CH_2); ^{13}C NMR (75 MHz, $CDCl_3$) δ: 179.2 (*C*=O), 172.4 (*C*=O), 156.0 (*C*=O), 143.8 (q*C*), 141.4 (q*C*), 131.5 (*C*H=CH_2), 127.7 (Ar-*C*H), 127.1 (Ar-*C*H), 125.1 (Ar-*C*H), 120.0 (Ar-*C*H), 119.0 (CH=CH_2), 67.1 (C(O)O*C*H_2CH), 66.1 (*C*H_2CH=CH_2), 53.9 (α-*C*), 47.2 (C(O)OCH_2*C*H), 33.8 (ζ-*C*), 32.5 (β-*C*), 28.6 (δ-*C*), 24.9 (γ-*C*), 24.4 (ε-*C*); IR (ν_{max}/cm^{-1}, neat) cm^{-1}: 3320, 2934, 1700, 1525, 1449, 1191; HRMS (*m/z*) $[M+H]^+$ calcd for $C_{26}H_{30}NO_6$, 452.2073; found 452.2082 (Δ 2.00 ppm); $[\alpha]_D^{20}$ -13.9 (c = 0.5 in $CHCl_3$).



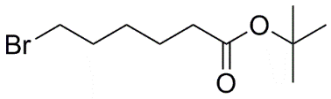
***N*-(2-methylprenyl)oxy phthalimide (54)** Tetramethylethylene (2.83 g, 33.6 mmol, 1 equiv.) was dissolved in CCl_4 (21 mL, 1.6 M) followed by the addition of *N*-bromosuccinimide (6.05 g, 33.6 mmol, 1 equiv.) and benzoyl peroxide (244 mg, 1.0 mmol, 0.03 equiv.). The resulting suspension was heated to reflux for 4 h before cooling to give a solution of 2-methylprenylbromide. *N*-hydroxyphthalimide (5.86 g, 36.0 mmol, 1.07 equiv.) was dissolved in DMSO (56 mL, 0.6 M) followed by the addition of K_2CO_3 (4.97 g, 36.0 mmol, 1.07 equiv.). The resulting mixture was vigorously stirred for 20 mins. To this mixture was added the solution of 2-methylprenylbromide in CCl_4 and the reaction

mixture was stirred at room temperature for 18 h. The resulting mixture was diluted with water (100 mL) and extracted with EtOAc (3 x 100 mL). The combined organic layers were washed with brine (300 mL), dried ($MgSO_4$) and the solvent removed *in vacuo*. Purification was carried out by flash column chromatography (20-40% EtOAc/Hexane). Fractions containing the product were combined and the solvent removed *in vacuo* to give *N*-(2-methylprenyl)oxy phthalimide (3.03 g, 37%) as a white crystalline powder. ^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.84-7.79 (2H, *m*, Ar-C*H*), 7.75-7.71 (2H, *m*, Ar-C*H*), 4.71 (2H, *s*, CH_2), 1.91 (3H, *s*, C(CH_3)=C(CH_3)$_2$), 1.76 (3H, *s*, C(CH_3)=C(CH_3)$_{cis}$(CH_3)), 1.70 (3H, *s*, C(CH_3)=C(CH_3)(CH_3)$_{trans}$). ^{13}C NMR (126 MHz, $CDCl_3$) δ: 163.8 (*C*=O), 136.3 (q*C*(CH_3)), 134.4 (Ar-*C*H), 129.2 (q*C*), 123.5 (Ar-*C*H), 122.3 (q*C*(CH_3)$_2$), 78.9 (*C*H_2), 21.2 (C(CH_3)=C(CH_3)(*C*H_3)$_{trans}$), 20.5 (C(CH_3)=C(*C*H_3)$_{cis}$(CH_3)), 17.7 (C(*C*H_3)=C(CH_3)$_2$). IR (ν_{max}/cm^{-1}, neat): 2908, 1783, 1724, 1609, 1466, 1380, 1354, 1230, 1182, 1152, 1080, 1064, 1013, 970, 919, 879, 788, 744, 700. HRMS-ESI (m/z) $[M+H]^+$ calcd for $C_{14}H_{16}NO_3$: 246.1130, found: 246.1131 (Δ 0.4 ppm).

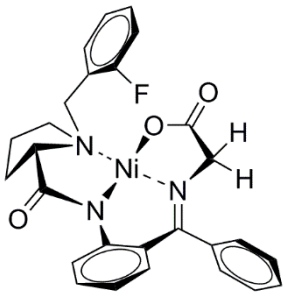


O-2-methylprenyl hydroxylamine (56) Hydrazine monohydrate (919 mg, 18.3 mmol, 3 equiv.) was dissolved in MeOH (60 mL, 0.3 M), followed by the addition of *N*-(2-methylprenyl)oxyphthalimide (1.5 g, 6.11 mmol, 1 equiv.) The solution was stirred at room temperature for 20 mins. 5% aqueous HCl (28 mL) was add to the mixture and stirred for a further 3 h. The precipitate was filtered and the filtrate diluted with water (90 mL), acidified (pH <2) and washed with Et_2O (90 mL). The organic layer was discarded. The aqueous layer was basified (pH >10), then extracted with Et_2O (2 x 100 mL). The combined ethereal layers were washed with brine (200 mL), dried ($MgSO_4$) and the solvent removed *in vacuo* to give *O*-2-methylprenyl hydroxylamine (694 mg, 99%) as a colourless oil. ^{1}H NMR (500 MHz, MeOD) δ: 4.86 (2H, *bs*, NH_2), 4.17 (2H, *s*, CH_2-1), 1.76 (3H, *s*, C(CH_3)=C(CH_3)$_2$), 1.71 (3H, *s*, C(CH_3)=C(CH_3)$_{cis}$(CH_3)), 1.70 (3H, *s*, C(CH_3)=C(CH_3)(CH_3)$_{trans}$). ^{13}C NMR (126 MHz, MeOD) δ: 133.0 (*C*(CH_3)=C(CH_3)$_2$), 124.9 (C(CH_3)=*C*(CH_3)$_2$), 77.9 (*C*H_2), 21.1 (C(CH_3)=C(CH_3)(*C*H_3)$_{trans}$), 20.5 (C(CH_3)=C(*C*H_3)$_{cis}$(CH_3)), 17.1

(C(*C*H$_3$)=C(CH$_3$)$_2$). IR (υ_{max}/cm^{-1}, neat): 2921, 2858, 1672, 1586, 1445, 1374, 1188, 1152, 980, 839, 731.

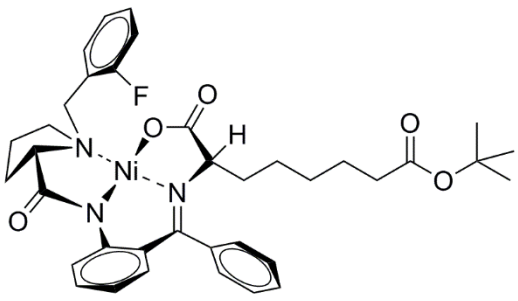


***Tert*-Butyl 6-bromohexanoate (58)** 6-Bromohexanoic acid (5.0 g, 25.6 mmol, 1, equiv.) was dissolved in dry dichloromethane (100 mL) at 0°C under a nitrogen atmosphere. *Tert*-butanol (12.3 mL, 128 mmol, 5 equiv.) was added to the solution followed by DMAP (313 mg, 2.56 mmol, 0.1 equiv.) were added and allowed to stir for 5 mins at 0 °C. DCC (5.82 g, 28.2 mmol, 1.1 equiv.) was added and the mixture warmed to room temperature, then stirred for a further 20h at room temperature. The mixture was filtered and the filtrate washed with water (100 mL). The two layers were separated and the organic layer was dried (MgSO$_4$), then the solvent removed *in vacuo*. Purification was carried out by flash column chromatography (0-20% ethyl acetate/hexane). Fractions containing the product were combined and the solvent removed *in vacuo* to give *tert*-butyl 6-bromohexanoate (4.97 g, 77%) as a colourless oil. ^{1}H NMR (500 MHz, CDCl$_3$) δ: 3.38 (2H, *t*, *J* = 6.8 Hz, ζ-C*H*$_2$), 2.20 (2H, *t*, *J* = 7.4, β-C*H*$_2$), 1.85 (2H, *tt*, *J* = 6.8 Hz, *J* = 7.3 Hz, ε-C*H*$_2$), 1.59 (2H, *tt*, *J* = 7.4 Hz, *J* = 7.8 Hz, γ-C*H*$_2$), 1.48-1.42 (2H, *m*, δ-C*H*$_2$), 1.42 (9H, *s*, C*H*$_3$). ^{13}C NMR (126 MHz, CDCl$_3$) δ: 172.9 (*C*=O), 80.2 (q*C*), 35.4 (β-*C*2), 33.6 (ζ-*C*), 32.5 (ε-*C*), 28.2 (*C*H$_3$), 27.7 (δ-*C*), 24.3 (γ-*C*).



(*S*)-({2-[1-(2-fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl} phenylmethylene)-glycinato-*N,N',N'',O*}nickel(II) (Gly-Ni-FBPB, 59) (*S*)-*N*-(2-Benzoylphenyl)-1-(2-fluorobenzyl)pyrrolidine-2-carboxamide (10 g, 24.8 mmol, 1 equiv.), nickel nitrate hexahydrate (14.5 g, 49.7 mmol, 2 equiv.) and glycine (3.7 g, 49.7 mmol, 2 equiv.) were dissolved in methanol (250 mL) at 50 °C. Powdered potassium hydroxide (9.8 g, 174 mmol, 7 equiv.) was added and the mixture heated to 70 °C for 1 hour. The reaction mixture was cooled and concentrated. The resulting residue was taken up in water (200.0 mL) and extracted with ethyl acetate (3 × 200 mL). The combined organic layers were

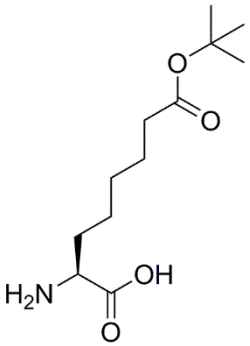
washed with saturated brine solution (3 × 600 mL), dried (MgSO4) and concentrated in vacuo to give (S)-({2-[1-(2-fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl}phenylmethylene)-glycinato-N,N',N'',O}nickel(II) (11.5 g, 90%) as a red crystalline solid. ^{19}F NMR (376 MHz, $CDCl_3$) δ: -113.6. ^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.34 (1H, *td*, *J* = 7.4 Hz, 1.6 Hz, Ar-C*H*), 8.30 (1H, *dd*, *J* = 8.7 Hz, 0.8 Hz, Ar-C*H*), 7.57-7.47 (3H, *m*, Ar-C*H*), 7.34-7.29 (1H, *m*, Ar-C*H*), 7.23 (1H, *td*, *J* = 3.7 Hz, *J* = 1.1 Hz, Ar-C*H*), 7.20 (1H, *ddt*, *J* = 8.6 Hz, *J* = 7.0 Hz, 1.6 Hz, Ar-C*H*), 7.14-7.08 (2H, *m*, Ar-C*H*), 7.00-6.96 (1H, *m*, Ar-C*H*), 6.82 (1H, *dd*, *J* = 8.3 Hz, *J* = 1.6 Hz, Ar-C*H*), 6.71 (1H, *ddt*, *J* = 8.1 Hz, *J* = 7.1 Hz, *J* = 1.1 Hz, Ar-C*H*), 4.46 (1H, *d*, *J* = 13.0 Hz, C*H*HAr), 3.92 (1H, *d*, *J* = 13.0 Hz, CH*H*Ar), 3.78 (1H, *d*, *J* = 20.1 Hz, α-C*H*), 3.68 (1H, *d*, *J* = 20.1 Hz, α'-C*H*), 3.66-3.60 (1H, *m*, δ-C*H* pyrrolidine), 3.46 (1H, *dd*, *J* = 10.8 Hz, *J* = 5.7 Hz, α-C*H* pyrrolidine), 3.42-3.30 (1H, *m*, γ-C*H*), 2.69-2.61 (1H, *m*, β-C*H* pyrrolidine), 2.52-2.41 (1H, *m*, β'-C*H* pyrrolidine), 2.17-2.05 (2H, m, γ'-C*H* pyrrolidine, δ'-C*H* pyrrolidine). ^{13}C NMR (126 MHz, $CDCl_3$) δ: 180.9 (*C*=O), 177.2 (*C*=O), 171.6 (*C*=N), 142.5 (q*C*), 134.5 (q*C*), 134.2 (q*C*), 133.2 (Ar-*C*H), 131.4 (Ar-*C*H), 129.7 (Ar-*C*H), 129.6 (Ar-*C*H), 129.3 (Ar-*C*H), 126.3 (Ar-*C*H), 125.6 (Ar-*C*H), 125.2 (Ar-*C*H), 124.5 (Ar-*C*H), 120.9 (Ar-*C*H), 120.6 (q*C*), 116.2 (Ar-*C*H), 116.0 (Ar-*C*H), 70.0 (α-*C* pyrrolidine), 61.3 (α-*C*H_2), 57.3 (δ-*C* pyrrolidine), 55.7 (*C*H_2Ar), 30.6 (β-*C* pyrrolidine), 23.7 (γ-*C* pyrrolidine). IR (ν_{max}/cm^{-1}, neat): 3056, 2974, 1679, 1633, 1589, 1547, 1489, 1471, 1440, 1365, 1332, 1303, 1257, 1173, 1063, 966, 929, 862, 812, 757, 727, 704, 688, 658. HRMS (ESI): $[M+Na]^+$ calcd for $C_{27}H_{24}N_3O_3F^{58}NaNi$: 538.1053, found: 538.1046 (Δ -1.3 ppm); $[M+H]^+$ calcd for $C_{27}H_{25}N_3O_3F^{58}Ni$: 516.1233, fround: 516.1240 (Δ +1.4 ppm); m.p.: 224-225 °C (hexane-ethyl acetate). $\alpha_{[D]}^{20}$: +2111.6 (c = 0.05, $CHCl_3$).



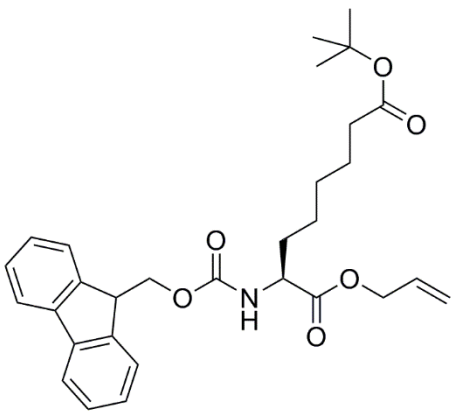
(*S*)-({2-[1-(2-fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl} phenylmethylene)-(*S*)-(tert-butyl hexanoate) glycinato-*N,N',N'',O*}nickel(II) (60) Potassium *tert*-butoxide (87 mg, 0.77 mmol, 4 equiv.) was taken up in dry acetonitrile (1.5 mL)

under a nitrogen atmosphere and stirred at room temperature for 10 mins. The reaction was then placed in an ice bath at 0 °C and stirred for a further 10 mins. (*S*)-({2-[1-(2-Fluorobenzyl)- pyrrolidine-2-carboxamide]phenyl}phenylmethylene)-glycinato- *N*,*N*',*N*'',*O*}nickel(II) (100 mg, 0.19 mmol, 1 equiv.) was added and stirred at 0 °C for 2 mins. The solution darkened in colour and the ice bath removed. A solution of ^{t}Bu 6-bromohexanoate (146 mg, 0.58 mmol, 3 equiv.) dissolved in acetonitrile (0.5 mL) was added to the reaction mixture. The solution was left to stir at room temperature for 15 mins. The reaction was quenched with 5% aqueous acetic acid solution (2 mL). The mixture was concentrated *in vacuo* and water (10 mL) was added to the residue. Dichloromethane (10 mL) was added and the two layers separated. The aqueous layer was extracted with dichloromethane (3 x 10 mL). The combined organic layers were washed with brine (3 x 300 mL) dried ($MgSO_4$) and concentrated *in vacuo*. Purification was carried out by flash column chromatography (20-70% ethyl acetate/dichloromethane). Fractions containing the product were combined and the solvent removed to give (*S*)-({2-[1-(2-fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl}phenylmethylene)-(S)-(tert-butyl hexanoate)glycinato-*N,N',N'',O*}nickel(II) (104 mg, 80%) as red crystalline powder. ^{19}F NMR (376 MHz, $CDCl_3$) δ: -113.9. ^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.27 (1H, *td*, *J* = 7.4 Hz, *J* = 1.7 Hz, Ar-C*H*), 8.13 (1H, *d*, *J* = 8.4 Hz, Ar-C*H*), 7.51-7.44 (2H, *m*, Ar-C*H*), 7.44-7.39 (1H, *m*, Ar-C*H*), 7.24-7.20 (1H, *m*, Ar-C*H*), 7.20-7.15 (1H, *m*, Ar-C*H*), 7.14-7.09 (2H, *m*, Ar-C*H*), 7.00 (1H, *ddt*, *J* = 9.7 Hz, *J* = 8.4 Hz, *J* = 1.2 Hz, Ar-C*H*), 6.89 (1H, *d*, *J* = 7.5 Hz, Ar-C*H*), 6.66-6.59 (2H, *m*, Ar-C*H*), 4.40 (1H, *d*, *J* = 13.1 Hz, C*H*HAr), 3.86 (1H, *dd*, *J* = 8.2 Hz, *J* = 3.4 Hz, α-C*H*), 3.80 (1H, *d*, *J* = 13.1 Hz, CH*H*Ar), 3.59-3.48 (1H, *m*, δ-C*H* pyrrolidine), 3.48-3.41 (2H, m, α-C*H* pyrrolidine, β-C*H* pyrrolidine), 2.85-2.76 (1H, *m*, γ-C*H* pyrrolidine), 2.58-2.48 (1H, *m*, γ'-C*H* pyrrolidine), 2.19-2.09 (3H, *m*, ζ-CH_2, δ-C*H* pyrrolidine), 2.09-1.99 (2H, *m*, γ-C*H*, β'-C*H* pyrrolidine), 1.99-1.88 (1H, *m*, β-C*H*) 1.63-2.54 (2H, *m*, β'-C*H*, γ'-C*H*), 1.54-1.46 (2H, *m*, ε-CH_2), 1.39 (9H, *s*, CH_3), 1.19-1.06 (2H, *m*, δ-CH_2). ^{13}C NMR (126 MHz, $CDCl_3$) δ: 180.1 (*C*=O), 179.3 (*C*=O), 172.9 (*C*=O), 170.4 (*C*=N), 142.3 (q*C*), 134.2 (Ar-*C*H), 133.8 (q*C*), 133.3 (Ar-*C*H), 132.2 (Ar-*C*H), 131.2 (Ar-*C*H), 129.7 (Ar-*C*H), 128.9 (Ar-*C*H), 127.6 (Ar-*C*H), 127.2 (Ar-*C*H), 126.6 (q*C*), 124.5 (Ar-*C*H), 123.7 (Ar-*C*H), 120.8 (Ar-*C*H), 120.4 (q*C*), 120.3 (q*C*), 116.2 (Ar-*C*H), 116.0 (Ar-*C*H),

80.0 (q*C*), 70.4 (α-*C*), 70.3 (α-*C* pyrrolidine), 56.8 (β-*C* pyrrolidine), 55.6 (*C*H$_2$Ar), 35.5 (ζ-*C*), 35.3 (β-*C*), 31.6 (q*C*(CH$_3$)$_3$), 30.7 (γ-*C* pyrrolidine), 28.7 (δ-*C*), 28.1 (*C*H$_3$), 25.1 (γ-*C*), 24.8 (ε-*C*), 23.7 (δ-*C* pyrrolidine). IR (ν_{max}/cm^{-1}, neat): 2933, 2862, 2161, 1722, 1673, 1637, 1587, 1544, 1492, 1440, 1365, 1335, 1255, 1163, 1144, 1111, 1062, 937, 847, 752, 708, 687. HRMS-ESI (m/z) [M+Na]$^+$ calcd for $C_{37}H_{43}N_3O_5F^{58}NaNi$: 708.2360, found: 708.2368 (Δ 1.1 ppm); [M+H]+ calcd for $C_{37}H_{43}N_3O_5F^{58}Ni$: 686.2540, found: 686.2518 (Δ -1.2 ppm); m.p.: 81-83°C (hexane-ethyl acetate); $\alpha_{[D]}^{20}$: +2042.6 (c = 0.05, CHCl$_3$).

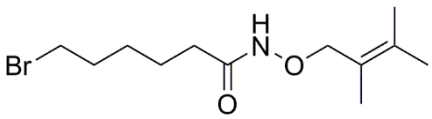


Aminosuberic acid 8-*tert*-butyl ester (61) (*S*)-({2-[1-(2-Fluorobenzyl)benzyl) pyrrolidine-2-carboxamide]-phenyl}phenylmethylene)-(*S*)-(tert-butyl hexanoate) glycinato-*N,N',N'',O*}nickel(II) (500 mg, 0.73 mmol, 1 equiv.) and 8-hydroxyquinoline (239 mg, 1.65 mmol, 2 equiv.) were taken up in acetonitrile (11 mL). Water (1.5 mL) was added and the mixture heated to 30 °C for 18 h. The mixture was filtered and the precipitate washed with water. The filtrate was extracted with ethyl acetate (3 x 20 mL). The aqueous layer was lyophilised to give (*S*)-2-amino-8-(*tert*-butoxy)-8-oxooctanoic acid (116 mg, 65%) as a white fluffy powder. ^{1}H NMR (500 MHz, D$_2$O) δ: 3.73 (1H, *t*, *J* = 6.13 Hz, α-C*H*), 2.31 (2H, *t*, *J* = 7.35 Hz, ζ-C*H*$_2$), 1.93-1.79 (2H, *m*, β-C*H*$_2$), 1.60 (2H, tt, *J* = 7.22 Hz, 7.22 Hz, ε-C*H*$_2$), 1.46 (9H, *s*, C*H*$_3$), 1.42-1.33 (4H, *m*, γ-C*H*$_2$, δ-C*H*$_2$). ^{13}C NMR (126 MHz, D$_2$O) δ: 176.6 (*C*=O), 174.9 (*C*=O), 82.4 (qC), 54.8 (α-*C*), 35.1 (ζ-*C*), 30.2 (β-*C*), 27.8 (γ-*C*), 27.2 (*C*H$_3$), 24.2 (ε-*C*), 24.0 (δ-*C*). IR (ν_{max}/cm^{-1}, neat): 2937, 2354, 2731, 1579, 1516, 1408, 1366, 1247, 1159, 850, 664. HRMS-ESI (m/z) [M+H]$^+$ calcd for $C_{12}H_{24}NO_4$: 246.1705, found: 246.1708 (Δ = 1.2 ppm); m.p.: 212-214 °C; $\alpha_{[D]}^{20}$: +2.87 (c = 0.05, H$_2$O).



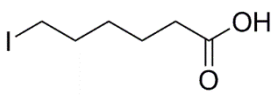
8-*tert*-Butoxy-2-{[(9*H*-fluoren-9-ylmethoxy)carbonyl]amino}-8-oxooctanoic acid (65) Aminosuberic acid 8-*tert*-butyl ester

(75 mg, 0.31 mmol, 1 equiv.) and potassium carbonate (85 mg, 0.61 mmol, 2 equiv.) were taken up in water (1 mL) and cooled to 0 °C. Fmoc succinimide (155 mg, 0.46 mmol, 1.5 equiv.) was dissolved in dioxane (2 mL) and added dropwise to the aqueous solution over 20 mins. The reaction was warmed to room temperature and left to stir for 24 h. Excess water was added to the mixture then extracted with ethyl acetate (3 x 10 mL). The combined organic layers were back extracted with sodium bicarbonate (3 x 20 mL) then theaqueous acidified to pH 1 with 3M hydrochloric acid. The aqueous fractions were then extracted with ethyl acetate (3 x 30 mL). The combined organic layers were washed with brine (3 x 30 mL), dried ($MgSO_4$) and concentrated *in vacuo*. Purification was carried out by flash column chromatography (0-7% methanol/dichloromethane + 0.1% acetic acid). Fractions containing the product were combined and the solvent removed *in vacuo* to give Fmoc-aminosuberic acid 8-*tert*-butyl ester (113 mg, 78%) as an off white powder. ^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.76 (2H, *d*, *J* = 7.5 Hz, Ar-C*H*), 7.62-7.56 (2H, *m*, Ar-C*H*), 7.40 (2H, *t*, *J* = 7.5 Hz, Ar-C*H*), 7.31 (2H, *t*, *J* = 7.5 Hz, Ar-C*H*), 4.45-4.36 (3H, *m*, α-C, C(O)OC*H*$_2$CH), 4.23 (1H, *t*, *J* =6.9 Hz, C(O)OCH$_2$C*H*), 2.21 (2H, *t*, *J* =7.4 Hz, ζ-C*H*$_2$), 1.95-1.86 (1H, *m*, β-C*H*), 1.77-1.67 (1H, *m*, β'-C*H*), 1.64-1.53 (2H, *m*, γ-C*H*$_2$), 1.44 (9H, *s*, C*H*$_3$), 1.43-1.25 (4H, *m*, δ-C*H*$_2$, ε-C*H*$_2$). ^{13}C NMR (126 MHz, $CDCl_3$) δ: 176.3 (*C*=O), 173.4 (*C*=O), 156.2 (*C*=O), 143.9 (q*C*), 141.5 (q*C*), 127.9 (Ar-*C*H), 127.2 (Ar-*C*H), 125.2 (Ar-*C*H), 120.1 (Ar-*C*H), 80.4 (qC(CH$_3$)$_3$), 67.2 (C(O)O*C*H$_2$CH), 53.8 (α-*C*), 47.3 (C(O)OCH$_2$*C*H), 35.5 (ζ-*C*), 32.3 (β-*C*), 28.7 (δ-*C*), 28.3 (*C*H$_3$), 25.0 (γ-*C*), 24.9 (ε-*C*). IR (ν_{max}/cm^{-1}, neat): 3333, 2935, 1705, 1523, 1450, 1367, 1247, 1149, 1102, 1053, 845, 758, 738, 667. $\alpha_{[D]}^{20}$: +2.81 (c = 0.05, $CHCl_3$).

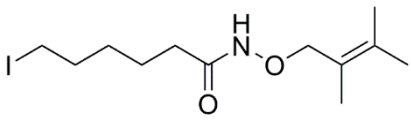


6-Bromo-*N*-((2-methylprenyl)oxy)hexanamide (70) 6-bromohexanoic acid (1.0 g, 5.2 mmol, 1 equiv.) was dissolved in dichloromethane (4 mL). To the solution was added thionyl chloride (751 µL, 10.4 mmol, 2 equiv.) and the resulting mixture heated to 50 °C for 1h. The mixture was cooled and concentrated *in vacuo* to give 6-bromohexanoyl chloride, which was used without further purification. *O*-(2-

methylprenyl)hydroxylamine (1.19 g, 10.4 mmol, 2 equiv.) and potassium hydroxide (581 mg, 10.4 mmol, 2 equiv.) were dissolved in tetrahydrofuran/water (8 mL, 1:1, v/v) and placed in an ice bath at 0 °C. 6-Bromohexanoyl chloride was added to the reaction. The mixture was brought to room temperature and stirred for 1h, then concentrated *in vacuo*. The residue was taken up in water (10 mL) and extracted with dichloromethane (3 x 10 mL). The combined organic layers were washed with brine (3 x 30 mL), dried ($MgSO_4$) and concentrated *in vacuo*. Purification was carried out by flash column chromatography (5-40% ethyl acetate/hexane). Fractions containing the product were combined and the solvent removed to give 6-bromo-*N*-((2-methylprenyl)oxy)hexanamide (137 mg, 47%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.13 (1H, bs, N*H*), 4.36 (2H, *bs*, C*H*$_2$C(CH_3)=C(CH_3)$_2$), 3.39 (2H, *t*, *J* = 6.6 Hz, ζ-C*H*$_2$), 2.07 (2H, *t*, *J* = 6.9 Hz, β-C*H*$_2$), 1.85 (2H, *tt*, *J* = 6.9 Hz, *J* = 6.9 Hz, ε-C*H*$_2$), 1.75 (6H, *s*, C(CH_3)=C(C*H*$_3$)$_2$), 1.70 (3H, *s*, C(C*H*$_3$)=C(CH_3)$_2$), 1.68-1.60 (2H, *m*, γ-C*H*$_2$), 1.56-1.38 (2H, *m*, δ-C*H*$_2$). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 170.4 (*C*=O), 132.0 (q*C*(CH*3*)), 123.9 (q*C*(CH_3)$_2$), 77.0 (*C*H$_2$C(CH_3)=C(CH_3)$_2$), 33.7 (β-*C*), 33.1 (ζ-*C*), 32.5 (ε-*C*), 27.7 (δ-*C*), 24.5 (γ-*C*), 21.1 (C(*C*H$_3$)=C(CH_3)$_2$), 20.5 (C(CH_3)=C*C*H$_{3cis}$), 17.0 (C(CH_3)=C*C*H$_{3trans}$).



6-iodohexanoic acid (76) 6-bromohexanoic acid (15 g, 77 mmol, 1 equiv.) and sodium iodide (58 g, 389 mmol, 5 equiv.) were taken up in acetone (125 mL) and heated to reflux for 18 h. The reaction was cooled, then ethyl acetate (125 mL) was added. The resulting mixture was washed with water (3 x 250 mL), brine (3 x 250 mL), dried ($MgSO_4$) and the solvent evaporated to give 6-iodohexanoic acid (16 g, 87%) as a white solid. ^{1}H NMR (300 MHz, $CDCl_3$) δ: 3.20 (2H, *t*, *J* = 7.1 Hz, ζ-C*H*$_2$), 2.39 (2H, *t*, *J* = 7.4 Hz, β-C*H*$_2$), 1.85 (2H, *quint*, *J* = 7.2 Hz, ε-C*H*$_2$), 1.67 (2H, *quint*, *J* = 7.2 Hz, γ-C*H*$_2$), 1.53-1.40 (2H, *m*, δ-C*H*$_2$). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 179.0 (*C*=O), 33.7 (β-C*H*$_2$), 33.2 (ε-C*H*$_2$), 30.0 (δ-C*H*$_2$), 23.7 (γ-C*H*$_2$), 6.5 (ζ-C*H*$_2$).



6-iodo-*N*-((2-methylprenyl)oxy)hexanamide (77) 6-iodohexanoic acid (2.0 g, 8.3 mmol, 1 equiv.) was dissolved in dichloromethane (5 mL). To the solution was added thionyl chloride (1.2 mL, 16.5 mmol, 2 equiv.) and the resulting mixture heated to 50 °C for 1h. The mixture was cooled and concentrated *in vacuo* to give 6-iodohexanoyl chloride, which was used without further purification. *O*-(2-methylprenyl)hydroxylamine (1.9 g, 16.5 mmol, 2 equiv.) and potassium hydroxide (927 mg, 16.5 mmol, 2 equiv.) were dissolved in tetrahydrofuran/water (12 mL, 1:1, v/v) and placed in an ice bath at 0 °C. 6-Iodohexanoyl chloride was added to the reaction. The mixture was brought to room temperature and stirred for 1h, then concentrated *in vacuo*. The residue was taken up in water (25 mL) and extracted with dichloromethane (3 x 25 mL). The combined organic layers were washed with brine (3 x 75 mL), dried ($MgSO_4$) and concentrated *in vacuo*. Purification was carried out by flash column chromatography (10-70% ethyl acetate/hexane). Fractions containing the product were combined and the solvent removed to give 6-iodo-*N*-((2-methylprenyl)oxy)hexanamide (1.8 g, 66%) as a colourless oil. ^{1}H NMR (500 MHz, $CDCl_3$) δ: 8.07 (1H, bs, N*H*), 4.44 (2H, *bs*, C*H*$_2$C(CH_3)=C(CH_3)$_2$), 3.21 (2H, *t*, *J* = 6.9 Hz, ζ-C*H*$_2$), 2.09 (2H, *m*, β-C*H*$_2$), 1.86 (2H, *t*, *J* = 7.5 Hz, ε-C*H*$_2$), 1.79 (6H, *s*, C(CH_3)=C(C*H*$_3$)$_2$), 1.70 (3H, *s*, C(C*H*$_3$)=C(CH_3)$_2$), 1.72-1.65 (2H, *m*, γ-C*H*$_2$), 1.51-1.39 (2H, *m*, δ-C*H*$_2$). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 170.4 (*C*=O), 134.3 (q*C*(CH3)), 124.0 (q*C*(CH_3)$_2$), 79.6 (*C*H$_2$C(CH_3)=C(CH_3)$_2$), 33.1 (β-*C*), 33.0 (ε-*C*), 30.0 (δ-*C*), 24.6 (γ-*C*), 21.0 (C(CH_3)=C*C*H$_{3cis}$), 20.4 (C(CH_3)=C*C*H$_{3trans}$), 14.3 (C(*C*H$_3$)=C(CH_3)$_2$), 6.6 (ζ-*C*).

5.3. Peptide syntheses

The peptides for Chapter 2 were synthesized on a CEM Liberty1 automated microwave-assisted solid-phase peptide synthesizer (CEM Corporation) using a 30 mL Teflon reactor vessel at either the 50 or 100 μmol scale using Rink Amide MBHA resin (substitution: 0.6-0.8 mmol/g).

The hydroxamic acid containing peptide for Chapter 3 was synthesised on a Biotage Initiator+ Alstra machine on Rink Amide ChemMatrix resin (1.05 mmol/g substitution).

Histone H3 tail and SNAIL-H3 peptides for Chapter 4 were synthesised on a Biotage Initiator+ Alstra machine on Rink Amide SpheriTide resin (1.05 mmol/g substitution).

5.3.1. General procedure for automated peptide synthesis

Fmoc-protected amino acids were made up as a solution of 0.2 M in DMF to give 5 equivalents relative to the resin once added to the reaction vessel. For the 50 µmol scale, the activator was made up to 0.25 M HCTU in DMF and the activator base was made up to 1 M DIPEA in NMP and for the 100 µmol scale, the activator was made up to 0.5 M HCTU in DMF and the activator base was made up to 2 M DIPEA in NMP. Both these sets of solutions allowed for 5 equivalents of activator and 10 equivalents of activator base relative to the resin. Deprotection solution was made up to a 20% v/v piperidine in DMF solution for peptides in Chapter 2. Deprotection solution for peptides for Chapters 3 and 4 was made up to a 20% v/v piperidine in DMF solution with 0.1M OxymaPure. Double coupling of Fmoc-Arg(Pbf)-OH was carried out at room temperature for 25 mins, then heated to 75 °C for 5 mins, drained and a second coupling step carried out at 75 °C for 5 mins. Room temperature double couplings were carried out for 60 mins. All other amino acids were coupled at 75 °C for 5 mins. Deprotections were carried out at 75 °C for 30 s, followed by a second deprotection at 75 °C for 3 mins.

5.3.2. General procedure for manual peptide synthesis

Fmoc-deprotection: Deprotection was carried out twice with a solution of 20 % v/v piperidine in DMF (or 20 % v/v piperidine in DMF with 0.1 M OxymaPure) with gentle rocking for 15 min, followed by sequential washes of the resin with DMF (3 x 3 mL), MeOH (3 x 3 mL) and DCM (3 x 3 mL).

Amino acid coupling: moc-protected amino acid (2 equivalents relative to the resin) was dissolved in DMF (0.05M) followed by addition of HCTU (2 equivalents relative to the resin) and DIPEA (4 equivalents relative to the resin). The resulting solution was allowed to activate for 5 min before addition to the prepared resin. The resin suspension was gently rocked for 2 h and then the resin drained and washed sequentially with DMF (3 x 3 mL), MeOH (3 x 3 mL) and DCM (3 x 3 mL), followed by gentle drying under a steady stream of nitrogen.

Cleavage test: A small number of dried resin beads were taken and cleaved with TFA/TES/water (95:2.5:2.5) for 1 h. The filtrate was drained, concentrated and diluted with 50% acetonitrile/water, then analysed by LCMS.

5.3.3. Grubbs ring closing metathesis

Peptide on resin (1 eq.) was swollen in DMF for 30 mins and drained. The resin was washed with three times each with DMF, MeOH and DCM. Grubbs' 1st generation catalyst (0.2 eq.) was dissolved in DCM (0.1 M) and added to the prepared resin. The resulting resin suspension was gently rocked for 2 h and drained. The resin was washed with DMF until the DMF ran clear. A second batch of newly dissolved Grubbs' 1st generation catalyst (0.2 eq.) in DCM (0.1 M) was added to the resin and gently rocked for 2 h then drained. The resin was washed three times each with DMF, MeOH and DCM before drying under a steady stream of nitrogen.

5.3.4. Synthesis of Ac-H4K16Hxa

Cleavage of Arg(Pbf)-His(Trt)-NH_2 from Sieber amide resin: Peptide on resin was treated with 1% TFA in DCM at room temperature for 3h. The filtrate was taken and the solvent removed. The residue was taken up in water and lyophilised.

Initial loading onto 2-chlorotityl-*N*-Fmoc- hydroxylamine resin: 2-Chlorotityl-*N*-Fmoc- hydroxylamine resin (0.69 mmol/g resin substitution) was swollen in DMF for 30 mins, filtered, then treated with 20% piperidine/DMF (2 x 30 min). The resin was washed with DMF (3 x 3 mL), MeOH (3 x 3 mL) and DCM (3 x 3 mL). Fmoc-oxooctanoic acid-OAllyl (1.2 equiv.) was taken up in DMF (0.1 M) followed by the addition of DIPEA (2.4 equiv.) and HBTU (1.2 equiv.). The resulting solution was allowed to activate for 2 mins before addition to the prepared resin. The resin suspension was shaken for 1h then filtered. The resin was washed with DMF (3 x 3 mL), MeOH (3 x 3 mL) and DCM (3 x 3 mL), then dried under a steady stream of nitrogen.

Allyl deprotection: Functionalised resin was swollen in DCM for 30 mins, filtered and taken up in DCM (0.1 M) and degassed for 2 mins. 1,3-Dimethylbarbituric acid (1.3 equiv.) was added and the resin suspension agitated before $Pd(PPh_3)_4$ (0.11 equiv.) was also added. The vessel was flushed with nitrogen and rocked for 1h. The resin was drained and washed with DCM until no more colour remained, then washed with DMF (3 x 3mL), and DCM (3 x 3mL), then dried under a steady stream of nitrogen.

Dipeptide coupling onto Fmoc-oxooctanoic acid-OH functionalised 2-chlorotityl-*N*-Fmoc-hydroxylamine resin: Functionalised resin was swollen in DMF for 30 mins then filtered. The resin was taken up in HBTU (1.1 equiv.) in DMF (0.5 M) with DIPEA (5 equiv.) and shaken for 5 mins. H-Arg(Pbf)-His(Trt)-NH_2 (1 equiv.) was added to the resin suspension and shaken for 1h. The resin was filtered and then washed with DMF (3 x 3 mL), MeOH (3 x 3 mL) and DCM (3 x 3 mL).

5.3.5. Cleavage from resin

Synthesized peptides were cleaved from the resin using a cleavage cocktail of TFA/TES/water (95:2.5:2.5) for 3 hours before being drained and the TFA blown off with a stream of nitrogen. The peptide was precipitated and washed

three times in cold diethyl ether and spun down to a pellet before the diethyl ether removed and the peptide dried under a steady stream of nitrogen.

5.3.6. Peptide purification

Crude peptides were purified by reverse-phase HPLC using a Dionex Ultimate 3000 system with a Phenomenex Gemini-NX 5µm C18 110Å AXIA packed column with dimensions 250 x 21.20 mm and purity confirmed by analytical reverse-phase HPLC using a Dionex Ultimate 3000 system with a Phenomenex Gemini-NX 5µm C18 110Å packed column with dimensions 150 x 4.60 mm (for the peptides in Chapter 2) or a Phenomenex Aeris 5µm Peptide XB-C18 100Å packed column with dimensions of 150 x 4.6mm (for all other peptides) and by LC-MS as above. Water was removed by lyophilisation using a FreeZone Benchtop Freeze Dry System. The purity of the peptides used in the biological assays were determined to be greater than 95 % in all instances.

5.3.7. Peptide characterisation

Table 4: Characterisation data for purified peptides. Analytical reverse-phase HPLC using a Dionex Ultimate 3000 system with a [+]Phenomenex Gemini-NX 5µm C18 110Å packed column with dimensions 150 x 4.60 mm or [#]Phenomenex Aeris 5µm Peptide XB-C18 100Å packed column with dimensions of 150 x 4.6mm, using the following gradients: [a]5-100% MeCN/H_2O (0.1% TFA), 30 min gradient, [b]5-50% MeCN/H_2O (0.1% TFA), 60 min gradient, [c]5-50% MeCN/H_2O (0.1% TFA), 15 min gradient, [d]5-200% MeCN/H_2O (0.1% TFA), 15 min gradient and [e]5-100% MeCN/H_2O (0.1% TFA), 15 min gradient.

Peptide	Calculated MW $[M+H]^+$ (Da)	Found MW $[M+H]^+$ (Da)	Δ MW (ppm)	R_t (min)	ES+ peaks (m/z)	Purity (%)	Yield (%)
H-VAESVLYYYLTK KN-NH_2 (native SMRT-DAD Helix 3 463-476, 7)	1689.9216	1689.9138	-4.6	13.753[a+]		>99	39
H-VAXSVLXYYLTKKN-NH_2 (stapled SMRT-DAD Helix 3 463-476 (E465S_5, Y469S_5), 8)	1647.9838	1647.9896	3.5	15.813[a+] 43.720[b+]		>99	14
H-VAEXVLYXYLTKKN-NH_2 (stapled SMRT-DAD Helix 3 463-476 (S466X, Y470X), 9)	1689.9944	1689.99933	-0.7	14.793[a+] 39.257[b+]		>99	12
H-VAESVLX(S_5)YYLX(S_5)KKN-NH_2 (stapled SMRT-DAD Helix 3 463-476 (Y469X(S_5), T473X(S_5)), 10)	1675.9787	1675.9790	0..2	14790[a+] 39370[b+]		>99	10

Peptide	Calculated MW $[M+H]^+$ (Da)	Found MW $[M+H]^+$ (Da)	Δ MW (ppm)	R_t (min)	ES+ peaks (m/z)	Purity (%)	Yield (%)
H-VAX(S_5)SVLX(S_5)YYLTKKN-NH_2 (Open Staple 463-476 (E465 X(S_5), Y469 X(S_5)), 11)	1676.0151	1676.0140	-0.7	20.873[a+]		87	13
H-VAX(S_5)SVLX(S_5)*p*YYLTKKN-NH_2 (stapled SMRT-DAD Helix 3 463-476 (E465X(S_5), Y469X(S_5), Y470*p*Y), 12)	1727.9501	1727.9540	2.3	16.537[a+] 26.093[b+]		>99	20
H-VAX(S_5)SVLX(S_5)Y*p*YLTKKN-NH_2 (stapled SMRT-DAD Helix 3 463-476 (E465X(S_5), Y469X(S_5), Y471pY), 13)	1727.9501	1727.9557	3.2	16.213[a+]		>99	15
Ac-KGGAHxRH-NH_2 (Ac-H4K16Hx, 51)	852.4804	852.4839	4.1	8.337[c#] 10.180[d#]		91	1
H-ARTK(Me1)QTARKSTGGKAPRKQLA-NH_2 (H3(1-21)K4me1, 79)	2267.3530			11.797[e#] 13.223[a#]	1134.5880 $[M+2H]^{2+}$ 756.7203 $[M+3H]^{3+}$ 567.8007 $[M+4H]^{4+}$ 454.4461 $[M+5H]^{5+}$ 378.6996 $[M+6H]^{6+}$	>99	45

Peptide	Calculated MW $[M+H]^+$ (Da)	Found MW $[M+H]^+$ (Da)	Δ MW (ppm)	R_t (min)	ES+ peaks (m/z)	Purity (%)	Yield (%)
H-ARTK(Me2)QTARKSTGGKAPRKQLA-NH_2 (H3(1-21)K4me2, 80)	2281.3687			11.860[e] 13.383[a]	1141.2145 $[M+2H]^{2+}$ 761.4611 $[M+3H]^{3+}$ 571.3414 $[M+4H]^{4+}$ 457.2676 $[M+5H]^{5+}$ 381.2220 $[M+6H]^{6+}$	>99	21
H-ARTK(Me1)QTARK(Ac)STGGKAPRKQLA-NH_2 (H3(1-21)K4me1K9ac, 81)	2309.3636			11.823[e] 13.300[a]	1155.7358 $[M+2H]^{2+}$ 770.8013 $[M+3H]^{3+}$ 578.3477 $[M+4H]^{4+}$ 462.8733 $[M+5H]^{5+}$ 385.7206 $[M+6H]^{6+}$	>99	23
H-ARTK(Me2)QTARK(Ac)STGGKAPRKQLA-NH_2 (H3(1-21)K4me2K9ac, 82)	2323.3793			11.890[e] 13.457[a]	1162.6917 $[M+2H]^{2+}$ 775.4574 $[M+3H]^{3+}$ 581.8487 $[M+4H]^{4+}$ 465.4766 $[M+5H]^{5+}$ 388.0627 $[M+6H]^{6+}$	>99	39

Peptide	Calculated MW $[M+H]^+$ (Da)	Found MW $[M+H]^+$ (Da)	Δ MW (ppm)	R_t (min)	ES+ peaks (m/z)	Purity (%)	Yield (%)
H-ARTK(Me2)QTARKSTGGK(Ac)APRKQLA-NH_2 (H3(1-21)K4me2K14ac, 83)	2323.3793			11.890[e] 13.447[a]	1162.7593 $[M+2H]^{2+}$ 775.4738 $[M+3H]^{3+}$ 581.8506 $[M+4H]^{4+}$ 465.6746 $[M+5H]^{5+}$ 388.0496 $[M+6H]^{6+}$	>99	23
H-PRSFLVARKSTGG-Aho (103)	1503.8760	1503.8820	4.0	12.787[e] 15.287[a]		>99	1
H-PRSFLVARKPSDPNR-Aho (104)	1640.9236	1640.9258	1.3	13.303[e] 18.193[a]		>99	7
H-PRSFLVARKSTGGKALPR-Aho (105)	1956.1613			17.603[e] 20.157[a]	979.0934$[M+2H]^{2+}$, 652.7234 $[M+3H]^{3+}$ 490.0418 $[M+4H]^{4+}$ 392.0247 $[M+5H]^{5+}$ 326.8578 $[M+6H]^{6+}$	>99	6

5.4. Circular Dichroism

Circular dichroism measurements were recorded using a Chriascan^{TM}-plus circular dichroism spectrometer (Applied Photophysics) with a 0.1 cm pathlength cuvette and a peptide concentration of 0.1 mg/mL. Circular dichroism spectra were obtained at 20 ˚C in deionised water and in 10% v/v TFE in deionised water. Each sample was done in triplicate from 180 to 260 nm and averaged. Melting curves were measured over a temperature range of 15-95 ˚C with temperature ramping of 2 ˚C and a temperature stabilization time of 3 min. The data was plotted using GraphPad Prism version 6.0 (GraphPad Software Inc.). Alpha-helical content of the peptides were calculated at 222 nm.

5.5. HDAC3/DAD expression and purification

The full length HDAC3 and residues 389-480 of SMRT, which corresponds to the DAD domain, were cloned into the pCDNA3 vector, which contains a cytomegalovirus (CMV) immediate early promoter for high level expression and an ampicillin resistance region to retain a high level of antibiotic resistance. An N-terminal 10*His-3*FLAG tag and a TEV protease cleavage site was incorporated into the DAD domain construct. Both constructs were co-transfected into HEK293F cells (Invitrogen), using 25 kDa branched polyethylenimine (PEI) (Sigma). In order to transfect 1 mL of cells at $1*10^6$ cells/mL final density, 0.5 µg DNA from each construct was diluted in 1 mL phosphate buffered saline (PBS) (Sigma), briefly vortexed, followed by the addition of 4 µL of 0.5 mg/mL PEI. It was briefly vortexed once again before incubation at room temperature for 20 min, followed by addition to the cells. In order to achieve larger transfection volumes, the protocol was scaled as necessary. The cells were harvested 48 h post transfection and used immediately or stored until required at -80 °C. Lysis of the cells was carried out by sonication in buffer containing 100 mM potassium acetate, 50 mM Tris pH 7.5, 5% v/v glycerol, 0.3% Triton X-100 and Roche complete protease inhibitor tablet (Buffer A). Insoluble material was removed by centrifugation and the protein complex was then bound to Anti-FLAG® tag resin consisting of a purified murine IgG_1 monoclonal antibody covalently linked to an agarose bead (Sigma)

at 4 °C for 1 h. The resin was washed three times with Buffer A, followed by three times with each of Buffer B (100 mM potassium acetate, 50 mM Tris pH 7.5, 5% v/v glycerol and 0.3% Triton X-100), Buffer C (300 mM potassium acetate, 50 mM Tris pH7.5 and 5% v/v glycerol) and Buffer D (50 mM potassium acetate, 50 mM Tris pH 7.5, 5% v/v glycerol and 0.5 mM TECEP). The protein complex was cleaved from the resin at 4°C overnight with Tobacco etch virus (TEV) protease in Buffer D. Following the concentration of the cleaved protein complex, it was further purified by gel filtration on a Superdex S200 column (GE Healthcare) in buffer containing 50 mM potassium acetate, 25 mM Tris pH 7.5 and 0.5 mM TECEP. The fractions were analysed by SDS-PAGE and showed the presence of a truncated form of HDAC3 along with the HDAC3/DAD protein complex. The fractions containing the protein complex were collected and concentrated and an accurate concentration determined by measuring the absorbance at 280 nm, in 6 M guanidine.

To remove the Ins(1,4,5,6)P_4 the protein complex was incubated in the presence of 1M NaCl. Protein was diluted to 0.5 µM in buffer containing 50 mM Tris pH 7.5, 5% v/v glycerol, 1 M NaCl and incubated at room temperature for 3h. The protein was then dialysed against buffer containing 50 mM NaCl, 50 mM Tris pH 7.5 and 5% v/v glycerol, at 4 °C overnight. The salt-stripped protein complex was then concentrated and the concentration accurately measured once more as above.

5.6. HDAC Activity Assays

A serial dilution of peptide was added to a 96-well plate in the presence of 50 nM final concentration of HDAC3/DAD either with or without 10 µM Ins(1,4,5,6)P_4 (in buffer containing 100 mM NaCl, 50 mM Tris pH 7.5 and 5% v/v glycerol). The 96-well plate was spun briefly then allowed to incubate at 37 °C for 30 min.

30 mM BOC acetyl-lysine substrate was diluted to 500 µM in deacetylase buffer and 10 µL aliquots of the substrate was simultaneously added to each well with the help of a multichannel pipette. The plate was spun briefly once again before further incubation at 37 °C for 30 min.

The deacetylase activity was quenched by addition of 50 µM of 2 µM Trichostatin A (TSA) in a 10µg/µL trypsin solution made up in buffer containing 50 mM Tris pH 7.5 and 100 mM NaCl. The samples were allowed to develop at room temperature for 10 min before measuring the fluorescence with an excitation wavelength of 360 nm and an emission wavelength of 470 nm on a Victor X5 plate reader (Perkin Elmer).

The HDAC activity assays were performed in triplicate and the data was plotted using GraphPad Prism version 6.0 (GraphPad Software Inc.).

6. References

1. C. Waddington, *Endeavour*, 1942, **1**, 18–20.
2. A. Eccleston, N. DeWitt, C. Gunter, B. Marte, and D. Nath, *Nature*, 2007, **447**, 395–395.
3. A. Bird, *Genes Dev.*, 2002, **16**, 6–21.
4. L. Stryer, J. M. Berg, and J. L. Tymoczko, *Biochemistry: International Edition*, W. H. Freeman, 6th edn., 2006.
5. L. Mariño-Ramírez, M. G. Kann, B. A. Shoemaker, and D. Landsman, *Expert Rev. Proteomics*, 2005, **2**, 719–729.
6. R. D. Kornberg, *Science*, 1974, **184**, 868–871.
7. R. D. Kornberg and J. O. Thonmas, *Science*, 1974, **184**, 865–868.
8. M. Vignali and J. L. Workman, *Nat. Struct. Biol.*, 1998, **5**, 1025–1028.
9. S. W. Harshman, N. L. Young, M. R. Parthun, and M. A. Freitas, *Nucleic Acids Res.*, 2013, **41**, 9593–9609.
10. D. Rossetto, N. Avvakumov, and J. Côté, *Epigenetics*, 2012, **7**, 1098–1108.
11. X. Zhang, M. Bolt, M. J. Guertin, W. Chen, S. Zhang, B. D. Cherrington, D. J. Slade, C. J. Dreyton, V. Subramanian, K. L. Bicker, P. R. Thompson, M. A. Mancini, J. T. Lis, and S. A. Coonrod, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 13331–13336.
12. M. A. Christophorou, G. Castelo-Branco, R. P. Halley-Stott, C. S. Oliveira, R. Loos, A. Radzisheuskaya, K. A. Mowen, P. Bertone, J. C. R. Silva, M. Zernicka-Goetz, M. L. Nielsen, J. B. Gurdon, and T. Kouzarides, *Nature*, 2014, **507**, 104–108.
13. A. J. Bannister and T. Kouzarides, *Cell Res.*, 2011, **21**, 381–395.
14. B. M. Turner, *Nat. Struct. Mol. Biol.*, 2005, **12**, 110–112.
15. S. S. Ng, W. W. Yue, U. Oppermann, and R. J. Klose, *Cell. Mol. Life Sci.*, 2009, **66**, 407–422.
16. P. G. Besant and P. V Attwood, *Methods Enzymol.*, 2010, **471**, 403–426.

17. Y. Zhang, *Genes Dev.*, 2003, **17**, 2733–2740.

18. J. Cao and Q. Yan, *Front. Oncol.*, 2012, **2**, 26.

19. I. L. Goldknopf, C. W. Taylor, R. M. Baum, L. C. Yeoman, M. O. Olson, A. W. Prestayko, and H. Busch, *J. Biol. Chem.*, 1975, **250**, 7182–7187.

20. M. Miranda and A. Sorkin, *Mol. Interv.*, 2007, **7**, 157–167.

21. F. Ikeda and I. Dikic, *EMBO Rep.*, 2008, **9**, 536–542.

22. B. E. Nickel, C. D. Allis, and J. R. Davie, *Biochemistry*, 1989, **28**, 958–963.

23. L. Hicke, *Nat. Rev. Mol. Cell Biol.*, 2001, **2**, 195–201.

24. G. L. Cuthbert, S. Daujat, A. W. Snowden, H. Erdjument-Bromage, T. Hagiwara, M. Yamada, R. Schneider, P. D. Gregory, P. Tempst, A. J. Bannister, and T. Kouzarides, *Cell*, 2004, **118**, 545–553.

25. R. B. Denman, *Bioessays*, 2005, **27**, 242–246.

26. V. G. Allfrey, R. Faulkner, and A. E. Mirsky, *Proc. Natl. Acad. Sci. U. S. A.*, 1964, **51**, 786–794.

27. C. Choudhary, C. Kumar, F. Gnad, M. L. Nielsen, M. Rehman, T. C. Walther, J. V Olsen, and M. Mann, *Science*, 2009, **325**, 834–840.

28. L. Zeng and M.-M. Zhou, *FEBS Lett.*, 2002, **513**, 124–128.

29. S. J. Conway, *ACS Med. Chem. Lett.*, 2012, **3**, 691–694.

30. P. Filippakopoulos and S. Knapp, *Nat. Rev. Drug Discov.*, 2014, **13**, 337–356.

31. J. Shi and C. R. Vakoc, *Mol. Cell*, 2014, **54**, 728–736.

32. C. J. Sherr, *Cell*, 2004, **116**, 235–246.

33. P. Guedes-Dias and J. M. A. Oliveira, *Biochim. Biophys. Acta*, 2013, **1832**, 1345–1359.

34. M. Dokmanovic, C. Clarke, and P. A. Marks, *Mol. Cancer Res.*, 2007, **5**, 981–989.

35. S. Ropero and M. Esteller, *Mol. Oncol.*, 2007, **1**, 19–25.

36. G. P. Delcuve, D. H. Khan, and J. R. Davie, *Clin. Epigenetics*, 2012, **4**, 5.

37. D. D. Leipe and D. Landsman, *Nucleic Acids Res.*, 1997, **25**, 3693–3697.

38. C. H. Arrowsmith, C. Bountra, P. V Fish, K. Lee, and M. Schapira, *Nat. Rev. Drug Discov.*, 2012, **11**, 384–400.

39. P. J. Watson, L. Fairall, and J. W. R. Schwabe, *Mol. Cell. Endocrinol.*, 2012, **348**, 440–449.

40. M. S. Finnin, J. R. Donigian, A. Cohen, V. M. Richon, R. A. Rifkind, P. A. Marks, R. Breslow, and N. P. Pavletich, *Nature*, 1999, **401**, 188–193.

41. J. R. Somoza, R. J. Skene, B. A. Katz, C. Mol, J. D. Ho, A. J. Jennings, C. Luong, A. Arvai, J. J. Buggy, E. Chi, J. Tang, B.-C. Sang, E. Verner, R. Wynands, E. M. Leahy, D. R. Dougan, G. Snell, M. Navre, M. W. Knuth, R. V Swanson, D. E. McRee, and L. W. Tari, *Structure*, 2004, **12**, 1325–1334.

42. P. J. Watson, L. Fairall, G. M. Santos, and J. W. R. Schwabe, *Nature*, 2012, **481**, 335–340.

43. A. Codina, J. D. Love, Y. Li, M. A. Lazar, D. Neuhaus, and J. W. R. Schwabe, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6009–6014.

44. P. J. Watson, C. J. Millard, A. M. Riley, N. S. Robertson, L. C. Wright, H. Y. Godage, S. M. Cowley, A. G. Jamieson, B. V. L. Potter, and J. W. R. Schwabe, *Nat. Commun.*, 2016, **7**, 10.1038/ncomms11262.

45. C. J. Millard, P. J. Watson, I. Celardo, Y. Gordiyenko, S. M. Cowley, C. V Robinson, L. Fairall, and J. W. R. Schwabe, *Mol. Cell*, 2013, **51**, 57–67.

46. B. W. Matthews, *Acc. Chem. Res.*, 1988, **21**, 333–340.

47. D. W. Christianson and W. N. Lipscomb, *Acc. Chem. Res.*, 1989, **22**, 62–69.

48. P. M. Lombardi, K. E. Cole, D. P. Dowling, and D. W. Christianson, *Curr. Opin. Struct. Biol.*, 2011, **21**, 735–743.

49. D. P. Dowling, S. L. Gantt, S. G. Gattis, C. A. Fierke, and D. W. Christianson, *Biochemistry*, 2008, **47**, 13554–13563.

50. A. Vannini, C. Volpari, P. Gallinari, P. Jones, M. Mattu, A. Carfí, R. De Francesco, C. Steinkühler, and S. Di Marco, *EMBO Rep.*, 2007, **8**, 879–

884.

51. D. Marchion and P. Münster, *Expert Rev. Anticancer Ther.*, 2007, **7**, 583–598.

52. J.-H. Choi, H. J. Kwon, B.-I. Yoon, J.-H. Kim, S. U. Han, H. J. Joo, and D.-Y. Kim, *Cancer Sci.*, 2001, **92**, 1300–1304.

53. A. J. Wilson, D.-S. Byun, N. Popova, L. B. Murray, K. L'Italien, Y. Sowa, D. Arango, A. Velcich, L. H. Augenlicht, and J. M. Mariadason, *J. Biol. Chem.*, 2006, **281**, 13548–13558.

54. K. Halkidou, L. Gaughan, S. Cook, H. Y. Leung, D. E. Neal, and C. N. Robson, *Prostate*, 2004, **59**, 177–189.

55. Z. Zhang, H. Yamashita, T. Toyama, H. Sugiura, Y. Ando, K. Mita, M. Hamaguchi, Y. Hara, S. Kobayashi, and H. Iwase, *Breast Cancer Res. Treat.*, 2005, **94**, 11–16.

56. Y. Toh, M. Yamamoto, K. Endo, Y. Ikeda, H. Baba, S. Kohnoe, H. Yonemasu, Y. Hachitanda, T. Okamura, and K. Sugimachi, *Oncol. Rep.*, 2003, **10**, 333–338.

57. B. H. Huang, M. Laban, C. H.-W. Leung, L. Lee, C. K. Lee, M. Salto-Tellez, G. C. Raju, and S. C. Hooi, *Cell Death Differ.*, 2005, **12**, 395–404.

58. J. Song, J. H. Noh, J. H. Lee, J. W. Eun, Y. M. Ahn, S. Y. Kim, S. H. Lee, W. S. Park, N. J. Yoo, J. Y. Lee, and S. W. Nam, *APMIS*, 2005, **113**, 264–268.

59. P. Zhu, E. Martin, J. Mengwasser, P. Schlag, K.-P. Janssen, and M. Göttlicher, *Cancer Cell*, 2004, **5**, 455–463.

60. Z. Zhang, H. Yamashita, T. Toyama, H. Sugiura, Y. Omoto, Y. Ando, K. Mita, M. Hamaguchi, S.-I. Hayashi, and H. Iwase, *Clin. Cancer Res.*, 2004, **10**, 6962–6968.

61. J. E. Bolden, M. J. Peart, and R. W. Johnstone, *Nat. Rev. Drug Discov.*, 2006, **5**, 769–784.

62. B. S. Mann, J. R. Johnson, M. H. Cohen, R. Justice, and R. Pazdur, *Oncologist*, 2007, **12**, 1247–1252.

63. R. M. Poole, *Drugs*, 2014, **74**, 1543–1554.

64. K. P. Garnock-Jones, *Drugs*, 2015, **75**, 695–704.

65. J. Hrabeta, M. Stiborova, V. Adam, R. Kizek, and T. Eckschlager, *Biomed. Pap. Med. Fac. Univ. Palacky. Olomouc. Czech. Repub.*, 2014, **158**, 161–169.

66. M. Mottamal, S. Zheng, T. L. Huang, and G. Wang, *Molecules*, 2015, **20**, 3898–3941.

67. A. M. Evens, S. Balasubramanian, J. M. Vose, W. Harb, L. I. Gordon, R. Langdon, J. Sprague, M. Sirisawad, C. Mani, J. Yue, Y. Luan, S. Horton, T. Graef, and N. L. Bartlett, *Clin. Cancer Res.*, 2016, **22**, 1059–1066.

68. MEI Pharma, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT01912274, Safety and Efficacy Study of Pracinostat With Azacitadine in Elderly Patients With Newly Diagnosed Acute Myeloid Leukemia (AML)*, 2013.

69. MEI Pharma, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT01993641, Phase 2 Study Adding Pracinostat to a Hypomethylating Agent (HMA) in Patients With MDS Who Failed to Respond to Single Agent HMA*, 2013.

70. 4SC AG, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 - Identifier NCT01277406, 4SC-201 (Resminostat) in Advanced Colorectal Carcinoma*, 2013.

71. Italfarmaco, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT00928707, Phase II Study of GIVINOSTAT (ITF2357) in Combination With Hydroxyurea in Polycythemia Vera*, 2009.

72. Italfarmaco, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT01761968, Long-term Study Evaluating the Effect of Givinostat in Patients With Chronic Myeloproliferative Neoplasms.*, 2012.

73. Italfarmaco, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT01901432, A Two-part*

Study to Assess the Safety and Preliminary Efficacy of Givinostat in Patients With Polycythemia Vera, 2013.

74. Curis, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT00728793, A Phase I Study of the Safety, Pharmacokinetics, and Anti-Tumor Activity of CUDC-101 in Patients With Advanced Solid Tumors*, 2008.

75. Curis, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT01171924, A Phase Ib Expansion Study Investigating the Safety, Efficacy, and Pharmacokinetics of Intravenous CUDC-101 in Subjects With Advanced H*, 2010.

76. Curis, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT01384799, Phase I Study of CUDC-101 With Cisplatin and Radiation in Subjects With Head & Neck Cancer*, 2011.

77. D. A. Yardley, R. R. Ismail-Khan, B. Melichar, M. Lichinitser, P. N. Munster, P. M. Klein, S. Cruickshank, K. D. Miller, M. J. Lee, and J. B. Trepel, *J. Clin. Oncol.*, 2013, **31**, 2128–2135.

78. A. Kuendgen, C. Strupp, M. Aivado, A. Bernhardt, B. Hildebrandt, R. Haas, U. Germing, and N. Gattermann, *Blood*, 2004, **104**, 1266–1269.

79. P. A. Marks, *Oncogene*, 2007, **26**, 1351–1356.

80. A. Jimeno and J. McDermott, *Drugs of Today*, 2014, **50**, 337.

81. H.-Z. Lee, V. E. Kwitkowski, P. L. Del Valle, M. S. Ricci, H. Saber, B. A. Habtemariam, J. Bullock, E. Bloomquist, Y. L. Shen, X.-H. Chen, J. Brown, N. Mehrotra, S. Dorff, R. Charlab, R. C. Kane, E. Kaminskas, R. Justice, A. T. Farrell, and R. Pazdur, *Clin. Cancer Res.*, 2015, **21**, 2666.

82. European Medicines Agency and Committee for Medicinal Products for Human Use, *Assessment Report: Istodax, Procedure No. EMEA/H/C/002122*, London, 2012.

83. R. Furumai, A. Matsuyama, N. Kobashi, K.-H. Lee, M. Nishiyama, H. Nakajima, A. Tanaka, Y. Komatsu, N. Nishino, M. Yoshida, and S. Horinouchi, *Cancer Res.*, 2002, **62**, 4916–4921.

84. J. J. Xiao, J. Byrd, G. Marcucci, M. Grever, and K. K. Chan, *Rapid Commun. Mass Spectrom.*, 2003, **17**, 757–766.

85. K. M. VanderMolen, W. McCulloch, C. J. Pearce, and N. H. Oberlies, *J. Antibiot. (Tokyo).*, 2011, **64**, 525–531.

86. B. Zhao and T. He, *Oncol. Rep.*, 2015, **33**, 304–310.

87. 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.

88. K. Venkatesan, J.-F. Rual, A. Vazquez, U. Stelzl, I. Lemmens, T. Hirozane-Kishikawa, T. Hao, M. Zenkner, X. Xin, K.-I. Goh, M. A. Yildirim, N. Simonis, K. Heinzmann, F. Gebreab, J. M. Sahalie, S. Cevik, C. Simon, A.-S. de Smet, E. Dann, A. Smolyar, A. Vinayagam, H. Yu, D. Szeto, H. Borick, A. Dricot, N. Klitgord, R. R. Murray, C. Lin, M. Lalowski, J. Timm, K. Rau, C. Boone, P. Braun, M. E. Cusick, F. P. Roth, D. E. Hill, J. Tavernier, E. E. Wanker, A.-L. Barabási, and M. Vidal, *Nat. Methods*, 2009, **6**, 83–90.

89. L. Bonetta, *Nature*, 2010, **468**, 851–854.

90. O. Keskin, A. Gursoy, B. Ma, and R. Nussinov, *Chem. Rev.*, 2008, **108**, 1225–1244.

91. D. P. Ryan and J. M. Matthews, *Curr. Opin. Struct. Biol.*, 2005, **15**, 441–446.

92. M. W. Gonzalez and M. G. Kann, *PLoS Comput. Biol.*, 2012, **8**, e1002819.

93. H. Yin and A. D. Hamilton, *Angew. Chem. Int. Ed. Engl.*, 2005, **44**, 4130–4163.

94. M. Guharoy and P. Chakrabarti, *Bioinformatics*, 2007, **23**, 1909–1918.

95. O. Keskin, B. Ma, and R. Nussinov, *J. Mol. Biol.*, 2005, **345**, 1281–1294.

96. A. L. Jochim and P. S. Arora, *ACS Chem. Biol.*, 2010, **5**, 919–923.

97. I. Saraogi and A. D. Hamilton, *Biochem. Soc. Trans.*, 2008, **36**, 1414–1417.

98. B. N. Bullock, A. L. Jochim, and P. S. Arora, *J. Am. Chem. Soc.*, 2011,

133, 14220–14223.

99. E. F. Lee, P. E. Czabotar, B. J. Smith, K. Deshayes, K. Zobel, P. M. Colman, and W. D. Fairlie, *Cell Death Differ.*, 2007, **14**, 1711–1713.

100. C. Tse, A. R. Shoemaker, J. Adickes, M. G. Anderson, J. Chen, S. Jin, E. F. Johnson, K. C. Marsh, M. J. Mitten, P. Nimmer, L. Roberts, S. K. Tahir, Y. Xiao, X. Yang, H. Zhang, S. Fesik, S. H. Rosenberg, and S. W. Elmore, *Cancer Res.*, 2008, **68**, 3421–3428.

101. L. T. Vassilev, B. T. Vu, B. Graves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, N. Fotouhi, and E. A. Liu, *Science*, 2004, **303**, 844–848.

102. B. L. Grasberger, T. Lu, C. Schubert, D. J. Parks, T. E. Carver, H. K. Koblish, M. D. Cummings, L. V LaFrance, K. L. Milkiewicz, R. R. Calvo, D. Maguire, J. Lattanze, C. F. Franks, S. Zhao, K. Ramachandren, G. R. Bylebyl, M. Zhang, C. L. Manthey, E. C. Petrella, M. W. Pantoliano, I. C. Deckman, J. C. Spurlino, A. C. Maroney, B. E. Tomczuk, C. J. Molloy, and R. F. Bone, *J. Med. Chem.*, 2005, **48**, 909–912.

103. J. D. Sadowsky, W. D. Fairlie, E. B. Hadley, H.-S. Lee, N. Umezawa, Z. Nikolovska-Coleska, S. Wang, D. C. S. Huang, Y. Tomita, and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 139–154.

104. E. F. Lee, J. D. Sadowsky, B. J. Smith, P. E. Czabotar, K. J. Peterson-Kaufman, P. M. Colman, S. H. Gellman, and W. D. Fairlie, *Angew. Chem. Int. Ed.*, 2009, **48**, 4318–4322.

105. T. Hara, S. R. Durell, M. C. Myers, and D. H. Appella, *J. Am. Chem. Soc.*, 2006, **128**, 1995–2004.

106. W. P. Nolan, G. S. Ratcliffe, and D. C. Rees, *Tetrahedron Lett.*, 1992, **33**, 6879–6882.

107. D. C. Horwell, W. Howson, G. S. Ratcliffe, and H. M. G. Willems, *Bioorg. Med. Chem.*, 1996, **4**, 33–42.

108. J. S. Albert and A. D. Hamilton, *Biochemistry*, 1995, **34**, 984–990.

109. V. Azzarito, K. Long, N. S. Murphy, and A. J. Wilson, *Nat. Chem.*, 2013, **5**, 161–173.

110. B. P. Orner, J. T. Ernst, and A. D. Hamilton, *J. Am. Chem. Soc.*, 2001, **123**, 5382–5383.

111. 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.

112. J. M. Rodriguez, L. Nevola, N. T. Ross, G. Lee, and A. D. Hamilton, *Chembiochem*, 2009, **10**, 829–833.

113. C. G. Cummings, N. T. Ross, W. P. Katt, and A. D. Hamilton, *Org. Lett.*, 2009, **11**, 25–28.

114. J. T. Ernst, J. Becerril, H. S. Park, H. Yin, and A. D. Hamilton, *Angew. Chem. Int. Ed.*, 2003, **42**, 535–539.

115. M. J. Adler, R. T. W. Scott, and A. D. Hamilton, *Chemistry*, 2012, **18**, 12974–12977.

116. S. M. Biros, L. Moisan, E. Mann, A. Carella, D. Zhai, J. C. Reed, and J. Rebek, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 4641–4645.

117. J. L. Yap, X. Cao, K. Vanommeslaeghe, K.-Y. Jung, C. Peddaboina, P. T. Wilder, A. Nan, A. D. MacKerell, W. R. Smythe, and S. Fletcher, *Org. Biomol. Chem.*, 2012, **10**, 2928–2933.

118. J. P. Plante, T. Burnley, B. Malkova, M. E. Webb, S. L. Warriner, T. A. Edwards, and A. J. Wilson, *Chem. Commun.*, 2009, 5091–5093.

119. V. Azzarito, P. Prabhakaran, A. I. Bartlett, N. S. Murphy, M. J. Hardie, C. A. Kilner, T. A. Edwards, S. L. Warriner, and A. J. Wilson, *Org. Biomol. Chem.*, 2012, **10**, 6469–6472.

120. F. Campbell, J. P. Plante, T. A. Edwards, S. L. Warriner, and A. J. Wilson, *Org. Biomol. Chem.*, 2010, **8**, 2344–2351.

121. K. Long, T. A. Edwards, and A. J. Wilson, *Bioorg. Med. Chem.*, 2013, **21**, 4034–4040.

122. S. Marqusee and R. L. Baldwin, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 8898–8902.

123. J. M. Scholtz, H. Qian, V. H. Robbins, and R. L. Baldwin, *Biochemistry*, 1993, **32**, 9668–9676.

124. M. Chorev, E. Roubini, R. L. McKee, S. W. Gibbons, M. E. Goldman, M. P. Caulfield, and M. Rosenblatt, *Biochemistry*, 1991, **30**, 5968–5974.

125. J. C. Phelan, N. J. Skelton, A. C. Braisted, and R. S. McDowell, *J. Am. Chem. Soc.*, 1997, **119**, 455–460.

126. J. K. Judice, J. Y. K. Tom, W. Huang, T. Wrin, J. Vennari, C. J. Petropoulos, and R. S. McDowell, *Proc. Natl. Acad. Sci.*, 1997, **94**, 13426–13430.

127. S. K. Sia, P. A. Carr, A. G. Cochran, V. N. Malashkevich, and P. S. Kim, *Proc. Natl. Acad. Sci.*, 2002, **99**, 14664–14669.

128. T. R. Geistlinger and R. K. Guy, *J. Am. Chem. Soc.*, 2003, **125**, 6852–6853.

129. D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh, and P. G. Schultz, *J. Am. Chem. Soc.*, 1991, **113**, 9391–9392.

130. F. Ruan, Y. Chen, and P. B. Hopkins, *J. Am. Chem. Soc.*, 1990, **112**, 9403–9404.

131. M. R. Ghadiri and A. K. Fernholz, *J. Am. Chem. Soc.*, 1990, **112**, 9633–9635.

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

133. A. Patgiri, A. L. Jochim, and P. S. Arora, *Acc. Chem. Res.*, 2008, **41**, 1289–1300.

134. S. A. Kawamoto, A. Coleska, X. Ran, C.-Y. Yang, and S. Wang, *J. Med. Chem.*, 2012, **55**, 1137–1146.

135. S. Kneissl, E. J. Loveridge, C. Williams, M. P. Crump, and R. K. Allemann, *Chembiochem*, 2008, **9**, 3046–3054.

136. I. L. Karle and P. Balaram, *Biochemistry*, 1990, **29**, 6747–6756.

137. P. Balaram, *Curr. Opin. Struct. Biol.*, 1992, **2**, 845–851.

138. C. E. Schafmeister, J. Po, and G. L. Verdine, *J. Am. Chem. Soc.*, 2000, **122**, 5891–5892.

139. S. Kumar and R. Nussinov, *Chembiochem*, 2002, **3**, 604–617.

140. M. E. Houston Jr., A. P. Campbell, B. Lix, C. M. Kay, B. D. Sykes, and R. S. Hodges, *Biochemistry*, 1996, **35**, 10041–10050.

141. L. Serrano, M. Bycroft, and A. R. Fersht, *J. Mol. Biol.*, 1991, **218**, 465–475.

142. C. Hunter, J. Singh, and J. Thornton, *J. Mol. Biol.*, 1991, **218**, 837–846.

143. R. Anjana, M. K. Vaishnavi, D. Sherlin, S. P. Kumar, K. Naveen, P. S. Kanth, and K. Sekar, *Bioinformation*, 2012, **8**, 1220–1224.

144. J. S. Albert and A. D. Hamilton, *Biochemistry*, 1995, **34**, 984–990.

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

146. J. Yang, K. Zhao, Y. Gong, A. Vologodskii, and N. R. Kallenbach, *J. Am. Chem. Soc.*, 1998, **120**, 10646–10652.

147. M. Siedlecka, G. Goch, A. Ejchart, H. Sticht, and A. Bierzynski, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 903–908.

148. R. M. Beesley, C. K. Ingold, and J. F. Thorpe, *J. Chem. Soc. Trans.*, 1915, **107**, 1080.

149. B. V Prasad and P. Balaram, *CRC Crit. Rev. Biochem.*, 1984, **16**, 307–348.

150. S. J. Miller and R. H. Grubbs, *J. Am. Chem. Soc.*, 1995, **117**, 5855–5856.

151. S. J. Miller, H. E. Blackwell, and R. H. Grubbs, *J. Am. Chem. Soc.*, 1996, **118**, 9606–9614.

152. 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.

153. F. Bernal, A. F. Tyler, S. J. Korsmeyer, L. D. Walensky, and G. L. Verdine, *J. Am. Chem. Soc.*, 2007, **129**, 2456–2457.

154. Y.-W. Kim, P. S. Kutchukian, and G. L. Verdine, *Org. Lett.*, 2010, **12**, 3046–3049.

155. Y.-W. Kim, T. N. Grossmann, and G. L. Verdine, *Nat. Protoc.*, 2011, **6**, 761–771.

156. G. H. Bird, F. Bernal, K. Pitter, and L. D. Walensky, *Methods Enzymol.*, 2008, **446**, 369–386.

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

158. H. E. Blackwell and R. H. Grubbs, *Angew. Chemie Int. Ed.*, 1998, **37**, 3281–3284.

159. Ø. Jacobsen, J. Klaveness, and P. Rongved, *Molecules*, 2010, **15**, 6638–6677.

160. Y.-W. Kim and G. L. Verdine, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 2533–2536.

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

162. G. J. Hilinski, Y.-W. Kim, J. Hong, P. S. Kutchukian, C. M. Crenshaw, S. S. Berkovitch, A. Chang, S. Ham, and G. L. Verdine, *J. Am. Chem. Soc.*, 2014, **136**, 12314–12322.

163. R. J. Platt, T. S. Han, B. R. Green, M. D. Smith, J. Skalicky, P. Gruszczynski, H. S. White, B. Olivera, G. Bulaj, and J. Gajewiak, *J. Biol. Chem.*, 2012, **287**, 20727–20736.

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

165. T. K. Pham, J. Yoo, and Y.-W. Kim, *Bull. Korean Chem. Soc.*, 2013, **34**, 2640–2644.

166. S. Y. Shim, Y.-W. Kim, and G. L. Verdine, *Chem. Biol. Drug Des.*, 2013, **82**, 635–642.

167. R. M. Williams and M. N. Im, *J. Am. Chem. Soc.*, 1991, **113**, 9276–9286.

168. Y. N. Belokon, V. I. Bakhmutov, N. I. Chernoglazova, K. A. Kochetkov, S. V. Vitt, N. S. Garbalinskaya, and V. M. Belikov, *J. Chem. Soc. Perkin Trans. 1*, 1988, 305–312.

169. Y. N. Belokon, V. I. Tararov, T. F. Maleev, Viktor I., Savel'eve, and M. G. Ryzhov, *Tetrahedron: Asymmetry*, 1998, **9**, 4249–4252.

170. B. Aillard, N. S. Robertson, A. R. Baldwin, S. Robins, and A. G. Jamieson, *Org. Biomol. Chem.*, 2014, **12**, 8775–8782.

171. Aileron Therapeutics Inc., *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT01775358, Phase 1 Safety Study of ALRN-5281 in Healthy Subjects*, 2013.

172. *Aileron Therapeutics Successfully Completes First-Ever Stapled Peptide Clinical Trial | Business Wire*, 2013.

173. Y. S. Chang, B. Graves, V. Guerlavais, C. Tovar, K. Packman, K.-H. To, K. A. Olson, K. Kesavan, P. Gangurde, A. Mukherjee, T. Baker, K. Darlak, C. Elkin, Z. Filipovic, F. Z. Qureshi, H. Cai, P. Berry, E. Feyfant, X. E. Shi, J. Horstick, D. A. Annis, A. M. Manning, N. Fotouhi, H. Nash, L. T. Vassilev, and T. K. Sawyer, *Proc. Natl. Acad. Sci.*, 2013, **110**, E3445–E3454.

174. *Aileron Therapeutics Publishes Data Demonstrating ATSP-7041 as First-in-class p53 Pathway Re-activator for the Treatment of Solid and Hematologic Cancers | Business Wire*, 2013.

175. R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149–2154.

176. S.-S. Wang, *J. Am. Chem. Soc.*, 1973, **95**, 1328–1333.

177. K. Barlos, D. Gatos, J. Kallitsis, G. Papaphotiu, P. Sotiriu, Y. Wenqing, and W. Schäfer, *Tetrahedron Lett.*, 1989, **30**, 3943–3946.

178. K. Barlos, D. Gatos, S. Kapolos, G. Papaphotiu, W. Schäfer, and Y. Wenqing, *Tetrahedron Lett.*, 1989, **30**, 3947–3950.

179. G. R. Matsueda and J. M. Stewart, *Peptides*, 1981, **2**, 45–50.

180. P. Sieber, *Tetrahedron Lett.*, 1987, **28**, 2107–2110.

181. W. C. Chan and P. D. White, Eds., *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*, Oxford University Press, New York, 2000.

182. C. A. Hood, G. Fuentes, H. Patel, K. Page, M. Menakuru, and J. H. Park,

J. Pept. Sci., 2008, **14**, 97–101.

183. J. M. Collins, in *Microwaves in Organic Synthesis*, eds. A. de la Hoz and A. Loupy, Wiley-VCH, Weinheim, 3rd edn., 2012, pp. 897–959.

184. J. K. Murray, J. Aral, and L. P. Miranda, in *Drug Design and Discovery: Methods and Protocols, Methods in Molecular Biology, Volume 716*, ed. S. D. Satyanarayanajois, Humana Press, New York, 2011, pp. 73–88.

185. H. M. Yu, S. T. Chen, and K. T. Wang, *J. Org. Chem.*, 1992, **57**, 4781–4784.

186. J. D. Ferguson, *Mol. Divers.*, 2003, **7**, 281–286.

187. B. Bacsa, B. Desai, G. Dibó, and C. O. Kappe, *J. Pept. Sci.*, 2006, **12**, 633–638.

188. B. Bacsa, K. Horváti, S. Bõsze, F. Andreae, and C. O. Kappe, *J. Org. Chem.*, 2008, **73**, 7532–7542.

189. C. Loffredo, N. A. Assunção, J. Gerhardt, and M. T. M. Miranda, *J. Pept. Sci.*, 2009, **15**, 808–817.

190. V. Santagada, F. Fiorino, E. Perissutti, B. Severino, V. De Filippis, B. Vivenzio, and G. Caliendo, *Tetrahedron Lett.*, 2001, **42**, 5171–5173.

191. J. K. Murray and S. H. Gellman, *Org. Lett.*, 2005, **7**, 1517–1520.

192. J. M. Collins, M. J. Collins, and R. C. Steorts, *Biopolymers*, 2003, **71**, 361.

193. J. M. Collins, K. A. Porter, S. K. Singh, and G. S. Vanier, *Org. Lett.*, 2014, **16**, 940–943.

194. T. Michels, R. Dölling, U. Haberkorn, and W. Mier, *Org. Lett.*, 2012, **14**, 5218–5221.

195. CEM, *New UV monitoring accessory now available for CEM microwave peptide synthesizers*, Matthews, NC, 2010.

196. CEM, *Improved Synthesis of a Difficult Peptide by UV Monitoring, Application Note B10-0011*, Matthews, NC, 2011.

197. R. H. Grubbs, P. L. Burk, and D. D. Carr, *J. Am. Chem. Soc.*, 1975, **97**, 3265–3267.

198. C. K. Ingold, *J. Chem. Soc., Trans.*, 1921, **119**, 305–329.

199. L. A. Boyer, R. R. Latek, and C. L. Peterson, *Nat. Rev. Mol. Cell Biol.*, 2004, **5**, 158–63.

200. M. H. Stenzel, *ACS Macro Lett.*, 2013, **2**, 14–18.

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

202. A. G. Jamieson, N. Boutard, D. Sabatino, and W. D. Lubell, *Chem. Biol. Drug Des.*, 2013, **81**, 148–65.

203. N. J. Greenfield, *Trends Anal. Chem.*, 1999, **18**, 236–244.

204. S. M. Kelly, T. J. Jess, and N. C. Price, *Biochim. Biophys. Acta - Proteins Proteomics*, 2005, **1751**, 119–139.

205. D. H. A. Corrêa and C. H. I. Ramos, *African J. Biochem. Res.*, 2009, **3**, 164–173.

206. Y.-H. Chen, J. T. Yang, and K. H. Chau, *Biochemistry*, 1974, **13**, 3350–3359.

207. R. Behrendt, P. White, and J. Offer, *J. Pept. Sci.*, 2016, **22**, 4–27.

208. H.-G. Chao, B. Leiting, P. D. Reiss, A. L. Burkhardt, C. E. Klimas, J. B. Bolen, and G. R. Matsueda, *J. Org. Chem.*, 1995, **60**, 7710–7711.

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

210. C. H. Douse, S. J. Maas, J. C. Thomas, J. A. Garnett, Y. Sun, E. Cota, and E. W. Tate, *ACS Chem. Biol.*, 2014, **9**, 2204–2209.

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

212. H. Jo, N. Meinhardt, Y. Wu, S. Kulkarni, X. Hu, K. E. Low, P. L. Davies, W. F. DeGrado, and D. C. Greenbaum, *J. Am. Chem. Soc.*, 2012, **134**, 17704–17713.

213. E. M. Bertino and G. A. Otterson, *Expert Opin. Investig. Drugs*, 2011, **20**, 1151–1158.

214. S. Subramanian, S. E. Bates, J. J. Wright, I. Espinoza-Delgado, and R. L. Piekarz, *Pharmaceuticals*, 2010, **3**, 2751–2767.

215. Y. Fujiwara, N. Yamamoto, Y. Yamada, K. Yamada, T. Otsuki, S. Kanazu, T. Iwasa, J. S. Hardwick, and T. Tamura, *Cancer Sci.*, 2009, **100**, 1728–1734.

216. M. Federico and L. Bagella, *J. Biomed. Biotechnol.*, 2011, **2011**, 1–12.

217. A. S. Madsen, H. M. E. Kristensen, G. Lanz, and C. A. Olsen, *ChemMedChem*, 2014, **9**, 614–626.

218. K. Chen, L. Xu, and O. Wiest, *J. Org. Chem.*, 2013, **78**, 5051–5055.

219. J. A. Jacobsen, J. L. Major Jourden, M. T. Miller, and S. M. Cohen, *Biochim. Biophys. Acta*, 2010, **1803**, 72–94.

220. S. Hanessian, V. Vinci, L. Auzzas, M. Marzi, and G. Giannini, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 4784–4787.

221. C. Temperini, A. Innocenti, A. Scozzafava, and C. T. Supuran, *Bioorg. Med. Chem. Lett.*, 2006, **16**, 4316–4320.

222. K. Vanommeslaeghe, S. Loverix, P. Geerlings, and D. Tourwé, *Bioorg. Med. Chem.*, 2005, **13**, 6070–6082.

223. K. M. Miller, J. V Tjeertes, J. Coates, G. Legube, S. E. Polo, S. Britton, and S. P. Jackson, *Nat. Struct. Mol. Biol.*, 2010, **17**, 1144–1151.

224. P. Kahnberg, A. J. Lucke, M. P. Glenn, G. M. Boyle, J. D. A. Tyndall, P. G. Parsons, and D. P. Fairlie, *J. Med. Chem.*, 2006, **49**, 7611–7622.

225. A. K. Chatterjee, T.-L. Choi, D. P. Sanders, and R. H. Grubbs, *J. Am. Chem. Soc.*, 2003, **125**, 11360–11370.

226. E. Hu, E. Dul, C.-M. Sung, Z. Chen, R. Kirkpatrick, G.-F. Zhang, K. Johanson, R. Liu, A. Lago, G. Hofmann, R. Macarron, M. de los Frailes, P. Perez, J. Krawiec, J. Winkler, and M. Jaye, *J. Pharmacol. Exp. Ther.*, 2003, **307**, 720–728.

227. A. Nikitjuka and A. Jirgensons, *Synlett*, 2012, **23**, 2972–2974.

228. M. A. Pilkington-Miksa, M. J. Writer, S. Sarkar, Q.-H. Meng, S. E. Barker, P. A. Shamlou, H. C. Hailes, S. L. Hart, and A. B. Tabor, *Bioconjug.*

Chem., 2007, **18**, 1800–1810.

229. G. Yuan, K. Lingfang, and L. Yingchun, *Method for preparing chiral alpha-unnatural amino acid by transition metal complex*, Patent China, 2013.

230. M. S. Masalovich, L. P. Ardasheva, and G. A. Shagisultanova, *Russ. J. Inorg. Chem.*, 2006, **51**, 1498–1503.

231. V. Prachayasittikul, S. Prachayasittikul, S. Ruchirawat, and V. Prachayasittikul, *Drug Des. Devel. Ther.*, 2013, **7**, 1157–1178.

232. Y. Shi, F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, R. A. Casero, and Y. Shi, *Cell*, 2004, **119**, 941–953.

233. F. Forneris, C. Binda, M. A. Vanoni, A. Mattevi, and E. Battaglioli, *FEBS Lett.*, 2005, **579**, 2203–2207.

234. F. Forneris, C. Binda, M. A. Vanoni, E. Battaglioli, and A. Mattevi, *J. Biol. Chem.*, 2005, **280**, 41360–41365.

235. C. Binda, A. Mattevi, and D. E. Edmondson, *J. Biol. Chem.*, 2002, **277**, 23973–23976.

236. F. Forneris, C. Binda, A. Adamo, E. Battaglioli, and A. Mattevi, *J. Biol. Chem.*, 2007, **282**, 20070–20074.

237. M. Paradís-Bas, J. Tulla-Puche, and F. Albericio, *Chem. Soc. Rev.*, 2016, **45**, 631–54.

238. P. Marek, A. M. Woys, K. Sutton, M. T. Zanni, and D. P. Raleigh, *Org. Lett.*, 2010, **12**, 4848–4851.

239. P. W. R. Harris, R. Kowalczyk, D. L. Hay, and M. A. Brimble, *Int. J. Pept. Res. Ther.*, 2012, **19**, 147–155.

240. *Synthesis design using pseudoproline dipeptides, Novabiochem Innovations (3/04)*, .

241. A. Abedini and D. P. Raleigh, *Org. Lett.*, 2005, **7**, 693–696.

242. R. Baron, C. Binda, M. Tortorici, J. A. McCammon, and A. Mattevi, *Structure*, 2011, **19**, 212–220.

243. W. Fiskus, S. Sharma, B. Shah, B. P. Portier, S. G. T. Devaraj, K. Liu, S. P. Iyer, D. Bearss, and K. N. Bhalla, *Leukemia*, 2014, **28**, 2155–2164.

244. S. Hamm, K. Kronthaler, U. Parnitzke, T. Prenzel, H. Kohlhof, D. Vitt, and R. Baumgartner, *Eur. J. Cancer*, 2016, **55**, S19–S22.

245. 4SC AG, *ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). 2000 Feb 29 -. Identifier NCT01344707, Oral Histone Deacetylase Inhibitor 4SC-202 in Patients With Advanced Hematologic Malignancies (TOPAS)*, 2015.

246. D. J. Smith, G. P. A. Yap, J. A. Kelley, and J. P. Schneider, *J. Org. Chem.*, 2011, **76**, 1513–1520.

247. A. S. Saghiyan, S. A. Dadayan, S. G. Petrosyan, L. L. Manasyan, A. V. Geolchanyan, S. M. Djamgaryan, S. A. Andreasyan, V. I. Maleev, and V. N. Khrustalev, *Tetrahedron: Asymmetry*, 2006, **17**, 455–467.

248. R. P. M. Storcken, L. Panella, F. L. van Delft, B. Kaptein, Q. B. Broxterman, H. E. Schoemaker, and F. P. J. T. Rutjes, *Adv. Synth. Catal.*, 2007, **349**, 161–164.

7. Appendix

Organic & Biomolecular Chemistry

PAPER

View Article Online
View Journal | View Issue

Robust asymmetric synthesis of unnatural alkenyl amino acids for conformationally constrained α-helix peptides†

Boris Aillard, Naomi S. Robertson, Adam R. Baldwin, Siobhan Robins and Andrew G. Jamieson*

Cite this: *Org. Biomol. Chem.*, 2014, 12, 8775

Received 27th August 2014,
Accepted 16th September 2014

DOI: 10.1039/c4ob01832j

www.rsc.org/obc

The efficient asymmetric synthesis of unnatural alkenyl amino acids required for peptide 'stapling' has been achieved using alkylation of a fluorine-modified Ni^{II} Schiff base complex as the key step.

Introduction

The asymmetric synthesis of enantiomerically pure unnatural α-amino acids using chiral auxiliaries or catalysts is an important field of bio-organic chemistry. These amino acids provide the starting materials required to generate peptidomimetic tools for the investigation of protein form and function. Conformationally constraining peptides to adopt the appropriate bioactive conformation decreases the entropic penalty of folding and can provide enhanced binding affinity to target receptors.[1] Employing synthetic, unnatural functionality as the constraining moiety has also been demonstrated to provide a range of other favourable physicochemical and pharmacological properties including improved stability toward peptidase degradation, increased bioavailability and membrane permeability.[2] Conformationally constrained peptides have thus come of age as tools for chemical biology and drug discovery.

The majority of research in this area focuses on the regulation of α-helix-mediated protein–protein interactions (PPIs). A plethora of synthetic approaches have therefore been developed to promote α-helical structure in peptides. Examples include: triazole bridge,[3] lactam bridge,[4] hydrogen bond surrogates,[5] disulphide bridges,[6] electrostatic interaction between side chains,[7] and incorporation of α,α-disubstituted amino acids.[8] Blackwell and Grubbs developed an *i*, *i* + 4 cross-linked peptide *via* cross metathesis of *O*-allyl functionalised residues.[9]

Verdine and co-workers expanded this method through the develop of an all hydrocarbon peptide "staple", determining the optimal chain length of α-methyl α-disubstituted unnatural amino acids bearing alkenes for *i*, *i* + 3, *i*, *i* + 4, and *i*, *i* + 7 staples.[10] These hydrocarbon stapled peptides were found to exhibit increased α-helicity and resistance to proteolysis.[11]

Recent literature suggests that peptides stapled with mono-substituted alkenyl amino acids can demonstrate similar improvements in α-helicity and proteolytic stability as their α-methyl α-disubstituted counterparts.[12,13]

During our research into the design and synthesis of peptide conformational constraints we required unnatural alkenyl α-methyl α-amino acids **1** & **2** (Fig. 1). The synthesis of these amino acids has previously been achieved by alkylation of alanine-derived oxazinones.[14] We disregarded this method due to the potential for over-reduction of the olefin moiety during the required metal reduction step.[15]

Belekon and co-workers have developed a more convenient method for the synthesis of α-methyl α-disubstituted amino

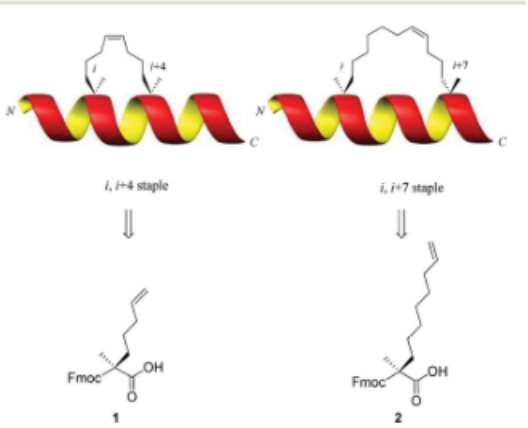


Fig. 1 (*S*)-Alkenyl α-methyl α-amino acids required for *i*, *i* + 4 and *i*, *i* + 7 peptide conformational constraints.

Department of Chemistry, University of Leicester, Leicester, LE1 7RH, UK. E-mail: andrew.jamieson@le.ac.uk; Fax: +44 (0)116252 3789; Tel: +44 (0)116 252 2105

†Electronic supplementary information (ESI) available: Full experimental procedures and spectroscopic data for compounds **1**–**16**. CCDC 1020152–1020157. For ESI and crystallographic data in CIF or other electronic format see DOI: 10.1039/c4ob01832j

acids that involves α-C-alkylation of chiral Ni^{II} Schiff base complexes derived from alanine. However, in our hands the key alkylation step gave low yield (42%) and poor diastereoselectivity (72% de). The diastereoselectivity of this reaction originates from the benzylproline moiety that sterically hinders one face of the complex. Attempts to increase selectivity have thus focused on modification of the benzylproline functionality within the Ni^{II} Schiff base complex. Introduction of the sterically bulky naphthylmethyl[16] or 2,4,6-trimethylbenzyl[17] functionality are reported to result in either poor chemical yield or complexes with limited solubility. Introduction of halogens onto the aromatic ring of the *N*-benzylproline moiety have been more successful in enhancing diastereoselectivity.[18] The aim of this work was to develop a robust asymmetric synthesis of enantiopure unnatural alkenyl amino acids required for peptide stapling.

Here we report the use of Ni^{II} Schiff base complexes, derived from a 2-fluorobenzyl ligand[18] for the diastereoselective synthesis of four alkenyl amino acids required for *i*, *i* + 4 and *i*, *i* + 7 peptide hydrocarbon staples.

The method is notable for the high diastereoselectivities of the alkylation reactions and the fluorinated substituent in the Ni^{II} Schiff base complexes facilitates purity and de analysis by ^{19}F NMR. Insights into the origin of the diastereoselectivity were observed from the X-ray crystal structure of the alkylated Ni^{II} Schiff base complex **9**.

Results and discussion

Our initial aim was the efficient synthesis of the chiral auxiliary (*S*)-*N*-(2-benzoylphenyl)-1-(2-fluorobenzyl)-pyrrolidine-2-carboxamide **4** (2-FBPB) from L-proline (Scheme 1).

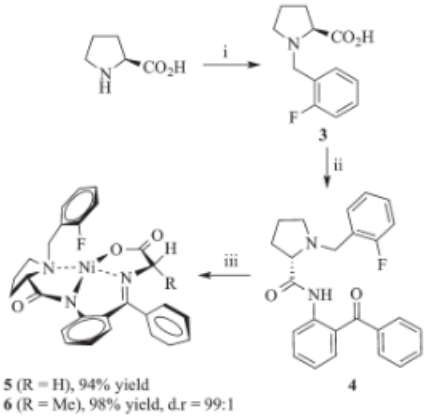
N-alkylation of L-proline with 2-fluorobenzyl bromide gave tertiary amine **3** in very good yield. Condensation of **3** with 2-aminobenzophenone using methanesulfonyl chloride and *N*-methylimidazole gave 2-FBPB **4** in respectable yield given the sterically cumbersome nature of the aniline nucleophile. However, the enantiopurity of 2-FBPB **4** was a disappointing 97% ee as determined by chiral HPLC and has not previously been reported. We therefore undertook a slow recrystallisation to enrich the desired enantiomer to >99% ee (see ESI†).

Complexation of **4** with nickel nitrate and either glycine or L-alanine under basic conditions gave a thermodynamic mixture of diastereomers of nickel Schiff base complexes (*S*)-Ni-Gly-2FBPB **5** and (*S*)-Ni-Ala-2FBPB **6**, respectively, in excellent yield as red crystals that could be stored in air without any significant decomposition (Scheme 1).

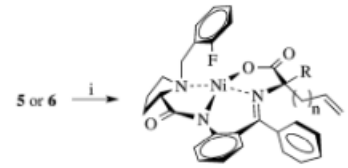
With the nickel Schiff base complexes **5** and **6** in hand, we turned our attention to asymmetric alkylation reactions with the appropriate electrophiles to provide the desired alkenyl amino acids required to synthesise stapled peptides.

Stereoselective alkylation of (*S*)-Ni-Gly-2FBPB **5** through formation of the enolate using sodium hydroxide as base followed by nucleophilic substitution reaction with 5-bromopentene gave complex **7** in excellent isolated yield and diastereoselectivity (>95 : 5 dr) (Scheme 2, entry 1). Under the same reaction conditions, (*S*)-Ni-Gly-2FBPB **5** also underwent efficient, stereoselective alkylation with 8-bromooctene (Scheme 2, entry 2). The α-proton of the amino acid of the glycine derived Schiff base moiety has a pK_a of around 11 and can be deprotonated relatively easily under basic reaction conditions. The second α-proton has a pK_a of around 15 and so is more difficult to deprotonate, however is labile using these reaction conditions. Addition of alkyl halide at room temperature prevented formation of α,α-dialkylation product.

Thus, the diastereomeric ratio of products obtained from these reactions reflects the position of the thermodynamic equilibrium.



Scheme 1 *Reagents and conditions*: i. 2-fluorobenzylbromide, KOH, ^{i}PrOH, 93%; ii. 2-aminobenzophenone, 1-Me-imidazole, MsCl, CH_2Cl_2, 50 °C, 55%; iii. glycine (**5**) or L-alanine (**6**), $Ni(NO_3)_2$, KOH, MeOH, 70 °C.



Entry	Initial complex	Alkyl bromide	Product	Isolated yield	d.r
1	**5**	5-Br-pentene	**7**	83%	>95:5
2	**5**	8-Br-octene	**8**	72%	>95:5
3	**6**	5-Br-pentene	**9**	62%	88:12
4	**6**	8-Br-octene	**10**	42%	89:11

Scheme 2 *Reagents and conditions*: i. $Br(CH_2)_nCH{=}CH_2$, NaOH, DMF, 0 °C to RT (R = H) or 50 °C (R = Me). Alkylation of chiral complexes **5** and **6** (dr *was determinated by* ^{19}F *NMR on crude materials*).

The nickel Schiff base complexes produced from these alkylation reactions were isolated in diastereomerically pure form, as red crystalline solids after flash column chromatography.

A further advantage of the 2-FBPB ligand **4** is that ^{19}F NMR spectroscopy can be used to monitor both reaction progression and also diastereoselectivity and constitutes an extremely useful tool functionality.

Alkylation of (*S*)-Ni-Ala-2FBPB **6** with 5-bromopent-1-ene or 8-bromopent-1-ene was also found to be successful when the enolate was formed at 0 °C and then alkylation at 50 °C to produce complexes **9** and **10** with good diastereoselectivity and yield (Scheme 2, entries 3 & 4). The α-methyl-α-substituted amino acid functionality produced during these alkylation reactions lacks an α-proton and so no epimerisation can occur, resulting in the kinetic diastereomer as the major product. (*S*)-Ni-Gly-2FBPB **5** gives better diastereoselectivity in comparison with (*S*)-Ni-Ala-2FBPB **6** in part, due to increased temperature of the alkylation reaction required for the latter.

The absolute stereochemistry of complexes **7**, **8** and **9** were determined from the X-ray crystal structures (Fig. 2 and ESI†). The crystal structure also gives an insight into the origin of the asymmetric induction. The 2-F-benzyl moiety is positioned across the *Re*-face of the complex and sterically restricts access to the electrophile, forcing reaction from the *Si*-face of the complex to give the 2*S* configuration at the amino acid α-carbon (Scheme 3). It has been proposed that an interaction between the halide atom and the central Ni-atom constrains the complex,[19] however the distance between these two atoms in the crystal structure of complex **9** (3.1 Å) exactly matches the sum of the van der Waals radii and suggests that no such attractive interaction exists in the solid state. Our preferred explanation for the increase in diastereoselectivity using the 2-FBPB ligand, over the BPB ligand is a displaced π–π stacking interaction between the *N*-benzyl functionality and the proline amide bond that constrains the complex. The fluorine atom creates a dipole with partial positive charge situated on the aromatic that may interact with ionised amide bond oxygen and so will facilitate this interaction.

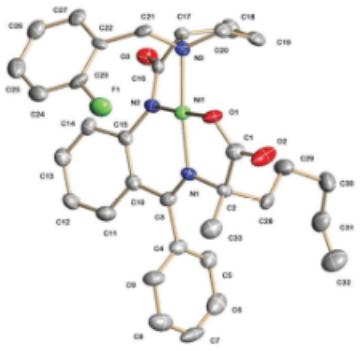
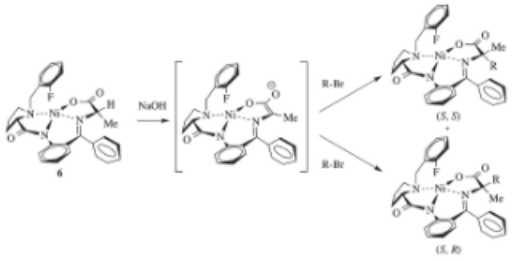
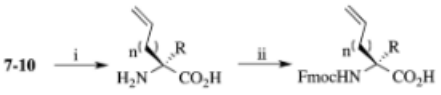


Fig. 2 Crystal structure of compound **9** (hydrogen atoms have been omitted from diagrams).



Scheme 3 Deprotonation and alkylation pathway of **6**.



11	R = H	n = 3	**15**
12	R = H	n = 6	**16**
13	R = Me	n = 3	**1**
14	R = Me	n = 6	**2**

Alkylated complex	Amino acid	Yield	Fmoc-AA	Yield
7	11	60%	15	65%
8	12	67%	16	63%
9	13	65%	1	82%
10	14	68%	2	62%

Scheme 4 *Reagents and conditions*: i. 3 M HCl–MeOH; ii. Fmoc-OSu, K_2CO_3, dioxane–H_2O. Decomposition of alkylated complexes **7–10** to amino acids **11–14** and Fmoc protection to compounds **1–2**, **15–16**.

Alkylated complexes **7–10** were readily decomposed under acidic conditions to give amino acids **11–14** and the chiral auxiliary **4**. Chiral ligand **4** was easily extracted from the crude mixture (enantiomerically pure as assessed by chiral HPLC) and can be reused, thus providing added value to this synthetic method. Purification *via* ion exchange chromatography gave amino acids **11–14**, which were subsequently treated with Fmoc succinimide under mildly basic conditions to afford Fmoc protected amino acids **1–2**, **15–16** (Scheme 4).

Conclusions

In summary, the outlined method enabled the diastereoselective synthesis of a range of enantiopure unnatural alkenyl amino acids required for peptide stapling in six steps from proline in up to 17% yield.

Incorporation of a fluorine atom on the chiral ligand of the nickel Schiff base complex increases diastereoselectivity of the key alkylation reactions and also facilitates chiral analysis by ^{19}F NMR. Moreover, we observed that the directing effect of the key alkylation step is the result of the electrophile approaching the enolate complex from the opposite face of the proline *N*-benzyl moiety. These features make this method

applicable to the large-scale synthesis of custom α,α-disubstituted amino acids. Efforts to establish in-depth structural understanding of the role played by the fluorine atom on diastereoselectivity are currently under active investigation in our laboratories. We are also currently using this method to synthesis other α,α-disubstituted amino acids to develop new peptide conformational constraints.

Experimental

(*S*)-1-(2-Fluorobenzyl)pyrrolidine-2-carboxylic acid (3)

L-Proline (100.0 g, 869.0 mmol) was added to a solution of potassium hydroxide (146.0 g, 2607.0 mmol, 3.0 equiv.) dissolved in isopropyl alcohol (1.0 L) at 40 °C. 2-Fluorobenzyl bromide (15.6 mL, 130 mmol, 1.0 equiv.) was added to the solution drop wise. The solution was allowed to stir at 40 °C for 18 hours and progress was monitored by TLC ($MeOH–CH_2Cl_2$, 1 : 4). Aqueous hydrochloric acid (37%) was added dropwise to the mixture until the solution reached pH 6–5 (as determined using a pH probe). The suspension was then cooled in an ice bath, filtered and the filtrate concentrated *in vacuo* to give (*S*)-1-(2-fluorobenzyl)pyrrolidine-2-carboxylic acid as a yellow solid (209.0 g, 93%). m.p: 78–80 °C (EtOH); $[\alpha]_D^{20}$ −24.1 (*c* 0.1 in MeOH); ^{1}H NMR (300 MHz; $CDCl_3$) δ 8.42 (1H, br s, OH), 7.52 (1H, td, *J* = 7.5, 1.3 Hz, Ar–C*H*), 7.35 (1H, m, Ar–C*H*), 7.12 (1H, t, *J* = 7.5 Hz, Ar–C*H*), 7.08 (1H, t, *J* = 8.8 Hz, Ar–C*H*), 4.37 (1H, d, *J* = 13.5 Hz, N–CH_2), 4.20 (1H, d, *J* = 13.5 Hz, N–CH_2), 3.74 (1H, dd, *J* = 9.2, 6.3 Hz, α-C*H*), 3.59 (1H, ddd, *J* = 10.6, 6.9, 3.8 Hz, δ-CH_2), 2.85 (1H, dt, *J* = 9.3, 8.5 Hz, δ-CH_2), 2.39–2.13 (2H, m, β-CH_2), 2.06–1.81 (2H, m, γ-CH_2); ^{13}C NMR (75 MHz; $CDCl_3$) δ 171.9 (s), 163.0 (s), 132.8 (d), 131.4 (d), 124.8 (d), 119.4 (s), 115.9 (d), 67.3 (d), 53.5 (t), 50.9 (t), 29.0 (t), 23.0 (t). Additional peaks arise from rotamers at 131.2, 119.2 and 115.6; ^{19}F NMR (282 MHz; $CDCl_3$) −116.3 (s); IR (ν_{max}/cm^{-1}, neat): 3458, 3013, 2970, 1736, 1618, 1443, 1369, 1229, 1109, 898, 758; HRMS-ESI (calcd for $C_{12}H_{15}NO_2F$ $[M + H]^+$) 224.1087, found 224.1091 (Δ = 1.8 ppm).

(*S*)-*N*-(2-Benzoylphenyl)-1-(2-fluorobenzyl)pyrrolidine-2-carboxamide ((2*S*)-FBPB) (4)

Methanesulfonyl chloride (1.0 mL, 13.4 mmol, 1.0 equiv.) was added to a solution of (*S*)-1-(2-fluorobenzyl)pyrrolidine-2-carboxylic acid **3** (3.0 g, 13.4 mmol) and *N*-methylimidazole (2.4 mL, 29.6 mmol, 2.2 equiv.) in CH_2Cl_2 (30.0 mL) at 0 °C. After 5 minutes 2-aminobenzophenone (2.4 g, 12.1 mmol, 0.9 equiv.) was added. The reaction mixture was heated to 50 °C for 12 hours, cooled and then saturated aqueous sodium hydrogen carbonate solution (30.0 mL) was added. The two layers were separated and the aqueous layer extracted with CH_2Cl_2 (3 × 30.0 mL). The combined organic layers were dried ($MgSO_4$) and concentrated *in vacuo*. Purification by flash column chromatography (15% ethyl acetate–hexane) followed by a recrystallisation with hexane and few drops of EtOAc gave the title compound as a pale yellow crystals (2.7 g, 55%). m.p: 88–90 °C (hexane–EtOAc); $[\alpha]_D^{20}$ −125.1 (*c* 0.25, MeOH); ^{1}H NMR (300 MHz; $CDCl_3$) δ 11.41 (1H, s, NH), 8.56 (1H, dd, *J* = 8.4, 1.0 Hz, Ar–C*H*), 7.79–7.73 (2H, m, Ar–C*H*), 7.61 (1H, td, *J* = 6.9, 1.1 Hz), 7.57–7.44 (5H, m, Ar–C*H*), 7.15–7.04 (2H, m, Ar–C*H*), 6.93 (1H, td, *J* = 7.4, 1.2 Hz, Ar–C*H*), 6.80 (1H, td, *J* = 9.0, 1.2 Hz, Ar–C*H*), 3.89 (1H, d, *J* = 13.3 Hz, N–CH_2), 3.73 (1H, dd, *J* = 13.3, 1.2 Hz, N–CH_2), 3.36 (1H, dd, *J* = 10.2, 4.7 Hz, α-C*H*), 3.24 (1H, m, β-CH_2), 2.48 (1H, m, β-CH_2), 2.25 (1H, m, δ-CH_2), 1.96 (1H, m, δ-CH_2), 1.89–1.73 (2H, m, γ-CH_2).; ^{13}C NMR (75 MHz; $CDCl_3$) δ 197.8 (s), 174.3 (s), 139.0 (s), 138.6 (s), 133.2 (d), 132.4 (d), 131.6 (d), 130.0 (d), 128.9 (d), 128.8 (d), 128.2 (d), 125.6 (s), 124.9 (s), 124.8 (s), 123.9 (d), 123.8 (d), 122.2 (d), 121.4 (d), 115.2 (d), 114.9 (d), 67.9 (d), 53.7 (t), 52.0 (t), 31.0 (t), 24.2 (t); ^{19}F NMR (376 MHz; $CDCl_3$) −117.5 (s); IR (ν_{max}/cm^{-1}, neat): 3276, 1682, 1647, 1577, 1512, 1444, 1286, 1266, 1102, 923, 768, 697; HRMS-ESI (calcd for $C_{25}H_{24}N_2O_2F$ $[M + H]^+$) 403.1822, found 403.1825 (Δ = 0.7 ppm); HPLC (OD-H column, hexane (5%)–^{i}PrOH isocratic): 15.2 min. X-rays: see ESI.†

General method of formation of Gly–Ni or Ala–Ni complexes (5)–(6)

(*S*)-*N*-(2-Benzoylphenyl)-1-(2-fluorobenzyl)pyrrolidine-2-carboxamide (2-FBPB) **4** (9.0 g, 22.4 mmol, 1.0 equiv.), $Ni(NO_3)_2{\cdot}6H_2O$ (13.0 g, 44.7 mmol, 2.0 equiv.) and glycine (3.4 g, 44.7 mmol, 2.0 equiv.) were dissolved in methanol (225.0 mL, 0.1 M) at 50 °C. Potassium hydroxide (8.8 g, 156.5 mmol, 7.0 equiv.) was added and the mixture was heated to 70 °C for 1 hour. The reaction mixture was cooled and then concentrated. The resulting residue was taken up in water (200.0 mL) and extracted with EtOAc (3 × 200.0 mL). The combined organic layers were washed with saturated brine solution (3 × 600.0 mL), dried ($MgSO_4$) and concentrated *in vacuo* to give the title compound as a red crystalline solid.

(*S*)-({2-[1-(2-Fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl}phenylmethylene)-glycinato-*N*,*N*′,*N*″,*O*}-nickel(II) (Ni-Gly-2-FBPB) (5). Yield 94%; m.p: 124–126 °C (hexane–EtOAc) (lit: 125–127 °C);[18] $[\alpha]_D^{20}$ +1897.9 (*c* 0.05, $CHCl_3$) (lit: +1300.0 (*c* 0.05, $CHCl_3$);[18] ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.40 (1H, dd, *J* = 7.8, 1.6 Hz, Ar–C*H*), 8.35 (1H, d, *J* = 9.0 Hz, Ar–C*H*), 7.63–7.50 (3H, m, Ar–C*H*), 7.37 (1H, m, Ar–C*H*), 7.33–7.23 (2H, m, Ar–C*H*), 7.21–7.10 (2H, m, Ar–C*H*), 7.03 (1H, m, Ar–C*H*), 6.87 (1H, d, *J* = 8.2 Hz, Ar–C*H*), 6.76 (1H, t, *J* = 7.8 Hz, Ar–C*H*), 4.52 (1H, d, *J* = 12.9 Hz, N–CH_2), 3.98 (1H, d, *J* = 12.9 Hz, N–CH_2), 3.82 (1H, d, *J* = 20.3 Hz, CH_2), 3.72 (1H, m, δ-CH_2), 3.70 (1H, d, *J* = 20.3 Hz, CH_2), 3.5 (1H, dd, *J* = 10.6, 5.9 Hz, α-CH_2), 3.41 (1H, m, γ-CH_2), 2.71 (1H, m, β-CH_2), 2.51 (1H, m, β-CH_2), 2.24–2.06 (2H, m, δ-CH_2); ^{13}C NMR (100 MHz, $CDCl_3$) δ 181.0 (s), 177.9 (s), 171.7 (s), 142.5 (s), 134.6 (s), 134.3 (s), 133.3 (d), 132.3 (d), 131.5 (d), 131.4 (d), 129.8 (d), 129.7 (d), 129.4 (d), 126.3 (d), 125.7 (s), 125.3 (d), 124.6 (d), 121.0 (d), 120.4 (s), 116.3 (d), 116.0 (d), 70.1 (d), 61.3 (t), 57.3 (t), 55.8 (t), 30.7 (t), 23.7 (t); ^{19}F NMR (376 MHz; $CDCl_3$) −113.6 (s).; IR (ν_{max}/cm^{-1}, neat): 2935, 2354, 1672, 1634, 1588, 1490, 1471, 1440, 1362, 1334, 1256, 1165, 1111, 1063, 962, 754, 723, 704.; HRMS-ESI (calcd for $C_{27}H_{24}N_3O_3F^{58}Ni$ $[M + H]^+$) 516.1233, found 516.1224 (Δ = −1.7 ppm); HPLC (OD-H column, hexane (50%)–^{i}PrOH isocratic): 18.76 min. X-rays: see ESI.†

(*S*)-({2-[1-(2-Fluorobenzyl)pyrrolidine-2-carboxamide]phenyl}-phenylmethylene)-(*S*)-alaninato-*N*,*N*′,*N*″,*O*}nickel(II) and (*S*)-({2-[1-(2-fluorobenzyl)pyrrolidine-2-carboxamide]phenyl}phenylmethylene)-(*R*)-alaninato-*N*,*N*′,*N*″,*O*}nickel(II) (Ni-Ala-2-FBPB) (6). Yield 98%, 99 : 1 dr; m.p: 284–286 °C (hexane–EtOAc) (lit: 283–285 °C);[18] $[\alpha]_D^{20}$ +3664.5 (*c* 0.05 in $CHCl_3$) (lit: +3126.6 (*c* 0.05 in $CHCl_3$);[18] ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.33 (1H, td, *J* = 7.3 Hz, Ar–C*H*), 8.13 (1H, dd, *J* = 8.6 Hz, Ar–C*H*), 7.53–7.49 (2H m, Ar–C*H*), 7.46 (1H, m, Ar–C*H*), 7.27–7.20 (2H, m, Ar–C*H*), 7.20–7.13 (2H, m, Ar–C*H*), 7.06 (1H, ddd, *J* = 9.8, 8.2 Hz, Ar–C*H*), 6.96 (1H, m, Ar–C*H*), 6.71–6.63 (2H, m, Ar–C*H*), 4.41 (1H, dd, *J* = 12.9, 1.2, N–C*H*H), 3.91 (1H, q, *J* = 7.0 Hz, α-C(Me)-*H*), 3.83 (1H, dd, *J* = 12.9, 0.8, N–CH*H*), 3.70 (1H, m, β(Pro)-C*H*H), 3.55–3.46 (2H, m, α(Pro)-C*H*, γ(Pro)-CH*H*), 2.38 (1H, m, δ(Pro)-C*H*H), 2.57 (1H, m, δ(Pro)-CH*H*), 2.22 (1H, m, γ(Pro)-CH*H*), 2.07 (1H, m, β(Pro)-CH*H*), 1.58 (3H, d, *J* = 7.0 Hz, C*H*$_3$); ^{13}C NMR (100 MHz, $CDCl_3$) δ 180.4 (s), 180.1 (s), 170.3 (s), 142.1 (s), 134.2 (d), 133.5 (s), 133.2 (d), 132.1 (d), 129.7 (d), 128.9 (d), 127.5 (d), 127.2 (d), 126.6 (s), 124.5 (d), 123.9 (d), 120.8 (d), 120.5 (s), 119.8 (s), 117.9 (d), 116.2 (d), 116.0 (d), 70.3 (d), 66.6 (d), 57.1 (t), 55.6 (t), 30.7 (t), 24.1 (t), 21.8 (q); ^{19}F NMR (376 MHz, $CDCl_3$) δ −113.9 (s); IR (ν_{max}/cm^{-1}, neat) 1680, 1626, 1593, 1440, 1359, 1330, 1302, 1266, 1231, 1168, 1064, 7621, 748, 710, 682; HRMS-ASAP (calcd for $C_{28}H_{27}N_3O_3F^{58}Ni$ [M + H]$^+$) 530.1390, found 530.1397 (Δ = 1.3 ppm), (calcd for $C_{28}H_{27}N_3O_3F^{60}Ni$ [M + H]$^+$) 532.1344, found 532.1381 (Δ = 7.0 ppm); HPLC: OD-H column, hexane (50%)–^{i}PrOH isocratic): 14.71 min. X-rays: see ESI.†

General method of alkylation of (5)–(6) with alkyl bromides

Freshly ground sodium hydroxide (0.54 g, 13.4 mmol, 4.0 equiv.) was taken up in DMF (20.0 mL) with stirring at 0 °C under an atmosphere of nitrogen. (*S*)-({2-[1-(2-Fluorobenzyl)-pyrrolidine-2-carboxamide]phenyl}phenylmethylene)-glycinato-*N*,*N*′,*N*″,*O*}nickel(II) 5 (1.73 g, 3.4 mmol, 1.0 equiv.) was added and stirred for 2 minutes; the solution darkened in colour and the ice bath was removed. A solution of 1-bromo-4-pentene (1.2 mL, 10.1 mmol, 3.0 equiv.) was added to reaction mixture. The solution was left to stir for 15 minutes at room temperature (for Gly-Ni-2-FBPB 5), or for 1 h at 50 °C (for Ala-Ni-2-FBPB 6) then quenched with the addition of water. The mixture was concentrated *in vacuo*, taken up in water (20.0 mL) and extracted with dichloromethane (3 × 25.0 mL). The combined organic extracts were washed with aqueous lithium chloride solution (5% v/v) (3 × 50.0 mL), brine (3 × 50.0 mL), dried ($MgSO_4$) and concentrated *in vacuo*. Purification by flash column chromatography (EtOAc–hexane 1 : 1) gave the title compound as a deep red-orange solid.

(*S*)-({2-[1-(2-Fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl}phenylmethylene)-(*S*)-pentenylglycinato-*N*,*N*′,*N*″,*O*}nickel(II) (Ni-α-pent-4-enyl gly-2-FBPB) (7). Yield 83%; m.p: 170–172 °C (hexane–EtOAc); $[\alpha]_D^{20}$ +2813.2 (*c* 0.05 in $CHCl_3$) (lit: +2172 (*c* 0.1, MeOH, 25 °C);[20] ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.31 (1H, t, *J* = 6.7 Hz, Ar–C*H*), 8.18 (1H, d, *J* = 8.6 Hz, Ar–C*H*), 7.57–7.41 (3H, m, Ar–C*H*), 7.32–7.20 (2H, m, Ar–C*H*), 7.16 (2H, t, *J* = 7.0 Hz, Ar–C*H*), 7.05 (1H, t, *J* = 8.8 Hz, Ar–C*H*), 6.92 (1H, d, *J* = 7.0 Hz, Ar–C*H*), 6.72–6.60 (2H, m, Ar–C*H*), 5.72 (1H, tdd, *J* = 13.3, 5.9, 4.3 Hz, C*H*═CH$_2$), 5.06–4.91 (2H, m, CH═C*H*$_2$), 4.44 (1H, d, *J* = 12.9 Hz, N–C*H*$_2$), 3.91 (1H, dd, *J* = 7.6, 3.3 Hz), 3.85 (1H, d, *J* = 12.9 Hz, N–C*H*$_2$), 3.61–3.40 (3H, m), 2.84 (1H, m), 2.56 (1H, m), 2.31–1.84 (6H, m), 1.78–1.55 (2H, m); ^{13}C NMR (100 MHz, $CDCl_3$) δ 180.0 (s), 179.2 (s), 170.4 (s), 160.3 (s), 142.2 (s), 137.7 (d), 134.2 (d), 133.8 (s), 133.3 (d), 132.1 (d), 131.2 (d), 129.7 (d), 128.9 (d), 128.8 (d), 127.6 (d), 127.1 (d), 126.6 (s), 124.5 (d), 123.6 (d), 120.7 (d), 120.4 (s), 116.2 (d), 115.9 (d), 115.2 (t), 70.4 (d), 56.7 (t), 55.6 (t), 34.7 (t), 33.3 (t), 30.6 (t), 24.5 (t), 23.6 (t); ^{19}F NMR (376 MHz; $CDCl_3$) −113.9 (s).; IR (ν_{max}/cm^{-1}, neat): 2920, 2854, 1645, 1333, 1256, 1168, 752; HRMS-ESI (calcd for $C_{32}H_{33}N_3O_3F^{58}Ni$ [M + H]$^+$) 584.1895, found 584.1893 (Δ = 4.6 ppm), (calcd for $C_{32}H_{33}N_3O_3F^{60}Ni$ [M + H]$^+$) 586.1814, found 586.1841(Δ = −1.7 ppm); HPLC (OD-H column, hexane (50%)–^{i}PrOH isocratic): 14.23 min. X-rays: see ESI.†

(*S*)-({2-[1-(2-Fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl}phenylmethylene)-(*S*)-octenylglycinato-*N*,*N*′,*N*″,*O*}nickel(II) (Ni-α-oct-7-enylgly-2-FBPB) (8). Yield 72%; m.p: 144–146 °C (hexane–EtOAc); $[\alpha]_D^{20}$ +3013.3 (*c* 0.05 in $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.31 (1H, td, *J* = 7.3, 1.8 Hz, Ar–C*H*), 8.18 (1H, d, *J* = 8.6 Hz, Ar–C*H*), 7.56–7.42 (3H, m, Ar–C*H*), 7.26–7.14 (4H, m, Ar–C*H*), 7.04 (1H, t, *J* = 9.4 Hz, Ar–C*H*), 6.92 (1H, d, *J* = 6.7 Hz, Ar–C*H*), 6.67 (2H, m, Ar–C*H*), 5.79 (1H, tdd, *J* = 10.2, 7.0, 3.1 Hz, C*H*═CH$_2$), 5.04–4.91 (2H, m), 4.45 (1H, d, *J* = 12.9 Hz), 3.92 (1H, dd, *J* = 7.8, 3.1 Hz), 3.86 (1H, d, *J* = 12.9 Hz), 3.62–3.44 (3H, m), 2.86 (1H, m), 2.57 (1H, m), 2.23–2.12 (2H, m), 2.11–1.99 (3H, m), 1.93 (1H, m), 1.73–1.57 (2H, m), 1.41–1.24 (4H, m), 1.23–1.07 (2H, m); ^{13}C NMR (100 MHz, $CDCl_3$) δ 180.0 (s), 179.4 (s), 170.3 (s), 160.4 (s), 142.3 (s), 138.9 (d), 134.2 (d), 133.9 (s), 133.3 (d), 132.1 (d), 131.2 (d), 139.7 (d), 129.7 (d), 128.8 (d), 127.6 (d), 127.2 (d), 126.7 (s), 124.5 (d), 123.7 (d), 120.7 (d), 120.4 (s), 116.2 (d), 116.0 (d), 114.3 (t), 70.5 (d), 56.8 (t), 55.6 (t), 35.3 (t), 33.7 (t), 30.7 (t), 29.2 (t), 28.9 (t), 28.8 (t), 25.3 (t), 23.6 (t). Additional peaks arise from rotamers at 131.1 (d), 128.9 (d), 124.4 (d), 120.3 (s), 70.4 (d); ^{19}F NMR (376 MHz; $CDCl_3$) 113.7 (s); IR (ν_{max}/cm^{-1}, neat): 3020, 1741, 1365, 1222; HRMS-ESI (calcd for $C_{35}H_{39}N_3O_3F^{58}Ni$ [M + H]$^+$) 626.2329, found 626.2321 (Δ = −1.3 ppm), (calcd for $C_{35}H_{38}N_3O_3FNa^{58}Ni$ [M + Na]$^+$) 648.2148, found 648.2155 (Δ = 1.1 ppm); HPLC (OD-H column, hexane (50%)–^{i}PrOH isocratic): 13.62 min. X-rays: see ESI.†

(*S*)-({2-[1-(2-Fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl}phenylmethylene)-(*S*)-pentenylalaninato-*N*,*N*′,*N*″,*O*}nickel(II) (Ni-α-pent-4-enyl Ala-2-FBPB) (9). Yield 62%; m.p: 190–192 °C (hexane–EtOAc); $[\alpha]_D^{20}$ +2271.2 (*c* 0.05, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.29 (1H, td, *J* = 7.4, 1.0 Hz, Ar–C*H*), 8.03 (1H, d, *J* = 8.5 Hz, Ar–C*H*), 7.51–7.44 (2H, m, Ar–C*H*), 7.38 (1H, m, Ar–C*H*), 7.33 (1H, m, Ar–C*H*), 7.29 (1H, m Ar–C*H*), 7.20 (1H, t, *J* = 7.4 Hz, Ar–C*H*), 7.16 (1H, ddd, *J* = 8.4, 6.2, 2.2 Hz, Ar–C*H*), 7.06 (1H, t, *J* = 9.1 Hz, Ar–C*H*), 6.97 (1H, d, *J* = 7.6 Hz, Ar–C*H*), 6.68–6.61 (2H, m, Ar–C*H*), 5.86 (1H, ddt, *J* = 17.0, 10.3, 6.5 Hz, C*H*═CH$_2$), 5.08 (1H, dd, *J* = 17.0, 1.0 Hz, CH═C*H*$_2$*cis*), 5.02 (1H, d, *J* = 10.3 Hz, CH═C*H*$_2$*trans*), 4.52 (1H, d, *J* = 13.1 Hz, N–C*H*H), 3.95 (1H, d, *J* = 13.1 Hz, N–CH*H*), 3.60 (1H, dd, *J* =

9.9, 6.5 Hz, α(Pro)-C*H*), 3.41 (1H, dd, *J* = 10.7, 6.4 Hz, δ(Pro)-C*H*H), 3.26 (1H, m, β(Pro)-C*H*H), 2.78 (1H, m, γ(Pro)-C*H*H), 2.52 (1H, m, γ(Pro)-CH*H*), 2.40 (1H, m, γ-C*H*H), 2.17–1.98 (5H, m, δ-CH_2, γ-CH*H*, δ(Pro)-CH*H*, β(Pro)-CH*H*), 1.75–1.62 (2H, m, β-CH_2), 1.23 (3H, s, CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 182.3 (s), 180.1 (s), 172.4 (s), 141.5 (s), 137.8 (d), 136.5 (s), 134.2 (d), 133.4 (d), 131.6 (d), 131.3 (d), 130.3 (d), 129.4 (d), 128.7 (s), 127.9 (d), 127.3 (s), 126.9 (d), 124.5 (d), 124.0 (d), 120.8 (d), 120.3 (s), 116.2 (d), 116.0 (d), 115.4 (t), 78.1 (s), 70.1 (d), 56.7 (t), 55.9 (t), 39.8 (t), 33.7 (t), 30.5 (t), 29.6 (q), 25.2 (t), 23.2 (t); ^{19}F NMR (376 MHz, $CDCl_3$) δ −113.7 (s); IR (ν_{max}/cm^{-1}, neat) 2927, 2854, 1662, 1635, 1435, 1354, 1334, 1254, 1234, 1163, 1111, 1064, 765, 745, 708, 698, 673; HRMS-ASAP (calcd for $C_{33}H_{35}N_3O_3F^{58}Ni$ [M + H]$^+$) 598.2016, found 598.2004 (Δ = −2.0 ppm), (calcd for $C_{33}H_{35}N_3O_3F^{60}Ni$ [M + H]$^+$) 600.1970, found 600.1999 (Δ = 4.8 ppm); HPLC (OD-H column, hexane (60%)–^{i}PrOH isocratic): 10.78 min. X-rays: see ESI.†

(*S*)-({2-[1-(2-Fluorobenzyl)benzyl)pyrrolidine-2-carboxamide]-phenyl}phenylmethylene)-(*S*)-octenylalaninato-*N*,*N*′,*N*″,*O*}nickel(II) (Ni-α-oct-4-enyl-Ala-2-FBPB) (10). Yield 42%; m.p: 70–72 °C (hexane–EtOAc); $[\alpha]_D^{20}$ +1978.1 (*c* 0.05, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.29 (1H, td, *J* = 7.5, 1.2 Hz, Ar–C*H*), 8.03 (1H, d, *J* = 9.1 Hz, Ar–C*H*), 7.50–7.44 (2H, m, Ar–C*H*), 7.38 (1H, m, Ar–C*H*), 7.34 (1H, m, Ar–C*H*), 7.29 (1H, m, Ar–C*H*), 7.20 (1H, td, *J* = 7.6, 1.2 Hz, Ar–C*H*), 7.15 (1H, ddd, *J* = 8.5, 6.2, 2.4 Hz, Ar–C*H*), 7.09 (1H, t, *J* = 9.1 Hz, Ar–C*H*), 6.95 (1H, d, *J* = 7.9 Hz, Ar–C*H*), 6.68–6.61 (2H, m, Ar–C*H*), 5.81 (1H, ddt, *J* = 17.1, 10.3, 6.8 Hz, C*H*═CH$_2$), 5.01 (1H, dq, *J* = 17.1, 1.6 Hz, CH═CH_2*cis*), 4.94 (1H, m, CH═CH_2*trans*), 4.52 (1H, d, *J* = 13.1 Hz, N–C*H*H), 3.96 (1H, d, *J* = 13.1 Hz, N–CH*H*),), 3.59 (1H, dd, *J* = 10.1, 6.6 Hz, α(Pro)-C*H*), 3.41 (1H, dd, *J* = 10.5, 6.6 Hz, δ(Pro)-C*H*H), 3.27 (1H, m, β(Pro)-C*H*H), 2.77 (1H, m, γ(Pro)-C*H*H), 2.52 (1H, m, γ(Pro)-CH*H*), 2.35 (1H, m, γ-C*H*H)), 2.12–1.97 (5H, m, δ-CH_2, γ-CH*H*, δ(Pro)-CH*H*, β(Pro)-CH*H*), 1.69 (1H, td, *J* = 13.3 Hz, η-C*H*H), 1.58 (1H, m, η-CH*H*), 1.48–1.37 (4H, m, β-CH_2, ε-CH_2),1.37–1.27 (2H, m, ζ-CH_2), 1.23 (3H, s, CH_3); ^{13}C NMR (126 MHz, $CDCl_3$) δ 182.5 (s), 180.1 (s), 172.4 (s), 141.5 (s), 138.9 (d), 136.5 (s), 134.2 (d), 133.4 (d), 131.6 (d), 131.3 (d), 130.4 (d), 129.4 (d), 128.5 (s), 128.0 (d), 127.3 (s), 126.9 (d), 124.6 (d), 124.0 (d), 120.8 (d), 120.5 (s), 116.3 (d), 116.1 (d), 114.4 (t), 78.3 (s), 70.1 (d), 56.7 (t), 55.9 (t), 40.3 (t), 33.8 (t), 30.6 (t), 29.7 (q), 29.6 (t), 29.2 (t), 28.9 (t), 20.0 (t), 23.22 (t); ^{19}F NMR (376 MHz, $CDCl_3$) δ −113.7 (s); IR (ν_{max}/cm^{-1}, neat) 2923, 2858, 1672, 1639, 1574, 1490, 1456, 1438, 1352, 1251, 1165, 1112, 1062, 903, 750, 711, 669; HRMS-ASAP: (calcd for $C_{36}H_{41}N_3O_3F^{58}Ni$ [M + H]$^+$) 640.2485, found 640.2485 (Δ = 0 ppm), (calcd for $C_{36}H_{41}N_3O_3F^{60}Ni$ [M + H]$^+$) 642.2440, found 642.2471 (Δ = 4.8 ppm); HPLC (OD-H column, hexane (50%)–^{i}PrOH isocratic): 8.34 min.

General method of formation of amino acids (11)–(14)

(*S*)-({2-[1-(2-Fluorobenzyl)pyrrolidine-2-carboxamide]phenyl}phenylmethylene)-(*S*)-pentenylglycinato-*N*,*N*′,*N*′,*O*}nickel(II) **5** (0.49 g, 0.87 mmol, 1.0 equiv.) was dissolved in methanol (52 mL) and added drop wise to 3 M hydrochloric acid (7.0 mL) at 70 °C, over 5 minutes. The solution was stirred for 30 minutes, a colour change of red to a transparent green/yellow was observed. The reaction was cooled to room temperature and methanol removed *in vacuo*. The residue was taken up in water (25.0 mL) and extracted with dichloromethane (3 × 25.0 mL). The organic extracts were combined, dried ($MgSO_4$) and concentrated *in vacuo* to reclaim compound **4** for re-use. The aqueous solution was added to a Dowex 50WX2 200 H$^+$ column (pre-washed with H_2O to pH 7) and washed with H_2O–EtOH (1 : 1, v/v, 250.0 mL) to remove the nickel salts. The amino acid product was eluted with 20% NH_4OH–EtOH (1 : 1 v/v, 350.0 mL). Fractions containing product were identified by staining with ninhydrin. The fractions where combined and concentrated *in vacuo*, triturated with Et_2O to give (*S*)-2-amino-hept-6-enoic acid as a white powder.

(*S*)-2-Aminohept-6-enoic acid (11). Yield 60%; m.p: 225–226 °C (Et_2O) (lit: 224–225 °C);[12] $[\alpha]_D^{30}$ +1.5 (*c* 0.1, H_2O) (lit: +9.5 (*c* 0.5, H_2O, 26 °C);[12] ^{1}H NMR (300 MHz, D_2O) δ 5.85 (1H, ddt, *J* = 17.5, 10.5, 6.7 Hz, C*H*═CH$_2$), 5.06 (1H, d, *J* = 19.3 Hz, CH═CH_2 *trans*), 4.99 (1H, d, *J* = 10.5 Hz, CH═CH_2 *cis*), 3.67 (1H, t, *J* = 6.1 Hz, C*H*CH$_2$CH$_2$), 2.08 (2H, q, *J* = 7.0 Hz, CH_2CH═CH$_2$), 1.92–1.70 (2H, m, CHCH_2CH$_2$), 1.56–1.32 (2H, m, CH$_2$CH_2CH$_2$). NH_2 and OH protons are missing from spectrum; ^{13}C NMR (75 MHz, $CDCl_3$) δ 175.5 (s, *C*═O), 138.6 (d, *C*H═CH$_2$), 115.0 (t, CH═*C*H$_2$), 54.8 (d, *C*HCH$_2$CH$_2$), 32.5 (t, *C*H$_2$CH═CH$_2$), 30.1 (t, CH*C*H$_2$CH$_2$), 23.6 (t, CH$_2$*C*H$_2$CH$_2$); IR (ν_{max}/cm^{-1}, neat): 2929 (br), 1579, 1514, 1407, 1355, 1322, 989, 912, 839; HRMS-ESI (calcd for $C_7H_{14}NO_2$ [M + H]$^+$) 144.1025, found 144.1028 (Δ = 2.1 ppm).

(*S*)-2-Aminodec-9-enoic-acid (12). Yield 67%; m.p: 255–257 °C (Et_2O); $[\alpha]_D^{30}$ +0.7 (*c* 0.03, MeOH); ^{1}H NMR (300 MHz, D_2O) δ 5.85 (1H, ddt, *J* = 17.1, 10.3, 6.7 Hz, C*H*═CH$_2$), 4.97 (1H, dq, *J* = 17.2, 2.3 Hz, CH═CH_2 *trans*), 4.90 (1H, ddt, *J* = 10.3, 2.5, 1.2 Hz, CH═CH_2 *cis*), 3.13 (1H, dd, *J* = 6.8, 6.1 Hz, C*H*CH$_2$CH$_2$), 1.97 (2H, qt, *J* = 7.0, 1.2 Hz, CH_2CH═CH$_2$), 1.57–1.41 (2H, m, CHCH_2CH$_2$), 1.37–1.18 (8H, m, CH$_2$(CH_2)$_4$CH$_2$). NH_2 and OH protons are missing from spectrum; ^{13}C NMR (75 MHz, $CDCl_3$) δ 184.0 (s, *C*═O), 140.5 (d, *C*H═CH$_2$), 114.0 (t, CH═*C*H$_2$), 56.1 (d, *C*HCH$_2$CH$_2$), 34.7 (t, *C*H$_2$CH═CH$_2$), 33.0 (t, CH*C*H$_2$CH$_2$), 28.5 (t, *C*H$_2$), 28.1 (t, *C*H$_2$), 28.0 (t, *C*H$_2$), 24.9 (t, *C*H$_2$); IR (ν_{max}/cm^{-1}, neat): 2929 (br), 2925, 2851, 1579, 1513, 1443, 1406, 1319, 1066, 911; HRMS-ESI (calcd for $C_{10}H_{20}NO_2$ [M + H]$^+$) 186.1494, found 186.1493 (Δ = −0.5 ppm).

(*S*)-2-Amino-2-methylhept-6-enoic acid (13). Yield 65%; m.p: 250–252 °C (Et_2O) (lit: >200 °C decomp);[21] $[\alpha]_D^{20}$ +3.22 (*c* 0.05, MeOH) (lit: +10.8 (*c* 0.83, MeOH, 25 °C);[21] ^{1}H NMR (300 MHz, D_2O) δ 5.84 (1 H, ddt, *J* = 17.0, 10.2, 6.7 Hz, C*H*═CH$_2$), 5.11–4.96 (2 H, m, CH═CH_2), 2.07 (2 H, q, *J* = 7.0 Hz, δ-CH_2), 1.86 (1 H, m, β-C*H*H), 1.71 (1 H, m, β-CH*H*), 1.54–1.40 (4 H, m, CH_3, γ-C*H*H), 1.31 (1 H, m, γ-CH*H*); ^{13}C NMR (100 MHz, D_2O) δ 177.0 (s, *C*═O), 138.7 (d, *C*H═CH$_2$), 115.1 (t, CH═*C*H$_2$), 61.6 (s, α-q*C*), 36.7 (t, β-CH$_2$), 32.8 (t, δ-CH$_2$), 22.5 (t, γ-*C*H$_2$), 22.5 (q, *C*H$_3$); IR (ν_{max}/cm^{-1}, neat) 3124, 3033, 2982, 1738, 1595, 1456, 1397, 1367, 1316, 1232, 912; HRMS-ESI (calcd for $C_8H_{16}NO_2$ [M + H]$^+$) 158.1181, found 158.1175 (Δ = −3.8 ppm).

(*S*)-2-Amino-2-methyldec-9-enoic acid (14). Yield 68%; m.p: 262–264 °C (Et_2O); $[\alpha]_D^{20}$ +3.59 (*c* 0.05, MeOH); ^{1}H NMR (500 MHz, D_2O) δ 5.89 (1 H, ddt, *J* = 17.1, 10.3, 6.8 Hz, C*H*—CH_2), 5.02 (1 H, d, *J* = 17.1 Hz, CH—CH_2*cis*), 4.95 (1 H, d, *J* = 10.3 Hz, CH—CH_2*trans*), 2.30 (2 H, q, *J* = 6.8 Hz, η-CH_2), 1.82 (1 H, m, β-C*H*H), 1.68 (1 H, m, β-CH*H*), 1.44 (3 H, s, CH_3), 1.39–1.26 (7 H, m, γ-C*H*H, δ-CH_2, ε-CH_2, ζ-CH_2), 1.19 (1 H, m, γ-CH*H*); ^{13}C NMR (126 MHz, D_2O) δ 177.2 (s, *C*—O), 140.3 (d, *C*H—CH_2), 114.0 (t, CH—*C*H_2), 61.7 (s, α-q*C*), 37.2 (t, β-CH_2), 33.0 (t, η-*C*H_2), 28.5 (t, δ-*C*H_2), 27.9 (t, ζ-*C*H_2), 27.8 (t, ε-*C*H_2), 23.1 (t, γ-*C*H_2), 22.5 (t, α-*C*H_3); IR (ν_{max}/cm^{-1}, neat) 3079, 2981, 2922, 2855, 1595, 1457, 1434, 1399, 1367, 1318, 1259, 993, 909, 791; HRMS-ESI: (calcd for $C_{11}H_{22}NO_2$ [M + H]$^+$) 200.1651, found 200.1651 (Δ = 0.0 ppm).

General method of formation of Fmoc protected amino acids (1)–(2), (15)–(16)

(*S*)-amino hept-6-enoic acid **11** (90 mg, 0.63 mmol, 1.0 equiv.) and potassium carbonate (174 mg, 1.26 mmol, 2.0 equiv.) was dissolved in water (2.0 mL) and cooled to 0 °C. 9-Fluorenylmethyl *N*-succinimidyl carbonate (320 mg, 0.95 mmol, 1.5 equiv.) was dissolved in dioxane (4.0 mL) and added drop wise to aqueous solution over 20 minutes. The reaction was warmed to room temperature and left for 24 hours. An excess volume of water was added and the mixture extracted (2 × 20.0 mL) with ethyl acetate. The combined organic fractions were back extracted with saturated bicarbonate solution and the aqueous layers acidified to pH 1 with 3 M HCl. The aqueous fractions were then extracted with EtOAc (3 × 20.0 mL). The organic layers were dried ($MgSO_4$) and concentrated *in vacuo*. Purification by column chromatography (SiO_2 eluted MeOH–CH_2Cl_2–AcOH (2 : 97 : 1)) gave the title compound as a white powder.

(*S*)-2((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)hept-6-enoic acid (15). Yield 65%; m.p: 125–127 °C (Et_2O); $[\alpha]_D^{30}$ +0.1 (*c* 0.1, $CHCl_3$) (lit: +3.1 (1.0, $CHCl_3$, 22.4 °C);[12] ^{1}H NMR (500 MHz, MeOD) δ 7.79 (2 H, d, *J* = 7.6 Hz, Ar–C*H*), 7.68 (2 H, d, *J* = 7.7 Hz, Ar–C*H*), 7.39 (2 H, t, *J* = 7.2 Hz, Ar–C*H*), 7.31 (2 H, t, *J* = 7.5 Hz, Ar–C*H*), 5.81 (1 H, ddt, *J* = 17.0, 10.2, 6.7 Hz, C*H*—CH_2), 5.03 (1 H, d, *J* = 17.5 Hz, CH—CH_2 *trans*), 4.96 (1 H, d, *J* = 9.5 Hz, CH—CH_2 *cis*), 4.35 (2 H, d, *J* = 7.2 Hz, CH_2OCONH), 4.23 (1 H, t, *J* = 6.8 Hz, C*H*CH$_2$OCONH), 4.14 (1 H, dd, *J* = 9.1, 4.8 Hz, C*H*CH$_2$CH$_2$), 2.14–2.05 (2 H, m, CH_2CH—CH_2), 1.85 (1 H, m, CHC*H*$_2$CH$_2$), 1.69 (1 H, m, CHC*H*$_2$CH$_2$), 1.56–1.44 (2 H, m, CH_2CH_2CH$_2$). NH and OH protons are missing from spectrum; ^{13}C NMR (125 MHz, MeOD) δ 176.1 (s), 158.8 (s), 145.6 (s), 142.6 (s), 139.4 (d), 128.8 (d), 128.2 (d), 126.3 (d), 120.9 (d), 115.4 (t), 68.0 (t), 55.2 (d), 48.3 (d, *identified by HMQC*), 34.3 (t), 32.2 (t), 26.3 (t); IR (ν_{max}/cm^{-1}, neat): 3336, 3017, 2971, 1739, 1681, 1532, 1446, 1366, 1229, 1217, 910, 758, 736; HRMS-ESI (calcd for $C_{22}H_{24}NO_4$ [M + H]$^+$) 366.1705, found 366.1705 (Δ = 0.0 ppm), (calcd for $C_{22}H_{23}NO_4Na$ [M + Na]$^+$) 388.1525, found 388.1500 (Δ = −6.4 ppm); HPLC (OD-H column, hexane–^{i}PrOH isocratic +0.1% AcOH): 9.56 min.

(*S*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)dec-9-enoic acid (16). Yield 63%; m.p: 225–226 °C (Et_2O); $[\alpha]_D^{30}$ +0.2 (*c* 0.02, $CHCl_3$); ^{1}H NMR (500 MHz, MeOD) δ 7.77 (2 H, d, *J* = 7.3 Hz, Ar–C*H*), 7.60 (2 H, d, *J* = 6.7 Hz, Ar–C*H*), 7.41 (2 H, t, *J* = 7.6 Hz, Ar–C*H*), 7.32 (2 H, t, *J* = 7.0 Hz, Ar–C*H*), 5.81 (1 H, ddt, *J* = 17.0, 10.2, 6.7 Hz, C*H*—CH_2), 5.03 (1 H, d, *J* = 17.5 Hz, CH—CH_2 *trans*), 4.96 (1 H, d, *J* = 9.5 Hz, CH—CH_2 *cis*), 4.35 (2 H, d, *J* = 7.2 Hz, CH_2OCONH), 4.23 (1 H, t, *J* = 6.8 Hz, C*H*CH$_2$OCONH), 4.14 (1 H, dd, *J* = 9.1, 4.8 Hz, C*H*CH$_2$CH$_2$), 2.03 (2 H, q, *J* = 6.7 Hz, CH_2CH—CH_2), 1.89 (1 H, m, CHC*H*$_2$CH$_2$), 1.67 (1 H, m, CHC*H*$_2$CH$_2$), 1.46–1.18 (8 H, m, CH_2(CH_2)$_4$CH$_2$). NH and OH protons are missing from spectrum; ^{13}C NMR (125 MHz, MeOD) δ 176.4 (s), 158.9 (s), 145.4 (s), 142.8 (s), 140.2 (d), 128.9 (d), 128.3 (d), 126.4 (d), 121.1 (d), 114.9 (t), 68.1 (t), 55.5 (d), 47.4 (d), 35.0 (t), 32.9 (t), 30.1 (3×t), 27.0 (t); IR (ν_{max}/cm^{-1}, neat): 3072, 2931, 2858, 2487, 1687, 1537, 1450, 1264, 1236, 1167, 1086, 1045, 992, 907, 738; HRMS-ESI (calcd for $C_{25}H_{30}NO_4$ [M + H]$^+$) 408.2175, found 408.2170 (Δ = −1.2 ppm), (calcd for $C_{25}H_{29}NO_4Na$ [M + Na]$^+$) 430.1994, found 430.1989 (Δ = −1.2 ppm); HPLC (OD-H column, hexane–^{i}PrOH isocratic +0.1% AcOH): 7.54 min.

(*S*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-2-methylhept-6-enoic acid (1). Yield 82%; $[\alpha]_D^{20}$ +3.5 (*c* 1.0 in MeOH); ^{1}H NMR (300 MHz; $CDCl_3$) δ 7.78 (2 H, d, *J* = 7.5 Hz, Ar–C*H*), 7.60 (2 H, d, *J* = 7.2 Hz, Ar–C*H*), 7.41 (2 H, t, *J* = 7.4 Hz, Ar–C*H*), 7.33 (2 H, t, *J* = 7.5, Ar–C*H*), 5.76 (1 H, m, CH_2C*H*—CH_2), 5.50 (1 H, s, N*H*), 5.09–4.89 (2 H, m, CH_2CH—CH_2), 4.53–4.33 (2 H, m, C(O)OCH_2CH), 4.23 (1 H, t, *J* = 6.6 Hz, C(O)OCH$_2$C*H*), 2.22–1.95 (2 H, m, CH_2CH—CH_2), 1.95–1.78 (1 H, m, C(Me)-CH_2CH_2), 1.60 (3 H, s, C(CH_3)CH$_2$), 1.50–1.26 (2 H, t, *J* = 7.1 Hz, C(Me)CH_2); ^{13}C NMR (75 MHz; $CDCl_3$) δ 178.3 (s), 154.8 (s), 143.8 (s), 141.4 (s), 138.0 (d), 127.7 (d), 127.1 (d), 125.0 (d), 120.0 (d), 115.1 (t), 66.5 (t), 59.7 (s), 47.2 (d), 36.2 (t), 33.4 (t), 30.9 (t), 23.3 (q); IR (ν_{max}/cm^{-1}, neat) 2932, 1701, 1507, 1450, 1334, 1253, 1074, 910, 758, 737; HRMS-ESI (calcd for $C_{23}H_{26}NO_4$ [M + H]$^+$) 380.1862, found 380.1860 (Δ = −0.5 ppm), (calcd for $C_{23}H_{25}NO_4Na$ [M + Na]$^+$) 402.1681, found 402.1692 (Δ = 2.7 ppm). HPLC (OD-H column, hexane (50%)–^{i}PrOH isocratic +0.1% AcOH): 8.07 min.

(*S*)-2-((((9*H*-Fluoren-9-yl)methoxy)carbonyl)amino)-2-methyldec-9-enoic acid (2). Yield 62%; $[\alpha]_D^{20}$ +5.89 (*c* 1.0, $CHCl_3$); ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.77 (2 H, d, *J* = 7.5 Hz, Ar–C*H*), 7.59 (2 H, d, *J* = 7.5 Hz, Ar–C*H*), 7.40 (2 H, t, *J* = 7.4 Hz, Ar–C*H*), 7.32 (2 H, t, *J* = 7.4 Hz, Ar–C*H*), 5.79 (1 H, ddt, *J* = 17.1, 10.2, 6.8 Hz, C*H*—CH_2), 5.46 (1 H, br s, N*H*), 4.98 (1 H, dd, *J* = 17.1, 1.5 Hz, CH—CH_2*cis*), 4.92 (1 H, d, *J* = 10.1 Hz, CH—CH_2*trans*), 4.41 (2 H, s, C(O)OCH_2CH), 4.22 (1 H, t, *J* = 6.8 Hz, C(O)OCH$_2$C*H*), 2.02 (2 H, q, *J* = 6.9 Hz, CH_2CH—CH_2), 1.82 (1 H, m, C*H*HCH$_2$CH—CH_2), 1.60 (3 H, s, CH_3), 1.39–1.10 (9 H, m, CH*H*CH$_2$CH—CH_2, CH_3C(CH_2)$_4$); ^{13}C NMR (100 MHz, $CDCl_3$) δ 179.3 (s), 157.7 (s), 143.9 (s), 141.4 (s), 139.0 (d), 127.7 (d), 127.0 (d), 124.9 (d), 120.0 (d), 114.2 (t), 66.5 (d), 59.8 (t), 47.2 (s), 36.8 (t), 33.7 (t), 29.3 (t), 28.9 (t), 28.8 (t), 23.9 (t), 23.2 (q); IR (ν_{max}/cm^{-1}, neat) 2926, 1705, 1506, 1450, 1338, 1251, 1105, 1086, 1052, 995, 909, 776, 758, 737, 674; HRMS-ESI (calcd for $C_{26}H_{32}NO_4$ [M + H]$^+$) 422.2331, found 422.2331 (Δ = 0.0 ppm);

HPLC (OD-H column, hexane (50%)–^{i}PrOH isocratic +0.1% AcOH): 7.28 min.

Acknowledgements

We thank Prof. Paul M. Cullis for insightful discussions. This work was funded by Engineering and Physical Sciences Research Council (EPSRC) (EP/L018152/1), University of Leicester Innovation Fellowship for A.G.J. The authors thank Mick Lee (mass spectrometry), Dr Gerry Griffith (NMR) and Kuldip Singh (X-ray crystallography) for technical assistance.

Notes and references

1 A. J. Wilson, *Chem. Soc. Rev.*, 2009, **38**, 3289–3300.
2 G. L. Verdine and G. J. Hilinski, *Methods Enzymol.*, 2012, **503**, 3–33.
3 S. A. Kawamoto, A. Coleska, X. Ran, H. Yi, C.-Y. Yang and S. Wang, *J. Med. Chem.*, 2012, **55**, 1137–1146.
4 C. Bracken, J. Gulyas, J. W. Taylor and J. Baum, *J. Am. Chem. Soc.*, 1994, **116**, 6431–6432.
5 A. Patgiri, K. K. Yadav, P. S. Arora and D. Bar-Sagi, *Nat. Chem. Biol.*, 2011, **7**, 585–587.
6 D. Y. Jackson, D. S. King, J. Chmielewski, S. Singh and P. G. Schultz, *J. Am. Chem. Soc.*, 1991, **113**, 9391–9392.
7 T. M. Klingler and D. L. Brutlag, *Protein Sci.*, 1994, **3**, 1847–1857.
8 P. Balaram, *Curr. Opin. Struct. Biol.*, 1992, **2**, 845–851.
9 H. E. Blackwell and R. H. Grubbs, *Angew. Chem., Int. Ed.*, 1998, **37**, 3281–3284.
10 C. E. Schafmeister, J. Po and G. L. Verdine, *J. Am. Chem. Soc.*, 2000, **122**, 5891–5892.
11 Y. W. Kim, T. N. Grossmann and G. L. Verdine, *Nat. Protoc.*, 2011, **6**, 761–771.
12 D. J. Yeo, S. L. Warriner and A. J. Wilson, *Chem. Commun.*, 2013, **49**, 9131–9133.
13 C. H. Douse, S. J. Maas, J. C. Thomas, J. A. Garnett, Y. Sun, E. Cota and E. W. Tate, *ACS Chem. Biol.*, 2014, DOI: 10.1021/cb500271c.
14 R. M. Williams and M. N. Im, *J. Am. Chem. Soc.*, 1991, **113**, 9276–9286.
15 W. Qiu, V. A. Soloshonok, C. Cai, X. Tang and V. J. Hruby, *Tetrahedron*, 2000, **56**, 2577–2582.
16 Y. N. Belekon, V. I. Maleev, M. B. Saporovskaya, V. I. Bakhmutov, T. V. Timofeeva, A. S. Batsanov, Y. T. Struchkov and V. M. Belikov, *Koord. Khim.*, 1988, **11**, 1565–1575.
17 A. Popkov, A. Gee, M. Nádvorník and A. Lyčka, *Transition Met. Chem.*, 2002, **27**, 884–887.
18 A. S. Saghyan, A. S. Dadayan, S. A. Dadayan, A. F. Mkrtchyan, A. V. Geolchanyan, L. L. Manasyan, H. R. Ajvazyan, V. N. Khrustalev, H. H. Hambardzumyan and V. I. Maleev, *Tetrahedron: Asymmetry*, 2010, **21**, 2956–2965.
19 A. S. Saghiyan, S. A. Dadayan, S. G. Petrosyan, L. L. Manasyan, A. V. Geolchanyan, S. M. Djamgaryan, S. A. Andreasyan, V. I. Maleev and V. N. Khrustalev, *Tetrahedron: Asymmetry*, 2006, **17**, 455–467.
20 A. D. Bautista, J. S. Appelbaum, C. J. Craig, J. Michel and A. Schepartz, *J. Am. Chem. Soc.*, 2010, **132**, 2904–2906.
21 R. P. M. Storcken, L. Panella, F. L. van Delft, B. Kaptein, Q. B. Broxterman, H. E. Schoemaker and F. P. J. T. Rutjes, *Adv. Synth. Catal.*, 2007, **349**, 161–164.

Reports in Organic Chemistry

Dovepress
open access to scientific and medical research

Open Access Full Text Article

REVIEW

Regulation of protein–protein interactions using stapled peptides

Naomi S Robertson
Andrew G Jamieson

Centre for Chemical Biology, Department of Chemistry, University of Leicester, Leicester, UK

Abstract: The targeting of protein–protein interactions (PPIs) that include secondary structure motifs such as the α-helix and β-sheet is a challenge in chemical biology. A new class of compounds called stapled peptides have been developed that mimic α-helix secondary structures involved in PPIs. These conformationally constrained peptides also show favorable physicochemical properties and so are promising lead compounds for drug discovery. This review focuses on the design aspects of hydrocarbon constrained α-helical proteomimetics and provides examples in which they target biologically relevant PPIs.

Keywords: protein–protein interactions, stapled helix, proteomimetics, peptide

Introduction

Protein–protein interactions (PPIs) are becoming increasingly popular drug targets because the chemical space required for inhibition is underexplored and so the intellectual property landscape is relatively barren.

However, the development of novel, cell-permeable molecules that regulate PPIs remains one of the major challenges in chemical biology due to the relatively large (typically ~1600 Å^2) and dynamic nature of the PPI interface. While these molecular recognition events involve interactions over large surface areas, the majority of the binding affinity and specificity has been found to originate from constellations of a few amino acid residues that are described as interaction "hotspots".[1] At a molecular level, these constellations of residues are frequently found to be specific protein secondary structures such as the α-helix or -strand. Although these motifs are usually thermodynamically favored in folded proteins, isolated peptides typically lack the ability to spontaneously adopt these bioactive conformations. Peptides also have a number of other therapeutically undesirable properties including poor bioavailability, limited stability toward peptidase proteolysis, and lack of membrane permeability that can potentially limit their use as drugs.

A challenge therefore exists to design molecules that occupy the chemical space large enough to modulate PPIs and yet small enough to have suitable drug-like physicochemical properties. Recent research efforts have focused on the design and synthesis of proteomimetics, molecules that mimic both the structure and/or function of proteins. Many different strategies have been used to interrogate PPIs and these have previously been reviewed.[2,3] Proteomimetics, and stapled peptides in particular, are proving useful for targeting PPIs.[4,5] The main focus of this review is the design, synthesis, and application of stapled peptides as molecular probes to disrupt PPIs.

Correspondence: Andrew G Jamieson
Centre for Chemical Biology, Department of Chemistry, University of Leicester, Leicester LE1 7RH, UK
Tel +44 116 252 2105
Email andrew.jamieson@leicester.ac.uk

submit your manuscript | www.dovepress.com
Dovepress
http://dx.doi.org/10.2147/ROC.S68161

Strategies for targeting PPIs

Many different strategies for targeting PPIs have been studied, including the use of small molecule peptidomimetic inhibitors, which include ABT-737[6] and ABT-263[7] from Abbott Laboratories, that were found to be inhibitors of the Bcl-x_L/Bak PPI (Figure 1). Additionally, the Nutlin[8] and benzodiazepinedione[9] families of compounds, developed by Hoffman-La Roche and Johnson & Johnson pharmaceuticals, respectively, were found to be inhibitors of the p53/mDM2 PPI. As well as small molecules, nonpeptidic scaffolds can be used to target PPIs such as the terphenyl scaffolds first developed by the Hamilton group.[10] Hara et al have developed polyproline-type helix peptoids that target the p53/mDM2 PPI.[11]

A large proportion of the PPIs that have been reported are mediated by α-helices.[12] Particular emphasis has therefore been placed on the design of conformational constraints that induce α-helix structure. Examples of α-helix constraints include salt bridges between charged amino acid side chain residues,[13,14] lactam bridges,[15,16] disulfide bridges,[17] hydrogen bond surrogates,[18,19] hydrophobic interactions,[20] metal ligation,[21,22] triazole staples synthesized from alkenyl and azido side chain residues,[23] photocontrollable macrocycles,[24] introduction of α,α-disubstituted amino acids,[25,26] and hydrocarbon staples[27] (Figure 2). In addition, short α-helical peptides have been nucleated with the help of capping groups[28–30] and using designed molecules to template the helix.[31,32]

Proteomimetic compounds of this type demonstrate a number of attractive physicochemical properties in comparison with small molecules, peptides, and biologics. For example, improved binding affinity is observed by overcoming the entropic penalty associated with peptide folding.

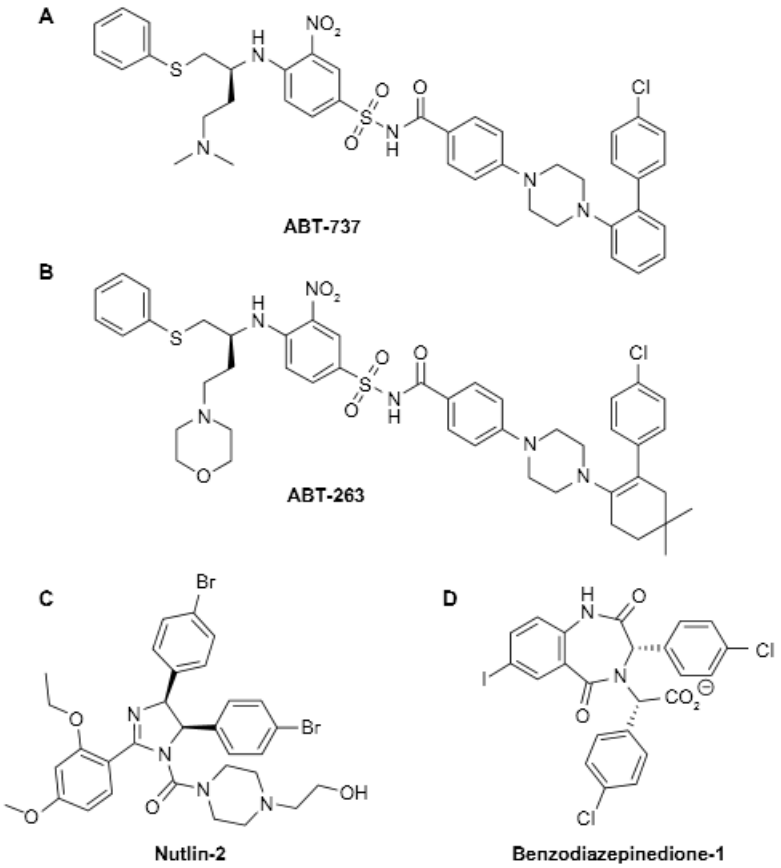


Figure 1 Structures of small molecule peptidomimetic inhibitors (**A**) ABT-737, (**B**) ABT-263 that target the Bcl-x_L/Bak PPII, (**C**) Nutlin-2, and (**D**) benzodiazepinedione-1 that target the p53/mDM2 PPI.

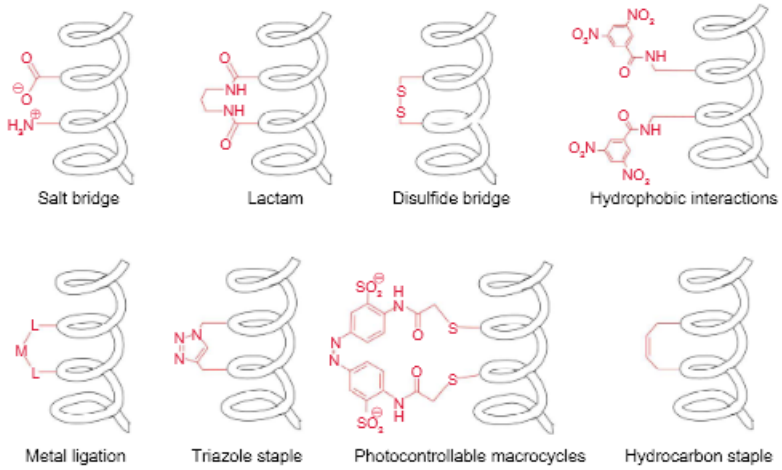


Figure 2 Strategies to constrain α-helical peptides.

Good target specificity relative to small molecules and comparable to biologics is achieved due to maintaining the original amino acid residues, with the correct orientation and spacing, to facilitate the molecular recognition event. Conformational constraints also prevent the required extended sheet conformation for recognition by peptidases and so proteomimetics generally have a significantly longer half-life in blood serum.

Of these techniques, hydrocarbon stapling is coming of age as a method to develop peptidomimetic drugs that regulate PPIs and so will be the main focus of this article.

Hydrocarbon peptide stapling

The first "stapled" peptide was reported by Miller et al[33,34] and Blackwell et al[35] in 1998. The effects of using ring closing metathesis (RCM) on peptide antibiotics containing two alkenyl side chains and rich in α-aminoisobutyric acid (Aib) residues are known to constrain 3_{10}-helices through the Thorpe–Ingold effect.[25,36,37]

More recently, Korsmeyer, Verdine, and Walensky have developed the rules for effective peptide stapling,[27,38] thus making the technology transferable to other PPIs. The social and economic value of this approach became apparent in 2005 with the foundation of Aileron Therapeutics, Inc. to develop and commercialize a drug-discovery pipeline based on stapled peptides.[39] In 2013, the company announced the successful completion of the first-in-human clinical trial of a stapled peptide drug, ALRN-5281 for the treatment of rare endocrine diseases such as adult growth hormone deficiency and HIV lipodystrophy. The main outcome of the trial showed ALRN-5281 had no incidence of adverse events in healthy subjects.[40] Additionally, Aileron Therapeutics is also developing a new stapled peptide drug, ATSP-7041, that targets intracellular PPIs. The *i*, *i+7* hydrocarbon staple feature of ATSP-7041 increases its helicity from 11% in its linear form to 70% helical in the stapled form at pH 7.0.[41] The study showed that ATSP-7041 is a potent dual inhibitor of MDM2 and MDMX, which are proteins overexpressed in cancers and deactivate p53 function. Thus, ATSP-7041 can reactivate the p53 pathway, including the inducement of cell-cycle arrest and apoptosis.[41]

Since the early works of Korsmeyer, Verdine, and Walensky, many other PPIs have been studied using stapled peptides. These include cancer targets such as p53,[42] MCL-1 BH3,[43] and PUMA BH3[44] and other therapeutic targets ranging from infectious diseases[45,46] to metabolism.[47,48] Obtaining crystal structures of stapled peptides bound to their targets is not trivial. In fact, there are less than 15 stapled peptide crystal structures available in the literature including an estrogen receptor beta binding stapled peptide[48] and a MCL-1 BH3 stapled peptide[43] (Figure 3). Interestingly, in these solid-state structures, the hydrocarbon staple functionality appears to form hydrophobic interactions with the surface of the target proteins and has been proposed to provide additional binding affinity.[43,47] Recent work by the Tate group has also revealed the crystal structures of stapled and hydrogen bond surrogate α-helical peptides in a fully buried binding site.[49] The fully buried stapled peptides present polar and hydrophobic functionalities on all sides of the peptide helix. These results, along with crystallographic data from Chang et al where

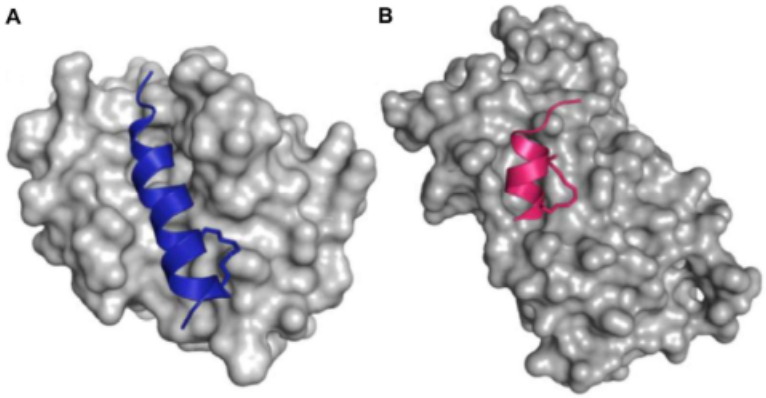


Figure 3 Crystal structures of (**A**) a BH3 stapled peptide (blue) bound to MCL-1 (gray) (PDB: 3MK8) and (**B**) a nuclear receptor helix stapled peptide (red) bound to estrogen receptor beta (gray) (PDB: 2YJD).

stapled peptide ATSP-7041 was bound to MDMX,[41] suggest the hydrophobic hydrocarbon staple can help in target binding and not just act as a nonparticipating constraint.

Stapled peptides have been used extensively to develop potential anticancer drugs through the regulation of PPIs involving the BCL-2 family of proteins implicated in cell apoptosis. Stapled peptide mimics of BH3 domain segments showed α-helical conformation when compared to the native sequences.[38] It was also found that stapling of the BH3 peptides gave protease resistance both in vitro and in vivo in comparison to the unmodified peptides.[38] This phenomenon is accounted for by the fact that proteases require peptides to be in an extended conformation in order to access the peptide backbone and hydrolyze the amide bond. However, by incorporation of a staple, the peptide is forced into the α-helical conformation and so the amide backbone is inaccessible.

Ring closing metathesis

Hydrocarbon stapling involves joining two alkenyl side chains to form a macrocycle using ruthenium-catalyzed RCM as the key step in the synthesis. On-resin peptide RCM was first carried out by Miller et al in 1996.[34] Due to their good functional group tolerance, ruthenium-based catalysts required for RCM are particularly well-suited for use on peptide substrates.

Miller et al's initial work involved the synthesis of cyclic peptides through incorporation of allyl groups onto the side chain of serine residues, followed by RCM to create peptides of varying ring sizes.[34] Following this, the efficient cross-linking of an Aib-containing heptapeptide was achieved by replacing two alanine residues at *i* and *i+4* positions with unbranched *O*-allyl serine or homoserine residues. The two terminal alkene moieties on the peptide were then treated with Grubbs' first generation catalyst to afford the macrocyclic peptide in high yield.[35] Catalytic hydrogenation then provided peptides with a 3_{10}-helix structure as determined by circular dichroism (CD) spectroscopy. The ease of RCM on these peptides has been attributed to the preorganization of a 3_{10}-helix on solid support facilitated by the Aib residues.

Structural requirements for stapling

The combined use of α,α-disubstituted alkenyl amino acid residues and RCM to produce all-hydrocarbon stapled peptides was first conceived by Schafmeister et al in 2000.[27] Initial work investigated the structural requirements needed for effective peptide stapling, including the length of the hydrocarbon bridge and the α-carbon stereochemistry of the α,α-disubstituted unnatural alkenyl amino acids (Figure 4).

A strategically useful feature of peptide stapling is that the hydrocarbon constraint can be incorporated at different positions along the peptide sequence. The most common types of peptide staple form a side chain–side chain bridge between the *i*, *i+3*, *i*, *i+4* or *i*, *i+7* residues. With shorter peptides a constraint across one turn of the helix (ie, *i*, *i+3* or an *i*, *i+4* staple) is employed. The *i*, *i+3* type staple requires one *R*- and one *S*-pentenyl alanine (Figure 4A), whereas two *S*-pentenyl alanine residues are required for an *i*, *i+4* type staple (Figure 4B). An *i*, *i+7* staple provides a hydrocarbon bridge spanning two turns of the helix and requires a *S*-pentenyl alanine and an *R*-octenyl alanine (Figure 4C). Again, the stereochemistry of these amino acids is very important to develop an effective constraint.

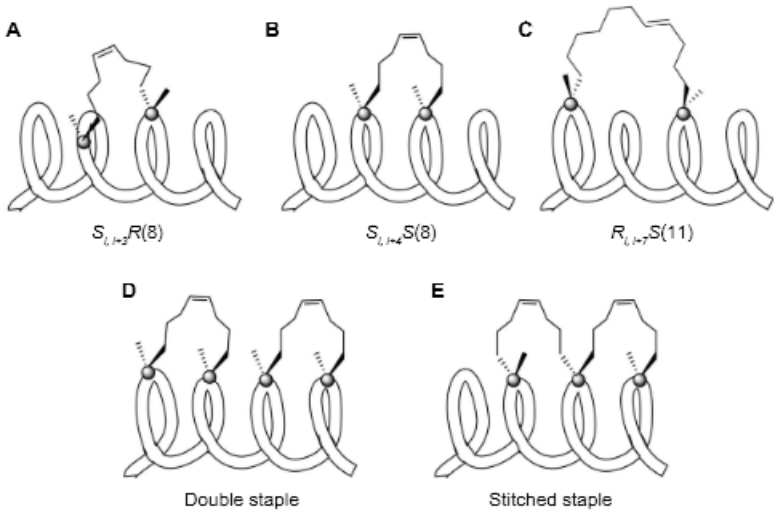


Figure 4 Representation of three types of stapled peptide constraints: (**A**) *i, i+3*; (**B**) *i, i+4*; (**C**) *i, i+7*; indicating the stereochemistry required for the stapled amino acids and the multiple stapling techniques used to constrain longer helices; (**D**) double staple; and (**E**) stitched staple.

Staple scanning

When limited structural information is available regarding a specific PPI, a staple scan can be used to determine the most effective position for the hydrocarbon constraint. This involves positioning an appropriate staple sequentially along the length of the peptide sequence. Staple scanning echoes the alanine scanning technique by also helping to determine if a residue is required for binding affinity.[50]

This strategy has been used to identify a selective inhibitor of MCL-1, a human cancer resistance factor.[43] Stewart et al found that the peptide, which had the highest α-helical character, as shown by CD, also had the highest binding affinity for MCL-1.[43]

An extension of the peptide stapling approach for longer helical peptide sequences involves the incorporation of multiple hydrocarbon bridges. These can take the form of multiple individual staples (Figure 4D) or "stitched" stapled peptides incorporating two alkenyl chains on one α-carbon of the central amino acid (Figure 4E).[51–54] One staple constraint has been shown to be ineffective for longer peptides; however, two staple constraints can significantly increase the helicity.[54] Stitched α-helical peptides can be thought to be the next generation in terms of hydrocarbon stapling, creating spiro-macrocycles, which are known to be stabilizing. Stitched peptides have been shown to give increased helicity, protease resistance, and cell permeability compared to the monostapled peptides.[54]

The size of the macrocycle formed upon RCM is important to achieve an increased degree of helicity in the constrained α-helical peptide. Different ring sizes are achieved depending on the length of the amino acid side chains used for the RCM. The macrocycle formed upon RCM of an *i–i+4* stapled peptide consists of 21 atoms, eight of which form the hydrocarbon constraint. This macrocycle is large in size and would generally mean that it would be problematic to ring close. However, the RCM reaction on these peptide substrates is relatively straightforward. This is due, in part, to the α-methyl-substituent of the alkenyl amino acid inducing the Thorpe–Ingold effect[55,56] on the peptide backbone, as well as the presence of a hydrophobic environment as a result of using nonpolar solvents, which promotes helix formation.[57]

Peptides incorporating the α,α-disubstituted amino acid Aib were first described by Karle and Balaram and were shown to adopt 3_{10}-helix structures.[25] The α-carbon dimethyl substitution of Aib restricts the torsion angles of the peptide bonds to φ=–49° and ψ=–26° in comparison to that of α-helix torsion angles, φ=–57° and ψ=–47°.[58] As a result of these bond angles, the hydrogen bonding network of the two types of helices differ (Figure 5). The α,α-disubstituted alkenyl amino acids required for peptide stapling also appear to provide this effect, resulting in the peptide substrates being folded ready to undergo RCM.

Monosubstituted hydrocarbon staples have been successfully synthesized using RCM on some peptide sequences.[43,59]

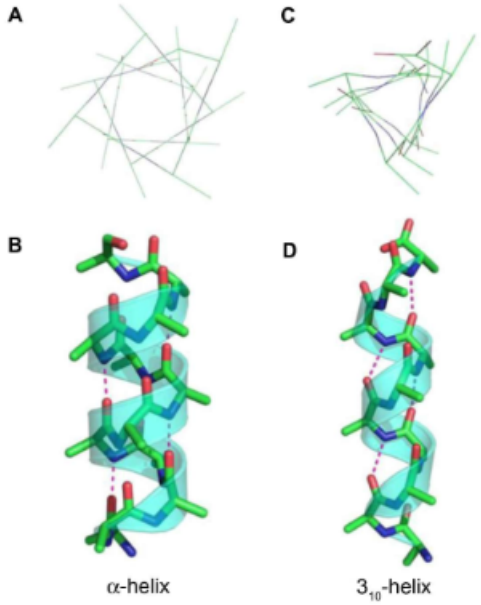


Figure 5 Representations of the (**A**) top view, (**B**) side view of an α-helix, (**C**) top view, and (**D**) side view of a 3_{10}-helix, showing the hydrogen bonding pattern (pink dotted lines) of the two types of helices.

However, it is preferable to incorporate an α,α-disubstituted amino acid, or some other helicity inducing substituent into the peptide to help promote the helical conformation and facilitate the RCM reaction. Although introduction of a hydrocarbon staple can improve the therapeutic properties of certain peptides,[38,43,52,60] it has been found that the increased degree of helicity obtained through the incorporation of a hydrocarbon staple does not necessarily correlate to an increase in its binding affinity.[61] This is partly because of the need for some flexibility to allow induced fit binding of some peptides to their binding partners.

Asymmetric α,α-disubstituted amino acid synthesis

The α,α-disubstituted alkenyl amino acids required for peptide stapling are commercially available, although they are expensive relative to standard fluorenylmethyloxycarbonyl (Fmoc) protected amino acids.

Asymmetric synthesis of the required alkenyl amino acids has been reported using Williams' glycine enolate alkylation method and alkylation of Belokon's nickel (II) Schiff base complex.[62–65] The Belokon method has come to the fore as the most effective strategy due to the harsh conditions with the potential for over reduction of the alkene moiety using the Williams method. Aillard et al recently reported a robust and efficient asymmetric synthesis of unnatural alkenyl amino acids. A fluorine-modified Ni^{II} Schiff base complex gave improved diastereoselectivity of the key alkylation reaction (>95: 5 dr).[66] This method now gives access to the required nonnative amino acids (Figure 6).

Design of the macrocycle constraint

Following the initial studies on peptide stapling by Schafmeister et al,[27] recent researches have focused on attempting to improve both the placement of the alkene in the staple and the lengths of the all-hydrocarbon staple. Notably, Pham et al have studied the placement of the alkene within the stapled macrocycle. A comparison of alkene

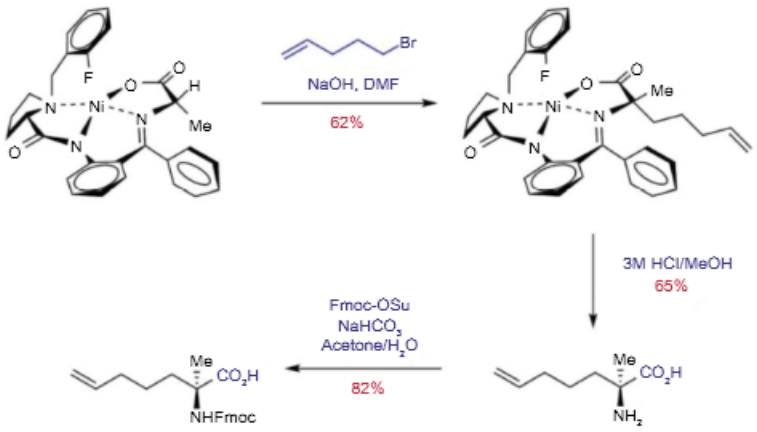


Figure 6 Asymmetric synthesis of Fmoc-S5-OH via fluorine-modified nickel (II) Schiff base complex.
Abbreviations: Fmoc, fluorenylmethyloxycarbonyl; OSu, O-succinimide; DMF, N,N-Dimethylformamide.

position within the *i*, *i+4* type stapled peptide was conducted using oct-2-enyl, oct-6-enyl, and the more usual oct-4-enyl macrocyclic constraint.[57]

Linear peptides incorporating a (*S*)-α-methyl, α-allylglycine ($\mathbf{S}_3$) at the *i* position and (*S*)-α-methyl, α-heptenylglycine ($\mathbf{S}_7$) at the *i+4* position provided access to both the oct-2-enyl staple ($\mathbf{S}_3$–$\mathbf{S}_7$) (Figure 7C) and oct-6-enyl staple ($\mathbf{S}_7$–$\mathbf{S}_3$) (Figure 7E). Both these peptides failed to undergo RCM to provide the corresponding macrocyclic peptides at room temperature. In fact, RCM on the $\mathbf{S}_3$–$\mathbf{S}_7$ analog only proceeded at an elevated temperature of 60°C, compared to the original oct-4-enyl staples ($\mathbf{S}_5$–$\mathbf{S}_5$) (Figure 7D), which proceed at room temperature. RCM of the $\mathbf{S}_7$–$\mathbf{S}_3$ substrate was very poor, even at 60°C and resulted in decomposition of the substrate peptide at the higher temperatures. The $\mathbf{S}_3$–$\mathbf{S}_7$ analog gave improved helicity in comparison to the native sequence; however, it still did not match the helicity shown by the $\mathbf{S}_5$–$\mathbf{S}_5$ stapled peptide.[57] These results demonstrate that the original oct-4-enyl *i*, *i*+4 type staple contains the alkene functionality in the optimal position on the hydrocarbon chain to provide an effective constraint. The *i*, *i+3* all-hydrocarbon staple has also been studied to determine the optimal length of hydrocarbon bridge across one turn of a helix. In general, for an effective *i*, *i+3* staple an $\mathbf{S}_5$ amino acid and its enantiomer, an $\mathbf{R}_5$ amino acid is required. However, Shim et al have studied the possibility of contracting the all-hydrocarbon bridge by removing two methylene units.[67]

Unsymmetrical contraction of the bridge was studied in order to analyze the differing positions of the alkene within the staple. Most of the linear peptide substrates failed to undergo RCM to give the corresponding macrocyclic peptide. However, two of the peptides did undergo metathesis, both of which had an *R*-configuration amino acid ($\mathbf{R}_3$ **or** $\mathbf{R}_5$) at the *i* position and an *S*-configuration amino acid ($\mathbf{S}_5$ **or** $\mathbf{S}_3$) at the *i+3* position (Figure 7F and G). However, the difference in the position of the staple within the macrocycle led to a pronounced difference in the α-helicity of the peptides, with the hex-4-enyl staple (Figure 7F) being much more helical (55%) than the hex-2-enyl staple (Figure 7G), which was 15% helical.[67]

Shim et al have investigated the ruthenium catalysts required form RCM including the Grubbs first generation, Grubbs second generation, and Hoveyda–Grubbs catalysts. RCM was performed at both 25°C and 65°C using the three different catalysts. The results showed that Grubbs first generation catalyst was the most efficient at both temperatures, but at elevated temperatures Grubbs second generation and Hoveyda–Grubbs catalysts have a significant increase in efficiency.[67]

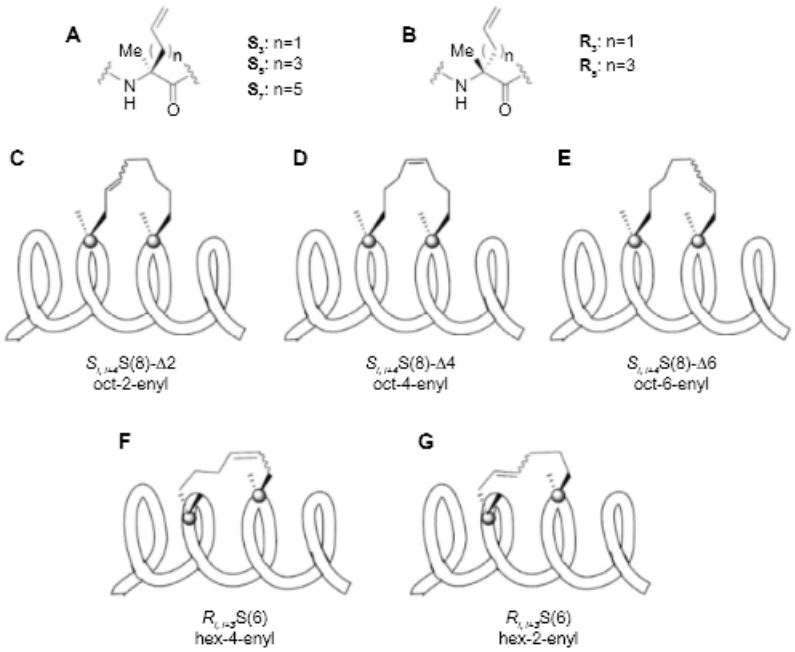


Figure 7 Structures of (**A**) the three **S**-alkenyl amino acids: $\mathbf{S}_3$, $\mathbf{S}_5$, and $\mathbf{S}_7$; (**B**) the two **R**-alkenyl amino acids: $\mathbf{R}_3$ and $\mathbf{R}_5$; and the representations of the different positions possible for the staple in an *i*, *i+4* staple; (**C**) oct-2-enyl; (**D**) oct-4-enyl; (**E**) oct-6-enyl; and in an *i*, *i+3* staple; (**F**) hex-4-enyl; and (**G**) hex-2-enyl.

Proposed mechanism of membrane permeability

A notable physicochemical property of stapled peptides is their ability to potentially promote cell penetration as a result of the hydrophobic hydrocarbon staple. Evidence has recently been reported for stapled peptides traversing the membrane and entering Jurkat leukemia cells[38] and human U2OS osteosarcoma cells[68] through a clathrin- and caveolin-independent, ATP-dependent pinocytotic mechanism.[5,38,42,68] Accumulation of the stapled peptides in ATP-depleted cells was reduced compared to normal cells. In addition, a significant reduction in the uptake of the stapled peptides into cells was observed after treatment of the cells with sodium chlorate, which blocks the sulfated proteoglycans presented on the cell surface. To confirm this finding, the researchers also treated proteoglycan deficient cells with the stapled peptides and found a reduction in cell uptake.[68] It has also been reported that cellular uptake can be impeded by the presence of serum for some peptides,[5] which could be as a result of competition with natural serum containing substrates for pinocytosis.[69]

Current and future directions

As with many new techniques, stapled peptides are surrounded by controversies, notably the recent works of two groups of researchers who contradict each other, and thus have initiated a debate on the effectiveness of peptide stapling in drug discovery. In 2012, LaBelle et al reported a stapled BIM peptide based on the BIM BH3 helix to target the BCL-2 family of proteins.[70] The stapled peptide, BIM $SAHB_A$, incorporated an *i*, *i+4* all-hydrocarbon crosslink and was found to have increased binding in comparison to the native BIM BH3 helix. In addition, they found BIM $SAHB_A$-induced cell apoptosis in leukemia by modulating the BCL-2 pathway.

Later that same year, Okamoto et al also reported a stapled peptide very similar to that of the previous study.[71] However, in their hands, the stapling of the peptide did not enhance its affinity or biological activity. The BIM BH3 all-hydrocarbon stapled peptide did not in fact induce apoptosis much more than the native sequence and was also not inherently cell permeable.[71] In addition, a lactam-stapled peptide based on the same helical sequence was found to have a greater percentage of helicity than the all-hydrocarbon staple.

Following this report, a rebuttal was published by Bird et al with the explanation that different lengths of BIM $SAHB_A$ were synthesized for specific studies.[72] A weaker binding BIM $SAHB_A$ (145–164) stapled peptide with moderate helicity compared to the BIM $SAHB_A$ (146–166) peptide was used in the in vivo and nuclear magnetic resonance experiments to increase the peptide's solubility and weaken its activation of BAX. BIM $SAHB_A$ (146–166) was three residues longer, had higher helicity and higher binding affinity than Bim $SAHB_A$ (145–164).

Okamoto et al subsequently reported data showing unstapled BIM SAHB peptides (containing two pentenylalanine residues at *i* and *i+4* positions, but have not been subjected to RCM) were as active as the stapled peptides toward two particular cancer cell lines.[73] Thus, it was concluded that the stapling was not responsible for the enhanced effect on activity, but rather, by altering the sequence of the peptides, the innate characteristics of that peptide are also altered.

The discrepancies in these studies demonstrate the subtleties in designing effectively constrained stapled peptides. Systematic rules for effective design are therefore urgently required and will be one of the main challenges for the community to address in coming years.

Recent peptide stapling techniques have returned to the original idea of using *O*-allyl homoserine residues, such as the stapled peptides used as inhibitors of the vitamin D receptor–coactivator interaction.[74] In addition, the use of all-hydrocarbon stapling is still prevalent, with many more targets being studied. PPIs involving membrane-bound proteins have also been studied, such as the transmembrane (TM) helices of small multidrug resistant proteins. These TM helices dimerize and oligomerize, forming PPIs, to extrude relatively large toxic substrates through channels created in the membrane, giving the resistant nature of these small multidrug resistant proteins.[75] Stapled peptide analogs of TM4 reduced the degradation of the peptides in blood plasma compared to a linear sequence. In addition, the stapled peptide also allowed the disruption of the dimerization domain between two TM4 helices, thus preventing efflux of the toxic substrates.[75]

Conclusion

Although still in its infancy, there is a bright future for stapled peptides. Current successes from stapled peptide clinical trials by Aileron Therapeutics are driving enthusiasm for the strategy in biomimetic drug development in the area of PPIs. In addition, stapled peptides are being used extensively as molecular probes in chemical biology to facilitate understanding of important biological mechanisms. In time, these examples are likely to lead to new drug discover efforts based on stapled peptides as lead compounds.

Future challenges include determining how stapled peptides penetrate the membrane, and establishing where they are located after entering the cell. A systematic set of rules are also required that outline the most effective position to locate a staple within a peptide sequence. The role, if any, of hydrocarbon staples on binding affinity has yet to be

determined, and a study focusing on nonspecific binding of stapled peptides is therefore required.

Acknowledgments

The authors thank the Engineering and Physical Sciences Research Council (EPSRC) (EP/L018152/1) and University of Leicester for financial support.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Keskin O, Ma B, Nussinov R. Hot regions in protein–protein interactions: the organization and contribution of structurally conserved hot spot residues. *J Mol Biol*. 2005;345(5):1281–1294.
2. Wilson AJ. Inhibition of protein-protein interactions using designed molecules. *Chem Soc Rev*. 2009;38(12):3289–3300.
3. Azzarito V, Long K, Murphy NS, Wilson AJ. Inhibition of α-helix-mediated protein-protein interactions using designed molecules. *Nat Chem*. 2013;5(3):161–173.
4. Verdine GL, Hilinski GJ. Stapled peptides for intracellular drug targets. *Methods Enzymol*. 2012;503:3–33.
5. Walensky LD, Bird GH. Hydrocarbon-stapled peptides: principles, practice, and progress. *J Med Chem*. 2014;57(15):6275–6288.
6. Lee EF, Czabotar PE, Smith BJ, et al. Crystal structure of ABT-737 complexed with Bcl-xL: implications for selectivity of antagonists of the Bcl-2 family. *Cell Death Differ*. 2007;14(9):1711–1713.
7. Tse C, Shoemaker AR, Adickes J, et al. ABT-263: a potent and orally bioavailable Bcl-2 family inhibitor. *Cancer Res*. 2008;68(9):3421–3428.
8. Vassilev LT, Vu BT, Graves B, et al. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science*. 2004; 303(5659):844–848.
9. Grasberger BL, Lu T, Schubert C, et al. Discovery and cocrystal structure of benzodiazepinedione HDM2 antagonists that activate p53 in cells. *J Med Chem*. 2005;48(4):909–912.
10. Orner BP, Ernst JT, Hamilton AD. Toward proteomimetics: terphenyl derivatives as structural and functional mimics of extended regions of an α-helix. *J Am Chem Soc*. 2001;123(22):5382–5383.
11. Hara T, Durell SR, Myers MC, Appella DH. Probing the structural requirements of peptoids that inhibit HDM2-p53 interactions. *J Am Chem Soc*. 2006;128(6):1995–2004.
12. Bullock BN, Jochim AL, Arora PS. Assessing helical protein interfaces for inhibitor design. *J Am Chem Soc*. 2011;133(36):14220–14223.
13. Marqusee S, Baldwin RL. Helix stabilization by Glu-…Lys+ salt bridges in short peptides of de novo design. *Proc Natl Acad Sci U S A*. 1987; 84(24):8898–8902.
14. Scholtz JM, Qian H, Robbins VH, Baldwin RL. The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry*. 1993;32(37):9668–9676.
15. Chorev M, Roubini E, McKee RL, et al. Cyclic parathyroid hormone-related protein antagonists: lysine 13 to aspartic acid 17 [i to (i + 4)] side chain to side chain lactamization. *Biochemistry*. 1991;30(24):5968–5974.
16. Phelan JC, Skelton NJ, Braisted AC, McDowell RS. A general method for constraining short peptides to an α-helical conformation. *J Am Chem Soc*. 1997;119(3):455–460.
17. Jackson DY, King DS, Chmielewski J, Singh S, Schultz PG. General approach to the synthesis of short α-helical peptides. *J Am Chem Soc*. 1991;113(24):9391–9392.
18. Chapman RN, Dimartino G, Arora PS. A highly stable short alpha-helix constrained by a main-chain hydrogen-bond surrogate. *J Am Chem Soc*. 2004;126(39):12252–12253.
19. Patgiri A, Jochim AL, Arora PS. A hydrogen bond surrogate approach for stabilization of short peptide sequences in alpha-helical conformation. *Acc Chem Res*. 2008;41(10):1289–1300.
20. Albert JS, Hamilton AD. Stabilization of helical domains in short peptides using hydrophobic interactions. *Biochemistry*. 1995;34(3):984–990.
21. Ruan F, Chen Y, Hopkins PB. Metal ion-enhanced helicity in synthetic peptides containing unnatural, metal-ligating residues. *J Am Chem Soc*. 1990;112(25):9403–9404.
22. Ghadiri MR, Fernholz AK. Peptide architecture. Design of stable α-helical metallopeptides via a novel exchange-inert ruthenium(III) complex. *J Am Chem Soc*. 1990;112(26):9633–9635.
23. Kawamoto SA, Coleska A, Ran X, Yang C-Y, Wang S. Design of triazole-stapled BCL9 α-helical peptides to target the β-catenin/B-cell CLL/lymphoma 9 (BCL9) protein–protein interaction. *J Med Chem*. 2012;55(3):1137–1146.
24. Kneissl S, Loveridge EJ, Williams C, Crump MP, Allemann RK. Photocontrollable peptide-based switches target the anti-apoptotic protein Bcl-xL. *Chembiochem*. 2008;9(18):3046–3054.
25. Karle IL, Balaram P. Structural characteristics of alpha.-helical peptide molecules containing Aib residues. *Biochemistry*. 1990;29(29): 6747–6756.
26. Balaram P. Non-standard amino acids in peptide design and protein engineering. *Curr Opin Struct Biol*. 1992;2(6):845–851.
27. Schafmeister CE, Po J, Verdine GL. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *J Am Chem Soc*. 2000;122(24):5891–5892.
28. Forood B, Feliciano EJ, Nambiar KP. Stabilization of alpha-helical structures in short peptides via end capping. *Proc Natl Acad Sci U S A*. 1993;90(3):838–842.
29. Zhou HX, Lyu PC, Wemmer DE, Kallenbach NR. Structure of a C-terminal.alpha.-helix cap in a synthetic peptide. *J Am Chem Soc*. 1994;116(3):1139–1140.
30. Yang J, Zhao K, Gong Y, Vologodskii A, Kallenbach NR. α-Helix nucleation constant in copolypeptides of alanine and ornithine or lysine. *J Am Chem Soc*. 1998;120(41):10646–10652.
31. Kemp DS, Curran TP, Boyd JG, Allen TJ. Studies of N-terminal templates for alpha.-helix formation. Synthesis and conformational analysis of peptide conjugates of (2S,5S,8S,11S)-1-acetyl-1,4-diaza-3-keto-5-carboxy-10-thiatricyclo[2.8.1.04,8]tridecane (Ac-Hel1-OH). *J Org Chem*. 1991;56(23):6683–6697.
32. Kemp DS, Curran TP, Davis WM, Boyd JG, Muendel C. Studies of N-terminal templates for alpha-helix formation. Synthesis and conformational analysis of (2S,5S,8S,11S)-1-acetyl-1,4-diaza-3-keto-5-carboxy-10-thiatricyclo[2.8.1.04,8]tridecane (Ac-Hel1-OH). *J Org Chem*. 1991;56(23):6672–6682.
33. Miller SJ, Grubbs RH. Synthesis of conformationally restricted amino acids and peptides employing olefin metathesis. *J Am Chem Soc*. 1995;117(21):5855–5856.
34. Miller SJ, Blackwell HE, Grubbs RH. Application of ring-closing metathesis to the synthesis of rigidified amino acids and peptides. *J Am Chem Soc*. 1996;118(40):9606–9614.
35. Blackwell HE, Grubbs RH. Highly efficient synthesis of covalently cross-linked peptide helices by ring-closing metathesis. *Angew Chemie Int Ed*. 1998;37(23):3281–3284.
36. Toniolo C, Benedetti E. Structures of polypeptides from α-amino acids disubstituted at the α-carbon. *Macromolecules*. 1991;24(14):4004–4009.
37. Demizu Y, Doi M, Kurihara M, et al. Conformational studies on peptides containing α,α-disubstituted α-amino acids: chiral cyclic α,α-disubstituted α-amino acid as an α-helical inducer. *Org Biomol Chem*. 2011;9(9):3303–3312.
38. Walensky LD, Kung AL, Escher I, et al. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science*. 2004;305(5689):1466–1470.
39. Sawyer TK. AILERON therapeutics. *Chem Biol Drug Des*. 2009;73(1): 3–6.
40. ClinicalTrials.gov [Internet]. *Phase 1 Safety Study of ALRN-5281 in Healthy Subjects*. Bethesda, MD: National Library of Medicine (US); 2013. [Identifier NCT01775358, February 29, 2000].
41. Chang YS, Graves B, Guerlavais V, et al. Stapled α-helical peptide drug development: a potent dual inhibitor of MDM2 and MDMX for p53-dependent cancer therapy. *Proc Natl Acad Sci U S A*. 2013;110(36): E3445–E3454.

42. Bernal F, Tyler AF, Korsmeyer SJ, Walensky LD, Verdine GL. Reactivation of the p53 tumor suppressor pathway by a stapled p53 peptide. *J Am Chem Soc*. 2007;129(9):2456–2457.
43. Stewart ML, Fire E, Keating AE, Walensky LD. The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. *Nat Chem Biol*. 2010;6(8):595–601.
44. Edwards AL, Gavathiotis E, LaBelle JL, et al. Multimodal interaction with BCL-2 family proteins underlies the proapoptotic activity of PUMA BH3. *Chem Biol*. 2013;20(7):888–902.
45. Long YQ, Huang SX, Zawahir Z, et al. Design of cell-permeable stapled peptides as HIV-1 integrase inhibitors. *J Med Chem*. 2013;56(13):5601–5612.
46. Chapuis H, Slaninová J, Bednárová L, Monincová L, Buděšínský M, Čeřovský V. Effect of hydrocarbon stapling on the properties of α-helical antimicrobial peptides isolated from the venom of hymenoptera. *Amino Acids*. 2012;43(5):2047–2058.
47. Phillips C, Roberts LR, Schade M, et al. Design and structure of stapled peptides binding to estrogen receptors. *J Am Chem Soc*. 2011;133(25):9696–9699.
48. Szlyk B, Braun CR, Ljubicic S, et al. A phospho-BAD BH3 helix activates glucokinase by a mechanism distinct from that of allosteric activators. *Nat Struct Mol Biol*. 2014;21(1):36–42.
49. Douse CH, Maas SJ, Thomas JC, et al. Crystal structures of stapled and hydrogen bond surrogate peptides targeting a fully buried protein-helix interaction. *ACS Chem Biol*. 2014;9(10):2204–2209.
50. Jamieson AG, Boutard N, Sabatino D, Lubell WD. Peptide scanning for studying structure-activity relationships in drug discovery. *Chem Biol Drug Des*. 2013;81(1):148–165.
51. Bird GH, Madani N, Perry AF, et al. Hydrocarbon double-stapling remedies the proteolytic instability of a lengthy peptide therapeutic. *Proc Natl Acad Sci U S A*. 2010;107(32):14093–14098.
52. Rao T, Ruiz-Gómez G, Hill TA, Hoang HN, Fairlie DP, Mason JM. Truncated and helix-constrained peptides with high affinity and specificity for the cFos coiled-coil of AP-1. *PLoS One*. 2013;8(3):e59415.
53. Bird GH, Boyapalle S, Wong T, et al. Mucosal delivery of a double-stapled RSV peptide prevents nasopulmonary infection. *J Clin Invest*. 2014;124(5):2113–2124.
54. Hilinski GJ, Kim YW, Hong J, et al. Stitched α-helical peptides via bis ring-closing metathesis. *J Am Chem Soc*. 2014;136(35):12314–12322.
55. Beesley RM, Ingold CK, Thorpe JF. The formation and stability of spiro-compounds. Part I. spiro-compounds from cyclohexane. *J Chem Soc Trans*. 1915;107:1080–1106.
56. Fürstner A, Langemann K. A concise total synthesis of dactylol via ring closing metathesis. *J Org Chem*. 1996;61(25):8746–8749.
57. Pham TK, Yoo J, Kim Y-W. Comparison of Oct-2-enyl and Oct-4-enyl staples for their formation and α-helix stabilizing effects. *Bull Korean Chem Soc*. 2013;34(9):2640–2644.
58. Arnott S, Dover SD. Refinement of bond angles of an α-helix. *J Mol Biol*. 1967;30(1):209–212.
59. Yeo DJ, Warriner SL, Wilson AJ. Monosubstituted alkenyl amino acids for peptide "stapling". *Chem Commun (Camb)*. 2013;49(80):9131–9133.
60. Danial NN, Walensky LD, Zhang CY, et al. Dual role of proapoptotic BAD in insulin secretion and beta cell survival. *Nat Med*. 2008;14(2):144–153.
61. Giordanetto F, Revell JD, Knerr L, et al. Stapled Vasoactive Intestinal Peptide (VIP) derivatives improve VPAC2 agonism and glucose-dependent insulin secretion. *ACS Med Chem Lett*. 2013;4(12):1163–1168.
62. Williams RM, Sinclair PJ, Zhai D, Chen D. Practical asymmetric syntheses of α-amino acids through carbon-carbon bond constructions on electrophilic glycine templates. *J Am Chem Soc*. 1988;110(5):1547–1557.
63. Williams RM, Im MN. Asymmetric synthesis of monosubstituted and α,α-disubstituted α-amino acids via diastereoselective glycine enolate alkylations. *J Am Chem Soc*. 1991;113(24):9276–9286.
64. Belokon YN, Bakhmutov VI, Chernoglazova NI, et al. General method for the asymmetric synthesis of α-amino acids via alkylation of the chiral nickel(II) Schiff base complexes of glycine and alanine. *J Chem Soc Perkin Trans 1*. 1988;(2):305–312.
65. Belokon YN, Tararov VI, Maleev VI, Savel'eve TF, Ryzhov MG. Improved procedures for the synthesis of (S)-2-[N-(N′-benzylprolyl)amino]benzophenone (BPB) and Ni(II) complexes of Schiff's bases derived from BPB and amino acids. *Tetrahedron Asymmetry*. 1998;9(23):4249–4252.
66. Aillard B, Robertson NS, Baldwin AR, Robins S, Jamieson AG. Robust asymmetric synthesis of unnatural alkenyl amino acids for conformationally constrained α-helix peptides. *Org Biomol Chem*. 2014;12(43):8775–8782.
67. Shim SY, Kim Y-W, Verdine GL. A new i, i + 3 peptide stapling system for α-helix stabilization. *Chem Biol Drug Des*. 2013;82(6):635–642.
68. Chu Q, Moellering RE, Hilinski GJ, et al. Towards understanding cell penetration by stapled peptides. *MedChemCommun*. 2014;6(1):111–119.
69. Commisso C, Davidson SM, Soydaner-Azeloglu RG, et al. Macropinocytosis of protein is an amino acid supply route in Ras-transformed cells. *Nature*. 2013;497(7451):633–637.
70. LaBelle JL, Katz SG, Bird GH, et al. A stapled BIM peptide overcomes apoptotic resistance in hematologic cancers. *J Clin Invest*. 2012;122(6):2018–2031.
71. Okamoto T, Zobel K, Fedorova A, et al. Stabilizing the pro-apoptotic BimBH3 helix (BimSAHB) does not necessarily enhance affinity or biological activity. *ACS Chem Biol*. 2013;8(2):297–302.
72. Bird GH, Gavathiotis E, Labelle JL, Katz SG, Walensky LD. Distinct BimBH3 (BimSAHB) stapled peptides for structural and cellular studies. *ACS Chem Biol*. 2014;9(3):831–837.
73. Okamoto T, Segal D, Zobel K, et al. Further insights into the effects of pre-organizing the BimBH3 helix. *ACS Chem Biol*. 2014;9(3):838–839.
74. Misawa T, Demizu Y, Kawamura M, Yamagata N, Kurihara M. Structural development of stapled short helical peptides as vitamin D receptor (VDR)–coactivator interaction inhibitors. *Bioorg Med Chem*. 2015;23(5):1055–1061.
75. Bellmann-Sickert K, Stone TA, Poulsen BE, Deber CM. Efflux by small multidrug resistance proteins is inhibited by membrane-interactive helix-stapled peptides. *J Biol Chem*. 2014;290(3):1752–1759.

nature COMMUNICATIONS

ARTICLE

Received 21 Oct 2015 | Accepted 7 Mar 2016 | Published 25 Apr 2016

DOI: 10.1038/ncomms11262 OPEN

Insights into the activation mechanism of class I HDAC complexes by inositol phosphates

Peter J. Watson[1,*], Christopher J. Millard[1,*], Andrew M. Riley[2], Naomi S. Robertson[3], Lyndsey C. Wright[1], Himali Y. Godage[2], Shaun M. Cowley[1], Andrew G. Jamieson[3], Barry V.L. Potter[2,4] & John W.R. Schwabe[1]

Histone deacetylases (HDACs) 1, 2 and 3 form the catalytic subunit of several large transcriptional repression complexes. Unexpectedly, the enzymatic activity of HDACs in these complexes has been shown to be regulated by inositol phosphates, which bind in a pocket sandwiched between the HDAC and co-repressor proteins. However, the actual mechanism of activation remains poorly understood. Here we have elucidated the stereochemical requirements for binding and activation by inositol phosphates, demonstrating that activation requires three adjacent phosphate groups and that other positions on the inositol ring can tolerate bulky substituents. We also demonstrate that there is allosteric communication between the inositol-binding site and the active site. The crystal structure of the HDAC1:MTA1 complex bound to a novel peptide-based inhibitor and to inositol hexaphosphate suggests a molecular basis of substrate recognition, and an entropically driven allosteric mechanism of activation.

[1] Henry Wellcome Laboratories of Structural Biology, Department of Molecular and Cell Biology, University of Leicester, Leicester LE1 9HN, UK. [2] Department of Pharmacy and Pharmacology, University of Bath BA2 7AY, UK. [3] Department of Chemistry, University of Leicester, Leicester LE1 7RH, UK. [4] Department of Pharmacology, University of Oxford, Oxford OX1 3QT, UK. * These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.W.R.S. (email: john.schwabe@le.ac.uk).

Class I histone deacetylases (HDACs) are enzymes involved in 'epigenetic' gene regulation through controlling the acetylation state of lysine sidechains in histone tails[1]. They act as the catalytic subunit of several large protein complexes that repress gene expression when targeted to the genome. Recent structural and functional studies of class I HDACs in complex with their cognate co-repressors have suggested that the activity of these complexes is regulated in the cell by inositol phosphates that are likely derived from membrane phospholipids[2–4]. Understanding the regulation of these complexes is important since they are promising targets for epigenetic therapies for a range of diseases[5]. These include numerous cancers as well as spinal muscular atrophy[6], Friedrich's ataxia[7], Alzheimer's disease[8] and HIV infection[9]. Five HDAC inhibitors are now variously licensed for use in the clinic for the treatment of cutaneous T-cell lymphoma, peripheral T-cell lymphoma[10,11] and multiple myeloma[12].

The class I HDAC family comprises of HDACs 1–3 and 8 (reviewed in ref. 13). HDACs 1–3 are assembled into at least five large multi-protein co-repressor complexes that are recruited to chromatin through interaction with repressive transcription factors or other silencing co-factors[14]. The enzymatic activity of HDACs 1–3 show significant enhancement when incorporated into their cognate co-repressor complexes[15–20]. HDAC8, however, sits alone as the only class I HDAC that is not recruited into a larger complex and is fully active in isolation[21,22]. HDACs 1 and 2 are found within several distinct co-repressor complexes including NuRD[23], Sin3A[24], CoREST[25] and MiDAC[4,26]. HDAC3, however, is exclusively recruited to the SMRT/NCoR co-repressor complex[20,27]. The regulation of these complexes by inositol phosphates was first suggested by the surprising discovery that inositol 1,4,5,6-tetrakisphosphate ($Ins(1,4,5,6)P_4$) was present in the HDAC3:SMRT crystal structure[2]. The $Ins(1,4,5,6)P_4$ is located at a binding pocket formed at the interface between HDAC3 and the co-repressor. The finding that the $Ins(1,4,5,6)P_4$ co-purified with the HDAC3 complex from mammalian cells suggests that it is likely to be a physiologically relevant activator of the complex. However, it is not possible to exclude the possibility that other inositol phosphates might also be able to activate the complex. Indeed, $Ins(1,4,5,6)P_4$ is only one of several higher order inositol phosphates which are produced in cells from $Ins(1,4,5)P_3$, the well-known second messenger that regulates Ca^{2+} release through binding to the inositol trisphosphate receptor ($InsP_3R$) (ref. 28).

Importantly, the key residues which coordinate the binding of $Ins(1,4,5,6)P_4$ to the HDAC3:SMRT complex were found to be conserved in several class I HDAC complexes, suggesting that these complexes may also be activated by inositol phosphates. However, it is notable that the key residues are not conserved in the Sin3A co-repressor. Indeed, the structure of the HDAC1:MTA1 complex confirmed that the inositol phosphate-binding pocket was present in other class I HDAC co-repressor complexes[3].

We initially proposed that $Ins(1,4,5,6)P_4$ serves as an 'inter-molecular glue', mediating interaction between HDAC3 and SMRT[2]. It later emerged that longer constructs of SMRT form a constitutive complex with HDAC3 and that the role of the $Ins(1,4,5,6)P_4$ is to activate the HDAC3 enzyme itself[3]. Intriguingly, we observed using mass-spectrometry, that the HDAC3:SMRT complex always co-purifies with $Ins(1,4,5,6)P_4$ and that the $Ins(1,4,5,6)P_4$ can only be removed using a high-salt wash (resulting in an inactive complex). In contrast, mass-spectrometry showed that the HDAC1:MTA1 complex does not co-purify with $Ins(1,4,5,6)P_4$ or any other inositol phosphates. However, the HDAC1:MTA1 complex is nevertheless robustly activated by exogenous $Ins(1,4,5,6)P_4$. The novel MiDAC complex has also been shown to be activated by exogenous $Ins(1,4,5,6)P_4$ (ref. 4).

The physiological importance of inositol phosphate activation of HDAC complexes is supported by the finding that mutants in the inositol phosphate-binding pocket of HDAC1 are unable to fully restore HDAC activity in HDAC1/2 knock-out ES cells and rescue their viability[29]. Furthermore, mice containing a mutation of one of the key inositol phosphate-binding residues in SMRT (Y470) exhibit increased local histone acetylation *in vivo*[30].

Whilst the structures of HDACs 1 and 3 in complex with their cognate co-repressors, along with functional studies, have established that these complexes are activated by inositol phosphates[3], the exact mechanism through which inositol phosphates activate HDACs remains unclear. To address this important issue we have taken a chemical biology approach to understand what are the important stereochemical features of inositol phosphates that are required to activate class I HDAC complexes. We have used inositol phosphate derivatives to directly investigate the binding of inositol phosphates to HDAC complexes *in vitro* and to demonstrate how further derivatives might be developed as tools to modulate HDAC activity. These approaches do not purport to identify which inositol phosphates are relevant *in vivo*. In addition, we have developed a novel peptide-based HDAC inhibitor. This is essentially a substrate mimic based on the histone H4 tail, but incorporating a hydroxamic acid functionality. Using a structural biology approach we have explored the details of substrate recognition, as well as the relationship between the binding of substrate and activation by inositol phosphates. These studies, together with kinetic and mutational analyses of these enzyme complexes, give insight into the mechanism underlying the inositol phosphate-mediated allostery.

Results

HDAC3 activation by inositol phosphates. To understand the stereochemical basis underlying the activation of HDAC3 by inositol phosphates we investigated the ability of a range of inositol phosphates and derivatives to enhance the activity of HDAC3 in deacetylase assays. To achieve this we used an assay based on the complex of HDAC3 with an extended SMRT construct which is stable in the absence of inositol phosphate[3]. In this assay, any endogenous inositol phosphate that has been co-purified with the complex is removed by treatment with high-ionic strength buffer, which results in a complex that shows little catalytic activity but can be readily activated by the addition of $Ins(1,4,5,6)P_4$ (Supplementary Fig. 1).

The crystal structure of $Ins(1,4,5,6)P_4$ bound to the HDAC3:SMRT complex suggests that each of the four phosphates on the inositol ring are recognized by distinct sites within the binding pocket. These sites, referred to as A, B, C and D, accommodate phosphates at positions 4, 5, 6 and 1, respectively, on the inositol ring (Fig. 1a). Positions 2 and 3 of the inositol ring are relatively unobstructed and the free hydroxyls do not appear to be recognized by the protein suggesting that higher order inositol phosphates might also be able to bind and activate the HDAC3:SMRT complex. Indeed, both $Ins(1,3,4,5,6)P_5$ and $InsP_6$ are able to activate the complex to the same degree as $Ins(1,4,5,6)P_4$ (Fig. 1b). Interestingly, they bind with higher apparent affinity than $Ins(1,4,5,6)P_4$; $Ins(1,3,4,5,6)P_5$ shows a twofold increase and $InsP_6$ a 16-fold increase (K_d values of 5.0 and 0.6 μM, respectively, compared with 10 μM). This higher apparent affinity is mostly likely due to the increased negative charge contributing to a higher on-rate for binding to the positively charged pocket of the HDAC3:SMRT complex. In

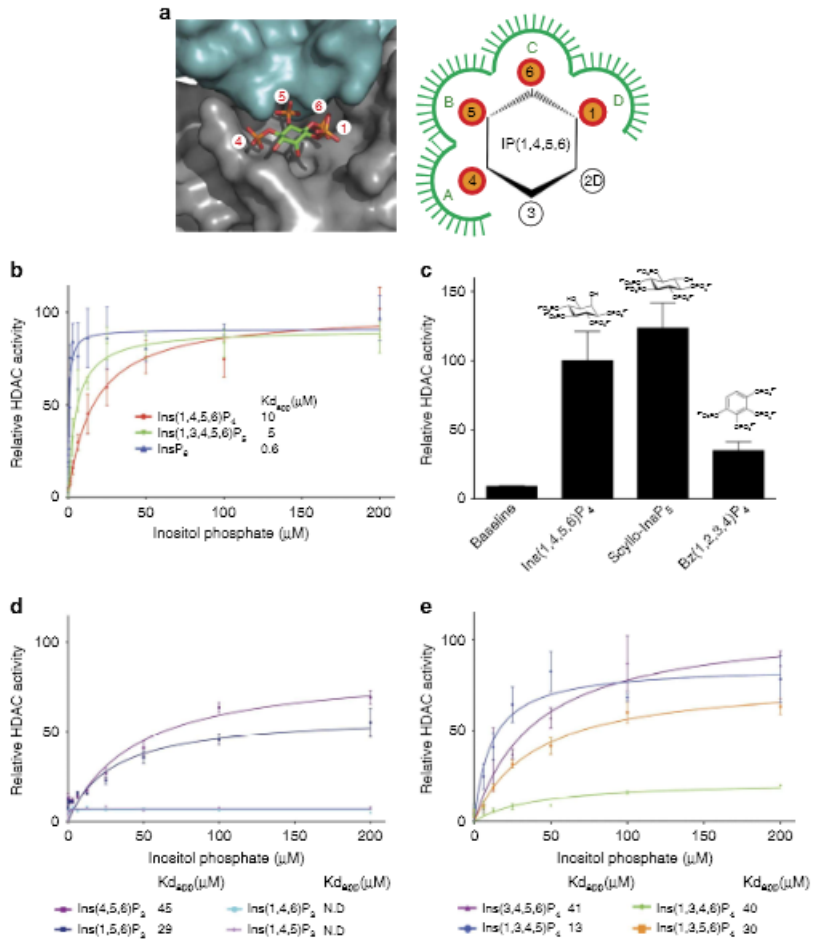


Figure 1 | The stereospecific requirements for activation of HDAC activity by various inositol phosphates. (**a**) Ins(1,4,5,6)P4 bound to the HDAC3:SMRT complex, with the HDAC3 shown as a grey surface and the SMRT as a cyan surface. Schematic representation of the inositol phosphate-binding pocket, with the binding sites for the four phosphates of Ins(1,4,5,6)P4 designated A, B, C and D as indicated. 2D indicates that the axial hydroxyl group on the second position of the inositol ring is pointing down in this view. (**b**,**d**,**e**) Activation of HDAC3 by inositol phosphates using a BOC-Lys-AMC assay. HDAC activity is expressed relative to the maximal Ins(1,4,5,6)P4-stimulated activity and is plotted against the inositol phosphate concentration (μM). (**c**) The ability of *scyllo*-InsP5 and the planar benzene 1,2,3,4-tetrakisphosphate to activate HDAC3:SMRT is compared with Ins(1,4,5,6)P4. HDAC activity was measured in the presence of 200 μM inositol phosphate/benzene tetrakisphosphate. HDAC activity is expressed relative to the maximal Ins(1,4,5,6)P4 activity. Error bars indicate ± s.e.m. ($n = 3$).

addition to being able to accommodate a phosphate group on the 2 position of the inositol ring, it also appears that the axial orientation is not important, since *scyllo*-InsP5 (ref. 31), which has an equatorial 2-hydroxyl, exhibits similar activity to Ins(1,4,5,6)P4 (Fig. 1c).

The stereochemistry of the Ins(1,4,5,6)P4-binding pocket suggests that the chair conformation of the inositol is important to position the phosphates correctly. This interpretation is supported by the finding that benzene 1,2,3,4-tetrakisphosphate[32] (four adjacent phosphates on a planar ring) fails to substantially activate the complex (Fig. 1c).

To fully rationalize the stereochemical requirements for activation of the HDAC3:SMRT complex, we explored the ability of eight tris- and tetrakis-inositol phosphates to enhance the deacetylase activity of the complex. These molecules varied in their ability to activate the HDAC3:SMRT complex and bound with a range of apparent dissociation constants (Fig. 1d,e). Interpreting these data is complicated by there being potentially multiple modes of binding for the various inositol phosphates. Furthermore, some of these modes might support binding to, but not activation of, the complex. Careful analysis allows us to conclude that a minimum of three adjacent phosphates is required for the substantial activation of HDAC3 in the complex. These phosphates must occupy sites A, B and C in the binding pocket (Fig. 1a). Full analyses of these data and conclusions are presented in the Supplementary Materials.

The requirement for three phosphates in positions A, B and C, fits well with the HDAC3:SMRT structure. The phosphates at

sites B and C are essentially completely buried at the interface of HDAC3 and the SMRT co-repressor. The requirement for site A to be occupied can also be rationalized, since the phosphate at this site forms a salt bridge with R265 in HDAC3 which has been shown to be essential for activation of the enzyme[2]. It is important to note that the well-established physiologically relevant inositol phosphate signalling molecule, $Ins(1,4,5)P_3$, which activates the $InsP_3R$ to open calcium channels, is completely unable to activate HDAC3. This can be explained as it cannot simultaneously fulfil sites A, B and C.

To complement our analysis based on the ability of different inositol phosphates to activate the HDAC3:SMRT complex, we also performed computational docking studies. We used Autodock to first examine the contribution to the overall binding energy of the phosphates at the 1 or 4 position of $Ins(1,4,5,6)P_4$ (Supplementary Fig. 4a–c). Removal of phosphate P1 (occupying site D) reduced the binding energy by just 1.2 kcal mol^{-1}; whereas removal of phosphate P4 (occupying site A) reduced the binding energy by 5.4 kcal mol^{-1}. This fits well with the activation data, which suggest that site A is more important than site D. We also freely docked $Ins(1,3,4,5)P_4$ into the complex, since this molecule activates to greater than 75% yet can only fulfil three adjacent phosphate sites. The three lowest energy-docked molecules adopt an orientation that positions phosphates P3, P4 and P5 in sites A, B and C with an average binding energy of -18.72 kcal mol^{-1} (Supplementary Fig. 4d). This energy is very similar to the calculated binding energy of $Ins(1,4,5,6)P_4$ (-18.86 kcal mol^{-1}). The majority of the other docking solutions also fulfil positions A, B and C, but with lower binding energies. None of the docking solutions positioned the phosphates in sites B, C and D, leaving the A site unoccupied. Taken together, these computational docking studies fit well with our conclusion that sites A, B and C are essential for both inositol phosphate binding and activation of the complex.

Activation of the HDAC3:SMRT complex by synthetic analogues. To explore further the distinction between molecules that activate the $InsP_3$ receptor and those that activate HDACs, we tested the ability of synthetic adenophostin A, a potent $InsP_3$ receptor agonist[33,34], to activate the HDAC3:SMRT complex. We observed no activity for adenophostin A even at the high concentration of 200 μM (Fig. 2a). This lack of activity likely results from the fact that the phosphate groups on the pyranose ring of adenophostin A mimic the 4 and 5 position phosphates of $Ins(1,4,5)P_3$ and hence cannot satisfy the sites required for activation.

The observation that $InsP_6$ is able to activate the HDAC complex, raised the question as to whether larger groups might be added on the 2 and 3 positions of the inositol ring. Accordingly, we synthesized and evaluated inositol phosphates with bulky substituents (benzyl groups) at the 2 and 3 positions (Fig. 2a). Strikingly, bulky groups at these positions can be tolerated with a minimal loss of HDAC activation.

It has recently been reported that inositol pyrophosphates may be important in *S. cerevisiae* for the regulation of the class 1 HDAC homologue Rpd3L (ref. 35). Since the inositol-binding residues identified in HDAC3:SMRT are also present in Rpd3 and the Snt1 co-repressor, we speculated that this regulation might be mediated through the same inositol phosphate-binding pocket. We therefore tested whether pyrophosphate analogues and pyrophosphate, 5-PP-$InsP_4$, might be able to activate the HDAC3 complex. The pyrophosphate mimic 1-PA-$InsP_5$ (ref. 36) (pyrophosphate mimic on position 1 of the inositol ring) has similar activity to that of $Ins(1,4,5,6)P_4$, whereas 5-PP-$InsP_4$ (pyrophosphate on position 5) has reduced activity (~60%) compared with $Ins(1,4,5,6)P_4$. In contrast, the pyrophosphate mimic at position 5 in 2-OH-5-PA-$InsP_4$ (ref. 37) is completely inert (Fig. 2b). This difference in activity may be due to differences in the possible binding modes of the pyrophosphates and their mimics. 1-PA-$InsP_5$ can bind in a way that sites A, B, C and D are all satisfied, whereas 5-PP-$InsP_4$ can only satisfy the sites A, B and C. 2-OH-5-PA-$InsP_4$, which contains a carbonyl in place of the pyrophosphate of 5-PP-$InsP_4$, cannot form the crucial salt bridge with R265 in HDAC3.

Inositol hexakisphosphate activation of HDAC1. We have previously shown that the HDAC1:MTA1 complex can also be activated by $Ins(1,4,5,6)P_4$ (ref. 3). Here we show that, as for the HDAC3:SMRT complex, the HDAC1:MTA1 complex can be activated by both $Ins(1,3,4,5,6)P_5$ and $InsP_6$, suggesting that the mode of inositol phosphate binding is likely to be similar (Fig. 3a). To investigate this, we soaked crystals of the HDAC1:MTA1 complex (which had been co-crystallized with a peptide-based inhibitor) with 225 mM $InsP_6$ before freezing for data collection. The structure was solved to 3.3 Å by molecular replacement, by using the structure of unliganded HDAC1:MTA1 (pdb code 4BKX) (Table 1). The structure shows that the $InsP_6$ is bound as predicted in the pocket formed at the interface of HDAC1 and MTA1, adjacent to the active site of HDAC1 (Fig. 3b, Supplementary Fig. 5). The occupancy of $InsP_6$ in the crystal was refined to 70%.

As might have been anticipated, there are some differences in the mode of $InsP_6$ binding when compared with $Ins(1,4,5,6)P_4$ binding to the HDAC3:SMRT complex. In particular the inositol ring is tipped away from the HDAC (towards MTA1), which seems to correlate with the sidechain of R270 (HDAC1) adopting a different position from that of the corresponding R265 in HDAC3 (Fig. 3c,d). The molecule is supported in this new position through a hydrogen bond between the hydroxyl of Y336 and the C3 phosphate group of the $InsP_6$.

Careful analysis of the orientation of the $InsP_6$ molecule at the inter-molecular interface showed that the electron density clearly fits best when the inositol phosphate orientation is rotated by 120° relative to that seen in the HDAC3:SMRT:$Ins(1,4,5,6)P_4$ complex (Fig. 3c,d). This rotation allows the phosphate on the C2 position of the inositol to be accommodated in the pocket, and compensates for small structural differences between the two complexes. Despite this difference in orientation, the sites A, B and C (see above) are occupied by phosphates in exactly the same position as in the HDAC3 complex with $Ins(1,4,5,6)P_4$; thus, supporting the importance of satisfying these positions so as to activate the enzyme. Interestingly, the phosphates in sites A and B overlap perfectly with the position of the ordered sulphates seen in the HDAC1:MTA1 complex in the absence of inositol phosphate[3].

A direct binding assay for inositol phosphates. The finding that inositol phosphate analogues derivatized on carbons 2 and 3 of the inositol ring are able to activate the HDAC3:SMRT complex, suggested that it might be possible to use a fluorescence anisotropy assay to directly measure the inositol phosphate binding to the complex. Accordingly, we used a fluorescent derivative of $Ins(1,3,4,5,6)P_5$ in which fluorescein was coupled via a linker to the oxygen on carbon 2 of the inositol ring (2-FAM-$Ins(1,3,4,5,6)P_5$) (Fig. 4a) (ref. 38). Importantly, the 2-FAM-$Ins(1,3,4,5,6)P_5$ activates the HDAC3:SMRT complex to a similar extent as $Ins(1,4,5,6)P_4$ (Supplementary Fig. 6).

Interestingly, the dissociation constant measured using this direct binding assay is 0.3 μM (±0.01 μM). This is ~10-fold tighter than the apparent K_d value observed in the activation assays, suggesting that binding does not necessarily equate to

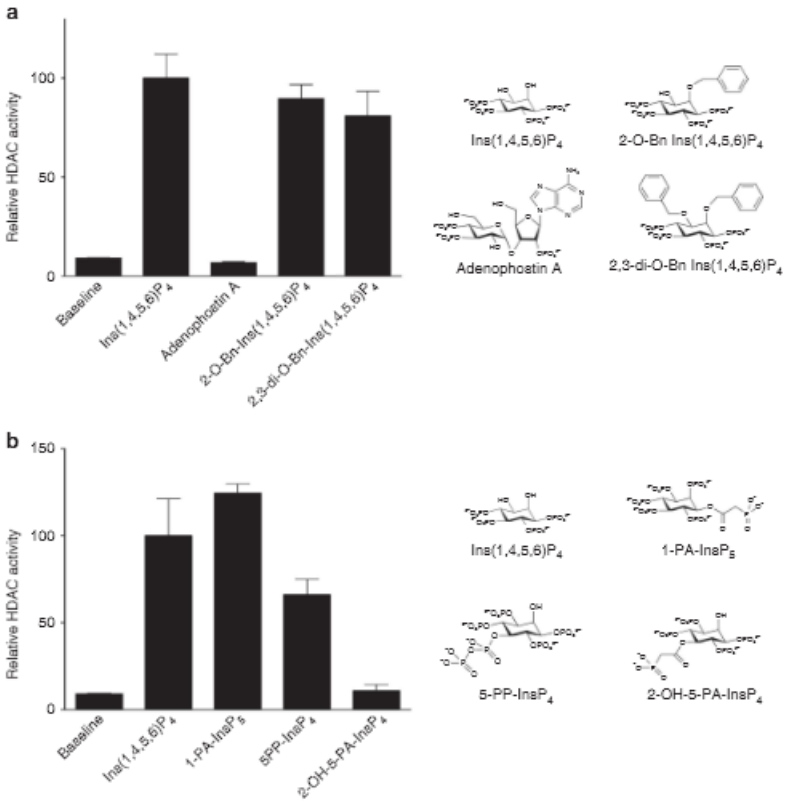


Figure 2 | Activation of the HDAC3:SMRT complex by inositol phosphate analogues and derivatives. (**a**) The ability of adenophostin A and inositol phosphates with bulky substituents at the 2 and 3 positions to activate HDAC3:SMRT compared with Ins(1,4,5,6)P_4. (**b**) Comparison of inositol pyrophosphate and pyrophosphate mimics with Ins(1,4,5,6)P_4. In all cases, compounds were tested at 200 μM for their ability to stimulate HDAC activity. HDAC activity is expressed as the percentage of maximal Ins(1,4,5,6)P_4 activity. Error bars indicate ± s.e.m. ($n = 3$).

activation. This is likely due to there being multiple modes of binding to the complex, such that some forms of the bound complex are not fully active. However, we cannot rule out that the fluorescein moiety might contribute in part to the binding.

The concept that binding does not necessarily equate to activation is supported by the results of competition assays. These indicate that certain inositol phosphates, such as Ins(1,3,4,6)P_4 and Ins(1,4,5)P_3, which do not activate the complex and cannot simultaneously occupy sites A, B and C, can nevertheless displace the 2-FAM-Ins(1,3,4,5,6)P_5 with IC_{50} values of 1.5 and 19 μM, respectively (Fig. 4b). We note that the InsP_3R agonist, adenophostin A[33,34], neither activates HDAC3, nor competes for inositol phosphate binding.

Inositol phosphates as allosteric regulators of Km and kcat. Lysine deacetylase assays using the minimal substrate Boc-Lys(Ac)-AMC show that both the HDAC3:SMRT and HDAC1:MTA1 complexes are reproducibly activated by inositol phosphates. However, in this assay, the HDAC3 complex appears to be much more sensitive to inositol phosphates giving a 10–15-fold increase in activity, compared with a 2–3-fold activation of the HDAC1 complex. To explore further the differences between the two complexes we established a more physiological, real-time kinetic assay using substrates based on the tails of histones H4 and H3. The peptides tested were; H4 12–18(K16ac), H3 23–29(K27ac) and H3 6–12(K9ac). Interestingly, there is a marked variation in Km, kcat and enzymatic efficiency (kcat/Km) for the various substrates. HDAC3:SMRT is most catalytically active against H3 23–29(K27ac) (Supplementary Fig. 7), whereas HDAC1:MTA1 is most catalytically active against H4 12–18(K16ac) (Fig. 4d). In both cases an approximate 3-fold greater kcat is observed for the preferred versus non-preferred substrate, in the absence of inositol phosphate. In general, inositol phosphates had a modest effect on Km (sometimes lowering, but sometimes increasing), and a more marked effect on kcat (always increasing). The net effect of both inositol phosphates tested was between a two and fivefold increase in catalytic efficiency (kcat/Km) (Fig. 4c,d, Supplementary Fig. 7).

A novel peptide-based inhibitor mimics substrate binding. To understand how a substrate, such as the histone H4 tail, interacts with HDAC1, we synthesized a novel inhibitor based on this peptide. This inhibitor (termed H4K16Hx) comprises residues 12–18 of histone H4 with K16 being replaced by a hydroxamic acid functionality (Supplementary Fig. 8). H4K16Hx is markedly different from the macrocyclic peptide inhibitors[39] (cyclic peptides derived from natural products), one of which (romidepsin) is licensed for the treatment of cutaneous and

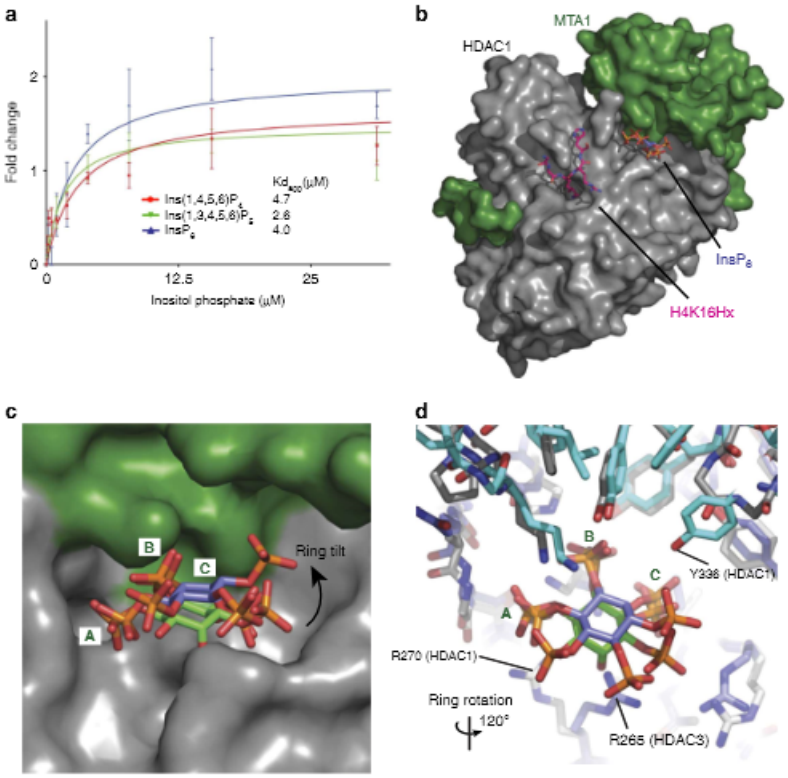


Figure 3 | Structural basis for the activation of HDAC1:MTA1 by inositol phosphates. (**a**) HDAC activity of HDAC1:MTA1 is enhanced by Ins(1,4,5,6)P_4, Ins(1,3,4,5,6)P_5 and InsP_6. A BOC-Lys-AMC assay was used - error bars indicate ± s.e.m. ($n=3$). (**b**) Crystal structure of HDAC1:MTA1 with InsP_6 bound at the interface between the two proteins. The modified histone H4 peptide H4K16Hx (pink) is bound to the active site. (**c**,**d**) Three adjacent phosphates of InsP_6 (purple) occupy sites A, B and C in an almost identical manner to the Ins(1,4,5,6)P_4 (green) molecule bound to the HDAC3:SMRT crystal structure. However, InsP_6 lies in a different orientation to Ins(1,4,5,6)P_4; the ring is rotated by 120° and is tilted towards MTA1.

peripheral T-cell lymphoma[10]. Inhibition assays show that this peptide inhibits the enzyme with an IC_{50} of 336 nM which is comparable with IC_{50} values reported for other hydroxamic acid-based inhibitors of Zn-dependent HDACs[40] (Fig. 5a).

We co-crystallized the HDAC1:MTA1 complex with this peptide before soaking it with InsP_6 (see above). The peptide was observed to be bound at the active site of the enzyme with nearly 100% occupancy, although residues Lys12 and Gly13 were not observed in the electron-density map suggesting that these are not constrained on substrate binding. The hydroxamic acid functionality is buried within the narrow active-site channel with the carbonyl oxygen coordinating the Zn^{2+} in a similar manner to that observed with other class I HDAC inhibitors (for example trichostatin A (TSA) and SAHA bound to HDAC8: pdb codes 1T69 and 1T64, respectively). Several backbone amides in the H4K16 peptide make complementary polar contacts with the sidechain of D99 at the rim of the HDAC active site (Fig. 5b, Supplementary Fig. 5). Interestingly, D99 is conserved in all class I and class II HDACs. Comparison with the structure in the absence of the histone peptide suggests that this residue, and its immediate neighbours, undergo a conformation change on peptide binding. The critical importance of D99 is supported by the finding that mutation to alanine results in total loss of catalytic activity (Fig. 5c).

Mechanism of class I HDAC activation by inositol phosphates. The finding that the sites A, B and C are essential for activation of the HDAC3:SMRT complex prompted us to investigate in more detail the mechanism through which inositol phosphates activate class I HDACs. We previously noted that the phosphate at site A in the HDAC3:SMRT structure makes an apparently important contact to the sidechain of R265, and that mutation of this arginine resulted in a significant loss in activity of the enzyme[2,3]. One possible explanation for this was that R265A mutant abolished binding of Ins(1,4,5,6)P_4 to the complex. To test this we measured the binding affinity of 2-FAM-Ins(1,3,4,5,6)P_5 with the complex. Surprisingly, the dissociation constant was only modestly increased from 0.3 μM to 0.6 μM (Fig. 6a), suggesting that binding of phosphates in sites B and C is the main thermodynamic driver of the interaction and that the interaction of the phosphate in site A with R265 plays a more important role in activation. Fitting with this hypothesis, the mutant complex is only weakly activated by the addition of inositol phosphate (24% of the activation seen for the wild-type complex; Fig. 6b). Furthermore the mutation significantly impairs the ability of the complex to repress transcription of a luciferase reporter gene (Fig. 6c).

Given the importance of R265 in the HDAC3:SMRT:Ins(1,4,5,6)P_4 complex, it is surprising that the equivalent residue,

Table 1 | Data collection and refinement statistics.

Data collection	
Wavelength (Å)	0.96862
Space group	P 3_2 2 1
Cell dimensions	
a, b, c (Å)	107.8, 107.8, 134.2
α, β, γ (°)	90, 90, 120
Resolution (Å)	93.4–3.3 (3.56–3.3)*
I/σI	2.5 (1.3)*
Completeness (%)	93.1 (92.0)*
Redundancy	2.4 (2.4)*
Rmerge (%)	38.3 (86.4)*
$CC_{1/2}$	0.795 (0.456)*
Refinement	
R_{work}/R_{free} (%)	25.5/29.9
Reflections	12,263
Number of atoms	
Protein	4,326
Zn ions, K ions and IP6	39
Water	0
B-factor (Å^2)	
Protein	34.9
Zn ions, K ions and acetate	29.7
R.m.s.d.	
Bond lengths (Å)	0.008
Bond angles (°)	1.228
Ramachandran plot	
Favoured (%)	92.3
Allowed (%)	7.7
Outliers (%)	0.0

*The highest resolution shell is shown in parentheses.

R270, in the HDAC1:MTA1:InsP6 complex has a different orientation, such that it interacts with the phosphate group in site C. Accordingly, the R270A mutation in the HDAC1:MTA1 complex does not have as pronounced an effect as that seen when R265 is mutated in the HDAC3:SMRT complex, but it still results in a twofold reduction in kcat (Supplementary Fig. 9). We previously observed that the R270Q mutation only modestly impairs the ability of HDAC1 to rescue cells in which both HDAC1 and 2 have been deleted. This mutation does, however, strongly synergise with other mutations in the inositol phosphate-binding pocket[29]. Interestingly, a survey of the available HDAC2 structures (pdb codes: 3MAX, 4LXZ, 4LY1) in the absence of an inositol phosphate ligand, suggests that this arginine is rather mobile and can adopt either the position seen for R265 in the HDAC3:SMRT:Ins(1,4,5,6)P_4 structure or the position seen in the HDAC1:MTA1:InsP_6 complex (Supplementary Fig. 10). In either position the backbone trajectory remains unchanged. It seems likely, therefore, that the important feature is that this sidechain becomes immobilized on inositol phosphate binding and that this stabilization is important for activation.

Notably we have also observed a measurable cross-talk between the active site of HDAC3 and the inositol phosphate-binding site. This is illustrated by the finding that binding of hydroxamic acid inhibitors, such as TSA, SAHA and MS-275 (a structurally distinct benzamide inhibitor) at the active site results in a threefold enhancement of the binding of 2-FAM-Ins(1,3,4,5,6)P_5 to the allosteric site (Fig. 6d). Interestingly, the inhibitor H4K16Hx has a more modest effect on the Kd for 2-FAM-Ins(1,3,4,5,6)P_5. Those inhibitors that particularly influence 2-FAM-Ins(1,3,4,5,6)P_5 binding position an aromatic group at the mouth of the active site.

To further investigate the effects of both inhibitors and inositol phosphates, we used a circular dichroism-based thermal-stability assay. The unliganded HDAC3:SMRT complex showed co-operative unfolding at 45 °C, which increased to 51 and 53 °C by SAHA and InsP_6, respectively. When the thermal denaturation was measured in the presence of both SAHA and InsP_6 the melting temperature further increased to 62 °C. This indicates that binding to the active site or inositol phosphate results in significant stabilization of the complex (Fig. 6e). Similar effects have also been observed on the stabilization of the HDAC1:MTA1 complex (Supplementary Fig. 11).

Discussion

Class I HDAC complexes have been implicated in the regulation of a wide range of biological processes from early development and X-chromosome inactivation to metabolism and circadian rhythms[41–45]. The finding that these complexes are regulated by inositol phosphates was unexpected and raises the question what is the physiological role of inositol phosphates; under what circumstances do their levels change and which biological processes are consequently regulated? We have not sought to answer these questions in this manuscript. Rather we seek to understand the molecular mechanisms through which inositol phosphates regulate the activity of HDAC complexes. Specifically, we wished to understand what are the important features of inositol phosphates that allow them to activate HDAC co-repressor complexes; what is the relationship between substrate and inositol phosphate binding and what is the mechanism underlying the allosteric activation? We have systematically tested the ability of a range of naturally occurring and synthetic inositol phosphates and their analogues to activate class I HDAC complexes and this has allowed us to elucidate the important stereochemical features of inositol phosphates that are required to allosterically activate class I HDAC complexes. The use of a fluorescently labelled inositol phosphate derivate has allowed the investigation of the direct binding of inositol phosphates to HDACs for the first time. This and other derivatives have effectively demonstrated that inositol phosphate derivatives may have potential use for modulating HDAC co-repressor complexes. The structure of HDAC1 in complex with the novel histone H4 12–18 tail inhibitor peptide has given insight into understanding both substrate recognition and its relationship with inositol phosphate binding. Taken together with kinetic and mutational studies these results allow us to begin to understand the mechanism underlying the allosteric activation by inositol phosphates.

Our investigations into HDAC3:SMRT activation have led us to propose a biological rationale for higher order inositol phosphate signalling to HDAC complexes. Ins(1,4,5)P_3 is a major precursor for higher inositol phosphates, yet is completely unable to activate HDAC3 since it cannot simultaneously occupy sites A, B and C (Fig. 1a). Addition of a phosphate to Ins(1,4,5)P_3 at either the 3 or 6 position creates a molecule that can robustly activate HDAC3. Such phosphorylation of Ins(1,4,5)P_3 is carried out by an inositol phosphate multikinase enzyme (called IPMK or IPK2) (ref. 46,47). We expect, therefore, that the activity of IPMK is critical for class I HDAC activity *in vivo*, since it is essential for the production of all of the inositol phosphates that are able to activate HDAC complexes. The predominant nuclear localization of IPMK fits well with this biological rationale[48]. Furthermore, the fact that Ins(1,4,5)P_3, as well as the potent InsP_3R agonist adenophostin A, are both completely inert in activating HDAC complexes indicates that the signalling role played by higher

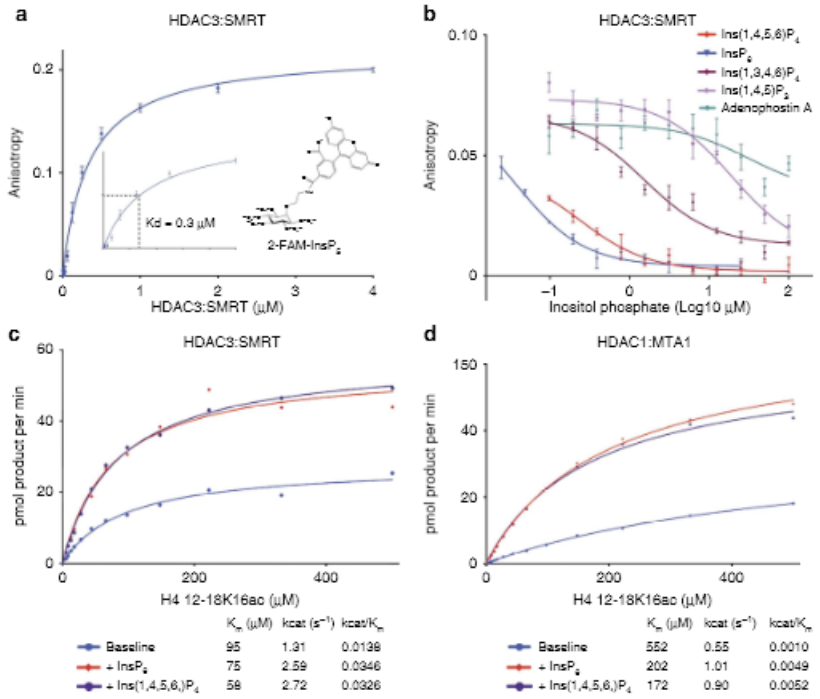


Figure 4 | Direct binding of inositol phosphates and effect of inositol phosphate on kinetic parameters. (**a**) Fluorescence anisotropy assay of the binding of 2-FAM-Ins(1,3,4,5,6)P_5 to HDAC3:SMRT. (**b**) Displacement of 2-FAM-Ins(1,3,4,5,6)P_5 by various inositol phosphates. IC_{50} values were calculated using Graphpad Prism. (**c**,**d**) Determination of enzyme kinetic parameters in the presence and absence of $InsP_5$ and Ins(1,4,5,6)P_4 for the HDAC3:SMRT and HDAC1:MTA1 complexes with H4 12-18K16ac peptide. Error bars indicate ± s.e.m. ($n = 3$).

order inositol phosphates in modulating the activity of class I HDACs is physiologically isolated from the role played by Ins(1,4,5)P_3 in Ca^{2+} signalling.

Support for the concept that IPMK plays a key role comes from previous findings that both IPMK and its yeast homologue, Arg82p, have been reported to play a role in gene activation[46,49–54]. However the role of the kinase activity of Arg82p/IPMK in gene activation is controversial. The kinase activity has been shown to be required for Pho5 transcription and chromatin remodelling[49], whereas other studies suggest that the kinase activity is not required for the regulation of arginine metabolism in yeast[53,54]. IPMK plays a role as a transcriptional co-activator in immediate early gene induction in mice through its interaction with the histone acetyltransferase CBP, though this interaction is not dependent on the kinase activity of IPMK[52]. Interestingly, Xu *et al.*[52] speculate that the products of the kinase activity of IPMK might be important for repression of gene transcription, whereas the non-catalytic activity stimulates histone acetylation and therefore gene activation.

Despite the uncertainty in the role of the kinase activity of IPMK, it is well-established that all the products of IPMK (higher order inositol phosphates) vary throughout the cell-cycle[55]. This is particularly interesting, since it is established that histone acetylation levels also vary during the cell-cycle[56], yet the levels of HDAC3 remain constant[57]. Furthermore, deletion of HDAC3 results in loss of the cell-cycle-dependent variation in histone acetylation[57]. If HDAC3 is responsible for cell-cycle regulation of histone acetylation yet its levels do not change, then some other factor (perhaps inositol phosphate levels?) must change.

Despite these suggestive observations it still remains unclear what is the physiologically relevant species that activates HDAC complexes *in vivo*. While Ins(1,4,5,6)P_4 co-purified with the complex from HEK293 cells, it has become clear that a range of higher order phosphates are able to activate *in vitro*. Importantly inositol phosphate kinases that lie downstream of IPMK and give rise to inositol pyrophosphates have been shown to be important for the activation of the yeast class I HDAC RPD3L and for implementing a stress response in yeast[35]. Indeed, Worley *et al.*[35] have shown that mutation of residues in the inositol-binding site on the yeast HDAC homologue Rpd3 results in similar effects on gene expression in the environmental stress response as deleting enzymes in the pyrophosphate synthesis pathway (including Arg82). These findings fit well with our observation that certain inositol pyrophosphates and pyrophosphate analogues are able to activate the HDAC3:SMRT complex *in vitro*.

It is striking that there were only very minor structural changes when HDAC1 binds $InsP_5$ and the synthetic substrate mimic. The most significant is the rearrangement of the sidechain of D99 that mediates essential interactions with the peptide backbone of the substrate. The importance of this residue in HDAC8 has been noted before[58]. This rearrangement appears to be the result of substrate binding, rather than an allosteric consequence of inositol phosphate binding. Importantly, although the kinetic analysis indicates that substrate binding is strongly influenced by inositol phosphate binding and that catalytic turnover is increased, there is little evidence for a substantive structural change mediating these allosteric effects. This contrasts with the classic dogma of allostery that requires concerted structural

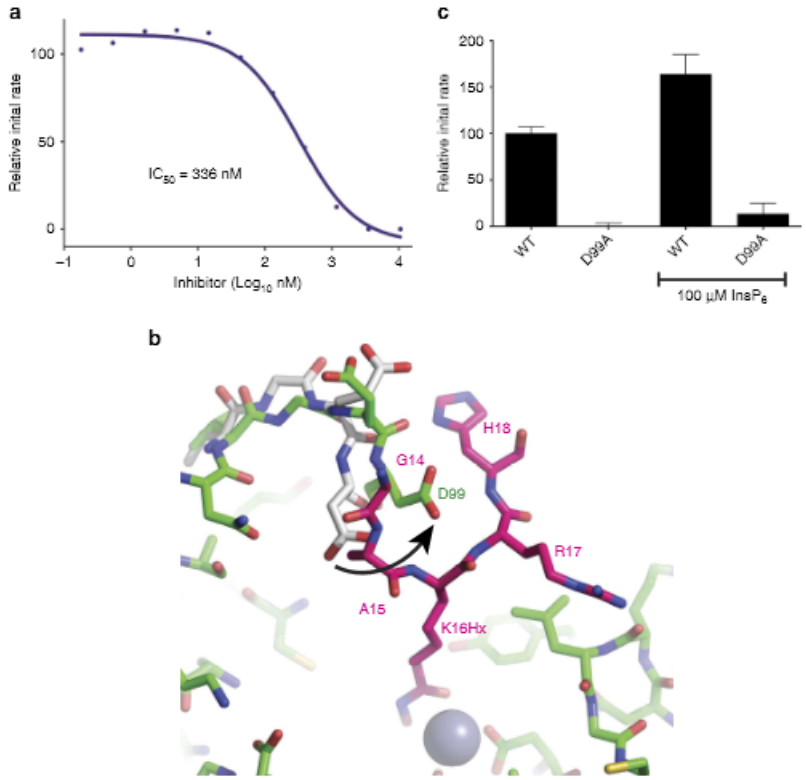


Figure 5 | Inhibition of HDAC1:MTA1 by a novel peptide-based inhibitor. (**a**) Inhibition curve of HDAC1:MTA1 by the histone H4 peptide H4K16Hx. (**b**)The position of H4K16Hx (pink) bound to the active site of HDAC1 (purple). The sidechain of HDAC1 residue D99 undergoes a significant rearrangement (indicated by arrow) to coordinate binding of the peptide (apo-HDAC1 shown in grey). (**c**) The D99A HDAC1 mutant does not show deacetylase activity and activity is not rescued on addition of $InsP_6$. HDAC assays were performed using an electrophoretic mobility shift assay on the caliper platform. Error bars indicate ± s.e.m. ($n = 3$).

changes to mediate thermodynamic or kinetic alterations to an enzyme mechanism[59]. The allosteric mechanism in the HDAC complexes would seem to fit much better with concepts of allostery being the result of changes in entropy (dynamics) following binding at the allosteric site influencing the activity at the active site[60]. This interpretation is supported by recent computational studies which suggest that Ins(1,4,5,6)P_4 binding induces a change in the dynamic behaviour of the complex in the vicinity of the inositol phosphate and active sites[61]. Furthermore, the thermal denaturation studies reported here show that the binding of inositol phosphate results in a significant stabilization of both the HDAC3:SMRT and HDAC1:MTA1 complexes.

Whilst it is recognized that dynamically driven allostery can occur over a large distance, the fact that the allosteric and active sites are in relatively close proximity provides a likely explanation of how inositol phosphate binding can influence the dynamics of the active site[59]. It is particularly notable that the inositol phosphate and the substrate–peptide interact on either side of a loop involving residues Q26-P29 in HDAC1. Indeed, both the substrate and inositol phosphate form hydrogen bonds with the same peptide bond—Q26-G27 (Fig. 6f).

HDACs have recently shown promise as therapeutic targets to treat a number of different diseases. However, one of the major challenges is that inhibitors of class I HDACs exhibit relatively modest subclass specificity and, furthermore, several different complexes, with diverse biological functions, contain a common catalytic HDAC subunit. The finding that class I HDACs behave quite differently when they are in complex with their cognate co-repressors, provides a new opportunity to develop inhibitors that are specific for individual complexes. Understanding the molecular details of substrate binding, and the allosteric mechanism of activation by inositol phosphates, is likely to be essential for rational drug design.

Methods

HDAC3:SMRT protein expression. Full-length HDAC3, and residues 350–480 of SMRT were cloned into pcDNA3 vectors, with a FLAG tag and TEV protease cleavage site in the SMRT construct. To express the complex both constructs were co-transfected into HEK293F cells (Invitrogen). To transfect 300 ml of cells, 150 μg of each construct was diluted in 30 ml PBS (Sigma), vortexed briefly and 1.5 ml of 0.5 mg ml^{-1} 25 kDa branched polyethyleniminc (Sigma) added. This mixture was vortexed briefly and incubated for 20 min at room temperature and then added to cells at a final density of 1×10^6 cells per ml. Cells were harvested after 48 h, resuspended in buffer A (100 mM potassium acetate, 50 mM Tris pH 7.5, 5% glycerol, 0.3% Triton X-100, Roche complete protease inhibitor tablet), lysed by sonication and then centrifuged to remove the insoluble material. The cleared supernatant was then incubated with FLAG resin (Sigma) for 1 h at 4 °C. The resin was then washed three times with buffer A, three times with buffer B (300 mM potassium acetate, 50 mM Tris pH 7.5, 5% glycerol), and three times with buffer C (50 mM potassium acetate, 50 mM Tris pH 7.5, 5% glycerol, 0.5 mM TCEP). The

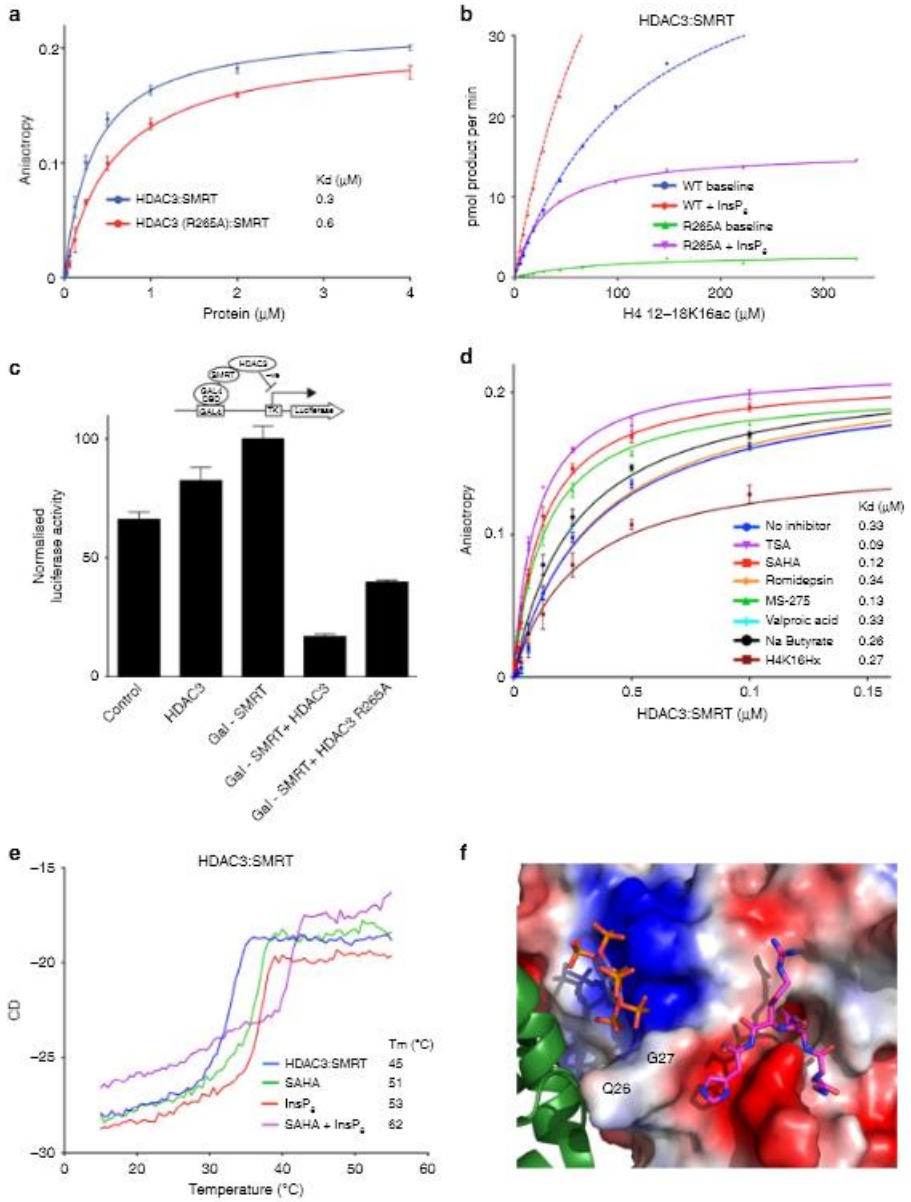


Figure 6 | Exploring the mechanism of activation by inositol phosphates. (**a**) Comparison of the binding of 2-FAM-Ins(1,3,4,5,6)P_5 to HDAC3:SMRT and HDAC3 (R265A):SMRT. (**b**) Measurement of Km and kcat in the presence and absence of $InsP_6$, for HDAC3 (R265A):SMRT. Data for the wild-type HDAC3:SMRT are also shown for comparison (N.B. the data are presented on a different scale from Fig. 4c). (**c**) Repression assay showing that repression mediated through HDAC3 recruitment by GAL4 - SMRT, is reduced by mutation of R265. (**d**) Comparison of the effect of various different HDAC inhibitors on the binding of 2-FAM-Ins(1,3,4,5,6)P_5 to HDAC3:SMRT. (**e**) CD denaturation curves of HDAC3:SMRT showing that $InsP_6$ and SAHA both stabilize the complex and have a greater effect when combined. Tm (°C) = melting temperature. Molar ellipticiy was monitored at 222 nm, and melting curves fitted using GraphPad Prism. (**f**) Close-up view of the H4K16Hx (pink sticks) bound in the active site of HDAC1 (electrostatic surface), $InsP_6$ (purple) is bound in close proximity to the active site at the interface with MTA1 (green cartoon). Residues involved in binding both $InsP_6$ and H4K16Hx are labelled. Error bars indicate ± s.e.m. ($n = 3$).

resin was incubated with TEV protease overnight at 4 °C in buffer C to elute the protein complex. Final purification of the eluted protein was performed by gel filtration in buffer containing 50 mM potassium acetate, 25 mM Tris pH 7.5, 0.5 mM TCEP, on a Superdex S200 column (GE healthcare). The intrinsically bound endogenous Ins(1,4,5,6)P_4 was removed by treatment with high-ionic strength buffer. 1 μM protein was incubated for 4 h at room temperature in buffer containing 50 mM Tris pH 7.5, 1 M NaCl, 5% glycerol, then dialysed overnight against buffer containing 50 mM Tris pH 7.5, 50 mM NaCl, 5% glycerol. Proteins were then concentrated using an Amicon concentrator and the protein concentration determined by A_{280} in 6 M guanidine hydrochloride.

HDAC1:MTA1 protein expression and structure determination. Full-length HDAC1 and residues 162–354 of MTA1 were cloned into pcDNA3 vectors, with a FLAG tag and TEV protease cleavage site in the MTA1 construct. Transfections and protein purification were performed as for the HDAC3:SMRT complex with the modification of buffer B to 200 mM potassium acetate, 50 mM Tris pH 7.5, 5% glycerol. Diffracting crystals were obtained by sitting-drop vapour diffusion at 20 °C against wells containing 0.1 M HEPES pH 7.5, 2 M ammonium sulfate and 5% PEG400 by mixing HDAC1:MTA1 (5 mg ml^{-1}) and peptide H4K16Hx at a 1:2 molar ratio. Crystals were frozen in mother liquor with the addition of $InsP_6$ (225 mM) and 15% glycerol (cryoprotectant). Data were collected at the Diamond synchrotron microfocus beamline I24 with use of the grid-scan tool to centre the crystals. Diffraction data from 4 crystals were processed using Mosflm and combined using Aimless. The structure was solved by molecular replacement with HDAC1:MTA1 (PDB code 4BKX) as a search model using Phaser. The HDAC1:MTA1 structure was built with Coot and refined using Refmac5 (ref. 62). The atomic coordinates for the structure of the HDAC1:MTA1: $InsP_6$:H4K16Hx complex have been deposited in the Protein Data Bank under accession code 5ICN.

Peptide synthesis. Details of the syntheses of peptide H4K16Hx (histone H4 residues 12–18 with hydroxamic acid functionality at K16) and fluorescein-labelled peptide H4K16Ac (FITC-H4 12–18 acetylated at K16) are given in the Supplementary Methods.

Inositol phosphates and derivatives. Adenophostin A[54], Ins(1,3,4,5)P_4 (ref. 63) *scyllo*-InsP_5 (ref. 31), Bz(1,2,3,4)P_4 (ref. 32), 1-PA-InsP_5 (ref. 36), 2-OH-5-PA-InsP_4 (ref. 37), 2-FAM-Ins(1,3,4,5,6)P_5 (ref. 38) were synthesized as previously reported. The identities and purities of these compounds were confirmed by ^{1}H and ^{31}P NMR spectroscopy. Ins(1,3,5,6)P_4, Ins(3,4,5,6)P_4, Ins(1,3,4,5,6)P_5, Ins(1,4,5)P_3, Ins(1,4,6)P_3, Ins(1,5,6)P_3, were purchased from Cayman Chemical Company. $InsP_6$ was purchased from Sigma. Details of the new syntheses of Ins(1,4,5,6)P_4, Ins(4,5,6)P_3, Ins(1,3,4,6)P_4, 2-*O*-Bn-Ins(1,4,5,6)P_4 and 2,3-di-*O*-Bn-Ins(1,4,5,6)P_4 are given in the Supplementary Methods.

HDAC assay. HDAC activity of the protein complex was measured using an HDAC assay with a BOC-Lys-AMC substrate. 50 nM HDAC3:SMRT or 50 nM HDAC1:MTA1 was incubated with 100 μM BOC-Lys-AMC substrate in a final volume of 50 μl buffer A (50 mM Tris pH 7.5, 50 mM NaCl, 5% Glycerol), in a black 96-well plate for 30 min at 37 °C. The assay was developed by the addition of 50 μl of developer solution (2 mM TSA, 10 mg ml^{-1} Trypsin, 50 mM Tris pH 7.5, 100 mM NaCl). Fluorescence was measured at 335/460 nm using a Victor X5 plate reader (Perkin Elmer).

To test the ability of inositol phosphates and derivatives to activate the complex, 50 nM HDAC3:SMRT (which had been stripped of its endogenous co-purified inositol phosphate) or 50 nM HDAC1:MTA1 was incubated with inositol phosphate or analogue for 30 min at 37 °C before the HDAC activity was measured. All measurements were performed in triplicate and data analysed using GraphPad Prism (version 6.0, GraphPad Software, Inc). In the case of titrations K_d values were calculated by nonlinear curve fitting with a one-site binding (hyperbola) model ($Y = B_{max}*X/(K_d + X)$).

Enzyme kinetics. Michaelis–Menten kinetics were determined using a HDAC assay on the Caliper EZ Reader II system (Caliper Life Sciences, http://www.caliperls.com). To assess the effect of inositol phosphate, 20 nM HDAC3:SMRT or 80 nM HDAC1:MTA1 were pre-mixed with $InsP_6$ before dilution and addition of fluorescein-labelled H4K16Ac substrate. Reactions were carried out in duplicate in 30 μl reaction volumes, performed at room temperature in 50 mM Tris pH 7.5, 50 mM NaCl, 5% glycerol. Data was analysed using GraphPad Prism (version 6.0).

Fluorescence anisotropy assays. Fluorescence anisotropy experiments were performed in a black 96-well plate (Corning), each well contained 10 nM 2-FAM-Ins(1,3,4,5,6)P_5 and an increasing concentration of HDAC3:SMRT protein, in assay buffer (50 mM NaCl, 50 mM Tris pH 7.5, 0.02% Tween-20), final volume of 25 μl. When required, assays were performed in the presence of HDAC inhibitor (20 mM final concentration). Plates were mixed by shaking for 1 s, and measurements taken using a Victor X5 plate reader (Perkin Elmer) at room temperature with an excitation wavelength of 480 nm and an emission wavelength of 535 nm. Experiments were performed in triplicate and data were analysed using GraphPad Prism (version 6.0), K_d values were calculated by nonlinear curve fitting using a one-site binding (hyperbola) model ($Y = Bmax*X/(K_d + X)$).

Displacement assays were performed essentially as above but with a fixed concentration of protein/2-FAM-Ins(1,3,4,5,6)P_5 and an increasing concentration of inositol phosphate. The concentration of protein was determined as that which gave 100% bound in the protein titration experiments. The data were analysed and IC_{50} values calculated using GraphPad Prism (version 6.0).

Reporter gene assays. HEK293T cells were transfected with MH-100-TK-luc reporter plasmid (containing Gal-binding sites), pCMX β-galactosidase, along with GAL-DBD-fused $SMRT_{350-480}$ (wild type and mutants) and untagged HDAC3, using polyethylenimine. Cells were lysed and assayed for reporter expression 48 h after transfection. Luciferase activity was determined using the Luciferase Assay Kit (Biovision) and normalized to the β-galactosidase activity. Measurements were carried out on a Victor X5 plate reader (Perkin Elmer).

Circular dichroism. A total of 14 μM of HDAC3:SMRT and 3.5 μM of HDAC1:MTA1 were incubated with 200 μM of $InsP_6$ or 40 μM SAHA, either individually or in combination. CD was monitored at 222 nm using a Applied Photophysics Chiroscan plus spectropolarimeter, and the sample temperature was increased from 10 °C to 90 °C (1 °C per minute). The data were analysed and melting temperatures were calculated using GraphPad Prism (version 6.0).

Computational docking studies. Docking studies were carried out with AutoDock4 (ref. 64), utilizing the AutoDockTools 1.5.6 GUI. Non-polar hydrogens and Gasteiger atomic charges were added to the HDAC3:SMRT DAD atomic coordinates (PDB ID: 4A69) in AutoDockTools. The inositol phosphate-binding site was as defined in ref. 2. Probes were calculated at every 0.375 Å grid position of a grid box (box size *x*, *y*, *z* = 21.406, 50.64, 23.036 Å, respectively), centred upon the inositol ring. The docking of the inositol phosphates was run using the Lamarckian genetic algorithm in AutoDock4. Other parameters were set to the default values.

References

1. de Ruijter, A. J. M., van Gennip, A. H., Caron, H. N., Kemp, S. & van Kuilenburg, A. B. P. Histone deacetylases (HDACs): characterization of the classical HDAC family. *Biochem. J.* **370,** 737–749 (2003).
2. Watson, P. J., Fairall, L., Santos, G. M. & Schwabe, J. W. R. Structure of HDAC3 bound to co-repressor and inositol tetraphosphate. *Nature* **481,** 335–340 (2012).
3. Millard, C. J. *et al.* Class I HDACs share a common mechanism of regulation by inositol phosphates. *Mol. Cell.* **51,** 57–67 (2013).
4. Itoh, T. *et al.* Structural and functional characterization of a cell cycle associated HDAC1/2 complex reveals the structural basis for complex assembly and nucleosome targeting. *Nucleic Acids Res.* **43,** 2033–2044 (2015).
5. West, A. C. & Johnstone, R. W. New and emerging HDAC inhibitors for cancer treatment. *J. Clin. Invest.* **124,** 30–39 (2014).
6. Sumner, C. J. *et al.* Valproic acid increases SMN levels in spinal muscular atrophy patient cells. *Ann. Neurol.* **54,** 647–654 (2003).
7. Herman, D. *et al.* Histone deacetylase inhibitors reverse gene silencing in Friedreich's ataxia. *Nat. Chem. Biol.* **2,** 551–558 (2006).
8. Kilgore, M. *et al.* Inhibitors of class 1 histone deacetylases reverse contextual memory deficits in a mouse model of Alzheimer's disease. *Neuropsychopharmacology* **35,** 870–880 (2009).
9. Shirakawa, K., Chavez, L., Hakre, S., Calvanese, V. & Verdin, E. Reactivation of latent HIV by histone deacetylase inhibitors. *Trends. Microbiol.* **21,** 1–9 (2013).
10. Wagner, J. M., Hackanson, B., Lübbert, M. & Jung, M. Histone deacetylase (HDAC) inhibitors in recent clinical trials for cancer therapy. *Clin. Epigenet.* **1,** 117–136 (2010).
11. Lee, H. Z. *et al.* FDA approval: belinostat for the treatment of patients with relapsed or refractory peripheral T-cell lymphoma. *Clin. Cancer Res.* **21,** 1–6 (2015).
12. Laubach, J. P., Moreau, P., San-Miguel, J. F. & Richardson, P. G. Panobinostat for the treatment of multiple myeloma. *Clin. Cancer Res.* **21,** 4767–4773 (2015).
13. Micelli, C. & Rastelli, G. Histone deacetylases: structural determinants of inhibitor selectivity. *Drug Discov. Today* **20,** 1–18 (2015).
14. Watson, P. J., Fairall, L. & Schwabe, J. W. R. Nuclear hormone receptor co-repressors: structure and function. *Mol. Cell. Endocrinol.* **348,** 440–449 (2012).
15. Zhang, Y. *et al.* Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes Dev.* **13,** 1924–1935 (1999).
16. Li, J. *et al.* Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J.* **19,** 4342–4350 (2000).
17. Lechner, T. *et al.* Sds3 (suppressor of defective silencing 3) is an integral component of the yeast Sin3 · Rpd3 histone deacetylase complex and is required for histone deacetylase activity. *J. Biol. Chem.* **275,** 40961–40966 (2000).

18. Zhang, J., Kalkum, M., Chait, B. T. & Roeder, R. G. The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Mol. Cell.* **9,** 611–623 (2002).
19. Guenther, M. G., Barak, O. & Lazar, M. A. The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. *Mol. Cell. Biol.* **21,** 6091–6101 (2001).
20. Wen, Y. D. *et al.* The histone deacetylase-3 complex contains nuclear receptor corepressors. *Proc. Natl Acad. Sci. USA* **97,** 7202–7207 (2000).
21. Hu, E. *et al.* Cloning and characterization of a novel human class I histone deacetylase that functions as a transcription repressor. *J. Biol. Chem.* **275,** 15254–15264 (2000).
22. Lee, H., Rezai-Zadeh, N. & Seto, E. Negative regulation of histone deacetylase 8 activity by cyclic AMP-dependent protein kinase A. *Mol. Cell. Biol.* **24,** 765–773 (2003).
23. Xue, Y. *et al.* NURD, a novel complex with both ATP-dependent chromatin-remodeling and histone deacetylase activities. *Mol. Cell.* **2,** 851–861 (1998).
24. Hassig, C. A., Fleischer, T. C., Billin, A. N., Schreiber, S. L. & Ayer, D. E. Histone deacetylase activity is required for full transcriptional repression by mSin3A. *Cell* **89,** 341–347 (1997).
25. You, A., Tong, J. K., Grozinger, C. M. & Schreiber, S. L. CoREST is an integral component of the CoREST-human histone deacetylase complex. *Proc. Natl Acad. Sci. USA* **98,** 1454–1458 (2001).
26. Bantscheff, M. *et al.* Chemoproteomics profiling of HDAC inhibitors reveals selective targeting of HDAC complexes. *Nat. Biotechnol.* **29,** 255–265 (2011).
27. Guenther, M. G. *et al.* A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness. *Genes Dev.* **14,** 1048–1057 (2000).
28. Irvine, R. F. & Schell, M. J. Back in the water: the return of the inositol phosphates. *Nat. Rev. Mol. Cell Biol.* **2,** 327–338 (2001).
29. Jamaladdin, S. *et al.* Histone deacetylase (HDAC) 1 and 2 are essential for accurate cell division and the pluripotency of embryonic stem cells. *Proc. Natl Acad. Sci. USA* **111,** 9840–9845 (2014).
30. You, S. H. *et al.* Nuclear receptor co-repressors are required for the histone-deacetylase activity of HDAC3 *in vivo*. *Nat. Struct. Mol. Biol.* **20,** 182–187 (2013).
31. Riley, A. M. *et al. scyllo*-inositol pentakisphosphate as an analogue of *myo*-inositol 1,3,4,5,6-pentakisphosphate: chemical synthesis, physicochemistry and biological applications. *Chembiochem* **7,** 1114–1122 (2006).
32. Mills, S. J. *et al.* Novel inositol phospholipid headgroup surrogate crystallized in the pleckstrin homology domain of protein kinase Bα. *ACS Chem. Biol.* **2,** 242–246 (2007).
33. Takahashi, S., Kinoshita, T. & Takahashi, M. Adenophostins A and B: potent agonists of inositol-1,4,5-trisphosphate receptor produced by Penicillium brevicompactum. Structure elucidation. *J. Antibiot.* **47,** 95–100 (1994).
34. Marwood, R. D., Correa, V., Taylor, C. W. & Potter, B. V. L. Synthesis of adenophostin A. *Tetrahedron: Asymmetry* **11,** 397–403 (2000).
35. Worley, J., Luo, X. & Capaldi, A. P. Inositol pyrophosphates regulate cell growth and the environmental stress response by activating the HDAC Rpd3L. *Cell. Rep.* **3,** 1476–1482 (2013).
36. Pulloor, N. K. *et al.* Human genome-wide RNAi screen identifies an essential role for inositol pyrophosphates in Type-I interferon response. *PLoS Pathog.* **10,** e1003981 (2014).
37. Riley, A. M., Wang, H., Weaver, J. D., Shears, S. B. & Potter, B. V. L. First synthetic analogues of diphosphoinositol polyphosphates: interaction with PP-InsP_5 kinase. *Chem. Commun.* **48,** 11292–11294 (2012).
38. Riley, A. M., Windhorst, S., Lin, H.-Y. & Potter, B. V. L. Cellular internalisation of an inositol phosphate visualised by using fluorescent InsP_5. *ChemBioChem* **15,** 57–67 (2014).
39. Mwakwari, S. C., Patil, V., Guerrant, W. & Oyelere, A. K. Macrocyclic histone deacetylase inhibitors. *Curr. Top. Med. Chem.* **10,** 1423–1440 (2010).
40. Hu, E. *et al.* Identification of novel isoform-selective inhibitors within class I histone deacetylases. *J. Pharmacol. Exp. Ther.* **307,** 720–728 (2003).
41. Knutson, S. K. *et al.* Liver-specific deletion of histone deacetylase 3 disrupts metabolic transcriptional networks. *EMBO J.* **27,** 1017–1028 (2008).
42. Feng, D. *et al.* A circadian rhythm orchestrated by histone deacetylase 3 controls hepatic lipid metabolism. *Science* **331,** 1315–1319 (2011).
43. McHugh, C. A. *et al.* The Xist lncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* **521,** 232–236 (2015).
44. Casas-Delucchi, C. S. *et al.* Histone acetylation controls the inactive X chromosome replication dynamics. *Nat. Commun.* **2,** 222–11 (2011).
45. Sun, Z. *et al.* Hepatic Hdac3 promotes gluconeogenesis by repressing lipid synthesis and sequestration. *Nat. Med.* **18,** 934–942 (2012).
46. Odom, A. R. A role for nuclear inositol 1,4,5-trisphosphate kinase in transcriptional control. *Science* **287,** 2026–2029 (2000).
47. Saiardi, A. *et al.* Mammalian inositol polyphosphate multikinase synthesizes inositol 1,4,5-trisphosphate and an inositol pyrophosphate. *Proc. Natl Acad. Sci. USA* **98,** 2306–2311 (2001).
48. Resnick, A. C. & Saiardi, A. Inositol polyphosphate multikinase: metabolic architect of nuclear inositides. *Front. Biosci. J. Virtual* **13,** 856–866 (2007).
49. Steger, D. J., Haswell, E. S., Miller, A. L., Wente, S. R. & O'Shea, E. K. Regulation of chromatin remodeling by inositol polyphosphates. *Science* **299,** 114–116 (2003).
50. El Alami, M., Messenguy, F., Scherens, B. & Dubois, E. Arg82p is a bifunctional protein whose inositol polyphosphate kinase activity is essential for nitrogen and PHO gene expression but not for Mcm1p chaperoning in yeast. *Mol. Microbiol.* **49,** 457–468 (2003).
51. Shen, X., Xiao, H., Ranallo, R., Wu, W.-H. & Wu, C. Modulation of ATP-dependent chromatin-remodeling complexes by inositol polyphosphates. *Science* **299,** 112–114 (2003).
52. Xu, R. *et al.* Inositol polyphosphate multikinase is a transcriptional coactivator required for immediate early gene induction. *Proc. Natl Acad. Sci. USA* **110,** 16181–16186 (2013).
53. Bosch, D. & Saiardi, A. Arginine transcriptional response does not require inositol phosphate synthesis. *J. Biol. Chem.* **287,** 38347–38355 (2012).
54. Dubois, E., Dewaste, V., Erneux, C. & Messenguy, F. Inositol polyphosphate kinase activity of Arg82/ArgRIII is not required for the regulation of the arginine metabolism in yeast. *FEBS Lett.* **486,** 300–304 (2000).
55. Barker, C. J., Wright, J., Hughes, P. J., Kirk, C. J. & Michell, R. H. Complex changes in cellular inositol phosphate complement accompany transit through the cell cycle. *Biochem. J.* **380,** 465–473 (2004).
56. Unnikrishnan, A., Gafken, P. R. & Tsukiyama, T. Dynamic changes in histone acetylation regulate origins of DNA replication. *Nat. Struct. Mol. Biol.* **17,** 430–437 (2010).
57. Bhaskara, S. *et al.* Hdac3 is essential for the maintenance of chromatin structure and genome stability. *Cancer Cell* **18,** 436–447 (2010).
58. Vannini, A. *et al.* Substrate binding to histone deacetylases as shown by the crystal structure of the HDAC8-substrate complex. *EMBO. Rep.* **8,** 879–884 (2007).
59. Tsai, C. J., del Sol, A. & Nussinov, R. Allostery: absence of a change in shape does not imply that allostery is not at play. *J. Mol. Biol.* **378,** 1–11 (2008).
60. Cooper, A. & Dryden, D. T. F. Allostery without conformational change. *Eur. Biophys. J.* **11,** 103–109 (1984).
61. Arrar, M., Turnham, R., Pierce, L., de Oliveira, C. A. F. & McCammon, J. A. Structural insight into the separate roles of inositol tetraphosphate and deacetylase-activating domain in activation of histone deacetylase 3. *Protein Sci.* **22,** 83–92 (2013).
62. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta. Crystallogr. D. Biol. Crystallogr.* **67,** 235–242 (2011).
63. Riley, A. M., Mahon, M. F. & Potter, B. V. L. Rapid synthesis of the enantiomers of *myo*-inositol-1,3,4,5-tetrakisphosphate by direct chiral desymmetrization of *myo*-inositol orthoformate. *Angewandte Chemie* **36,** 1472–1474 (1997).
64. Pospisil, P. *et al.* Computational and biological evaluation of quinazolinone prodrug for targeting pancreatic cancer. *Chem. Biol. Drug Des.* **79,** 926–934 (2012).

Acknowledgements

We thank Stuart Conway and Sonia Diab for supplying Ins(1,4,5,6)P_4; Stephen Mills for supplying Bz(1,2,3,4)P_4; Adam Baldwin for help with the synthesis of the peptide-based inhibitor; John Challiss, Louise Fairall and Peter Moody for helpful comments and discussion. We thank Diamond Light Source for access to beamline I24 that contributed to the results presented here. This work was supported by the Wellcome Trust (grants WT085408 and WT100237 to J.W.R.S., WT082837 to A.M.R. and B.V.L.P. and WT101010 to B.V.L.P.) and the BBSRC (grants BB/J009598/1 and BB/N002954/1) to S.M.C. and J.W.R.S. S.M.C. is the recipient of a Medical Research Council Senior Fellowship MR/J009202/1. J.W.R.S. and B.V.L.P. are both Wellcome Trust Senior Investigators. J.W.R.S. is a Royal Society Wolfson Research Merit Award holder.

Author contributions

J.W.R.S., P.J.W. and C.J.M. conceived the project and designed the experiments with help from A.M.R. and B.V.L.P. (experiments involving various inositol phosphates and derivatives including FAM-IP5), and A.G.J. (design of the peptide-based inhibitor). P.J.W. prepared the HDAC3:SMRT complex, established the methodology and performed most of the biochemical assays. C.J.M. determined the structure of HDAC1:MTA1 with InsP_4 and inhibitor and performed the biochemical assays with this complex. L.C.W., with support from S.M.C., performed the transcriptional repression assays. A.M.R. synthesized many of the inositol phosphates and derivatives. N.S.R. synthesized the peptide-based inhibitor. H.Y.G. synthesized Ins(1,3,4,6)P_4. A.G.J. and B.V.L.P. supervised the various chemical syntheses. P.J.W., C.J.M. and J.W.R.S. analysed and interpreted the data with assistance from A.M.R., B.V.L.P. and A.G.J. in relation to the chemical aspects. P.J.W., C.J.M. and J.W.R.S. wrote the manuscript, which was critically reviewed and approved by all the authors.

Additional information

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Watson, P. J. *et al.* Insights into the activation mechanism of class I HDAC complexes by inositol phosphates. *Nat. Commun.* 7:11262 doi: 10.1038/ncomms11262 (2016).