A study on side chain linked peptides, toward the development of talin inhibitors using β3 integrin peptide analogues

A thesis

submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy in Chemistry

at

The University of Adelaide

School of Physical Sciences

by

Kelly Lee Keeling

THE UNIVERSITY of ADELAIDE

March 2018

Table of contents

Abstract

This thesis discusses the design and production of peptides with side chain linkers that are intended to bind to the F3 domain of talin. The talin F3 domain was targeted as it is involved in the activation of integrin membrane proteins present in platelets. The over activation of these integrins can result in clotting within the blood vessels causing heart disease, however, current medication targeting integrin have negative side effects.

The design and synthesis of short peptides based on the sequence of the β3 integrin tail that binds to the F3 domain of talin is presented. The binding affinity of peptides to the talin F3 domain was tested using NMR titrations to reveal the ideal location for the linker in the production of potential therapeutics which target integrin activation. Side chain linked peptides with high helical content have previously been shown to improve binding affinity. This drove investigation of side chain constrained peptides to increase their helical content, and thus, their binding affinity to talin F3 domain and cellular uptake.

It is demonstrated that side chain linkers are effective in stabilising the helical structure of the short peptides. When incorporated in the β3 integrin sequence in specific locations, lactam linkers improved binding affinity of these peptides to the talin F3 domain. Additionally, all-hydrocarbon and triazole linkers enhanced the peptide's cellular uptake when compared to the native peptide of this sequence. The position and type of side chain linkers were investigated. The result of which showed that the position of the linker had a significant impact on the binding affinity to talin. The lactam linker between residues in positions 725 and 729 created a peptide (**7**) with the highest binding affinity.

The cell penetration of peptides with different linker types was tested using NIH 3T3 mouse cells, and HEK298 cells. A number of side chain linkers were tested with the triazole linker producing the most α-helical peptide, and the all-hydrocarbon linker producing peptides with the greatest cellular uptake.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint-award of this degree.

I acknowledge that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I also give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Kelly Lee Keeling Date

Publications

Work in this thesis has appeared in the following publication:

Keeling, K. L.; Cho, O.; Scanlon, D. B.; Booker, G. W.; Abell, A. D.; Wegener, K. L., The key position: influence of staple location on constrained peptide conformation and binding. Org Biomol Chem 2016, 14 (41), 9731-9735.

Acknowledgements

Alone we can do so little;

together we can do so much.

Helen Keller

I would like to acknowledge the assistance, support and guidance given by

My supervisor Andrew Abell and Kate Wegener for their unwavering support and suggestions on this journey. Denis Scanlon, Jade Cottam, Sabrina Heng and Michelle Zhang for advice in the laboratory and assistance with equipment maintenance. Aimee Horsfall for working with me on protein synthesis. Claire Weekly and Briony Gliddon for assistance with developing cellular uptake experiments, and Emma Parkenson-lawrance for her assistance with the CD spectrometer. All the members of the Abell group, past and present for being there and accepting me as part of the group. Family and friends for being just what I needed.

Abbreviations

Ala (A)	alanine
Arg (R)	arginine
Asn (N)	asparagine
Asp (D)	aspartic acid
Boc	*tert*-butoxycarbonyl
Cys (C)	cysteine
DCM	dichloromethane
DIC	*N*,*N*'-diisopropylcarbodiimide
DIPEA	*N*,*N*-diisopropylethylamine
DMF	*N*,*N*-dimethylformamide
DODT	3,6-dioxa-1,8-octanedithiol
DTT	dithiothreitol
equiv.	equivalent
ESI-MS	electrospray ionisation mass spectrometry
Fmoc	9-fluorenylmethyloxycarbonyl
Gln (Q)	glutamine
Glu (E)	glutamic acid
Gly (G)	glycine
HATU	2-(7-aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HBTU	1-[Bis(dimethylamino) methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
HCTU	2-(6-chloro-1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate
His (H)	histidine
HOBt	*N*-hydroxybenzotriazole
hArg	homoarginine
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
Hr	Hour
Hrs	Hours
Hz	hertz (in NMR)
Ile (I)	isoleucine
LCMS	liquid chromatography mass spectrometry
Leu (L)	leucine
lit.	literature value
Lys (K)	lysine
m/z	mass to charge ratio
Met (M)	methionine
MHz	megahertz (in NMR)
NMP	*N*-methylpyrrolidine
NMR	nuclear magnetic resonance
Phe (F)	phenylalanine
ppm	parts per million
Pro (P)	proline
RP-HPLC	reverse phase high performance liquid chromatography
rt	room temperature
semi-prep	semi-preparative
Ser (S)	serine
SPE	solid phase extraction
SPPS	solid phase peptide synthesis

*t*Bu	*tert*-butyl
TFA	trifluoroacetic acid
TFE	trifluoroethanol
THF	tetrahydrofuran
Thr (T)	threonine
TIPS	triisopropylsilane
TLC	thin-layer chromatography
Tris	tris(hydroxymethyl)amino methane
Trp (W)	tryptophan
Trt	trityl
Tyr (Y)	tyrosine
Val (V)	valine
w/v	mass per unit volume

Chapter 1

Introduction

1.1 Protein-Protein Interactions

Protein-protein interactions play a central role in all biological processes, mediating molecular events that define life, growth, reproduction, ageing, disease and death[1]. The structure of a protein is determined by its three-dimensional structure, which arises through highly ordered folding of the polypeptide backbone. An understanding of protein structure and function with regards to the modulation of protein-protein interactions has the potential to i) facilitate the development of a better understanding of biological systems by providing tools to study complex biological interactions; ii) provide novel approaches for diagnostics of health and disease; iii) most importantly, provide a basis for new molecular therapeutics to fight disease.

1.2 Secondary Structure

Protein secondary structure, as defined by the ordered conformational arrangement of amino acids in localised regions of a polypeptide or protein, plays a key role in protein-protein interactions[2]. The α-helix together with the β-sheet account for over 60% of the predominant structure at the interface of protein-protein interactions[2]. The structure of a protein can be described by the set of dihedral angles (Figure 1-1) phi (ϕ) and psi (ψ), which define the spatial orientation of the peptide backbone and the presence of specific hydrogen bonds between backbone carbonyl and amide functional groups.

Almost half a century ago Ramachandran, Sasisekharan, and Ramakrishnan identified that the energy landscape of a blocked peptide unit can be mapped in two dimensions[3]. The Ramachandran plot (Figure 1-1) demonstrated that steric clashes alone eliminate ¾ of ϕ, ψ- space. Regions free from steric clashes are the 'allowed' regions, and are occupied by common secondary structures, such as α-helices, 3_{10}-helicies, β-sheets and β-turns.

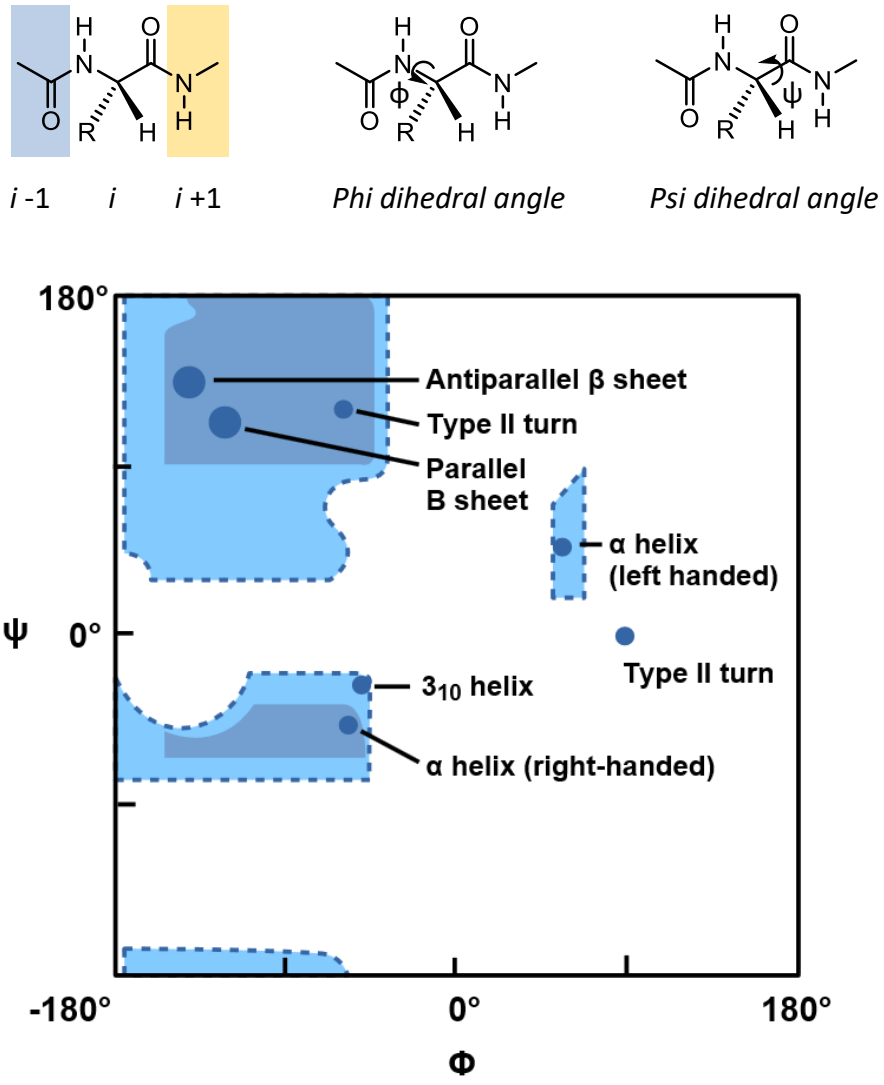


Figure 1-1: Top: '*i*' is a single residue within a peptide chain where R is the amino acid side chain. Connection via amide bond to the adjacent residue toward the *N*-terminal *i*-1 is highlighted in blue, while the residue toward the *C*-terminal *i*+1 is highlighted in orange. Dihedral angles phi (ϕ) and psi (Ψ) are indicated on the residue in the *i* position. Bottom: Ramachandran plot of the steric energies of interaction as a function of the ϕ and angles Ψ first developed by Ramachandran and Sasisekharan (1968) to represent the sterically allowed conformations of a polypeptide chain. The space occupied by geometry of certain secondary structures defined by their phi and psi angles (blue), and disallowed regions (white). (Adapted from Van Holde et.al[12])

1.2.1 α-Helix

The α-helix, first described by Linus Pauling[4], is the most common secondary structure found in proteins, and often forms fundamental recognition elements in protein-protein interactions[2]. α-Helices are observed in almost 40% of natural polypeptides[5], and form compact rod like structures with 3.6 amino acids per turn, with a length of 5.4 Å (Figure 1-2B) and a radius, not including side chains, of 2.3 Å (Figure 1-2D). The average length of helical domains in proteins is relatively short, spanning two to three helical turns (8-12 residues)[6]. The backbone arrangement of the α-helix places the *i*, *i*+4, *i*+7 and *i*+11 side chains on the same face of the folded structure (Figure 1-2C). The principal geometry for the α-helix is $\phi = -57°$ and $\psi = -47°$, with hydrogen bonds between the backbone of carbonyl (C=O) and amine (NH) groups of the *i* and *i*+4 residues. All of the backbone amide groups of an α-helix are involved in intra-chain hydrogen bonds, therefore the interactions of helices with other peptide domains or small molecules occurs exclusively through side-chain interactions.

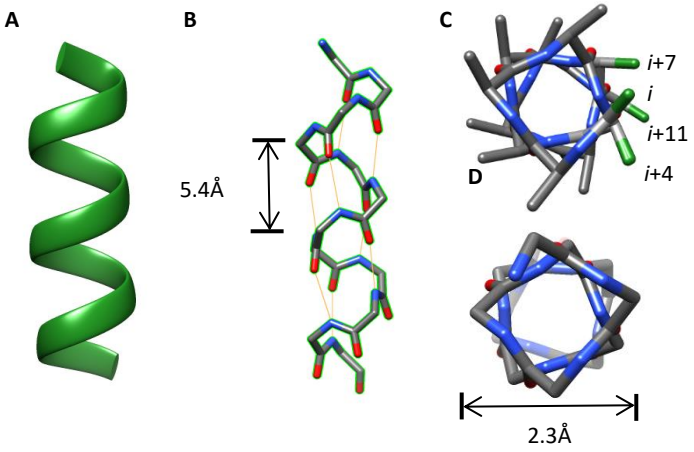


Figure 1-2: Depictions of α-helical structure. **A**: ribbon. **B**: backbone (carbon grey, nitrogen blue, oxygen red) indicating hydrogen bonding between NH of the *i* and the C=O of the *i*+4 residues. **C**: end view of α-helix with alanine side chains indicating the *i*, *i*+4, *i*+7 and *i*+11 side chains in green, **D**: end view without side chains, indicating diameter of 2.3 Å.

1.2.2 3_{10}-Helices

3_{10}-Helices are significantly less common than α-helices, however they still play important roles in native peptide structure[7]. 3_{10}-Helices are often found as an extension of α-helices, in loops, and as connectors between β-sheets. A 3_{10}-helix is tighter than the α-helix, with dihedral angles of ϕ = -74° and ψ = -4°. Whereas α-helices display hydrogen bonds between the *i* and *i*+4 residues, hydrogen bonding in 3_{10}-helices occur between the *i* and *i*+3 residues (Figure 1-3). The backbone arrangement of a 3_{10}-helix

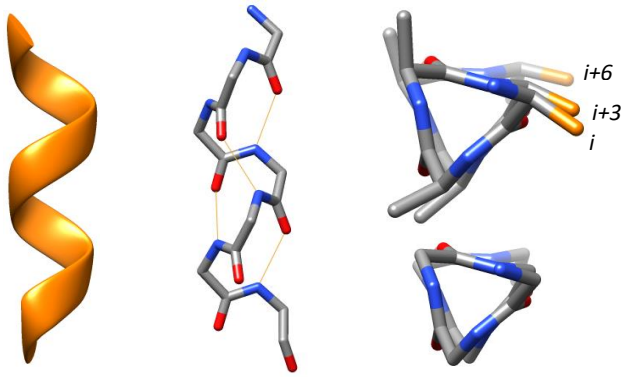


Figure 1-3: Depictions of 3_{10}-helical structure. Left: ribbon. Middle: backbone (carbon grey, nitrogen blue, oxygen red) indicating hydrogen bonding between NH of the *i* and the C=O of the *i*+3 residues. Right: end view of 3_{10}-helix (top) with alanine side chains indicating the *i*, *i*+3 and *i*+6 side chains in , (bottom) without side chains.

places side chains of the *i*, *i*+3 and *i*+6 residues on the same face of the helix (Figure 1-3).

1.2.3 β-Sheets and Turns

A β-sheet is the second most common secondary structure in proteins, and together with the α-helix account for over 60% of the predominant structure at the interface of protein-protein interactions[2]. Hydrogen bonding between β-strands results in self-

association to give β-sheet secondary structure, which adopts either parallel (Figure 1-4A) or antiparallel (Figure 1-4B) twisted pleated arrangements[8].

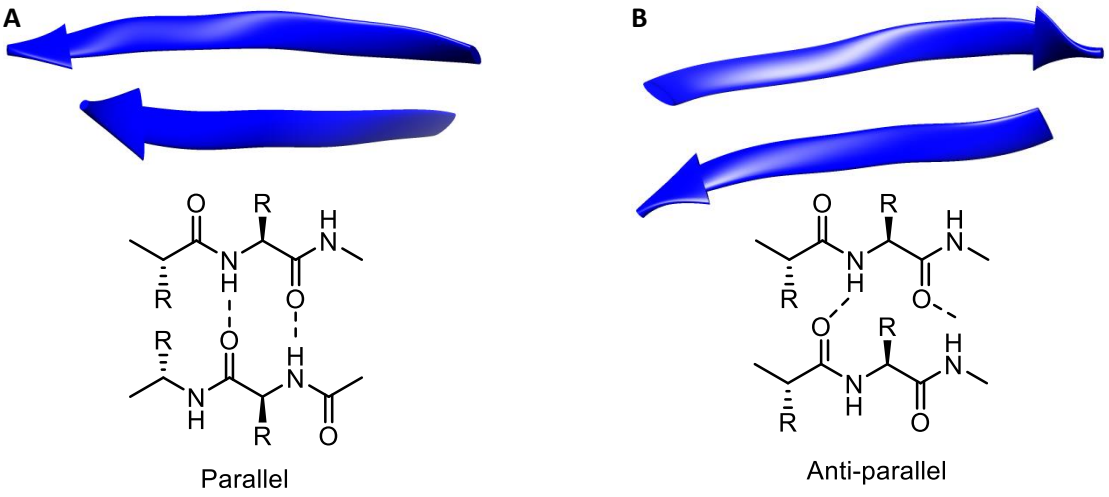


Figure 1-4: A) Ribbon representation of parallel (left, PDB 2B1L) and antiparallel (right, 2B1L) β-sheet arrangement. B) Structural arrangement of residues, indicating hydrogen bonding interactions between strands (dashed lines) in parallel (left) and antiparallel (right) β-sheet.

β-Turns represent sites where proteins reverse their overall chain direction. These structures are also known as reverse turns or hairpin turns and are the most common turn structure. These turn structures consist of four residues stabilised by hydrogen bonding between *i* and *i*+3 residues[9] (Figure 1-5). Dihedral angles ϕ and ψ of *i*+1 and *i*+2 residues are used to classify the type of β-turn.

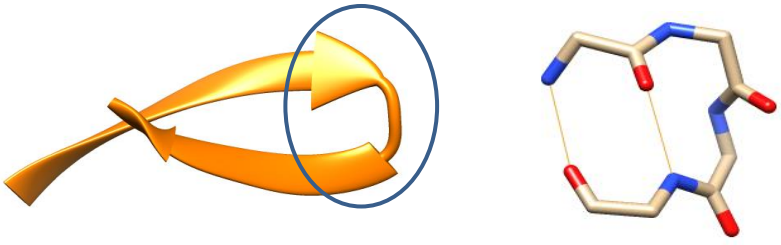


Figure 1-5: Ribbon representation of β-turn (circled, left), and structural representation (right) showing hydrogen bonding interactions (orange lines) between *i* and *i*+3 residues.

1.3 Techniques for the Characterisation of Secondary Structure

Identifying the structural elements of protein interfaces is important for designing modulators of protein-protein interactions. Many techniques can be used to identify secondary structure within peptides and proteins, either in solution or in their crystalline state. Characterisation of secondary structure in solution is often performed using circular dichroism (CD), and nuclear magnetic resonance (NMR) spectroscopy. CD spectroscopy gives an indication of the percentage of secondary structure in peptides and proteins. NMR spectroscopy allows identification of the presence, and location of secondary structure.

1.3.1 Circular Dichroism Spectroscopy

Circular dichroism (CD) spectroscopy produces characteristic spectra for different types of secondary structure and can be used to estimate the proportion of secondary structure present in a peptide when in solution. The major chromophore in a peptide or protein is the amide group, and the asymmetric secondary structures formed by polypeptides give rise to distinct CD spectra (Figure 1-6). A negative minima between 215-230 nm, commonly observed for α-helices and β-sheets, is attributed to $n \rightarrow \pi^*$ transition[10]. A positive maximum at 190 nm and negative minimum at 208 nm characterise the structure as α-helix rather than β-sheet. Absorbance's at 190 nm, and 208 nm arise from excitation splitting of the NV_1 transition. The well-defined geometry of the α-helix results in the NV_1 transition being split by the electric dipole transition moments among amides[11]. The ratio of ellipticity at θ_{208} to θ_{220} gives an indication of the type of helical structure. A θ_{208} /θ_{222} ratio of 0.9 is indicative of α-helical structure, and greater than 1 is indicative of a coiled coil structure. A 3_{10}-helix can be identified in CD spectroscopy by a θ_{208} /θ_{222} ratio of less than 0.4[7].

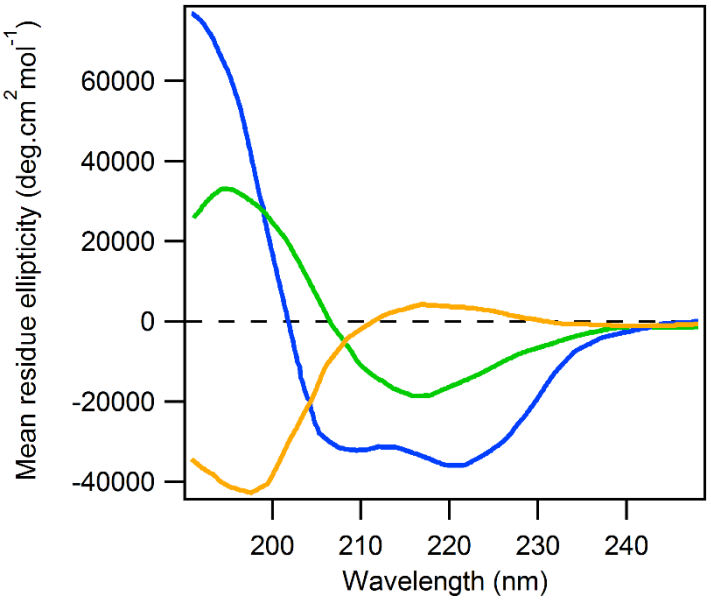


Figure 1-6: CD spectra for random coil (orange), β-sheet (green) and α-helix (blue).

1.3.2 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy allows identification of secondary structure in peptides, and small proteins (<35-40kDa) in solution. Amino acids have characteristic resonances resulting from the protons on the amine (NH), α carbon (Hα), and the side chains, identified in Figure 1-7. Scalar coupling constants and chemical shifts are commonly used to identify secondary structure within polypeptides. In α-helical structures scalar coupling between the amide and alpha protons of greater than 6 Hz ($^3J_{NHCH\alpha}$ < 6 Hz), and up-field shifts for the alpha proton (δ(Hα)) relative to random coil values are normally observed[12]. Nuclear Overhauser effect spectroscopy (NOESY) (or rotating frame Overhauser effect spectroscopy (ROESY) for short peptides) identifies short inter-proton distances (<5 Å) allowing for identification of sequential connectivity, through intra-residue NH-NH (d_{NN}) and NH-CH ($d_{\alpha N}$, $d_{\beta N}$ etc.) cross peaks of nuclei in close proximity (Figure 1-7). For helical structure, non-sequential medium range $d_{\alpha N}(i,i+4)$, $d_{\alpha N}(i,i+3)$ and $d_{\alpha\beta}(i,i+3)$ NOEs are indicative of α-helical structure (Figure 1-7). Presence of $d_{\alpha N}(i,i+4)$ versus weak $d_{\alpha N}(i,i+2)$ NOE's indicates α-helicity rather than 3_{10}-helicity or turn conformations. NOESY spectra indicate the presence of 3_{10}-helical structure by clear $d_{\alpha N}(i,i+2)$ NOE's, $d_{\alpha N}(i,i+3)$ $d_{\alpha\beta}(i,i+3)$ NOE's[13, 14].

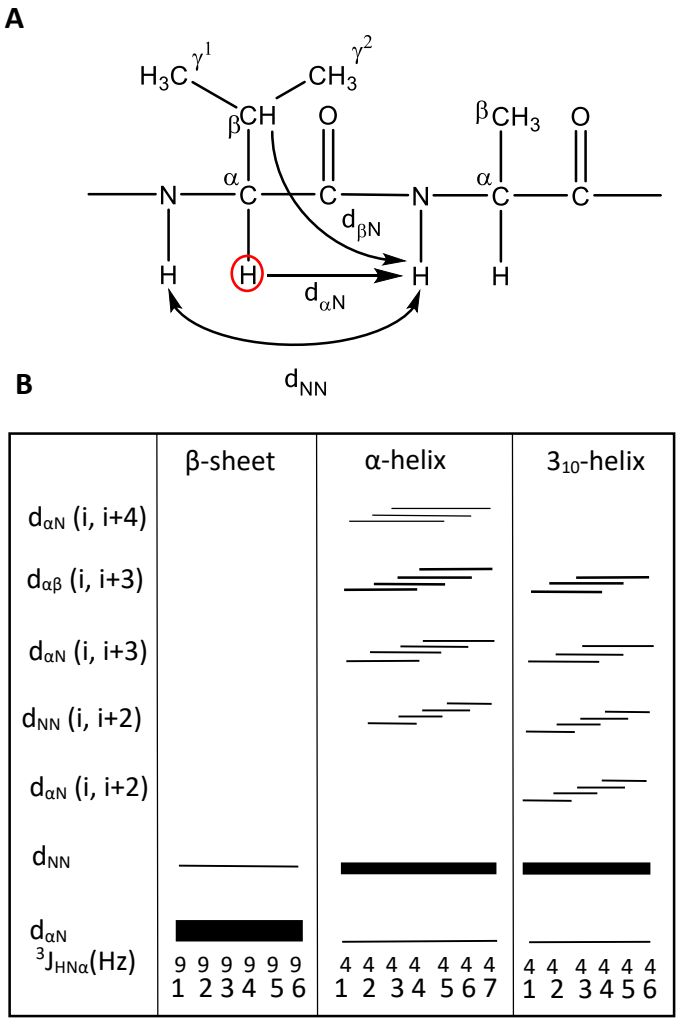


Figure 1-7: A) NOE interactions between *i* and *i*+1 residues. α, β and γ carbon are labelled, protons bound correspond to labelled carbon, i.e. Hα is highlighted by a red circle. B) Summary diagram of NOE interactions in common secondary structures.

1.4 Modulating Protein-Protein Interactions

Where proteins contain deep hydrophobic pockets for binding small-molecule substrates (Figure 1-8A), protein-protein interfaces are often very large polar surfaces, with only shallow ligand-binding hydrophobic clefts (Figure 1-8B). The competitive inhibition of protein-protein interactions thus requires a molecule to make discontinuous, non-covalent contacts over a much larger surface area and with a less defined shape compared to a conventional substrate-binding cavity[15]. In the past, many protein-protein interactions were thought to be “undruggable”, due to the challenges

associated with identifying effective approaches to develop compounds that modulate or inhibit these interactions.

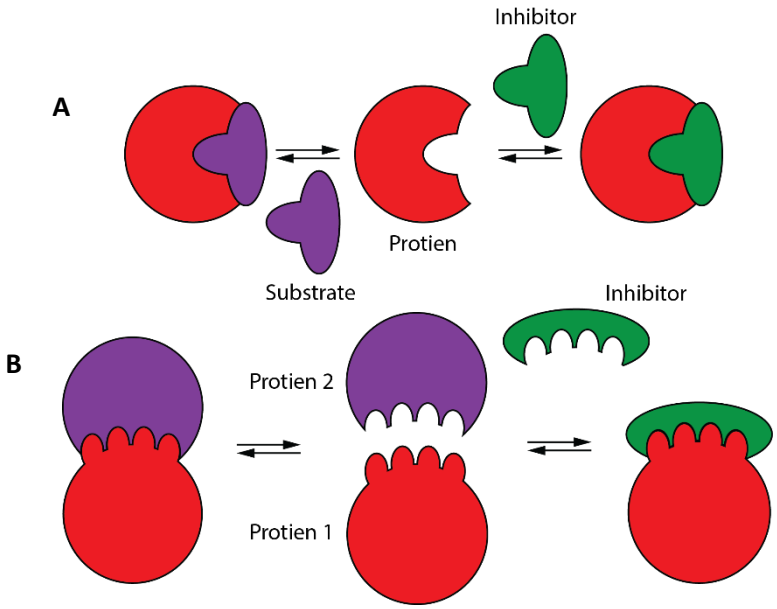


Figure 1-8: A) Schematic depiction of enzyme substrate binding and protein-protein interaction, identifying the required inhibitor design for each of these interactions. (Adapted from Wilson *et.al*[15])

Figure 1-9 displays an example of small molecules interaction with a protein binding site and a protein-protein interaction. Figure 1-9A is an image from the X-ray crystal structure of a small molecule substrate, biotinyl-5'AMP, bound to the biotin protein ligase of *Staphylococcus aureus*[16]. The natural substrate adopts a U-conformation on binding to the hydrophobic active site, which is partially enclosed by a loop structure shown in blue, protecting the substrate from the aqueous environment. In contrast Figure 1-9B is an image from the average NMR solution structure of β3 integrin cytoplasmic tail interacting with talin[17]. This interaction occurs over a large surface with few hydrophobic pockets and is characteristic of protein-protein interactions.

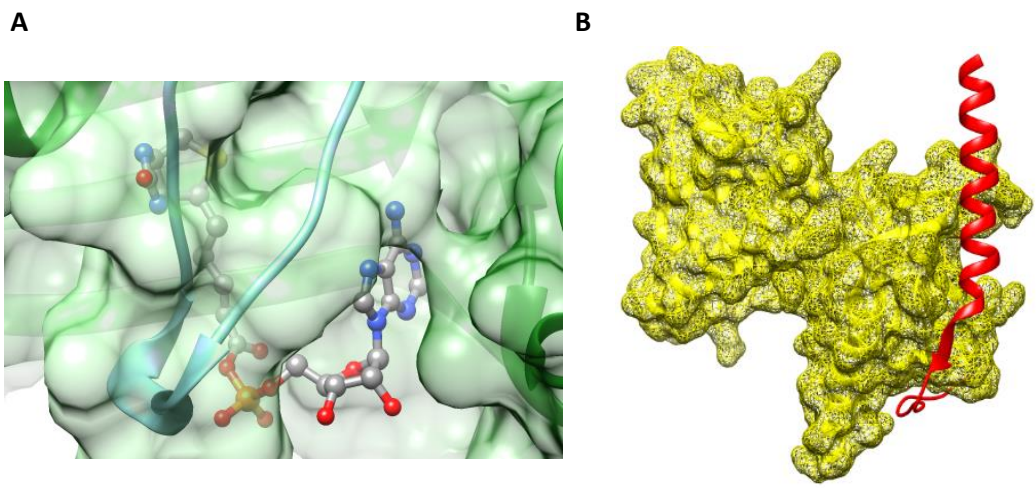


Figure 1-9: A) Biotnyl-5'-AMP (coloured by element) bound to biotin protein ligase of *Staphylococcus Aureus* (ribbon: green/blue, surface: green)[16]; B) β3 integrin cytoplasmic tail (ribbon: red) interacting with talin (surface: yellow), a critical interaction required for platelet

The past decade has seen significant progress in the development of small molecules, that obey the rule of 5[18], as inhibitors of protein-protein interactions[1,19,20]. Nutlin-2 is an example of a small molecule identified by high through-put screening. This compound competitively inhibits p53-*h*DM2 interaction (Figure 1-10A), and leads to cell cycle arrest, apoptosis and inhibition of tumour growth in mice[21]. Biologically inspired approaches have also been developed to target protein-protein interactions, e.g. therapeutic antibodies such as Efalizumab that binds to the αLβ2 integrin receptor[22] (Figure 1-10B). It is evident that small molecules, and large biologically-derived compounds can be produced to target protein-protein interactions, and compounds obtained through both these strategies have reached clinical trials. Compounds and scaffolds applicable to a broad range of topographies have been designed to target protein-protein interaction identified through screening of chemical libraries[15] and rational design. Collectively these compounds result in a heterogeneous class of compounds known as peptidomimetics[23].

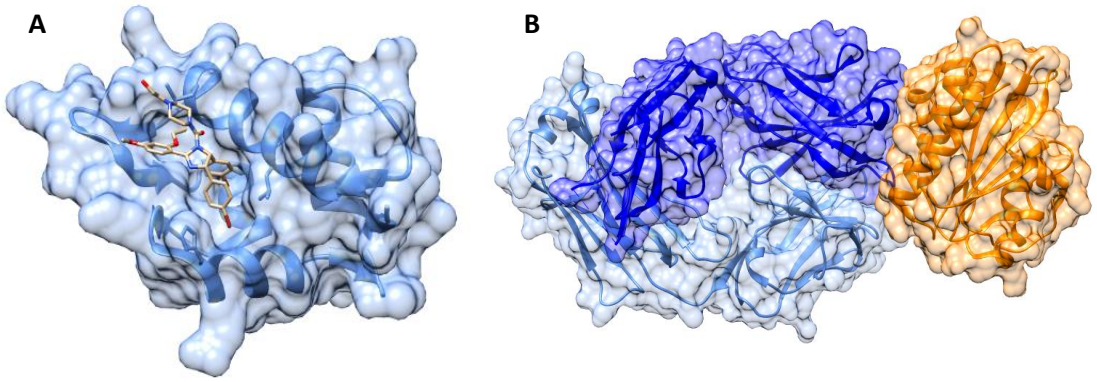


Figure 1-10: **A**) Nutlin-2 (small molecule coloured by element) bound to p53-*h*DM2 (blue)[22]; **B**) Structure of Efalizumab Fab complex (blue) with LFA-1, α_L 1 domain (orange)[23].

1.4.1 Peptidomimetics

The design of inhibitors for conventional targets are often based endogenous ligands. A similar approach can be applied to protein interfaces using a section of the protein which binds to the interface. This sequence-based approach has limitation as native peptide sequences derived from proteins are not ideal as drugs due to their poor bioavailability, and low stability under physiological conditions. Stable isolated peptides, with less than 15 amino acids, rarely adopt defined conformations in isolation, and loss of structure diminishes the ability of the peptide to bind the target interface.

Protein-protein interactions are also critically dependent on a few residues at the interface, which make a dominant contribution to the free energy of binding. The key residues that contribute the majority of the binding affinity are known as 'hot-spot' residues. The concept of hot-spot residues was first identified, by Wells and Clackson, in the course of alanine mutagenesis studies involving the interaction of human growth hormone with its receptor[24].

Approaches to broadly target protein-protein interactions have endeavoured to overcome the issues associated with native peptides by stabilising peptide structure, i) using synthetic modification within the peptide sequence such as side chain linkers (1.5), ii) and by developing scaffolds mimicking secondary structure topology, to which 'hot spot' residues can be inserted. For example, cyclotides are backbone cyclised

peptides containing a cysteine knot motif built from six conserved cysteine residues (Figure 1-11[25]). The high sequence variability, stability, and cell penetrating properties of cyclotides make them potential scaffolds, to be used to graft known active peptides[26].

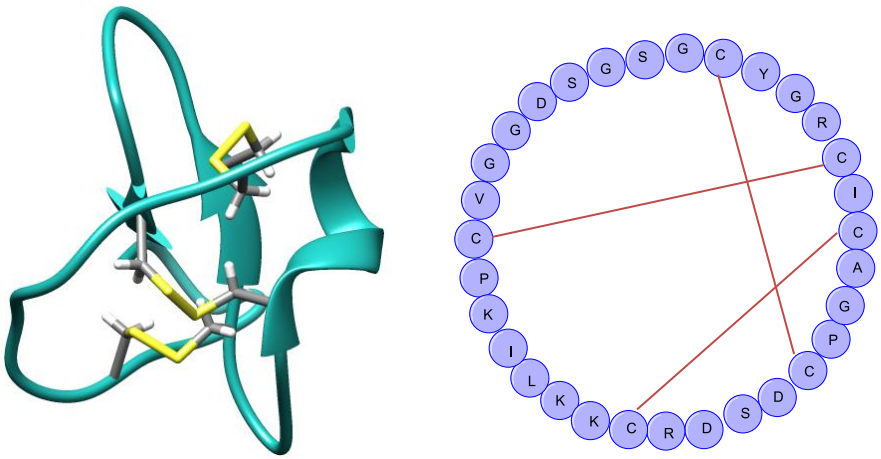


Figure 1-11: Left: NMR solution structure of circular protein, MCoTI-II, a cyclotide, a novel trypsin inhibitor isolated from *Momordica cochinchinensis*, a member of the *Cucurbitaceae* plant family[25]. Right: Schematic of cyclic peptide indicating links between cysteine residues which result in a cysteine knot structure.

1.5 Helix Stabilisation: Side Chain Covalently Constrained α-Helices

The average helical domain within a protein spans only 8-12 residues and as such, short helices can be developed to participate in selective interactions with biomolecules. However, such peptides rarely retain their structural conformation once isolated from the native sequence, and therefore their associated binding affinity for a target is often diminished. The pre-organisation of peptide residues can restore and even enhance the binding affinity relative to the native interaction. The development of numerous mimics of this important structural topology has been inspired by the prevalence of α-helical secondary structure in peptides and proteins. Approaches used to generate helical secondary structure topology can be divided into three general categories: helix stabilisation, helix foldamers, and helical surface mimetics[27-33]. Strategies for helix stabilisation are detailed below.

A highly promising approach for helix stabilisation is the incorporation of a linker between amino acids designed to nucleate α-helical structure. Side chain constraints are obtained by covalently linking the side chains of amino acids on the same face of the helix, i.e. *i* and *i*+4, or *i* and *i*+7. This stabilises α-helical motifs in biologically important polypeptides[34]. Such constrained α-helical peptides have been shown to enhance resistance to proteolysis, improve affinity for the target protein interface, and increase cell permeability relative to their unconstrained counterpart [35, 36].

There are numerous methodologies for covalently linking side chains to give a helical side chain constrained peptide[37, 38]. The most extensively investigated constraint strategies include lactam bridges[39] and all-hydrocarbon constraints[40]. Developed more recently, triazole[41-43] and perfluorobenzyl[36] constraint strategies have shown promise in nucleating helicity and enhancing cell permeability respectively (Figure 1-12).

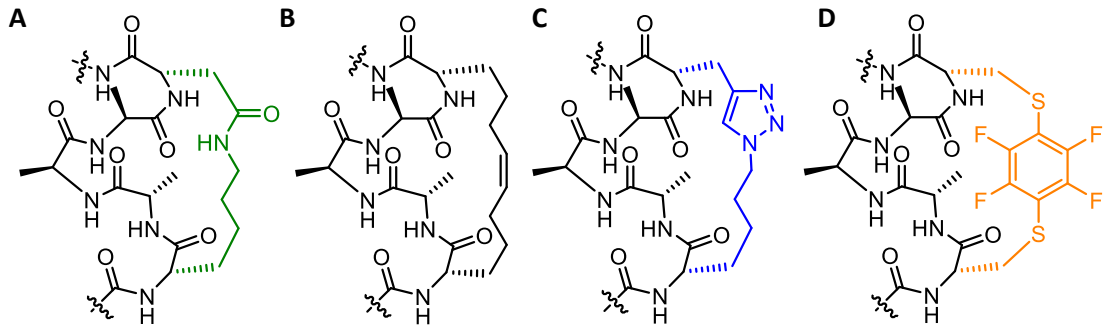


Figure 1-12: Structures indicating backbone constrained into helical conformation by A: lactam linker, B: all-hydrocarbon linker, C: triazole linker, D: perfluorobenzyl linker.

1.5.1 Lactam Bridge Constraints

Lactam bridge linkers have received much attention since first being reported to stabilise α-helical structure in a segment of human growth hormone releasing factor (GRF), by Felix and co-workers in 1988[44]. A lactam bridge linking *i* and *i*+4 residues was introduced by coupling the side chains of lysine and aspartic acid. The peptide analogues with lactam bridge linkers were found to have increased helical content compared to the unconstrained peptide analogue. These analogues were also shown to be highly potent, potentially due to stabilisation of a preferred bioactive conformation. Optimisation studies of lactam bridge linkers investigated the linker length, and amide

position within the bridge for inducing helicity [39,45]. For linkers between *i* and *i*+4 residues, a 7 atom linker was identified as optimal for inducing α-helicity. This linker was formed by coupling of lysine and aspartic acid side-chains. Shorter linkers destabilised α-helical structure, while longer linkers only stabilised helical secondary structures in some peptides. The order in which lysine and aspartic acid are incorporated into the peptide sequence also influences helical content. The linkers formed between lysine in the i^{th} position, and aspartic acid in the i+4th position (KD linker, Figure 1-13)[39] produced greater helix stabilisation than the alternative linker between aspartic acid (*i*) and lysine (*i*+4) (DK linker), rationalised to be due to a sterically preferred positioning of the amide bond in relationship to the peptide backbone. Lactam bridge linkers produced between glutamic acid and lysine were optimised with glutamic acid in the i^{th} position and lysine in i+4th position (EK)[45].

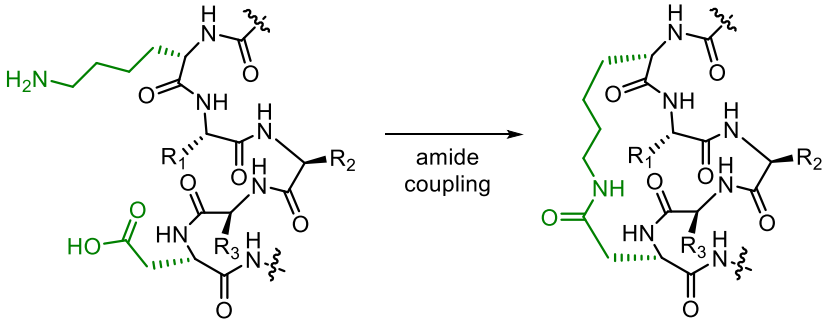


Figure 1-13: Schematic of the formation of a KD lactam bridge linker. The lactam bridge is formed using an amide coupling reaction, linking lysine an aspartic acid side chains of residues incorporated in *i*, and *i*+4 positions respectively.

Lactam bridge linkers have been reported to increase the stability of glucagon-like peptide-1 (GLP-1) to enzyme degradation[46], and have been used to determine the preferred structural conformation of secretin analogues in the development of effective receptor antagonists for the secretin receptor[47]. Modified peptides containing two and three non-overlapping lactam bridge linkers displayed increased helicity, while retaining significant flexibility. The introduction of overlapping lactam bridges has resulted in stable, regular α-helical structures that were highly resistant to thermal degradation.

1.5.2 Constraints introduced by Ring Closing Metathesis

Blackwell and Grubbs[48] were the first to design helix inducing linkers to peptides using olefin ring closing metathesis (RCM). Peptides were designed with two *O*-allyl serine residues, or homo *O*-allyl serine residues in the *i* and *i*+4 positions of a short hydrophobic peptide sequence[48]. CD of the resulting peptides suggest neither linker enhanced α-helical content, however NMR and crystallography identified a significant increase in rigidity of the native structure[49]. Based on this work, Verdine and co-workers[50] developed an all-hydrocarbon linker produced via RCM of α-methylated amino acids bearing olefin side-chains. An 8 atom linker between *i* and *i*+4 residues with *S* configuration (Figure 1-14) was highly effective at increasing α-helical content of peptides derived from Bcl-2 homology domain[51]. Additionally, incorporation of the all-hydrocarbon linker increased cellular uptake of peptides through an energy dependent endocytosis mechanism[51, 52].

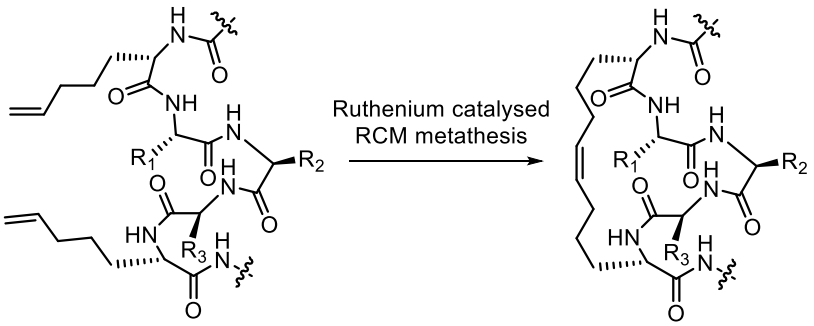


Figure 1-14: Schematic of the formation of an all-hydrocarbon linker using ring closing metathesis, linking *i*, and *i*+4 residues containing olefin groups.

All-hydrocarbon linkers have been used to enhance the binding of peptides that are known to modulate the activity of several biological targets involved in a variety of human diseases, including cancer[51, 53-60], diabetes[61], human immunodeficiency virus (HIV)[35, 62-64], and atherosclerosis[65]. Notably, a peptide derived from the transcriptional co-activator protein, incorporating an all-hydrocarbon linker, directly disrupted formation of the Notch transcription factor complex[59], inhibiting leukemic progression in a murine model of T-cell acute lymphoblastic leukemia. Baek and co-workers reported a peptide (SAH-p53-8) with an all-hydrocarbon linker that interacted with the

protein surface, augmenting the binding interface[66] (Figure 1-15). This study introduced the possibility of designing peptide linkers, which can interact with the target protein to improve binding affinity.

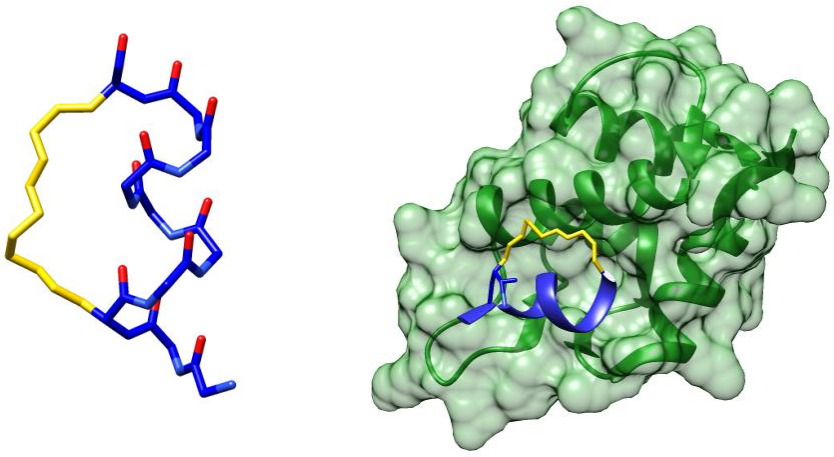


Figure 1-15: Left: X-ray crystal structure of all-hydrocarbon linked peptide, SAH-p53-8. Hydrocarbon linker in yellow, backbone in blue, oxygen of carbonyl groups displayed in red. (Right) X-ray crystal structure of SAH-p53-8 bound to E3 ubiquitin protein ligase[63].

1.5.3 Triazole Linker

An α-helical turn was shown to be stabilised in cyclopeptides containing 1,4-disubstituted [1,2,3] triazolyl-linkers generated though copper(I)-mediated Huisgen 1,3-dipolar cycloaddition (the 'click' reaction) reactions[67]. Cantel et al. accessed these triazole linkers by reaction of the side chain azido and alkynyl moieties of azido-L-norleucine [L-Nle (εN_3)] and L-propargylglycine (L-Pra) in *i* and *i*+4 positions (Figure 1-16). Triazole constraints were used to produce analogues of the peptide inhibitor for parathyroid hormone-related peptide[67]. CD and NMR studies revealed that the triazole linked peptides displayed some variation in backbone structure when compared to that of the lactam bridged derivative, however both linkers organised residues within the helical turn in the same spatial orientation. Scrima[41] and subsequently Kawamoto[43] optimised triazole linkers between *i* and *i*+4 residues. Both studies revealed an optimum linker length of 8 atoms (a total of 5 CH_2 flanking the triazole ring) for inducing helicity. Azido-L-norleucine [L-Nle(εN_3)] and D-propargylglycine (D-Pra) at *i* and *i*+4 positions were found to produce the optimum linkers for enhancing helicity and affinity of analogues

derived from a helical segment of B-cell CLL/lymphoma 9 (BCL9) for β-catenin[43]. Linkers derived from D-propargylglycine (D-Pra) were only marginally better than the L-propargylglycine (L-Pra), and subsequent analysis of alanine rich pentapeptides, suggest that α-helicity was enhanced to a greater extent in peptides with linkers derived from L -propargylglycine (L-Pra)[68], and there is variation in the resulting helical structured of these linked peptides, which may be dependent on the sequence of the peptide.

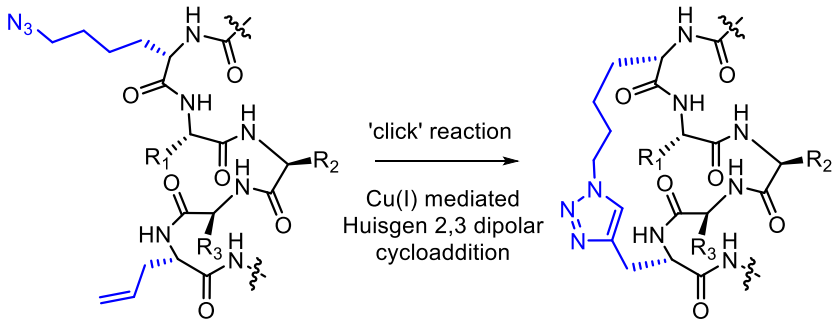


Figure 1-16: Schematic of the formation of a triazole linker using 'click' reaction, linking L-norleucine (εN_3), and L-propargylglycine residues *i*, and *i*+4 positions.

Kawamoto also developed peptides containing two triazole bridges[43]. The peptides with two triazole bridges had a higher percentage of helical structure in CD studies, compared to peptides with a single constraint. The peptides were also highly potent for targeting β-catenin, a protein that plays an important role in several types of cancer. Incorporation of two linkers also improved protease stability relative to the linear peptide sequence. Kawamoto proposed that the generation of peptides containing two triazole linkers may be possible in scenarios where incorporating two hydrocarbon linkers could lead to significant amounts of mixed linkers (various combinations of side chains forming linkers)[43]. Jacobsen and co-workers applied the triazole side chain constraint strategy to stabilising 3_{10}-helical structure[42]. Constraints were formed between side chains of azido-L-norvaline and *O*-propynyl-L-serine in the *i* and *i*+3 positions for an aminoisobutyric acid (Aib) rich sequence[42]. NMR, and X-ray diffraction were used to analyse the peptide structure, and results suggest the constrained peptide adopted a more 'ideal' 3_{10}-helix.

1.5.4 Cysteine Perfluoroarylation

Spokoyny and co-workers developed the cysteine arylation platform as a mild synthetic method for cyclisation of unprotected peptides with cysteine residues in *i* and *i*+4 positions[36]. The linker is introduced by nucleophilic aromatic substitution of hexafluorobenzene (Figure 1-17) and decafluorobiphenyl species, to give two different perfluoroaryl linkers. These perfluoroaryl links were incorporated into short peptides derived from the glycosaminoglycan (Gag) polyprotein, which is involved in HIV. The cellular uptake of the perfluoroaryl linked peptide analogues was comparable to that of the all-hydrocarbon bridge analogue previously investigated[35]. CD of both single ring and double ring species were complicated as the linkers gave rise to an absorbance maximum at 222 nm. α-Helical content was estimated by subtracting the absorbance of the cysteine disubstituted ring species from CD spectra of the constrained peptides. The single ring species had an estimated increase in helicity of 40%, while the double ring species displayed an increase of 20%.

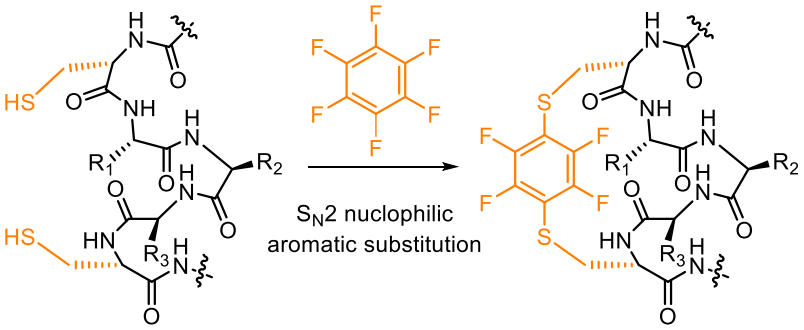


Figure 1-17: Schematic of the formation of a perfluorobenzyl linker using nucleophilic aromatic substitution reaction, linking cysteine residues *i*, and *i*+4 position by reaction with hexafluorobenzene.

1.6 Chemical Synthesis of Peptides

The original approach to peptide synthesis was to carry out reactions in solution[69]. After each synthetic step, the product needed to be isolated and purified. The solution phase peptide synthesis strategy is therefore only amenable to short peptide sequences. Convergent strategies were developed linking blocks of amino acids together to access longer peptide sequences. The development of solid phase peptide synthesis (SPPS) was the breakthrough that enabled widespread access to synthetic peptide sequences,

without the need for extended isolation or purification procedures after each step. Comparison of solution phase and solid phase techniques, in Table 1-1, briefly summarises the benefits and disadvantages of each method.

Table 1-1: Comparison of chemical synthetic methods of peptide synthesis, solution phase peptide synthesis and solid phase peptide synthesis. Adapted from Kent[70].

	Method	
	Solution Phase	**Solid Phase**
No. amino acids	Less than 10	10-40
Strategy	Segment/convergent	Stepwise
Cost of raw material	Low	Higher
Solubility problems	Frequent/serious	Occasional/minor
Work-up required	Maximal	Minimal
Toxic solvent waste	Moderate	High
Chemistry	Highly specialised	Simple/general
Automation	No	Yes
Purity	Excellent	Acceptable

1.6.1 Solid Phase Peptide Synthesis

Solid phase peptide synthesis (SPPS) involves sequential addition of amino acids to an insoluble polymeric support (resin). The peptide chain is grown from *C*-terminus to *N*-terminus by addition of appropriately protected amino acids. Amino acids are temporarily protected with *N*α-protecting groups. These groups are removed once the amino acid is linked to the resin, so addition of the subsequent amino acid can occur. Functional groups of reactive side chains are protected with 'semi-permanent' protecting groups, which generally remain through the entire synthesis, and are usually removed during or after the peptide's cleavage from the resin (Figure 1-18).

There are two common *N*α-protecting group strategies employed during solid phase peptide synthesis, acid labile tert-butyloxycarbonyl (Boc) protecting group strategy, or the base labile 9-fluorenylmethoxycarbonyl (Fmoc) protecting group strategy. The Fmoc

protecting group strategy reduces issues associated with sequential acid treatments, while also requiring milder conditions for the removal of the peptide from the solid support. The Fmoc protecting group strategy is often favoured over the Boc protecting group strategy, due to milder conditions for peptide removal, and the Boc strategy is generally reserved for the synthesis of non-natural peptide analogues which are base sensitive.

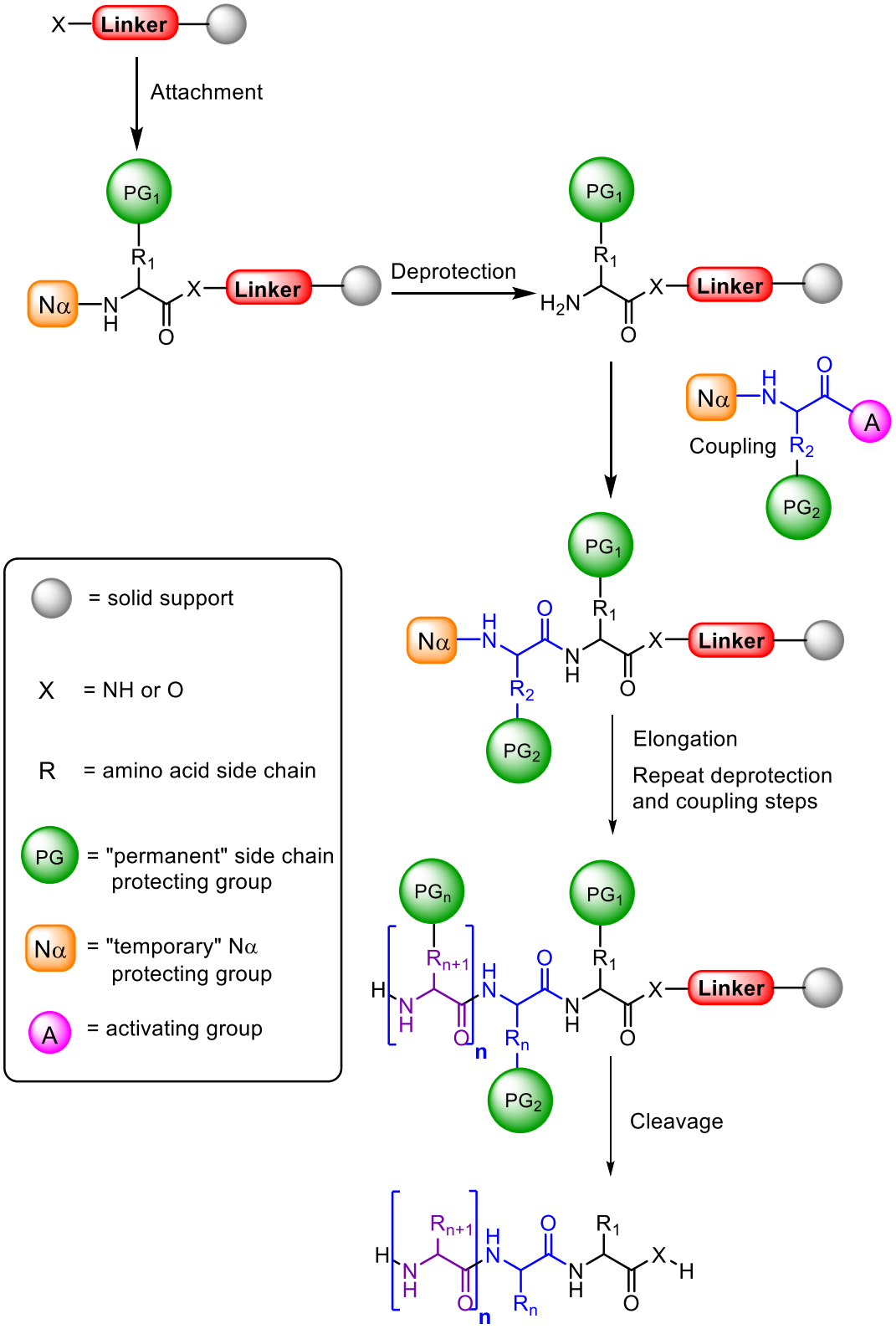


Figure 1-18: General schematic showing the steps involved in solid phase peptide synthesis. Peptides are assembled on a solid support by attachment through the linker which makes up the resin. Repetition of coupling and deprotection steps result in elongation of the peptide chain. Once the peptide is complete it is cleaved from the resin. During cleavage the side chain, protecting groups are often removed.

1.6.2 9-Fluorenylmethoxycarbonyl (Fmoc) Protecting Group Strategy

Solid phase peptide synthesis using the Fmoc protecting group strategy involves the assembly of amino acids with Fmoc *N*α-protection, and acid labile side chain protecting groups[71]. The process of peptide assembly involves two major steps, coupling and deprotection (Figure 1-18). The Fmoc group is removed by addition of piperidine to give a free amine. Resin tests can be utilised to establish that the reactions involved in coupling and de-protection steps have reached completion[72-74]. Once the full length peptide has been synthesised it is cleaved from the resin, the conditions required for cleavage often result in the side chain protecting groups being removed. Solid phase peptide synthesis can be performed manually or using automation. The development of automated microwave peptide synthesis has allowed rapid assembly of peptide sequences using heating by microwave radiation, to improve the efficiency of coupling and deprotection steps[75].

1.7 Defining a relevant target

1.7.1 Coronary Heart Disease

Coronary heart disease accounts for the largest single cause of death in the world[76], and as such, new tools for treatment of heart disease are in high demand. Improved treatments for disease can reduce fatalities and improve quality of life. The application of methodologies that target protein-protein interactions provides the potential to target novel pathways.

Coronary heart disease presents in two clinical forms, myocardial infarction, and angina. Acute myocardial infarction (a heart attack) and angina are caused by obstructions to the flow of blood to the heart, usually caused by coronary artery thrombosis. Thrombosis formation results from a disease known as atherosclerosis, which is the chronic inflammation of arteries, developed over decades in response to the biological effects of risk factors. As a result of atherosclerosis, plaques are formed within arteries, which can limit blood flow (causing stable angina). When physical disruption (rupture) of the plaque occurs, this exposes pro-coagulant material within the core of the plaque to coagulation proteins on platelets, triggering thrombosis[77].

1.7.1.1 The Role of Platelets

Platelets play a central role in the coagulation cascade, recognised to potentiate clot formation and inflammation[78]. Platelets circulate through the vessels in an inactive state, which is supported by healthy endothelium. Rupture or erosion of the fibrous cap, which covers plaques, results in platelet activation and aggregation, the formation of fibrin rich thrombi, and the partial or total occlusion of the coronary artery. Plaque rupture or erosion results in the release of tissue factor and exposure of sub-endothelial collagen and von Willebrand factor (vWF). Circulating platelets adhere to the exposed collagen and vWF and undergo cytoskeletal rearrangement, generation of cyclooxygenase I (COX-I) dependent thromboxane A_2 (TxA_2), and release the contents of stored granules including adenosine diphosphate (ADP). Interaction of thromboxane A_2 with the thromboxane A_2 receptor and ADP with its receptor (P2Y12) initiates positive feedback activating additional platelets in a regional manner, and promoting the expression of additional αIIbβ3 integrin receptors which bind fibrinogen crosslinking platelets[79] (Figure 1-19).

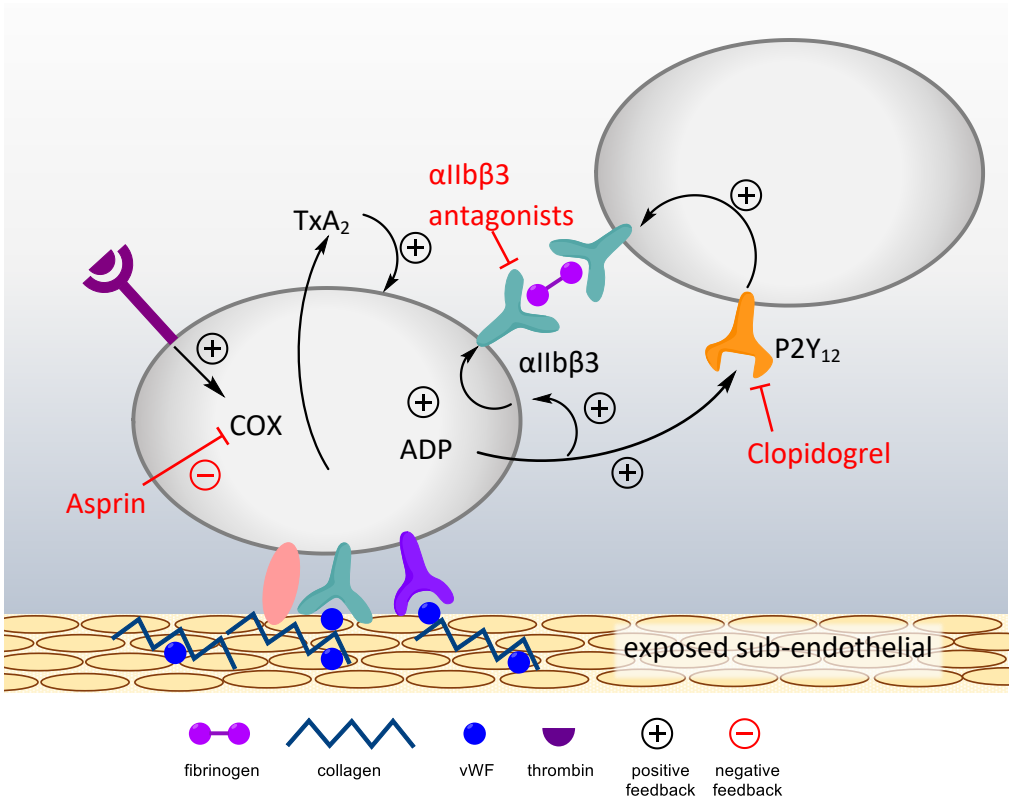


Figure 1-19: Platelet activation, indicating key receptors, their substrates, and key antiplatelet therapeutics which target platelet receptors (adapted from Lilly et.al[74]).

Current therapeutics target key pathways of platelet activation include preventing thromboxane A_2 synthesis, ADP-mediated signalling, and integrin αIIbβ3 activation. The most common antiplatelet agents include aspirin, clopidogrel and integrin αIIbβ3 antagonists[80] (Figure 1-19). Patients have variable responses to all currently available therapeutics, consequently novel antiplatelet agents are being sought which over-come issues of currently available medication[81-83], with bleeding (often gastrointestinal bleeding) potentially the most significant side effect[84].

1.7.1.2 Targeting αIIbβ3 Integrin

Modulation of integrin αIIbβ3 activity using an antagonist is an extremely promising approach, due to the key role of αIIbβ3 in platelet aggregation, particularly as integrin αIIbβ3 is specific to platelets. Inhibitors of αIIbβ3 integrin approved by the FDA include an antibody fragment (abciximab), a cyclic peptide (eptifibatide) and a small molecule (tirofiban), which are all administered intravenously. Abciximab (c7E3 Fab) is a chimeric human-murine monoclonal antibody Fab (fragment antigen binding) fragment[85]. Eptifibatide is a cyclic heptapeptide containing mercaptopropionyl (MPA) and six amino acid residues (MPA, hArg, Arg, Gly, Asp, Trp, Pro and Cys)[86]. The peptide is cyclised by formation of an intra-chain disulfide bond between the cysteine and the mercaptopropionyl moieties (Figure 1-20A). Tirofiban is a tyrosine derived small molecule Arg-Gly-Asp (RGD) mimetic. Many integrins recognise the RGD sequences in their ligands[87] (Figure 1-20B). The uses of current pharmacological inhibitors targeting αIIbβ3 integrin have unfortunately been limited, due to an associated increased likelihood of severe bleeding[88, 89], and a lack of oral bioavailability. Attempts to develop oral antagonists for more convenient administration have been unsuccessful[90]. Never-the-less, αIIbβ3 integrin antagonists have important benefits in the treatment of coronary heart disease, and are incorporated in treatment plans of many patients immediately after surgery[85].

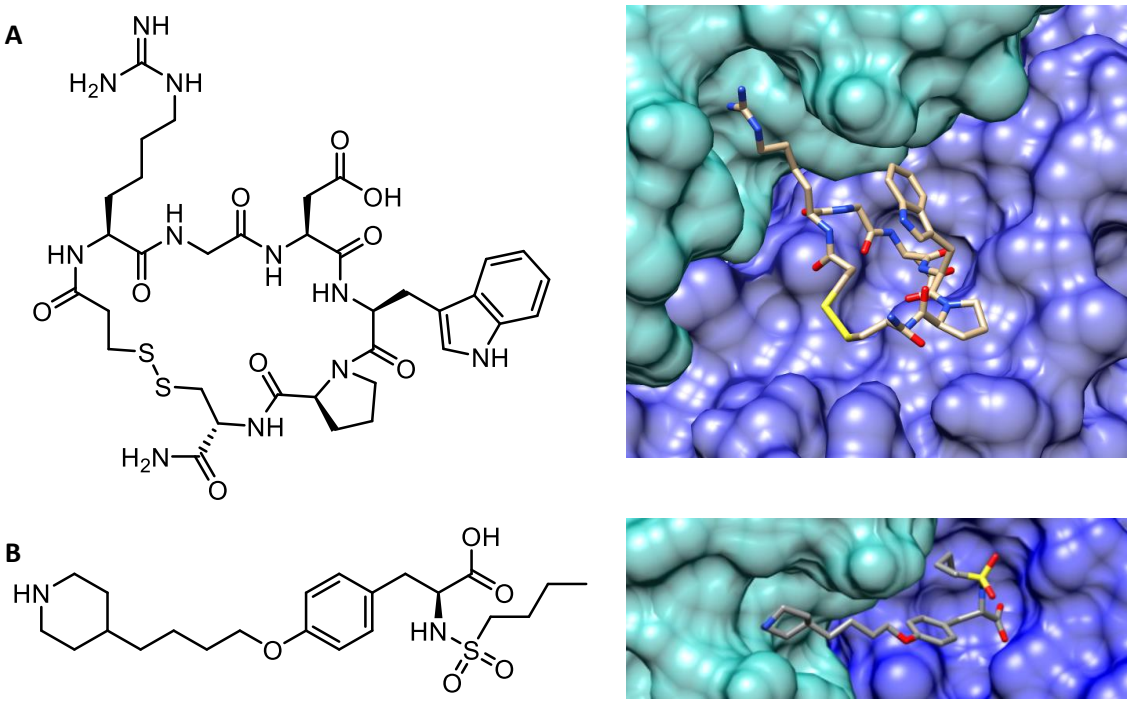


Figure 1-20: A) Left, the structure of Eptifibatide, is a cyclic heptapeptide. Right: Eptifibatide bound to extracellular pocket of αIIbβ3 integrin. B) left, the structure of Tirofiban, a small molecule based on the Arg-Gly-Asp (RGD) sequence. Right: Tirofiban bound to extracellular pocket of αIIbβ3 integrin[83].

Structurally, integrins are made up of two subunits, α and β, each of which consist of a short cytoplasmic tail, a single transmembrane helix, and a large extracellular domain[91] (Figure 1-21). The large extracellular domain binds large protein ligands that can bridge platelets together (fibrinogen and von Willebrand factor), or modulate aggregation (vitronetin and fibronectin)[92]. Integrins are expressed on the surface of cells in a low affinity binding conformation, or inactive state. Activation of the integrin complex involves the transition from a low to a high affinity state for binding ligands. Once activated, the αIIbβ3 integrin extracellular domains extend and attach to circulating proteins, thereby initiating aggregation. Integrins are true receptors, involved in bidirectional signalling across the plasma membrane[92, 93]. Ligand binding to the extracellular domains is transduced to the cytoplasm to regulate intracellular activities, a process referred to as outside-in signalling. On the other hand, inside-out activation[94], is stimulated within the cytoplasm when cells are activated, signalling integrin activation.

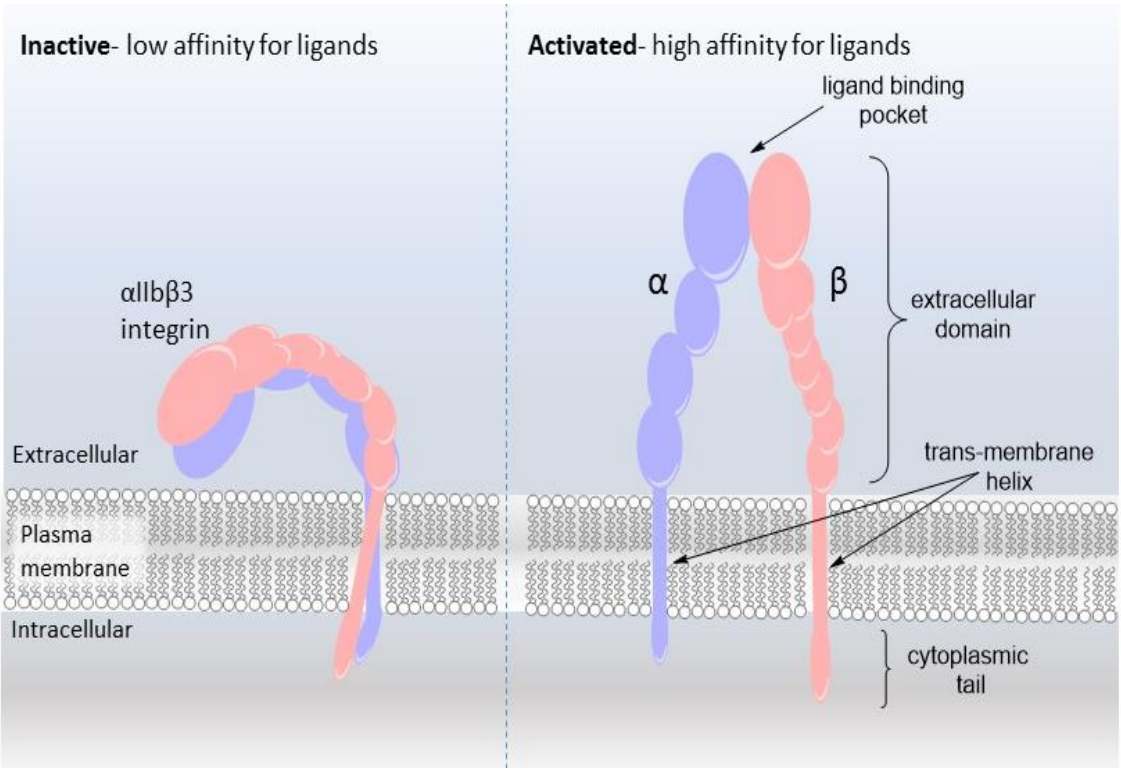


Figure 1-21: Schematic depiction of the αIIbβ3 integrin membrane-spanning complex. Left: displays the receptor in its inactive conformation: Right: displays the activated state of αIIbβ3 integrin indicating the α and β subunits, short cytoplasmic tail, a single transmembrane helix and a large extracellular domain. (Adapted from Shattil *et al*.[91])

Alternative methods for modulating αIIbβ3 integrin have the potential to overcome the issues associated with current therapeutics. All current pharmacologic inhibitors of αIIbβ3 integrin target the extracellular binding pocket to impaired platelet aggregation, but these compounds are associated with increased likelihood of severe bleeding[88, 89], and lack oral bioavailability. The function of the αIIbβ3 integrin in platelet aggregation is dependent on its ability to undergo activation[95] (Figure 1-21). Alternate methods of targeting αIIbβ3 integrin focusing on preventing activation, rather than directly binding to the extracellular pocket, may result in therapeutics which overcome issues with current drugs[96]. Investigation of modulators of the protein-protein interactions involved in αIIbβ3 integrin activation is therefore a promising avenue for the development of probes and possible novel therapeutics targeting platelet aggregation.

1.8 Overview of Thesis

This thesis investigates the stabilisation of an α-helical conformation in peptides derived from the cytoplasmic tail of β3 integrin. Covalent side chain linkers, incorporated through linking side chains of *i* and *i*+4 residues, are used to stabilise α-helical conformation. Peptides are designed to prevent the activation of αIIbβ3 integrin through competitive binding to talin F3 domain. The synthetic peptides are prepared by solid phase peptide synthesis. The binding of peptide derivatives of the β3 integrin tail is investigated to determine the effect of linker location on affinity, and to determine the ideal location for linker incorporation. The effect of linker type on peptide helicity and cellular uptake is also studied to identify the effect of polarity and overall charge of peptides entering the cell.

Chapter 2 is concerned with the design and synthesis of a panel of lactam linked peptides, which target talin by binding at the integrin-talin binding interface. The identification of hot spot residues from computational analysis is used to identify the integrin target sequence. The effect of linker location on helicity and structural characteristics is investigated using circular dichroism (CD) spectroscopy and nuclear magnetic resonance (NMR) spectroscopy. NMR titrations to determine binding affinity for talin will then be compared with structural information to determine the requirements determining the linker location to achieve the greatest binding affinity.

Chapter 3 investigates the effect of linker type on the conformation and cellular uptake properties of linked peptides. The α-helical content of peptides derived from the β3 integrin cytoplasmic tail with lactam, all-hydrocarbon and triazole linkers incorporated are to be compared using CD spectroscopy. Cellular uptake of these peptides are determined to facilitate direct comparison between different constraint strategies. The structural characteristics and cellular uptake of a perfluoroaryl linker will also be investigated using CD spectroscopy and confocal microscopy. Additionally, the effect of overall negative, neutral and positive charge of peptides on cellular uptake using all-hydrocarbon linked peptides is investigated.

1.9 References

1. Wells, J. A.; McClendon, C. L., Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007,** *450* (7172), 1001-1009.
2. Guharoy, M.; Chakrabarti, P., Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein–protein interactions. *Bioinformatics* **2007,** *23* (15), 1909-1918.
3. Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V., Stereochemistry of polypeptide chain configurations. *Journal of Molecular Biology* **1963,** *7* (1), 95-99.
4. Pauling, L.; Corey, R. B.; Branson, H. R., The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences* **1951,** *37* (4), 205-211.
5. Otaki, J. M.; Tsutsumi, M.; Gotoh, T.; Yamamoto, H., Secondary Structure Characterization Based on Amino Acid Composition and Availability in Proteins. *Journal of Chemical Information and Modeling* **2010,** *50* (4), 690-700.
6. Barlow, D. J.; Thornton, J. M., Helix geometry in proteins. *Journal of Molecular Biology* **1988,** *201* (3), 601-619.
7. Toniolo, C.; Polese, A.; Formaggio, F.; Crisma, M.; Kamphuis, J., Circular Dichroism Spectrum of a Peptide 310-Helix. *Journal of the American Chemical Society* **1996,** *118* (11), 2744-2745.
8. Eisenberg, D., The discovery of the α-helix and β-sheet, the principal structural features of proteins. *Proceedings of the National Academy of Sciences of the United States of America* **2003,** *100* (20), 11207-11210.
9. Venkatachalam, C. M., Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. *Biopolymers* **1968,** *6* (10), 1425-1436.
10. Woody, R. W.; Koslowski, A., Recent developments in the electronic spectroscopy of amides and α-helical polypeptides. *Biophysical Chemistry* **2002,** *101–102* (0), 535-551.
11. Moffitt, W., Optical Rotatory Dispersion of Helical Polymers. *The Journal of Chemical Physics* **1956,** *25* (3), 467-478.
12. Van Holde, K. E.; Johnson, W. C.; Ho, P. S., *Principles of Physical Biochemistry*. Pearson/Prentice Hall: 2006.
13. Merutka, G.; Morikis, D.; Bruschweiler, R.; Wright, P. E., NMR evidence for multiple conformations in a highly helical model peptide. *Biochemistry* **1993,** *32* (48), 13089-13097.
14. Gratias, R.; Konat, R.; Kessler, H.; Crisma, M.; Valle, G.; Polese, A.; Formaggio, F.; Toniolo, C.; Broxterman, Q. B.; Kamphuis, J., First Step Toward the Quantitative Identification of Peptide 310-Helix Conformation with NMR Spectroscopy: NMR and X-ray Diffraction Structural Analysis of a Fully-Developed 310-Helical Peptide Standard. *Journal of the American Chemical Society* **1998,** *120* (19), 4763-4770.
15. J. Wilson, A., Inhibition of protein-protein interactions using designed molecules. *Chemical Society Reviews* **2009,** *38* (12), 3289-3300.
16. Soares da Costa, T. P.; Tieu, W.; Yap, M. Y.; Pendini, N. R.; Polyak, S. W.; Sejer Pedersen, D.; Morona, R.; Turnidge, J. D.; Wallace, J. C.; Wilce, M. C.; Booker, G. W.; Abell, A. D., Selective inhibition of biotin protein ligase from Staphylococcus aureus. *J Biol Chem* **2012,** *287* (21), 17823-32.

17. Wegener, K. L.; Partridge, A. W.; Han, J.; Pickford, A. R.; Liddington, R. C.; Ginsberg, M. H.; Campbell, I. D., Structural basis of integrin activation by talin. *Cell* **2007,** *128* (1), 171-82.
18. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J., Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1. *Advanced Drug Delivery Reviews* **2001,** *46* (1–3), 3-26.
19. Berg, T., Modulation of Protein–Protein Interactions with Small Organic Molecules. *Angewandte Chemie International Edition* **2003,** *42* (22), 2462-2481.
20. Surade, S.; Blundell, Tom L., Structural Biology and Drug Discovery of Difficult Targets: The Limits of Ligandability. *Chemistry & Biology* **2012,** *19* (1), 42-50.
21. Vassilev, L. T.; Vu, B. T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E. A., In Vivo Activation of the p53 Pathway by Small-Molecule Antagonists of MDM2. *Science* **2004,** *303* (5659), 844-848.
22. Li, S.; Wang, H.; Peng, B.; Zhang, M.; Zhang, D.; Hou, S.; Guo, Y.; Ding, J., Efalizumab binding to the LFA-1 αL I domain blocks ICAM-1 binding via steric hindrance. *Proceedings of the National Academy of Sciences* **2009,** *106* (11), 4349-4354.
23. Vagner, J.; Qu, H.; Hruby, V. J., Peptidomimetics, a synthetic tool of drug discovery. *Current opinion in chemical biology* **2008,** *12* (3), 292-296.
24. Clackson, T.; Wells, J., A hot spot of binding energy in a hormone-receptor interface. *Science* **1995,** *267* (5196), 383-386.
25. Felizmenio-Quimio, M. E.; Daly, N. L.; Craik, D. J., Circular proteins in plants: solution structure of a novel macrocyclic trypsin inhibitor from Momordica cochinchinensis. *J Biol Chem* **2001,** *276* (25), 22875-82.
26. Gould, A.; Ji, Y.; Aboye, T. L.; Camarero, J. A., Cyclotides, a novel ultrastable polypeptide scaffold for drug discovery. *Current Pharmaceutical Design* **2011,** *17* (38), 4294-4307.
27. Garner, J.; Harding, M. M., Design and synthesis of [small alpha]-helical peptides and mimetics. *Organic & Biomolecular Chemistry* **2007,** *5* (22), 3577-3585.
28. Davis, J. M.; Tsou, L. K.; Hamilton, A. D., Synthetic non-peptide mimetics of [small alpha]-helices. *Chemical Society Reviews* **2007,** *36* (2), 326-334.
29. Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F., Foldamers as versatile frameworks for the design and evolution of function. *Nature chemical biology* **2007,** *3* (5), 10.1038/nchembio876.
30. Murray, J. K.; Gellman, S. H., Targeting protein–protein interactions: Lessons from p53/MDM2. *Peptide Science* **2007,** *88* (5), 657-686.
31. Cheng, R. P.; Gellman, S. H.; DeGrado, W. F., β-Peptides: From Structure to Function. *Chemical Reviews* **2001,** *101* (10), 3219-3232.
32. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S., A Field Guide to Foldamers. *Chemical Reviews* **2001,** *101* (12), 3893-4012.
33. Wu, C. W.; Kirshenbaum, K.; Sanborn, T. J.; Patch, J. A.; Huang, K.; Dill, K. A.; Zuckermann, R. N.; Barron, A. E., Structural and Spectroscopic Studies of Peptoid Oligomers with α-Chiral Aliphatic Side Chains. *Journal of the American Chemical Society* **2003,** *125* (44), 13525-13530.
34. Pelay-Gimeno, M.; Glas, A.; Koch, O.; Grossmann, T. N., Structure-Based Design of Inhibitors of Protein–Protein Interactions: Mimicking Peptide Binding Epitopes. *Angewandte Chemie International Edition* **2015,** *54* (31), 8896-8927.

35. Zhang, H.; Zhao, Q.; Bhattacharya, S.; Waheed, A. A.; Tong, X.; Hong, A.; Heck, S.; Curreli, F.; Goger, M.; Cowburn, D.; Freed, E. O.; Debnath, A. K., A Cell-penetrating Helical Peptide as a Potential HIV-1 Inhibitor. *Journal of Molecular Biology* **2008,** *378* (3), 565-580.
36. Spokoyny, A. M.; Zou, Y.; Ling, J. J.; Yu, H.; Lin, Y.-S.; Pentelute, B. L., A Perfluoroaryl-Cysteine SNAr Chemistry Approach to Unprotected Peptide Stapling. *Journal of the American Chemical Society* **2013,** *135* (16), 5946-5949.
37. Lau, Y. H.; de Andrade, P.; Wu, Y.; Spring, D. R., Peptide stapling techniques based on different macrocyclisation chemistries. *Chemical Society Reviews* **2015,** *44* (1), 91-102.
38. Rezaei Araghi, R.; Keating, A. E., Designing helical peptide inhibitors of protein–protein interactions. *Current Opinion in Structural Biology* **2016,** *39*, 27-38.
39. Taylor, J. W., The synthesis and study of side-chain lactam-bridged peptides. *Biopolymers* **2002,** *66* (1), 49-75.
40. Verdine, G. L.; Hilinski, G. J., All-hydrocarbon stapled peptides as Synthetic Cell-Accessible Mini-Proteins. *Drug Discovery Today: Technologies* **2012,** *9* (1), e41-e47.
41. Scrima, M.; Le Chevalier-Isaad, A.; Rovero, P.; Papini, A. M.; Chorev, M.; D'Ursi, A. M., CuI-Catalyzed Azide–Alkyne Intramolecular i-to-(i+4) Side-Chain-to-Side-Chain Cyclization Promotes the Formation of Helix-Like Secondary Structures. *European Journal of Organic Chemistry* **2010,** *2010* (3), 446-457.
42. Jacobsen, Ø.; Maekawa, H.; Ge, N.-H.; Görbitz, C. H.; Rongved, P. l.; Ottersen, O. P.; Amiry-Moghaddam, M.; Klaveness, J., Stapling of a 310-Helix with Click Chemistry. *The Journal of Organic Chemistry* **2011,** *76* (5), 1228-1238.
43. Kawamoto, S. A.; Coleska, A.; Ran, X.; Yi, H.; 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. *Journal of Medicinal Chemistry* **2011,** *55* (3), 1137-1146.
44. Felix, A. M.; Heimer, E. P.; Wang, C. T.; Lambros, T. J.; Fournier, A.; Mowles, T. F.; Maines, S.; Campbell, R. M.; Wegrzynski, B. B.; Toome, V.; et al., Synthesis, biological activity and conformational analysis of cyclic GRF analogs lactam. *Int J Pept Protein Res* **1988,** *32* (6), 441-54.
45. Houston, M. E., Jr.; Gannon, C. L.; Kay, C. M.; Hodges, R. S., Lactam bridge stabilization of alpha-helical peptides: ring size, orientation and positional effects. *J Pept Sci* **1995,** *1* (4), 274-82.
46. Murage, E. N.; Gao, G.; Bisello, A.; Ahn, J.-M., Development of Potent Glucagon-like Peptide-1 Agonists with High Enzyme Stability via Introduction of Multiple Lactam Bridges. *Journal of Medicinal Chemistry* **2010,** *53* (17), 6412-6420.
47. Dong, M.; Te, J. A.; Xu, X.; Wang, J.; Pinon, D. I.; Storjohann, L.; Bordner, A. J.; Miller, L. J., Lactam Constraints Provide Insights into the Receptor-Bound Conformation of Secretin and Stabilize a Receptor Antagonist. *Biochemistry* **2011,** *50* (38), 8181-8192.
48. Blackwell, H. E.; Grubbs, R. H., Highly Efficient Synthesis of Covalently Cross-Linked Peptide Helices by Ring-Closing Metathesis. *Angewandte Chemie International Edition* **1998,** *37* (23), 3281-3284.
49. Blackwell, H. E.; Sadowsky, J. D.; Howard, R. J.; Sampson, J. N.; Chao, J. A.; Steinmetz, W. E.; O'Leary, D. J.; Grubbs, R. H., Ring-closing metathesis of olefinic peptides: design, synthesis, and structural characterization of macrocyclic helical peptides. *J Org Chem* **2001,** *66* (16), 5291-302.

50. Schafmeister, C. E.; Po, J.; Verdine, G. L., An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *Journal of the American Chemical Society* **2000,** *122* (24), 5891-5892.
51. Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R. D.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J., Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science* **2004,** *305* (5689), 1466-70.
52. Chu, Q.; Moellering, R. E.; Hilinski, G. J.; Kim, Y.-W.; Grossmann, T. N.; Yeh, J. T. H.; Verdine, G. L., Towards understanding cell penetration by stapled peptides. *MedChemComm* **2015,** *6* (1), 111-119.
53. Walensky, L. D.; Pitter, K.; Morash, J.; Oh, K. J.; Barbuto, S.; Fisher, J.; Smith, E.; Verdine, G. L.; Korsmeyer, S. J., A Stapled BID BH3 Helix Directly Binds and Activates BAX. *Molecular Cell* **2006,** *24* (2), 199-210.
54. Gavathiotis, E.; Reyna, D. E.; Davis, M. L.; Bird, G. H.; Walensky, L. D., BH3-Triggered Structural Reorganization Drives the Activation of Proapoptotic BAX. *Molecular Cell* **2010,** *40* (3), 481-492.
55. Gavathiotis, E.; Suzuki, M.; Davis, M. L.; Pitter, K.; Bird, G. H.; Katz, S. G.; Tu, H.-C.; Kim, H.; Cheng, E. H. Y.; Tjandra, N.; Walensky, L. D., BAX activation is initiated at a novel interaction site. *Nature* **2008,** *455* (7216), 1076-1081.
56. Stewart, M. L.; Fire, E.; Keating, A. E.; Walensky, L. D., The MCL-1 BH3 helix is an exclusive MCL-1 inhibitor and apoptosis sensitizer. *Nat Chem Biol* **2010,** *6* (8), 595-601.
57. Bernal, F.; Tyler, A. F.; Korsmeyer, S. J.; Walensky, L. D.; Verdine, G. L., Reactivation of the p53 Tumor Suppressor Pathway by a Stapled p53 Peptide. *Journal of the American Chemical Society* **2007,** *129* (9), 2456-2457.
58. Bernal, F.; Wade, M.; Godes, M.; Davis, T. N.; Whitehead, D. G.; Kung, A. L.; Wahl, G. M.; Walensky, L. D., A Stapled p53 Helix Overcomes HDMX-Mediated Suppression of p53. *Cancer Cell* **2010,** *18* (5), 411-422.
59. Moellering, R. E.; Cornejo, M.; Davis, T. N.; Bianco, C. D.; Aster, J. C.; Blacklow, S. C.; Kung, A. L.; Gilliland, D. G.; Verdine, G. L.; Bradner, J. E., Direct inhibition of the NOTCH transcription factor complex. *Nature* **2009,** *462* (7270), 182-188.
60. Phillips, C.; Roberts, L. R.; Schade, M.; Bazin, R.; Bent, A.; Davies, N. L.; Moore, R.; Pannifer, A. D.; Pickford, A. R.; Prior, S. H.; Read, C. M.; Scott, A.; Brown, D. G.; Xu, B.; Irving, S. L., Design and Structure of Stapled Peptides Binding to Estrogen Receptors. *Journal of the American Chemical Society* **2011,** *133* (25), 9696-9699.
61. Danial, N. N.; Walensky, L. D.; Zhang, C.-Y.; Choi, C. S.; Fisher, J. K.; Molina, A. J. A.; Datta, S. R.; Pitter, K. L.; Bird, G. H.; Wikstrom, J. D.; Deeney, J. T.; Robertson, K.; Morash, J.; Kulkarni, A.; Neschen, S.; Kim, S.; Greenberg, M. E.; Corkey, B. E.; Shirihai, O. S.; Shulman, G. I.; Lowell, B. B.; Korsmeyer, S. J., Dual role of proapoptotic BAD in insulin secretion and beta cell survival. *Nat Med* **2008,** *14* (2), 144-153.
62. Bhattacharya, S.; Zhang, H.; Debnath, A. K.; Cowburn, D., Solution Structure of a Hydrocarbon Stapled Peptide Inhibitor in Complex with Monomeric C-terminal Domain of HIV-1 Capsid. *Journal of Biological Chemistry* **2008,** *283* (24), 16274-16278.
63. Zhang, H.; Curreli, F.; Zhang, X.; Bhattacharya, S.; Waheed, A.; Cooper, A.; Cowburn, D.; Freed, E.; Debnath, A., Antiviral activity of alpha-helical stapled peptides designed from the HIV-1 capsid dimerization domain. *Retrovirology* **2011,** *8* (1), 28.
64. Bird, G. H.; Madani, N.; Perry, A. F.; Princiotto, A. M.; Supko, J. G.; He, X.; Gavathiotis, E.; Sodroski, J. G.; Walensky, L. D., Hydrocarbon double-stapling remedies

the proteolytic instability of a lengthy peptide therapeutic. *Proceedings of the National Academy of Sciences* **2010,** *107* (32), 14093-14098.
65. Sviridov, D. O.; Ikpot, I. Z.; Stonik, J.; Drake, S. K.; Amar, M.; Osei-Hwedieh, D. O.; Piszczek, G.; Turner, S.; Remaley, A. T., Helix stabilization of amphipathic peptides by hydrocarbon stapling increases cholesterol efflux by the ABCA1 transporter. *Biochemical and Biophysical Research Communications* **2011,** *410* (3), 446-451.
66. Baek, S.; Kutchukian, P. S.; Verdine, G. L.; Huber, R.; Holak, T. A.; Lee, K. W.; Popowicz, G. M., Structure of the Stapled p53 Peptide Bound to Mdm2. *Journal of the American Chemical Society* **2011,** *134* (1), 103-106.
67. Cantel, S.; Le Chevalier Isaad, A.; Scrima, M.; Levy, J. J.; DiMarchi, R. D.; Rovero, P.; Halperin, J. A.; D'Ursi, A. M.; Papini, A. M.; Chorev, M., Synthesis and Conformational Analysis of a Cyclic Peptide Obtained via i to i+4 Intramolecular Side-Chain to Side-Chain Azide−Alkyne 1,3-Dipolar Cycloaddition. *The Journal of Organic Chemistry* **2008,** *73* (15), 5663-5674.
68. de Araujo, A. D.; Hoang, H. N.; Kok, W. M.; Diness, F.; Gupta, P.; Hill, T. A.; Driver, R. W.; Price, D. A.; Liras, S.; Fairlie, D. P., Comparative α-Helicity of Cyclic Pentapeptides in Water. *Angewandte Chemie International Edition* **2014,** *53* (27), 6965-6969.
69. Merrifield, R. B., Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide. *Journal of the American Chemical Society* **1963,** *85* (14), 2149-2154.
70. Kent, S. B., Chemical synthesis of peptides and proteins. *Annu Rev Biochem* **1988,** *57*, 957-89.
71. Coin, I.; Beyermann, M.; Bienert, M., Solid-phase peptide synthesis: from standard procedures to the synthesis of difficult sequences. *Nat. Protocols* **2007,** *2* (12), 3247-3256.
72. Kaiser, E.; Colescott, R. L.; Bossinger, C. D.; Cook, P. I., Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal Biochem* **1970,** *34* (2), 595-8.
73. Hancock, W. S.; Battersby, J. E., A new micro-test for the detection of incomplete coupling reactions in solid-phase peptide synthesis using 2,4,6-trinitrobenzene-sulphonic acid. *Analytical Biochemistry* **1976,** *71* (1), 260-264.
74. Christensen, T., Qualitative test for monitoring coupling completeness in solid phase peptide synthesis using chloranil. *Acta Chem. Scand. Ser. B* **1979,** *33*, 763-766.
75. Pedersen, S. L.; Tofteng, A. P.; Malik, L.; Jensen, K. J., Microwave heating in solid-phase peptide synthesis. *Chemical Society Reviews* **2012,** *41* (5), 1826-1844.
76. Organisation, W. H. *Global status report on noncommunicable disease 2014: Attending the nine global noncommunicable disease targets; a shared responsibility*; WHO Press: 20 Avenue Appia, 1211 Geneva 27, Switzerland 2014.
77. Lilly, L. S.; School, H. M., *Pathophysiology of Heart Disease: A Collaborative Project of Medical Students and Faculty*. Wolters Kluwer/Lippincott Williams & Wilkins: 2011.
78. Waksman, R.; Serruys, P. W.; Schaar, J., *The Vulnerable Plaque, Second Edition*. CRC Press: 2007.
79. Lilly, S. M.; Wilensky, R. L., Emerging therapies for acute coronary syndromes. *Frontiers in Pharmacology* **2011,** *2*.
80. Michelson, A. D., Antiplatelet therapies for the treatment of cardiovascular disease. *Nat Rev Drug Discov* **2010,** *9* (2), 154-169.

81. Wiviott, S. D.; Braunwald, E.; McCabe, C. H.; Horvath, I.; Keltai, M.; Herrman, J.-P. R.; Van de Werf, F.; Downey, W. E.; Scirica, B. M.; Murphy, S. A.; Antman, E. M., Intensive oral antiplatelet therapy for reduction of ischaemic events including stent thrombosis in patients with acute coronary syndromes treated with percutaneous coronary intervention and stenting in the TRITON-TIMI 38 trial: a subanalysis of a randomised trial. *The Lancet 371* (9621), 1353-1363.
82. Wang, T. H.; Bhatt, D. L.; Topol, E. J., Aspirin and clopidogrel resistance: an emerging clinical entity. *European heart journal* **2006,** *27* (6), 647-654.
83. Cox, D.; Brennan, M.; Moran, N., Integrins as therapeutic targets: lessons and opportunities. *Nat Rev Drug Discov* **2010,** *9* (10), 804-820.
84. Guthrie, R., Review and management of side effects associated with antiplatelet therapy for prevention of recurrent cerebrovascular events. *Advances in therapy* **2011,** *28* (6), 473-82.
85. Boersma, E.; Akkerhuis, K. M.; Théroux, P.; Califf, R. M.; Topol, E. J.; Simoons, M. L., Platelet Glycoprotein IIb/IIIa Receptor Inhibition in Non–ST-Elevation Acute Coronary Syndromes: Early Benefit During Medical Treatment Only, With Additional Protection During Percutaneous Coronary Intervention. *Circulation* **1999,** *100* (20), 2045-2048.
86. Harrington, R. A., Design and Methodology of the PURSUIT Trial. *American Journal of Cardiology 80* (4), 34B-38B.
87. McClellan, K. J.; Goa, K. L., Tirofiban. A review of its use in acute coronary syndromes. *Drugs* **1998,** *56* (6), 1067-80.
88. Coller, B. S., Anti-GPIIb/IIIa Drugs: Current Strategies and Future Directions. *Thrombosis and Haemostasis* **2001,** *86* (7), 427-43.
89. Serebruany, V. L.; Malinin, A. I.; Eisert, R. M.; Sane, D. C., Risk of bleeding complications with antiplatelet agents: Meta-analysis of 338,191 patients enrolled in 50 randomized controlled trials. *American Journal of Hematology* **2004,** *75* (1), 40-47.
90. Patrono, C.; Coller, B.; FitzGerald, G. A.; Hirsh, J.; Roth, G., Platelet-active drugs: the relationships among dose, effectiveness, and side effects: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. *Chest* **2004,** *126* (3 Suppl), 234s-264s.
91. Hynes, R. O., Integrins: A family of cell surface receptors. *Cell* **1987,** *48* (4), 549-554.
92. Hynes, R. O., Integrins: Bidirectional, Allosteric Signaling Machines. *Cell* **2002,** *110* (6), 673-687.
93. Luo, B. H.; Carman, C. V.; Springer, T. A., Structural basis of integrin regulation and signaling. *Annu Rev Immunol* **2007,** *25*, 619-47.
94. Anthis, N. J.; Wegener, K. L.; Ye, F.; Kim, C.; Goult, B. T.; Lowe, E. D.; VakonakisIoannis; Bate, N.; Critchley, D. R.; Ginsberg, M. H.; CampbellIain, D., The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. *EMBO* **2009,** *28*, 3623–3632.
95. Shattil, S. J.; Kim, C.; Ginsberg, M. H., The final steps of integrin activation: the end game. *Nat Rev Mol Cell Biol* **2010,** *11* (4), 288-300.
96. Petrich, B. G.; Fogelstrand, P.; Partridge, A. W.; Yousefi, N.; Ablooglu, A. J.; Shattil, S. J.; Ginsberg, M. H., The antithrombotic potential of selective blockade of talin-dependent integrin alpha IIb beta 3 (platelet GPIIb-IIIa) activation. *J Clin Invest* **2007,** *117* (8), 2250-9.

Chapter 2

Lactam Linker

2.1 Integrin activation

Modulation of the process by which αIIbβ3 integrin is activated presents as a promising target for the development of novel therapeutics for platelet aggregation[1]. Integrins are the main receptors that cells use to bind and respond to the extracellular matrix, and the regulation of integrin activation is important for a wide range of anchorage dependant cellular events. Activation of integrins is tightly regulated through inside-out signalling. The binding of a cytoskeletal protein, talin, to the β integrin cytoplasmic tail is the final step in the activation process[2, 3]. Talin binds to the β tail of integrin, causing separation of the cytoplasmic α and β tails, and activation[4] (Figure 2-1Figure).

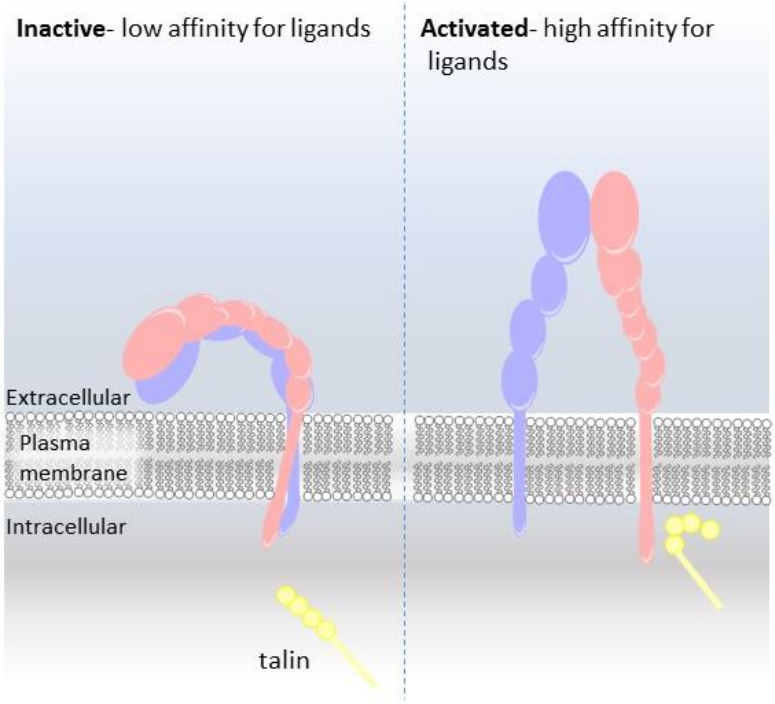


Figure 2-1: Activation of integrin by talin binding the final step in the activation process.

Talin is a 270 kDa protein, with an *N*-terminal globular head, and a *C*-terminal rod like tail. The rod like tail contains integrin, vinculin[5] and actin[6] binding sites, as well as binding sites for proteins involved in recruiting talin to the membrane to activate integrins[7]. Talins head contains an atypical FERM domain with the addition of an F0 domain to F1, F2 and F3 domains, which adopt an extended structure[8]. The model emerging from studies is that the plasma membrane interactions are mediated via the

F1, F2 and F3 subdomains[9-11]. The F3 domain binds the β3 integrin cytoplasmic tail over an extensive interface, which can be divided into two parts[12-14], the membrane proximal (MP) region which interacts with the helical section of the β3 tail close to the cell membrane, and the membrane distal (MD) region which interacts with the relatively unstructured section of the β3 tail. Wegener and co-workers defined the structure of a complex between talin and the membrane-proximal integrin β3 cytoplasmic domain, using NMR spectroscopy[15]. NMR solution structures of talin F3 bound to a chimeric peptide comprised of the membrane proximal (MP) region of the β3 tail and *C*-terminal region of PIPKIγ are shown in Figure 2-2A. The chimeric peptide was used in this study as it displayed tighter binding to the talin F3 domain than the native peptide. Importantly mutagenesis studies revealed integrin activation could be markedly diminished with β3 integrin variants lacking key residues F(730) and F(727) in the cytoplasmic tail domain, but which retained the ability to bind talin. The first crystal structure of an authentic full length β integrin cytoplasmic tail (1B) bound to talin F2-F3 domains (talin2 isoform)[13] was subsequently reported by Anthis et al.[4] (Figure 2-2B). This structural information provides a strong basis for the development of compounds which can modulate integrin activation.

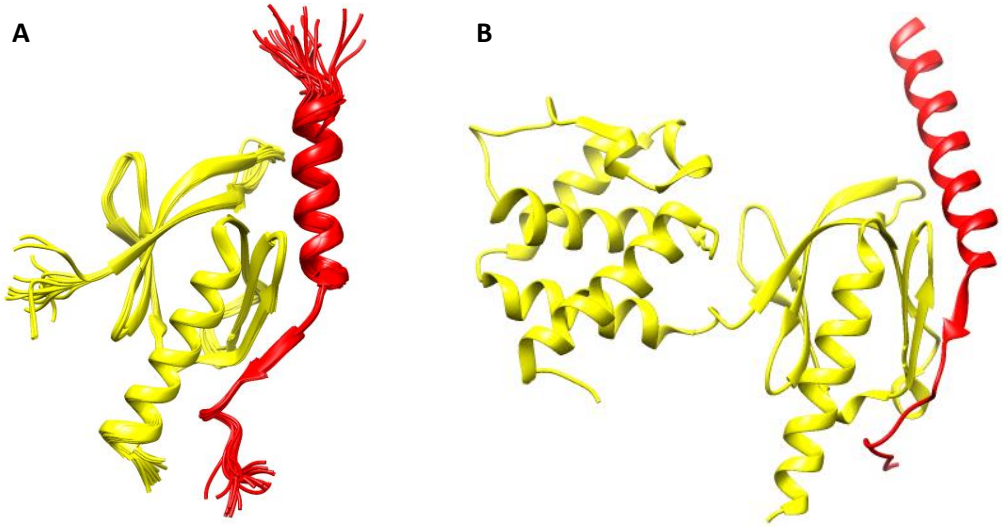


Figure 2-2: **A**: Averaged NMR solution structure of β3 bound talin F3 domain[14], **B**: Crystal structure of 1B integrin bound to talin F2-F3 domains[3].

2.2 Talin: A target for inhibiting integrin activation

Initial investigations of the talin-integrin αIIbβ3 interaction determined that overexpression of talin head domain lead to activation of αIIbβ3 receptor[16]. Activation was found to be dependent on the presence of talin head domain, and integrin β3 tail. Petrich and co-workers performed mutation studies in mice to investigate the effect of disrupting interactions with the β3 tail[1]. All mice displayed pathological bleeding with complete loss of β3 tail. Less than 5% of mice displayed pathological bleeding due to mutation to the β3 tail that disrupted talin binding alone. These mice also received protection from microvascular thrombosis. In mice with a mutation disrupting the binding of both talin and other cytoplasmic proteins, microvascular thrombosis was prevented, however over 50% of mice displayed evidence of pathological bleeding. The study established αIIbβ3 integrin activation was dependent on interaction of talin with the β3 tail, in vivo. The results also suggest that pathological bleeding is associated with the degree of loss of integrin activity. Selectively disrupting the interaction between talin and the β3 integrin cytoplasmic tail in this study appeared to prevent thrombosis formation, with reduced risk of pathological bleeding.

2.3 Helical constrained peptides derived from the β3 integrin cytoplasmic tail

2.3.1 Optimisation of lactam linked peptides

The initial objective of this study was to determine if peptides derived from the β3 tail would display high helicity once lactam linkers were incorporated at positions A and B. A panel of short peptides (shown in Table 2-1) with sequences derived from the helical section of MP region of the β3 tail, from H(722)-A(735), were synthesised using fluorenylmethyloxycarbonyl (Fmoc) solid phase peptide synthesis protocols[17] (see chapter 4.2 for experimental details). The peptides were designed with *C*-terminal amide groups rather than the native carboxylic acid group as this is known to improve the stability of peptides to protease degradation[18]. Rink amide resin was selected for synthesis as it produces peptides with *C*-terminal amides upon cleavage.

Table 2-1: Peptide sequences of derived from the integrin β3 tail. Lactam linked residues are indicated by blue lines between residues.

Peptide	Sequence	Linker Position
β3MP(722-735)	H-HDRKEFAKFEEERA-NH_2	None
1	H-HDRKEFAKFEEDRA-NH_2	A
2	H-HDRKEFAKFEEERA-NH_2	A
3	H-HDRKEFADFEEERA-NH_2	B
4	H-HDRDEFAKFEEERA-NH_2	B
5	H-HDREEFAKFEEERA-NH_2	B

The linear peptide β3MP(722-735) was synthesised using an automated microwave peptide synthesiser from R(734) to R(724). Manual coupling of the first amino acid to the resin resulted in higher purity of synthesised peptides, likely due to the resin undergoing a longer swelling time, and minimising incomplete couplings by monitoring the reaction with the TNBS test[19] for free amines. The final amino acids, D(723) and H(722), were also coupled manually due to formation of aspartimide after treatment of D(723) with piperidine to remove the *N*-Fmoc protection. Aspartimide formation was avoided by addition of 1-hydroxybenzotriazole (HOBt) to the deprotection solution in manual synthesis.

Lactam linked peptides **1**-**5** were synthesised manually from A(735) to D(723). The linkers were formed between the side chains of natural amino acids, lysine and aspartic acid or glutamic acid (Table 2-1, linked residues). Lysine residue side chains were protected with highly acid sensitive protecting groups, methoxytrityl (Mmt), and 2-phyenyl isopropyl oxy (Opip) protecting groups for aspartic acid and glutamic acid[20]. The Fmoc protected peptides were then treated with aliquots of 2% TFA in DCM for 2 minutes to selectively remove the acid labile Mmt, and Opip protection groups. Successful removal of protecting groups was indicated by the lack of yellow chromophore in the filtrate, produced by the presence of the Mmt group in solution.

Conditions required for lactamisation were identified by trialling different coupling reagents, (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP), *N*,*N*'-diisopropylcarbodiimide (DIC), *N*,*N*'-dicyclohexylcarbodiimide (DCC), bromotripyrrolidinophosphonium hexafluorophosphate (PyBroP) and *N*-[(dimethylamino)-1H-1,2,3-triazolo-[4,5-b]pyridin-1-ylmethylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide (HATU). Reactions were performed in DMF under basic conditions, achieved by addition of DIPEA. The reaction times for cyclisation of **1** (Table 2-1) were compared after addition of DIC, DCC, PyBroP and HATU to resin samples respectively (Table 2-2). Cyclisation with PyBroP, DIC and DCC produced a positive result using the TNBS test for detecting the presence of free amine after 48 hrs. PyBOP resulted in a negative TNBS test after 18 hrs and HATU exhibited a negative TNBS test for free amines after 1 hr. This suggests that cyclisation was incomplete for all samples other than those treated with HATU and PyBOP, which was confirmed using low resolution mass spectroscopy. While HATU required the shortest coupling time of 1 hr, cyclisation of the peptide under these conditions did not always result in the desired lactam linked peptide, rather it produced both the desired cyclic peptide and a by-product. Mass spectrometry of the crude sample suggested the presence of two major ions, one corresponding to the mass of the desired cyclic peptide, and the other to the mass of the linear peptide with the addition of 99 amu; which corresponds to addition of a guanidine group to the lysine residue required for cyclisation of the peptide, a common impurity from SPPS[21]. Guanidine capping of an amine is a result of the reaction of an amine with HATU. A proposed mechanism for the addition of the guanidine group is illustrated in Figure 2-3. The concentration of HATU to free carboxylic acid groups, in this case determined by the moles of unprotected aspartic acid side chains, is critical to ensure that the reaction results in cyclisation rather than capping of the lysine amine. Optimisation of the cyclisation conditions for the coupling reaction using HATU was not found to be possible due to variations in the moles of carboxylic acid groups available. The variations in the number of moles of peptide present for an individual reaction were likely a result of the resin loading after the addition of the first amino acid, and in the amount of resin due to the removal of resin for TNBS testing. PyBOP was therefore deemed to be the coupling reagent of

choice for reproducible results, despite the long reaction time required for cyclisation, and additionally the reaction could be successfully monitored by the TNBS resin test.

Table 2-2: Cyclisation conditions trialled using peptide **1** to determine most effective cyclisation reagent.

Coupling reagent	Time for cyclisation (hrs)	Result of TNBS test
PyBOP	18	Negative
DIC	48	Positive
DCC	48	Positive
PyBroP	48	Positive
HATU	1	Negative

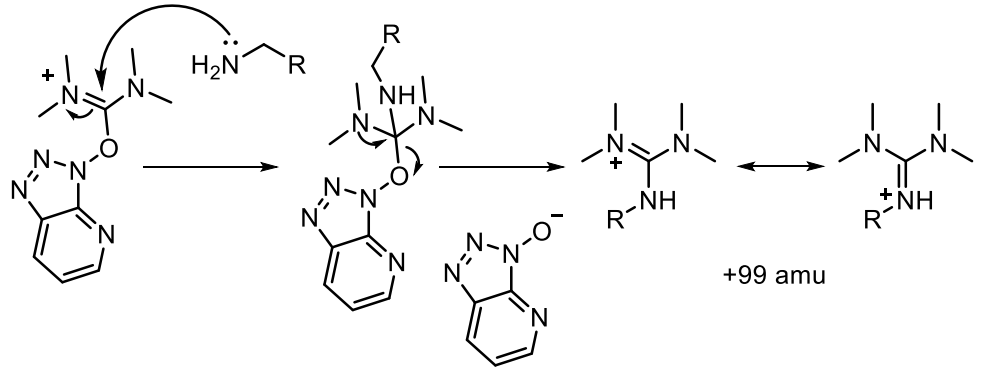


Figure 2-3: Mechanism for the addition of the guanidine group to a free amine, resulting in a by-product with an additional mass of 99 amu.

Peptides **1**-**5** were cyclised using optimised conditions identified from the trial reactions discussed above. The peptides were thus cyclised on treatment PyBOP, and DIPEA in DMF, for 18 hrs. Following cyclisation, the Fmoc group was removed on treatment with piperidine and the *N*-terminal histidine (His) residue was coupled using HATU as the coupling agent. Histidine was couped after cyclisation to prevent unintentional removal of the trityl (Trt) side chain protecting group, known to be acid sensitive[20] (Figure 2-4).

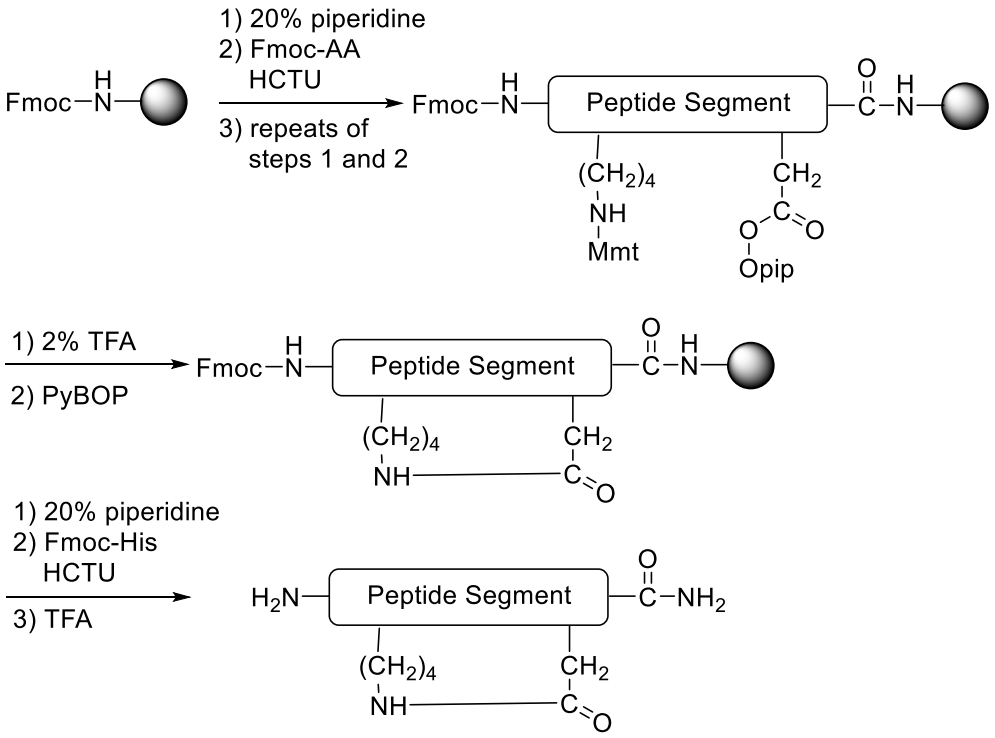


Figure 2-4: Scheme for the synthesis and cyclisation of lactam linked peptides **1-5**

Identification of the linker that would produce the peptide with the greatest helicity was determined through characterisation of the structure of peptides **1**-**5** using CD spectroscopy. Spectra were obtained of purified samples of the unconstrained β3MP(722-735) and linked peptides in phosphate buffer, at pH 6.5 (Figure 2-5), this was to ensure consistency with the conditions required for NMR titrations with the talin F3 domain discussed later in this chapter. The ratio of absorbance minima at 222 nm and 208 nm ($\theta_{222}/\theta_{208}$) obtained from the CD spectra (Table 2-3), and the position of the absorbance minima were used to identify the peptide with the greatest α-helical structure. Peptides displaying helical content have a $\theta_{222}/\theta_{208}$ of close to 1 and a minimum close to 208 nm. All linked peptides displayed spectra characteristic of α-helical structure, while the unconstrained peptide lacked an absorbance minimum at 222 nm, indicating that the linkers increased the structural content of the peptides.

Table 2-3: Peptides synthesised and analysed by CD spectroscopy.

Peptide	Linker Position	Linker	Ratio ($\theta_{222}/\theta_{208}$)	Minima (nm)
β3MP(722-735)	None	-	N/A	202
1	A	KD	0.75	207
2	A	KE	0.44	203
3	B	KD	0.35	203
4	B	DK	0.24	203
5	B	EK	0.26	202

Comparison of the linker position, either positioned at site A or B, and linker residue combination, with lysine and aspartic acid, or glutamic acid in the i or i+4th position, was explored to determine the arrangement that produced the greatest helicity. The two different linker positions tested were position A and B (Table 2-1). The KD linker was tested in both positions A and B, peptides **1** and **3** respectively, which produced peptides with the greatest helical content. Peptide **1** displayed the greatest α-helical content of all tested peptides with a $\theta_{222}/\theta_{208}$ ratio of 0.7, and minima at 207 nm. The $\theta_{222}/\theta_{208}$ ratio for all other linked peptides were 0.44 or below, indicating the presence of 3_{10} helicity, or a mixture of α-helicity and random coil structure.

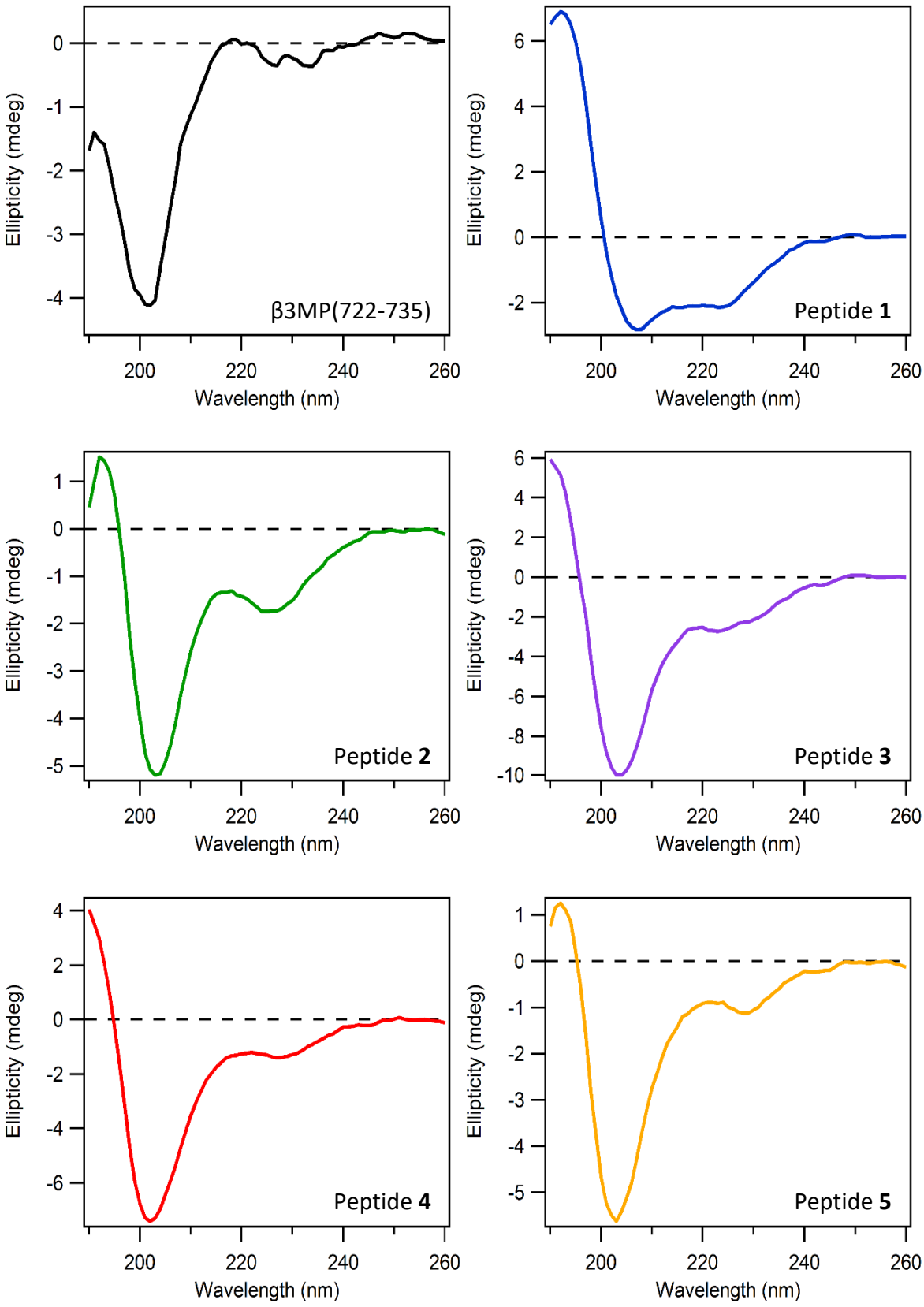


Figure 2-5: CD spectra of peptides β3MP(722-735) and **1**-**5** derived from the integrin β3 tail. (~25µM concentration, 10mM phosphate buffer, pH 6.5, 298K)

The most suitable linker length was determined by comparison of linker residues in position A. Peptides **1** and **2** link the lysine and glutamic acid residues (KE) or lysine and aspartic acid residues (KD). The KD linker is one atom shorter in length than the KE linker. It was hypothesised that the KE linker in peptide **2,** which replaces the potential electrostatic attraction between lysine and glutamic acid with a covalent bond, would enhance helicity as it conserves the peptides original sequence. The KE linker does appear to increase helical content, however replacing the glutamic acid with an aspartic acid residue, resulting in the shorter KD linker, in peptide **1**, resulted in the peptide with the greater helical content.

The results suggest that the optimum linker for enhancing α-helicity is between lysine and aspartic acid (KD), in the *i* and *i*+4 positions, which is in agreement with previous comparative studies for lactam bridged peptides[22, 23]. This linker was therefore chosen for use in further investigations.

2.3.2 Design of linked peptide: sequence and linker position

Compounds were designed to disrupt the talin-β3 integrin interaction, motivated by their potential application in preventing platelet aggregation. Peptides were developed to bind to talin and disrupt interaction with the β3 integrin cytoplasmic tail without disrupting any other known talin functions. The talin MP region was identified as the target region[13, 24] (Figure 2-6), the β3 tail is its only known ligand. The β3 tail binds in a helical conformation at the MP region of talin, and therefore peptides were designed based on the native MP β3 tail sequence, with helical structure stabilised using a side chain constraint approach (See Chapter 1.5).

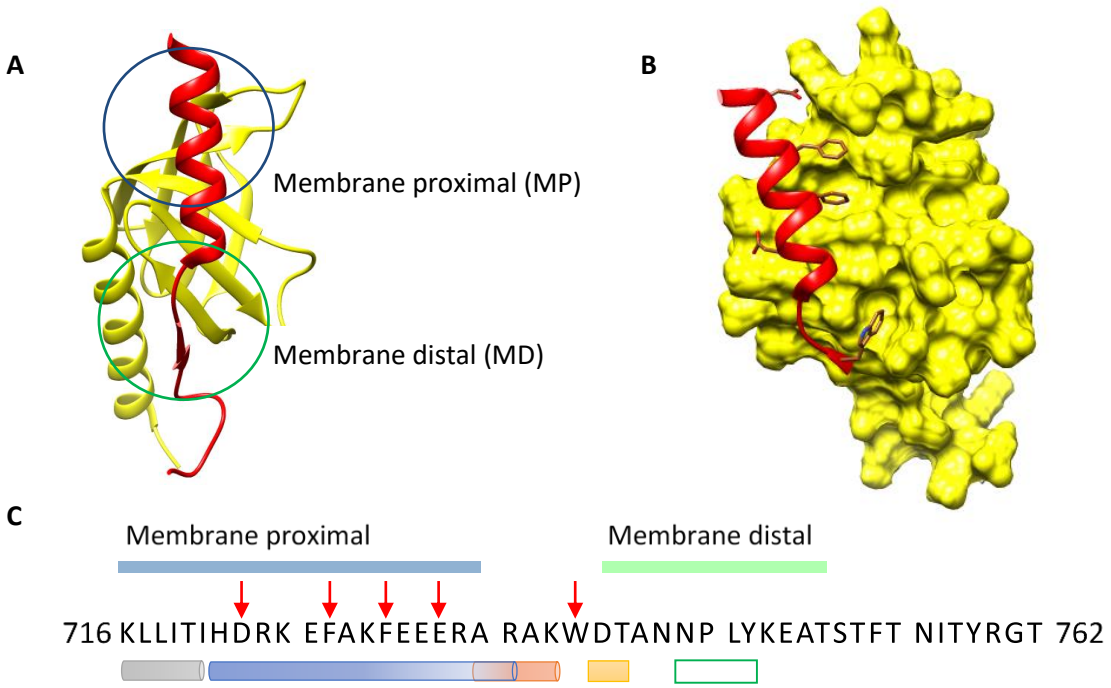


Figure 2-6: **A**: Ribbon representation of averaged NMR solution structures showing talin F3 domain (yellow) bound to chimeric β3 tail (red). Membrane proximal (MP, blue) and distal (MD, green) regions are circled[14]. **B**: Averaged NMR solution structure of talin F3 domain (yellow surface) bound to β3 tail (red ribbon) showing hotspot residues side chains. **C**: The native sequence of the β3 tail in one letter code. Above the sequence the MP region is indicated in blue and MD region in green. Below the sequence secondary structures identified by NMR and crystallography are shown. The grey regions are α-helical and buried in the membrane, blue is α-helical, and orange region was found to be 3_{10} helical in the crystal structure of the 1B tail[12], MD region of β3 integrin is relatively unstructured with a small region of β-sheet (orange) and a turn structure[16] (green outline), the crystal structure of 1B tail is unstructured in the MD region.

A Robetta online alanine scan[25] was performed on the NMR solution structure of talin F3 domain bound to chimeric β3 tail (PDB 2H7E). Energetically important residues at the protein-protein interface with a stabilising effect on ΔΔG of more than 1 kcal/mol were identified as hot-spot residues. Robetta successfully identified known hot spot residues within the MP region, F(727) and F(730). These residues had previously been determined to contribute significantly to binding from experimental alanine mutation studies [15,13] (Table). The energy contribution of residue E(733) has not been reported, however perturbation of integrin ligand binding and integrin activation has been linked to mutation at this site[26]. The tryptophan residue W(739) between the MP region and

the MD region was also identified by Robetta, which predicted it would have the greatest contribution to binding of all the residues in the chimeric β3 tail. The order of residue energy contribution identified by Robetta was in agreement with experimental results. In addition, D(723) was identified as a hotspot residue, while this is not predicted from either the computational alanine scan or experimental mutation studies, evidence of salt bridge formation with talin residue K(327) was identified in the crystal structure of the closely related integrin 1D bound to talin F2-F3[13].

Table 2-4: Values of ΔΔG calculated by Robetta online alanine server, and experimentally determine values previously reported for ΔΔG of the full length β3 tail determined by mutation of residues to alanine and analysis of binding to talin F3 domain by NMR.

Residue	Computational alanine scan (ΔΔG kcal/mol)*	Experimental alanine mutation (ΔΔG kJ/mol)#
F(727)	1.62	4.79 (1.14)^
F(730)	2.46	
E(733)	1.15	N/D
W(739)	4.39	8 (1.9)

*Determined from Robetta online alanine scanner

#Determine from NMR binding analysis [15,13]

^Calculated from combine mutation of F(727) and F(730) to alanine

The determination of hot-spot residues led to the identification of two target sequences from the β3 tail. The first target sequence from H(722) to A(735) (Figure 2-6 C), contained three of the four hot-spot residues identified, and incorporated residues from the helical MP region only. The second sequence extended to the start of the MD region: H(722) to W(739). Extending the sequence to the hot-spot residue W(739) allowed the incorporation of all four hot-spot residues, however extended past the helical region suggested by NMR solution structures (Figure 2-6).

The locations of the side chain linkers were selected on the opposite side to the interaction interface between the β3 integrin tail and the talin F3 domain. The linker between side chains prevents these side chains from interactions with the protein interface. The linker may also alter the structures interacting at the binding interface as the linker runs between these residues. Linker locations were selected which would minimise the linker's potential disruptive effects on binding of peptides to the talin F3

domain, while stabilising the helical region of the β3 tail. Figure 2-7 shows three positions for the location of the linker, B, C and D, which were identified that would not interfere with binding; site A replaced E(733) with the linker, and was selected to determine the effect of disrupting the interaction of E(733) with talin F3, as predicted by the Robetta scan (Table 2-4).

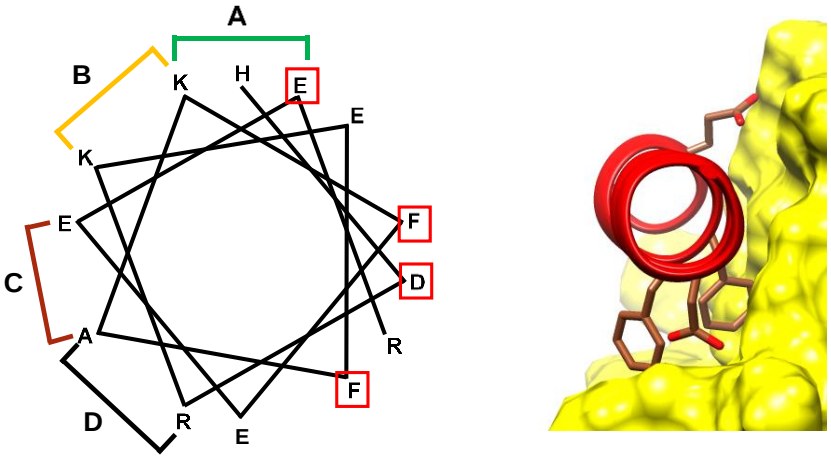


Figure 2-7: Left, helical wheel diagram of the α-helical region of the β3 integrin cytoplasmic tail, indicating location of constraints introduced by replacing residues with lysine and aspartic acid in the *i* and *i*+4th positions respectively. Right, image from the averaged NMR structure of the MP region bound to talin, with the structure of hotspot residues E(733), F(730), D(723), F(727) indicated by red squares.

2.3.3 Structure and affinity of linear β3-integrin peptide sequences

Two linear peptides of different lengths, β3MP(722-735) and β3MP(722-739), were tested for their structure and affinity to the talin F3 domain. A shorter peptide would speed up synthesis but may sacrifice binding affinity. The shorter peptide β3MP(722-735) tested consisted of 14 residues which comprised only the helical section of the peptide. A longer peptide of 18 residues, β3MP(722-739), was also tested which comprised both the helical section and an unstructured section of peptide which may increase affinity to the talin F3 domain.

The 18-residue peptide, β3MP(722-739), was synthesized by standard Fmoc solid-phase peptide synthesis protocols. Automated peptide synthesis was employed for the

addition of residues K(738)-R(724). With the first residue, tryptophan (W), and residues D(723) and H(722) coupled manually. The peptide was then removed from the resin and purified by semi-prep HPLC. Manual coupling of the final residues, D(723) and H(722), was used in order to prevent aspartimide formation, as previously discussed in the synthesis of β3MP(722-735). The 14-residue peptide, β3MP(722-735), was prepared as discussed in section 2.3.1.

β3MP peptides were characterised using CD spectroscopy in order to determine the structure present in the linear peptide (Figure 2-8). Quantification of helicity was based on the molar ellipticity at λ = 222 nm[27] (see chapter 4.1.2.2, equation 2). β3MP(722-735) did not display an absorption minimum at 222 nm, suggesting the structure of this peptide is predominately a random coil. In comparison, β3MP(722-739) gave a weak minimum at 222 nm which equates to a helical content of 23%. A calculated ratio of molar ellipticity ($[\theta]_{222}/[\theta]_{208}$) for β3MP(722-739) of 0.5 suggests that the structure may be a combination of α-helical, 3_{10} helical structure, and random coil[28].

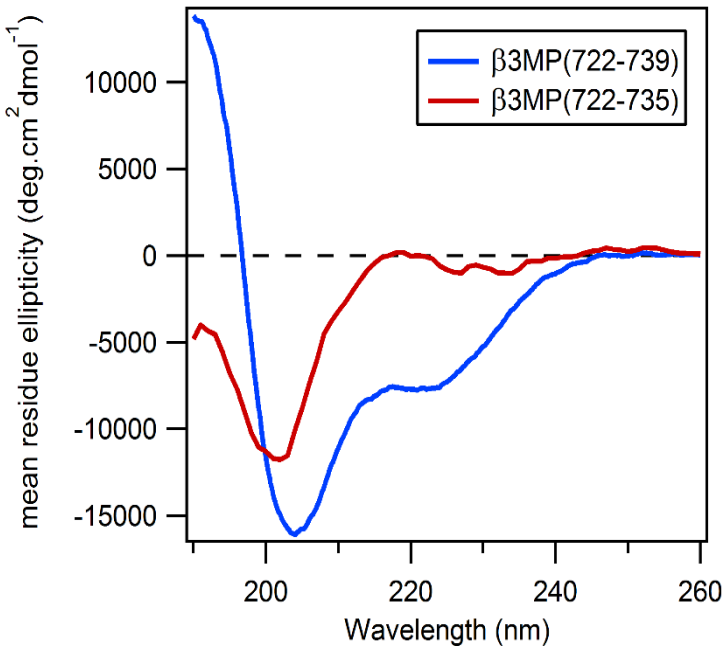


Figure 2-8: CD spectra of β3MP(722-735) and β3MP(722-739). (10mM phosphate buffer, pH 6.5, 298K)

The binding affinity of linear peptides β3MP(722-735) and β3MP(722-739) for the talin F3 domain was determined using two dimensional ^{15}N-^{1}H heteronuclear single-quantum

correlation (HSQC) spectra by analysing the resonance shifts in the F3 talin domain upon peptide binding. All NMR spectra were obtained and analysed in collaboration with Dr Kate Wegener, who guided sample preparation and performed NMR spectroscopy and analysis. HSQC spectra of a solution containing ^{15}N labelled talin F3 domain were recorded with increasing concentrations of each linear peptide β3MP(722-735) and β3MP(722-739). The binding affinity of β3MP(722-735) for the talin F3 domain was not determined as the peptide began to precipitate from solution before the required concentration for accurate determination of the dissociation constant (Kd) was reached. The binding affinity of β3MP(722-739) to talin F3 domain was determined to be 5.4 ± 0.2mM (Chapter 4.2.12 Figure 4-12), with the pattern of talin F3 resonance shifts indicating the peptide was interacting with the talin F3 domain in a similar way to longer fragments of β3-integrin tail, or the chimeric β3-integrin/PIPKIγ peptide (Figure 2-9). The binding affinity of the longer linear peptide was in the millimolar range, as expected since the binding affinity of the full length β3 integrin tail to the talin F3 domain was

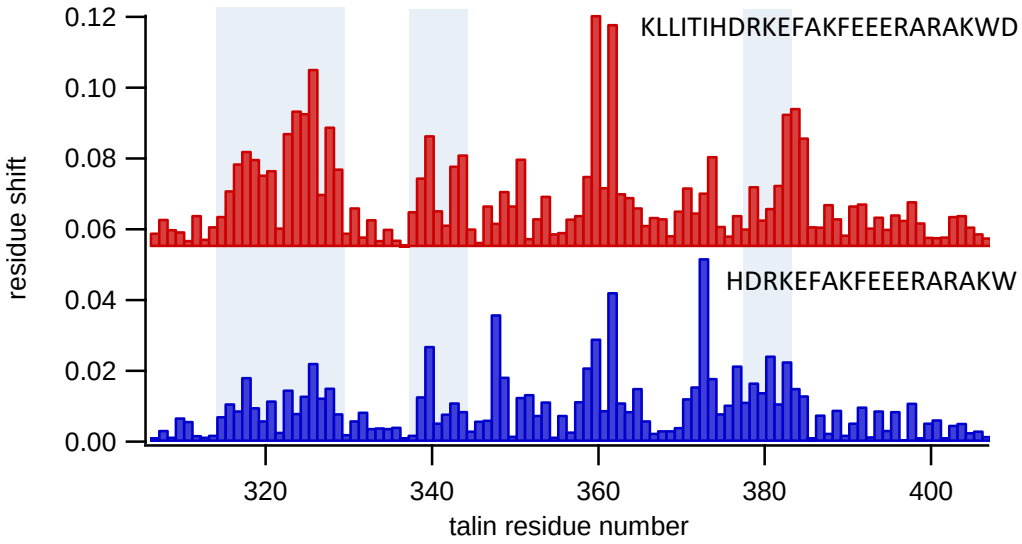


Figure 2-9: Plot of maximum shift of talin F3 resonance peaks in HSQC spectra with (top: red) addition of a peptide corresponding to the membrane proximal region (bottom: blue) β3MP(722-739) peptide. The regions highlighted in blue refer to the residues affected by engagement of the β3-integrin helix.

previously determined to be between 273 - 600μM[13,29].

2.3.4 Optimisation of constraint location using lactam bridged peptides

The affinity of the short peptide, β3MP(722-735), for the talin F3 domain was expected to be lower than the β3MP(722-739). The low binding affinity of β3MP(722-739) of 5.4 ± 0.2mM therefore prompted the use of this sequence spanning from H(722)-W(739), rather than the shorter β3MP(722-735) sequence for the incorporation of lactam linkers. A panel of lactam bridged peptides (see Table 2-5) was designed based on the β3MP(722-739) sequence to determine the ideal location to nucleate helicity and enhance binding of peptides to the talin F3 domain. Lactam constraints were introduced by linking the side chains of lysine and aspartic acid residues in the *i* and *i*+4 positions. The covalent constraints were introduced in positions A, B, C and D (Table 2-5, Figure 2-7). Positions were chosen so as to not interfere with hot spot residues, only position A would disrupt potential binding of E(733) recognised by the Robetta results as the weakest interaction of all the identified hotspot residues (refer to section 1.4.1).

Table 2-5: Sequences, helical character, and affinity of β3 integrin cytoplasmic tail derived peptides

Peptide	Sequence	Bridge position	Helicity	$[\theta]_{222}/[\theta]_{208}$
β3MP(722-735)	HDRKEFAKFEEERA	N/A	0%	0
β3MP(722-739)	HDRKEFAKFEEERARAKW	N/A	23%	0.55
6	HRDKEFA**K***FEE**D***RARAKW	A	48%	0.81
7	HRD**K***EFA**D***FEEERARAKW	B	57%	0.76
8	HDRKEF**K***KFE**D***ERARAKW	C	42%	0.77
9	HD**K***KEF**D***KFEEERARAKW	D	51%	0.75

Peptides **6**, **7**, **8** and **9** were synthesised using Fmoc solid-phase peptide synthesis[17]. Residues from K(738) to the residue before those involved in the linker were coupled using an automated microwave peptide synthesiser, with all other residues coupled manually as the linker residues required lower concentrations and longer coupling times. The success of the coupling reactions were also monitored by resin tests and mass spectrometry to ensure coupling reactions reached completion before elongation was continued. Linkers were introduced at positions A – D (Figure 2-7) in peptides **6**-**9** by incorporation of lysine and aspartic acid residues in the *i*, and *i*+4 positions indicated

by the bold resides in Table 2-5. These lysine and aspartic residues were incorporated during peptide synthesis with acid labile protecting groups' methoxytrityl (Mmt) and 2-phyenyl isopropyl oxy (2-PhiPr) respectively. This allowed selective removal of these side chain protecting groups while the peptide remained bound to the resin, which was achieved by consecutive additions of 1% TFA. Cyclisation was then achieved on resin with addition of PyBOP in DIPEA for 18 hrs using the optimised conditions previously describe for peptide **1** (see section 2.3.1), to produce cyclic peptide **6**, **7**, **8**, and **9**. Peptides were characterised using mass spectrometry to confirm cyclisation was successful, and their purity was confirmed by analytical HPLC.

CD spectra of **6**, **7**, **8**, and **9** revealed characteristic spectral features associated with α-helical structure, specifically minima at 208 and 222 nm, as apparent in Figure 2-10. These linked peptides demonstrated helical stabilisation, with increased helical content of between 15 and 28% compared to the linear peptide β3MP(722-739) (Table 2-5). Peptide **7** with the linker at position B displayed the greatest helical content.

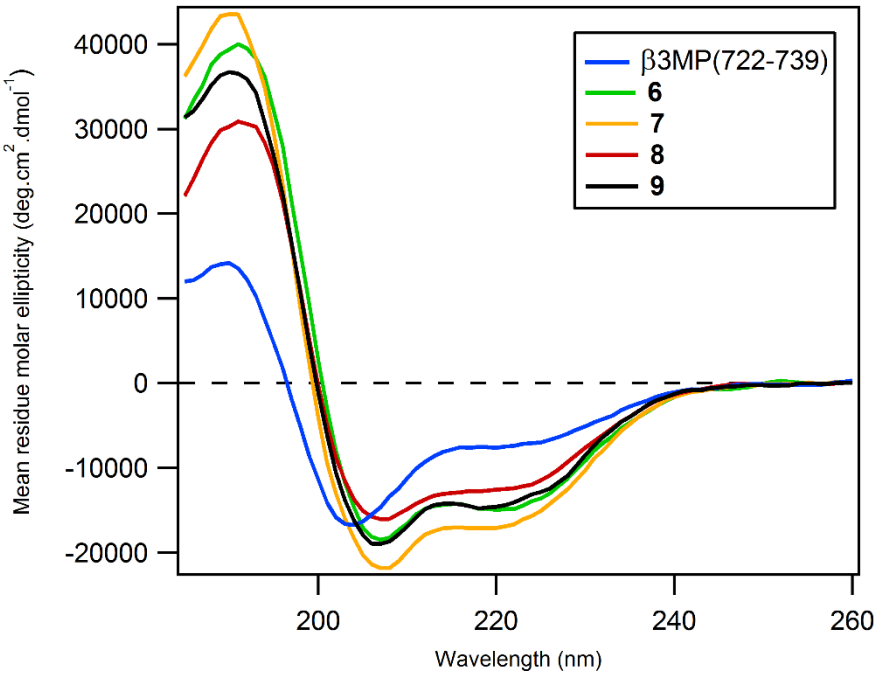


Figure 2-10: CD spectra for the linear and four lactam bridged peptides illustrating the effect of bridge and bridging position on helical content. (10mM phosphate buffer, pH 6.5, 298K)

Two dimensional ^{15}N-HSQC spectra of the ^{15}N labelled talin F3 domain, titrated with increasing amounts of peptides **6** – **9**, were used to generate binding curves (Figure 4-12) through observation of key residue shifts to determine Kd values (Table 2-6). The binding of β3MP(722-739), and peptides **6** - **9** resulted in a pattern of shifted talin resonances, comparable to those identified previously for the β3 integrin tail[15] (Figure 2-11). This suggests that peptides **6** – **9** bind to talin at the integrin binding site[15]. The Kd values determine the effect of bridge position on affinity of peptides for the talin F3 domain (Table 2-6). The binding affinities of peptides **7** and **9** were calculated to be 0.9 ± 0.02 and 2.7 ± 0.1 mM respectively, indicating improved binding affinity to that of the unconstrained β3MP(722-739) peptide which had a Kd of 5.4 ± 0.2 mM. A constraint at position B, peptide **7**, gave rise to a 5-fold increase in binding affinity compared to β3MP(722-739), while position D, peptide **9**, resulted in a more moderate increase in affinity. Constraints at positions A and C (**6** and **8**) resulted in a decrease in affinity compared to β3MP(722-739). The reduced affinity of **6** would lend support to the earlier Robetta scan results indicating contribution from E(733) to binding affinity for talin. However, the observed reduced affinity of **8** was unexpected, since the bridge does not interfere with the binding residues, suggesting that other factors needed to be considered for linker placement. These results prompted further structural analysis of the peptides to determine differences in the peptides structure that may explain the results for their binding affinities.

Table 2-6: displays the Kd of linear peptides and lactam linked peptides for binding to the talin F3 domain.

Peptide	Linker position	(mM)
β3MP(722-735)	none	N/D
β3MP(722-739)	none	5.4 ± 0.2
6	A	6.5 ± 0.3
7	B	0.9 ± 0.02
8	C	7.0 ± 0.4
9	D	2.7 ± 0.1

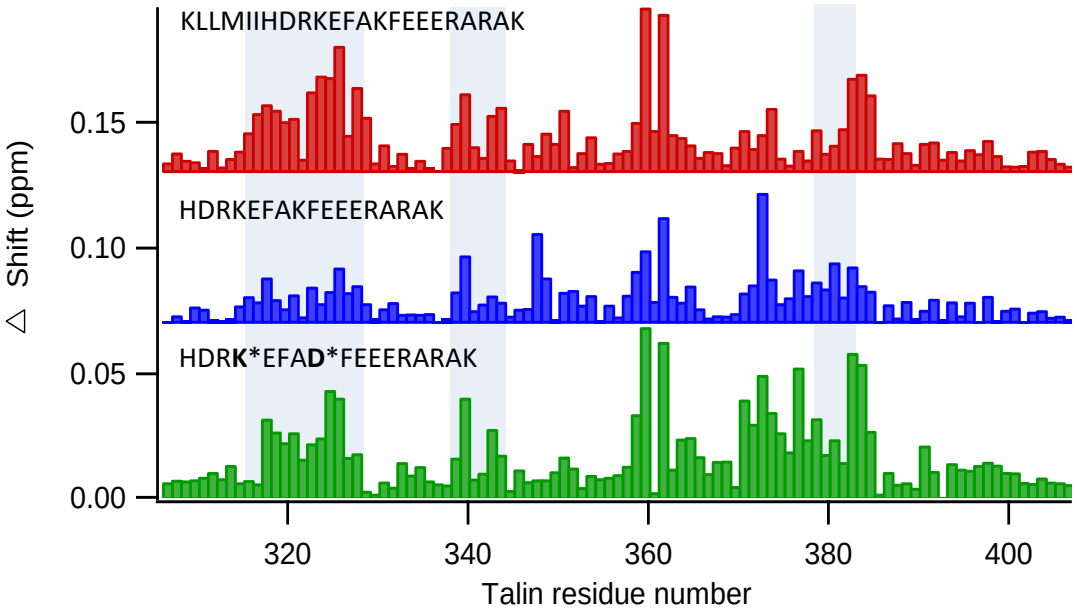


Figure 2-11: Residue shifts of ^{15}N-HSQC spectrum of ^{15}N labelled talin F3 with maximum peptide concentration.

2.3.5 Structural characterisation

CD spectra of the lactam bridge peptides **6** – **9**, show characteristic minima that provide compelling evidence that the lactam bridge enhances helical structure relative to β3MP(722-739), however this does not allow for detailed analysis of the structure of peptides at the atomic level. Thus ^{1}H homo-nuclear NMR spectroscopy studies were performed on β3MP(722-739) and cyclic peptides **6**-**9** in order to better define the effect of linker position on structure. This allows the determination of medium and long-range NOEs, coupling constants, temperature dependence of amide chemical shifts, and α-proton chemical shifts. Spectra were recorded for β3MP(722-739) and peptides **6**-**9** in PBS buffer (50mM phosphate, 100mM NaCl, pH 6.1) with 10% added D_2O to more closely emulate biological conditions. Peptide resonances were assigned, and a range of secondary structure indicators were determined as summarised in Figure 2-12.

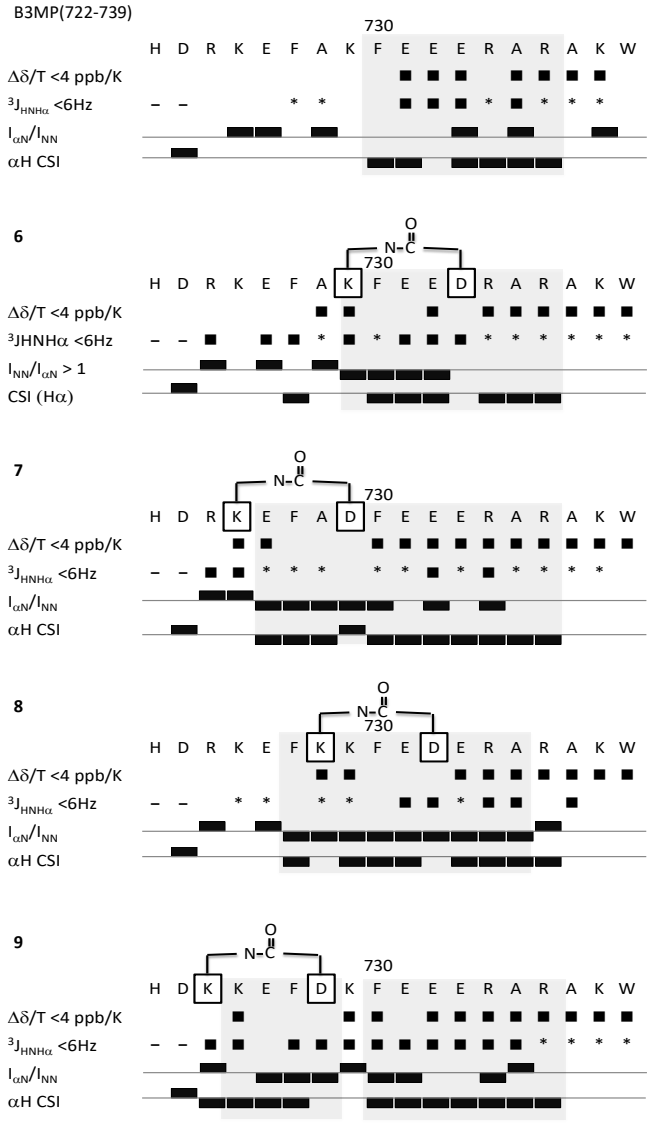


Figure 2-12: NMR summary diagrams. Δδ/T< 4ppb/K is the measure of the temperature dependence of the amide proton chemical shift. ^{3}JHNHα coupling constant of less than 6Hz (^{3}JHNHα < 6Hz) indicates helical structure. Where (▪) represents values which meets the criteria for helical structure, (-) for undetermined values due to structure or spectral overlap, (*) for values of ^{3}JHNHα < 6Hz with spectral overlap and no symbol is for values determined which are outside the range for helicity. Ratios of sequential NN(i,i+1) to αN(i,i+1) NOE intensities of less than one (INN/IαN < 1), are indicated. Here, the horizontal line indicates a value of one, with boxes above the line for values >1, and boxes below for values <1. Chemical shift indices (CSI) for αH resonances are also indicated, where boxes above and below the line signify differences from random coil values of > 0.1 or < 0.1, respectively.

2.3.5.1 NMR analysis of parent peptide structure

^{1}H homonuclear NMR spectroscopy was used to gain a better understand of the structure of unconstrained peptide β3MP(722-739) to determine how the linker location impacted the structures of linked peptides **6** – **9**. The β3MP(722-739) displayed upfield shifted resonances for ^{1}Hα compared to random coil values[30]. These shifts equated to αH CSI (Chemical Shift Index) values of less than -0.1 for residue 730 to 736 indicative of α-helical structure [31, 32]. Helical structure in this region was further supported by $^{3}J_{HNH\alpha}$ couplings of less than 6Hz for four of the five measurable residue resonances (Figure 2-12). This equates to ϕ angles between -70° and -30°which is typical of α-helices[33, 34]

The amide protons (NH) temperature coefficients for seven of the β3MP(722-739) peptide residues were measured to be -4 ppb/K indicating these residues are involved in hydrogen-bonding[35, 36]. All these residues are located in the *C*-terminus of the peptide, between residues 731-738, consistent with the presence of helical structure in this region[35, 36]. Helical structure is also supported by the presence of αβ(i,i+3) and αN(i,i+3) NOE connectivities (4.2.12, Figure 4-13).

The NMR data described above provides evidence for helical structure, particularly in the *C*-terminal region of β3MP(722-739). However the intensities of NN(i,i+1) NOEs, relative to their respective αN(i,i+1) NOEs (INN/IαN) are greater than 1 (Figure 11 and 4.2.12, Figure 4-13), which is inconsistent with helical structure. Similarly, the large intensities of the αN(i,i+1) to αN(i,i) (Iαα/Iαα) NOEs (4.2.12, Figure 4-13) are also indicative of more extended structure. This is in closer agreement with the earlier CD results for β3MP(722-739) (see section 2.3.3) indicating a greater proportion of extended, or random coil, structure present across the length of the peptide[37]. These two opposing sets of evidence are likely due to dynamic reorientation of the peptide in solution between extended and helical forms.

2.3.5.2 NMR analysis of the structure of lactam bridged peptides

The observed NOEs, coupling constants, temperature dependence of amide chemical shifts, and α-proton chemical shifts for lactam bridged peptides **6**, **7**, **8** and **9** are consistent with increased helical structure compared to β3MP(722-739) (Figure 2-12).

For each peptide, nucleation extends from the point of linker attachment toward the *C*-terminal, with residues in these regions displaying temperature coefficients greater than -4ppb/K indicative of a greater number of hydrogen-bonded residues (Figure 2-12). The absence of hydrogen bonding between residues on the *N*-terminal side of the linkers of the lactam linker peptides as well as the absence of other indicators suggests helicity is not being nucleated towards the *N*-terminal of the peptide. Unlike the data for β3MP(722-739), the lactam linked peptides gave NOE intensity ratios (IαN/INN) and (Iαα/Iαα) less than 1 for residues in the helical regions, supporting CSI, ^{3}JHNHα and amide temperature coefficients. This suggests more stable helical structures for the linked peptides relative to linear peptides. This matches the CD data discussed earlier in section 2.3.4.

NMR structural calculations were performed using the program ARIA[38] (Ambiguous Restraints for Iterative Assignment) in order to further understand the effect of the linker on constraining the structure of each peptide (Figure 2-13). These calculations, performed by Dr Kate Wegener, resulted in 20 refined structures for each of the peptides studied, β3MP(722-739) and peptides **6** – **9**, based on NOE, scalar coupling and hydrogen bond restraints (4.2.12, Figure 4-13). The calculated structures for β3MP(722-739) and peptides **6** – **9** displayed greater helical content (Figure 2-13) than their experimental NMR measurements (Figure 2-12). This suggests that the calculated structures generate a more static view of a highly dynamic system compared to NMR measurements. This is particularly apparent for peptide **8**. The calculated peptide structures do not depict the level of disorder present in solution, however when coupled with the analysis of the raw NMR measurements they produce a useful understanding of the structure of the peptides.

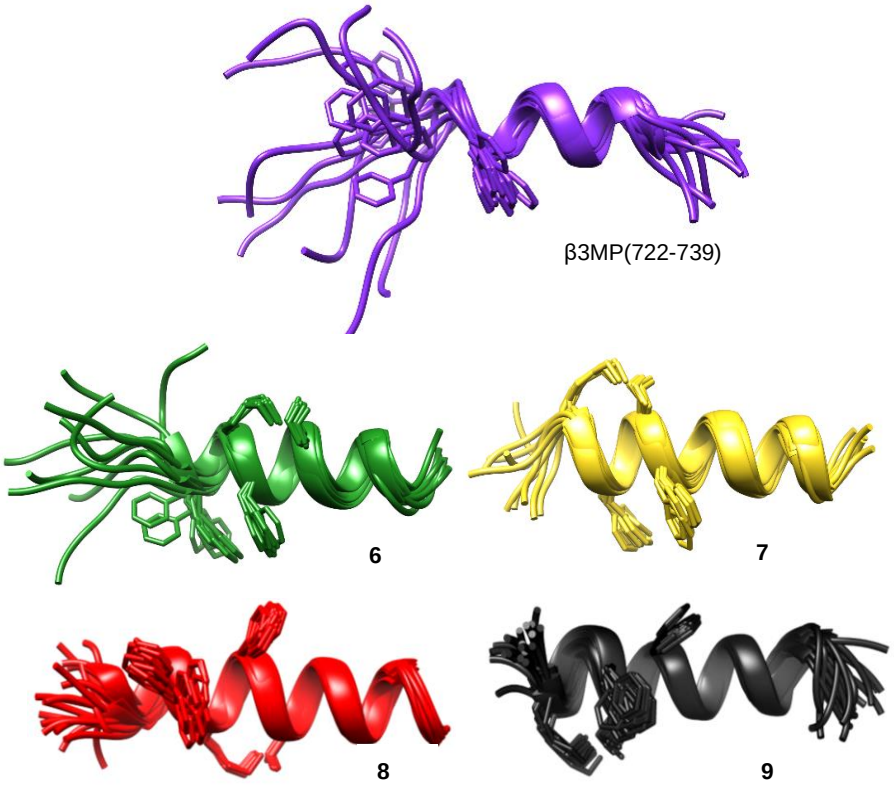


Figure 2-13: NMR structure calculations for β3MP(722-739) (purple), **6** (green) **7** (yellow) ,**8** (red), and **9** (black).[39]

The calculated structures of β3MP(722-739) depicted in Figure 2-13 suggest that the *C*-terminal region is dominated by α-helical structure, while the *N*-terminal region is largely unstructured[39]. Also depicted in Figure 2-13 are the calculated structures of peptides **6** – **9** which reveal unbroken helicity from the linker to the *C*-terminus. This was not supported by the NMR measurements, with chemical shifts for αH CSI values downfield of the expected values for the aspartate linker residue indicting helical structure. The disruption of helical structure at this point is not supported by $^{3}J_{HNH\alpha}$ coupling constants or IαN/INN values. The evidence above suggests asparagine or aspartate residues are not appropriate random coil reference residues, or that the proximity of the amide to the αH of this aspartic acid reside could be affecting the αH chemical shift. The aspartate linker residues of linked peptides also display large negative NH temperature coefficients suggest hydrogen bonding was not present. In contrast, the calculated structures consistently exhibited hydrogen bonding at this residue.

The apparent contradiction between structure calculations and NMR measurements could be due to the presence of a shortened hydrogen bond for the aspartate linker residues, or shielding of the residue by the lactam linker. The length of hydrogen bonds involving the NH proton is known to influence the NH temperature coefficients and NH secondary shifts[35, 36]. Shorter hydrogen bonds of less than 2Å are associated with downfield shifted NH resonances and temperature coefficients of less than -4ppb/K. The calculated structures of lactam bridged peptides **6** – **9** reveal hydrogen bond lengths of less than 2Å along the backbone between the residues. However, other hydrogen bonds along the backbone had lengths of less than 2Å but did not display NH shifts and temperature coefficients outside of the helical range. It is possible that the lactam amide group from the lactam bridge linker influences the chemical shifts and temperature coefficients associated with the aspartic acid residue. The downfield shifted NH resonance of aspartic acid residue suggests a deshielding effect for the linker, however, the large negative NH temperature coefficient is more consistent with a shielding effect. Based on this contradiction, the length of the hydrogen bond would better explain the observed measurements.

The lysine resides involved in the linker for linked peptides, **6** - **8**, display NH temperature coefficients greater than -4ppb/K suggest hydrogen bonding along the backbone. Peptide **9** displayed a temperature coefficient less than -4ppb/K indicating there was no hydrogen bonding along the backbone, attributable to the residue being particularly close to the N-terminus. In contradiction, the structure calculations could not reliably identify hydrogen bond acceptors, a discrepancy due to the previously discussed static structure calculation of a highly dynamic system within solution.

Decreased disorder in the *C*-terminal region after the linker was incorporated, as suggested by NMR structure calculations, appear to provide an insight into the binding affinities of the linked peptides to the talin F3 domain. The linkers of peptides **6** and **8** are closest to the *C*-terminus and structure calculations in these cases suggest a more disordered *N*-terminus, with minimal stabilisation outside of the linker region. Peptides **7** and **9** produced enhanced affinity in comparison to β3MP(722-739), a result supported by NMR assertion that the α-helical structure is apparent across the entire length of the peptides.

Disorder for the key hotspot *N*-terminal residues, phenylalanine residues F(727) and F(730), was significantly reduced for peptide **7** and **9** compared to **6** and **8**. Figure 2-14 shows an overlay of structure calculations of peptide **7** with β3 integrin/PIP kinase chimeric peptide bound to talin 1 (Figure 2-14, left) and the crystal structure of 1D and talin 2 (Figure 2-14, right). The structure calculations show good agreement in the locations of the phenylalanine residues in the crystal structure, however, a similar arrangement is seen with peptide **9**, and therefore the difference in affinity between peptides **7** and **9** remains unclear. Residue K729 in peptide **9** displays ^{1}αH chemical shifts similar to those of lysine in a random coil structure, and intensities of the NN(i,i+1) NOE, relative to its respective αN(i,i+1) NOEs (INN/IαN), to be greater than 1. These measurements suggest residue K729, located between key phenylalanine residues, lacks indicators of helicity. Further investigation would be required to verify this observation, which, in the future, may be utilised in studies to determine if an optimum distance for the placement of a lactam bridge linker from the native structured region of a peptide can be determined.

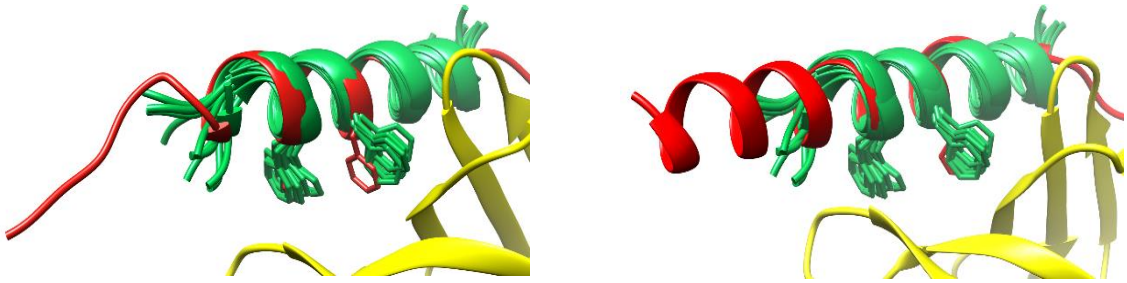


Figure 2-14: (B) Left Peptide **7** overlaid with Talin1/b3 peptide (2H7E). Right Peptide **7** (green) overlaid with Talin2 (yellow)/b1D peptide (red) (3G9W).

2.3.6 Protease stability

Protease require the substrate to adopt an extended β-sheet geometry for protease binding, and thus they do not hydrolyse residues within a helical structure. The hydrolysis of peptides due to protease binding can provide an indication of the stability of peptides to cellular conditions, and may suggest the regions of helical structure that cannot be accessed the protease. α-Chymotrypsin is known to cleave peptide bonds on the C-terminal side of phenylalanine residues, providing two predicted cleavage sites for the hydrolysis of integrin peptides, at F(727) and F(730). Liquid chromatography

coupled with mass spectrometry was used to monitor the chymotrypsin catalysed hydrolysis of the parent peptides β3MP(722-739) over time. As expected, treatment of the linear peptide β3MP(722-739) with α-chymotrypsin at an enzyme/substrate ratio of approximately 1/2500 resulted in rapid hydrolysis of the linear peptide, with a half-life of only 14 minutes. The breakdown products were identified by LCMS-ESI, with the major ions identified being the products of cleavage adjacent to F(727), however there was no evidence of cleavage at F(730). Under the same conditions, just under 90% (as measured by ion abundance) of peptide **7** remained after 24 hrs exposure to α-chymotrypsin. The linker within peptide **7** spans the established cleavage site for α-chymotrypsin, F(727), with the alternative cleavage site of F(730) immediately adjacent to the linker, see Figure 2-12. The increase in protease stability observed for peptide **7** relative to β3MP(722-739) may be due to the linker blocking entry of the peptide to the active site, as well as increased helical structure and a consequential decrease in β-sheet geometry required for hydrolysis.

Peptide **6** was similarly treated with chymotrypsin to determine the stability of a linked peptide when the cleavage site was not between the linked residues. While the lactam bridge linker does not span or involve a residue close to F(727) the half-life for peptide **6** was significantly extended to 43 minutes with comparison to the linear peptide, β3MP(722-739) with a half-life of 14 minutes. These results suggest that the secondary structure induced by the linker hinders the proteases ability to bind the peptides in the extended β-strand structure.

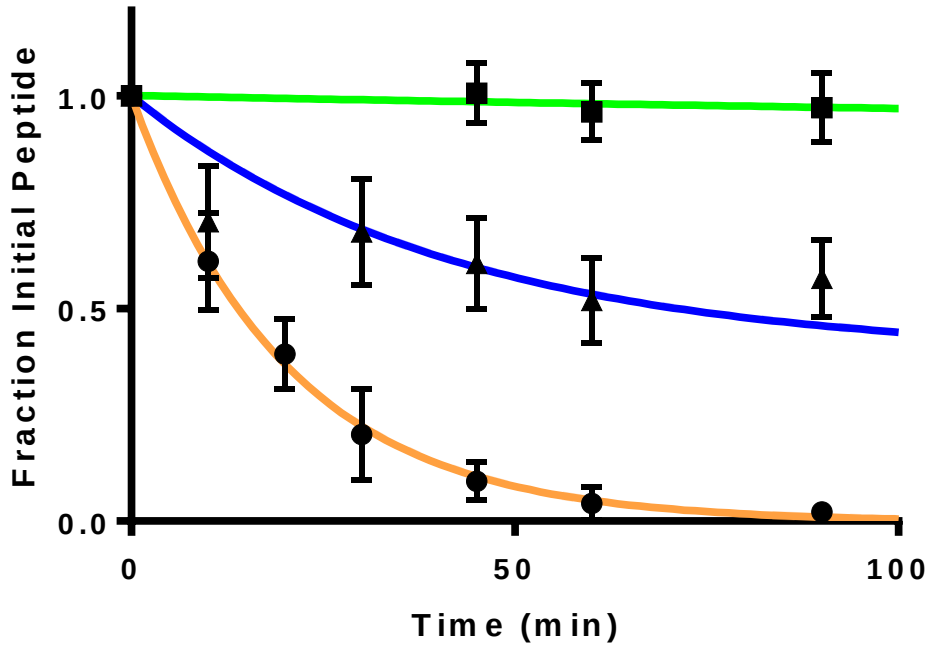


Figure 2-15 Time course of β3 peptide cleavage by chymotrypsin. β3MP(722-739) (●), peptide 7 (■) and peptide 6 (▲). Curves indicate the fit to non-linear regression exponential one phase decay.

These results confirm that a suitably placed lactam bridged linker increases resistance to proteolysis. The preferred site of cleavage for α-chymotrypsin in the linear peptide, β3MP(722-739), was F(727) suggesting that this cleavage site was more accessible to α-chymotrypsin than F(730) which is closer to the C-terminal end. This may be due to native helical structure present in this region of the linear peptide, β3MP(722-739), as revealed by the NMR studies discussed in section 2.3.5.

2.3.7 Conclusions

The studies undertaken in this chapter reveal that linker location has significant impact on both peptide helicity and on determining the binding affinity peptides and the talin F3 domain. While all peptides with linkers incorporated resulted in increased structural content when compared to the linear peptide β3MP(722-739) , this was not the case for

binding affinity. Stabilisation of helical structure did not necessarily produce enhanced affinity, even when bridged residues are on the opposite face to the binding residues. Detailed structural analysis using NMR spectroscopy of both the parent peptide (β3MP(722-739)) and linked peptides revealed nucleation occurred only in a single direction, on the *C*-terminal side of the linker, appearing to extend toward the native tendency for helical structure present at the *C*-terminal end of β3 integrin peptide. These findings demonstrate the importance of identifying the structural characteristics of partially structured patent peptides, as a means to provide insight into the optimum location for incorporation of a helical inducing linker. Helical structure at the *C*-terminal of the linear peptide (β3MP(722-739)) is supported by the proteolysis results, with residue F(730) protected from cleavage. The difference in the structure of peptides **7** and **9** are subtler, and yet results in a clear difference in affinity to F3 talin. Further investigation into the structural complexities of these peptides may produce finding that will assist with the determination of the optimum position for side chain linkers in the future. The evidence gathered here suggests that a constraint should be located in a region of low structural propensity avoiding key residues, which can be identified with Robetta online alanine scan server. In this case, a constraint on the *N*-terminal side of these residues resulted in helix stabilisation. Future work is required to produce optimised inhibitors selective for the integrin binding site of talin, with these peptides providing important leads in the development of a new class of anti-thrombotic therapeutic with a novel mechanism of action.

2.4 References

1. Petrich, B. G.; Fogelstrand, P.; Partridge, A. W.; Yousefi, N.; Ablooglu, A. J.; Shattil, S. J.; Ginsberg, M. H., The antithrombotic potential of selective blockade of talin-dependent integrin alpha IIb beta 3 (platelet GPIIb-IIIa) activation. *J Clin Invest* **2007,** *117* (8), 2250-9.
2. Tadokoro, S.; Shattil, S. J.; Eto, K.; Tai, V.; Liddington, R. C.; de Pereda, J. M.; Ginsberg, M. H.; Calderwood, D. A., Talin binding to integrin beta tails: a final common step in integrin activation. *Science* **2003,** *302* (5642), 103-6.
3. Iwamoto, D. V.; Calderwood, D. A., Regulation of integrin-mediated adhesions. *Current Opinion in Cell Biology* **2015,** *36*, 41-47.
4. Anthis, N. J. Structural Studies of Integrin Activation. University of Oxford, Oxford, 2009.
5. Gingras, A. R.; Ziegler, W. H.; Frank, R.; Barsukov, I. L.; Roberts, G. C.; Critchley, D. R.; Emsley, J., Mapping and consensus sequence identification for multiple vinculin binding sites within the talin rod. *J Biol Chem* **2005,** *280* (44), 37217-24.
6. Critchley, D. R., Biochemical and structural properties of the integrin-associated cytoskeletal protein talin. *Annual review of biophysics* **2009,** *38*, 235-54.
7. Goult, B. T.; Zacharchenko, T.; Bate, N.; Tsang, R.; Hey, F.; Gingras, A. R.; Elliott, P. R.; Roberts, G. C.; Ballestrem, C.; Critchley, D. R.; Barsukov, I. L., RIAM and vinculin binding to talin are mutually exclusive and regulate adhesion assembly and turnover. *J Biol Chem* **2013,** *288* (12), 8238-49.
8. Goult, B. T.; Bouaouina, M.; Elliott, P. R.; Bate, N.; Patel, B.; Gingras, A. R.; Grossmann, J. G.; Roberts, G. C.; Calderwood, D. A.; Critchley, D. R.; Barsukov, I. L., Structure of a double ubiquitin-like domain in the talin head: a role in integrin activation. *Embo j* **2010,** *29* (6), 1069-80.
9. Kalli, A. C.; Wegener, K. L.; Goult, B. T.; Anthis, N. J.; Campbell, I. D.; Sansom, M. S., The structure of the talin/integrin complex at a lipid bilayer: an NMR and MD simulation study. *Structure* **2010,** *18* (10), 1280-8.
10. Kim, C.; Schmidt, T.; Cho, E. G.; Ye, F.; Ulmer, T. S.; Ginsberg, M. H., Basic amino-acid side chains regulate transmembrane integrin signalling. *Nature* **2012,** *481* (7380), 209-13.
11. Kim, C.; Ye, F.; Hu, X.; Ginsberg, M. H., Talin activates integrins by altering the topology of the beta transmembrane domain. *J Cell Biol* **2012,** *197* (5), 605-11.
12. Wegener, K. L.; Basran, J.; Bagshaw, C. R.; Campbell, I. D.; Roberts, G. C. K.; Critchley, D. R.; Barsukov, I. L., Structural Basis for the Interaction between the Cytoplasmic Domain of the Hyaluronate Receptor Layilin and the Talin F3 Subdomain. *Journal of Molecular Biology* **2008,** *382* (1), 112-126.
13. Anthis, N. J.; Wegener, K. L.; Ye, F.; Kim, C.; Goult, B. T.; Lowe, E. D.; VakonakisIoannis; Bate, N.; Critchley, D. R.; Ginsberg, M. H.; CampbellIain, D., The structure of an integrin/talin complex reveals the basis of inside-out signal transduction. *EMBO* **2009,** *28*, 3623–3632.
14. Anthis, N. J.; Wegener, K. L.; Critchley, D. R.; Campbell, I. D., Structural Diversity in Integrin/Talin Interactions. *Structure* **2010,** *18* (12), 1654-1666.
15. Wegener, K. L.; Partridge, A. W.; Han, J.; Pickford, A. R.; Liddington, R. C.; Ginsberg, M. H.; Campbell, I. D., Structural basis of integrin activation by talin. *Cell* **2007,** *128* (1), 171-82.

16. Calderwood, D. A.; Zent, R.; Grant, R.; Rees, D. J. G.; Hynes, R. O.; Ginsberg, M. H., The Talin Head Domain Binds to Integrin β Subunit Cytoplasmic Tails and Regulates Integrin Activation. *Journal of Biological Chemistry* **1999,** *274* (40), 28071-28074.
17. Sheppard, R., The fluorenylmethoxycarbonyl group in solid phase synthesis. *Journal of Peptide Science* **2003,** *9* (9), 545-552.
18. Powell, M. F.; Stewart, T.; Otvos, L., Jr.; Urge, L.; Gaeta, F. C.; Sette, A.; Arrhenius, T.; Thomson, D.; Soda, K.; Colon, S. M., Peptide stability in drug development. II. Effect of single amino acid substitution and glycosylation on peptide reactivity in human serum. *Pharmaceutical research* **1993,** *10* (9), 1268-73.
19. Antoni, G.; Presentini, R.; Neri, P., A simple method for the estimation of amino groups on insoluble matrix beads. *Analytical Biochemistry* **1983,** *129* (1), 60-63.
20. Isidro-Llobet, A.; Alvarez, M.; Albericio, F., Amino acid-protecting groups. *Chem Rev* **2009,** *109* (6), 2455-504.
21. Arttamangkul, S.; Arbogast, B.; Barofsky, D.; Aldrich, J. V., Characterization of synthetic peptide byproducts from cyclization reactions using on-line HPLC-ion spray and tandem mass spectrometry. *Letters in Peptide Science* **1997,** *3* (6), 357-370.
22. Houston, M. E., Jr.; Gannon, C. L.; Kay, C. M.; Hodges, R. S., Lactam bridge stabilization of alpha-helical peptides: ring size, orientation and positional effects. *J Pept Sci* **1995,** *1* (4), 274-82.
23. Taylor, J. W., The synthesis and study of side-chain lactam-bridged peptides. *Biopolymers* **2002,** *66* (1), 49-75.
24. Garcia-Alvarez, B.; de Pereda, J. M.; Calderwood, D. A.; Ulmer, T. S.; Critchley, D.; Campbell, I. D.; Ginsberg, M. H.; Liddington, R. C., Structural determinants of integrin recognition by talin. *Mol Cell* **2003,** *11* (1), 49-58.
25. Kortemme, T.; Kim, D. E.; Baker, D., Computational alanine scanning of protein-protein interfaces. *Science's STKE : signal transduction knowledge environment* **2004,** *2004* (219), pl2.
26. Saltel, F.; Mortier, E.; Hytönen, V. P.; Jacquier, M.-C.; Zimmermann, P.; Vogel, V.; Liu, W.; Wehrle-Haller, B., New PI(4,5)P2- and membrane proximal integrin–binding motifs in the talin head control β3-integrin clustering. *The Journal of Cell Biology* **2009,** *187* (5), 715-731.
27. Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P., Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *Journal of the American Chemical Society* **2005,** *127* (9), 2974-2983.
28. Banerjee, R.; Sheet, T., Ratio of ellipticities between 192 and 208 nm (R1): An effective electronic circular dichroism parameter for characterization of the helical components of proteins and peptides. *Proteins: Structure, Function, and Bioinformatics* **2017,** *85* (11), 1975-1982.
29. Moore, D. T.; Nygren, P.; Jo, H.; Boesze-Battaglia, K.; Bennett, J. S.; DeGrado, W. F., Affinity of talin-1 for the β3-integrin cytosolic domain is modulated by its phospholipid bilayer environment. *Proceedings of the National Academy of Sciences* **2012,** *109* (3), 793-798.
30. Kjaergaard, M.; Poulsen, F. M., Sequence correction of random coil chemical shifts: correlation between neighbor correction factors and changes in the Ramachandran distribution. *J Biomol NMR* **2011,** *50* (2), 157-65.

31. Wishart, D. S.; Sykes, B. D.; Richards, F. M., The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* **1992,** *31* (6), 1647-51.
32. Wishart, D.; Bigam, C.; Holm, A.; Hodges, R.; Sykes, B., Random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *Journal of Biomolecular NMR* **1995,** *5* (1), 67-81.
33. Bradley, E. K.; Thomason, J. F.; Cohen, F. E.; Kosen, P. A.; Kuntz, I. D., Studies of synthetic helical peptides using circular dichroism and nuclear magnetic resonance. *J Mol Biol* **1990,** *215* (4), 607-22.
34. Barlow, D. J.; Thornton, J. M., Helix geometry in proteins. *Journal of Molecular Biology* **1988,** *201* (3), 601-619.
35. Cierpicki, T.; Otlewski, J., Amide proton temperature coefficients as hydrogen bond indicators in proteins. *Journal of Biomolecular NMR* **2001,** *21* (3), 249-261.
36. Wagner, G.; Pardi, A.; Wuethrich, K., Hydrogen bond length and proton NMR chemical shifts in proteins. *Journal of the American Chemical Society* **1983,** *105* (18), 5948-5949.
37. Jimenez, M. A.; Bruix, M.; Gonzalez, C.; Blanco, F. J.; Nieto, J. L.; Herranz, J.; Rico, M., CD and 1H-NMR studies on the conformational properties of peptide fragments from the C-terminal domain of thermolysin. *European journal of biochemistry / FEBS* **1993,** *211* (3), 569-81.
38. Rieping, W.; Habeck, M.; Bardiaux, B.; Bernard, A.; Malliavin, T. E.; Nilges, M., ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* **2007,** *23* (3), 381-2.
39. Keeling, K. L.; Cho, O.; Scanlon, D. B.; Booker, G. W.; Abell, A. D.; Wegener, K. L., The key position: influence of staple location on constrained peptide conformation and binding. *Org Biomol Chem* **2016,** *14* (41), 9731-9735.

Chapter 3

Linker Type

3.1 Comparison of α-Helical Side Chain Linkers

Comparative studies on peptides containing different side chain linkers provide an opportunity to streamline the design of helical peptides for biological applications. To date, studies have compared different types of functional groups joining peptide side chains based on the resulting peptides helicity. A study by Frankiewicz et al[1] compared four peptides with linkers connecting the *i* and *i*+4 residues of sequences derived from vasoactive intestinal peptide (VIP). These peptides were linked by amide bond formation (lactam KE linker, A(ii)), Huisgen azide-alkyne cycloaddition (triazole linker, B) and olefin ring closing metathesis (saturated and unsaturated diether linkers, C and D respectively) (Figure 3-1). An analysis of the circular dichroism (CD) of each revealed that the triazole linked peptide had the greatest α-helical content of the four peptides.

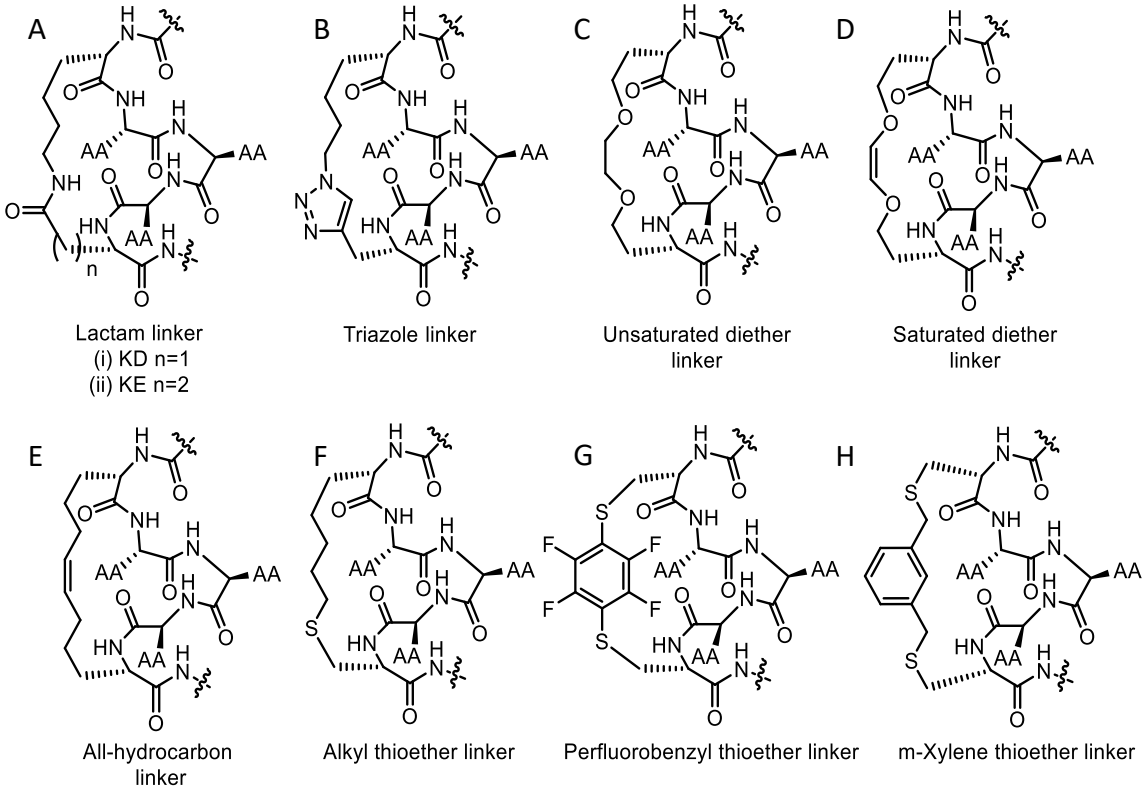


Figure 3-1: Side chain-side chain linkers, **A** (i) lactam (KD) (ii) lactam (KE), **B** triazole, **C** unsaturated diether, **D** saturated diether, **E** all-hydrocarbon, **F** alkyl thioether, **G** perfluorobenzyl thioether (PF benzyl) and **H** m-xylene thioether, which link *i* and *i*+4 residue side chains.

Studies of individual linker types which compared length and positioning[2-7] indicate that the linkers incorporated by Frankiewicz et al[1] were not optimal for generating peptides the greatest α-helical content. Linker types later investigated by de Araujo et.al[8] were selected from these optimised linkers. Model peptides were synthesised with lactam (KD) (A(i)), triazole (B), all-hydrocarbon (E), alkyl thioether (F), perfluorobenzyl thioether (PF benzyl, G) and m-xylene thioether (H) linkers, see Figure 3-1. The linkers joined the *i* and *i*+4 residues in pentapeptides with alanine residues between linking residues, and in 18-residue peptides. The 18-residue sequence contained predominantly alanine, with addition of arginine residues to ensure the peptides would be soluble in water. Analysis of CD and NMR spectra of these peptides showed that the lactam-linked pentapeptide (A(i)) had the greatest α-helical content, with the all-hydrocarbon (C) and triazole (B) linked peptides displaying slightly lower helicity. The other peptides showed comparatively lower helical content. The longer 18 residue peptides containing lactam (A(i)), all-hydrocarbon (C) and triazole (B) linkers gave similar results, with the percent helicity for the lactam linked peptide greater than that for peptides with the triazole and all-hydrocarbon linkers. However, it is unknown if these findings are transferable to other peptide sequences. Lactam, triazole and all-hydrocarbon linked peptides based on the β3 integrin sequence were therefore prepared in order to see if the lactam bridge also displayed the highest helical content in these cases.

3.2 Cellular Uptake

Currently available therapeutic peptides predominantly target extracellular proteins, rather than intracellular proteins. However, Verdine and co-workers[9] have shown that, for some peptide sequences, the introduction of an all-hydrocarbon linker causes the peptide to be taken up into cells. Current data suggests that increasing peptide positive charge also enhances cellular uptake, with a lesser role apparent for other peptide attributes, such as helicity and specific residues[10]. Investigations of all-hydrocarbon linked peptides conclude that cellular uptake is facilitated by the hydrophobicity of the component linker[10]. Cellular uptake of these peptides likely occurs through a clathrin-

and caveolin-independent endocytosis pathway that involves, in part, sulfated cell surface proteoglycan[11].

Peptides containing a PF benzyl linker are reported to show comparable cellular uptake properties to all-hydrocarbon linked peptides[12 13]. Results of fluorescence-activated cell sorting (FACS) analysis and confocal images of HEK293T cells treated with fluorescently-tagged macrocyclic peptides suggest that peptides containing all-hydrocarbon and PF benzyl linkers display similar and favourable cellular uptake properties[12]. However, other peptide linkers are yet to be included in such comparative studies, studies that are required to fully exploit linked-peptide therapeutic potential. The cellular uptake of peptides that incorporated lactam, triazole and all-hydrocarbon linkers, based on the β3 integrin sequence, were therefore investigated and compared in this thesis.

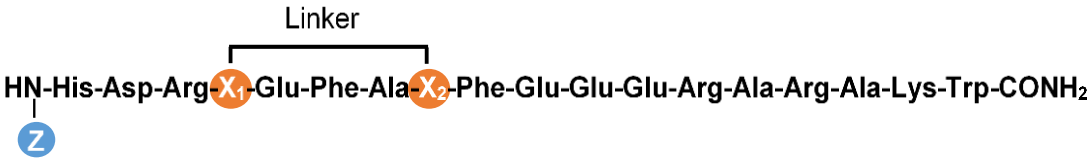
3.3 Design of Peptides for Comparative Studies of Helicity and Cellular Uptake

The ability of different linker types to induce helicity and influence cellular uptake has rarely been compared using a single peptide sequence, due in part to difficulties in optimising cyclisation protocols for each linking strategy[14]. The report by de Araujo et al[8], discussed in section 3.1, compared the helical content of model peptides containing common side chain linkers. This study found that lactam (KD), all-hydrocarbon and triazole linkers (Figure 1) produced highly helical peptides, with the lactam linker producing the greatest α-helical content[8]. However, it is not known if the lactam linker confers 'optimal' helicity in other sequences, particularly biologically relevant sequences. In addition, the effect of these linkers on cellular uptake was not investigated. In this chapter, four linkers were selected for comparison with peptides incorporating these linkers investigated for their helical content and cellular uptake. The selected linkers were either chosen for their ability to confer high helical content or allow for cellular uptake. Lactam, triazole and all-hydrocarbon linkers were selected based on the high helical content displayed for linked peptides in the study by de Araujo et al[8]. All-hydrocarbon and PF benzyl linkers were selected as peptides incorporating these linkers showed positive cellular uptake properties in previous studies[12 13]. The sequence selected for this study was the membrane proximal region of the β3 integrin

tail, as it may be used to target an intercellular protein, talin, and is therefore biologically relevant. This sequence is also very different to previously studied sequences as it contains few alanine residues.

The sequence selected from the β3 integrin tail contained 18 residues spanning from H722 to W739. The native sequence is shown for the linear peptide: β3MP(722-739) (Table 3-1, Entry 1). Linkers were incorporated between the side-chains of residues 725 and 729, as this was identified to be the optimal lactam linkage site based on helicity and affinity for the talin F3 domain (see Chapter 2). The target peptides for linker analysis were thus: linear peptide, β3MP(722-739), lactam linked peptide, **7**, triazole linked peptide **10**, and all-hydrocarbon linked peptide, **11** (Table 3-1, Entries 1-4).

Table 3-1: Peptides used in the current study, derived from the β3 integrin cytoplasmic tail. Residues 725 and 729 represented by X_1 and X_2 respectively (orange). The presence or absence of an N-terminal fluorescent label is indicated by Z, where Z is carboxyfluorescein (CF).



Peptide	Linker	Z	X_1	X_2
β3MP(722-739)	None	H	Lys	Lys
7	Lactam	H	Lys	Asp
10	Triazole	H	Nle (εN_3)[a]	Pra[b]
11	All-hydrocarbon	H	S(5)[c]	S(5)[c]
Fβ3MP(722-739)	None	β Ala- CF	Lys	Lys
F7	Lactam	β Ala- CF	Lys	Asp
F10	Triazole	β Ala- CF	Nle (εN_3)[a]	Pra
F11	All-hydrocarbon	β Ala- CF	S(5)[c]	S(5)[c]

a 6-azido L norleucine, b propargylglycine,
c (S)-2-(((9H–fluoren-9-yl)methoxy)carbonylamino)-2-methyl-hept-6-enoic acid

The effect of linker type on peptide cellular uptake was investigated using fluorescent tagged (carboxyfluorescein) linked peptides. Tagged peptides included: linear peptide, Fβ3MP(722-739), lactam linked peptide, **F7**, triazole linked peptide, **F10**, and all-hydrocarbon linked peptide, **F11**, (Table 3-1, entries 5-8).

Comparison between the cellular uptake of these peptides and their helical structure will help determine which of helical structure or linker chemistry has the greater influence on cellular uptake. This detailed analysis of peptide helicity and cellular uptake is expected to streamline the choice of strategies for future development of helical peptide therapeutics.

3.3.1 Synthesis of Triazole, All-Hydrocarbon and Lactam Linked Peptides, **7**, **10** and **11**

Peptides containing lactam, triazole and all-hydrocarbon linkers, **7**, **10** and **11** respectively in Table 3-1 (entries 2-4), with linkers shown in Figure 3-2B, were synthesised to answer the first of the questions posed in section 3.3: "Which of the linkers would result in the greatest helical structure in a biologically relevant sequence?"

Peptides β3MP(722-739), **7**, **10** and **11** (see Table 3-1 for the sequences) were synthesised using Fmoc solid phase peptide synthesis protocols[15] detailed in the experimental section (4.3). The synthesis of β3MP(722-739) and cyclisation of peptide **7** is described in 2.3. Cyclisation of peptide **10** was carried out in solution after cleavage from resin. The formation of the triazole linker in peptide **10** was achieved using Huisgen cycloaddition of the azide and alkyne groups on the side chains of residues Nle(εN3) and Pra (Figure 3-2A) with copper (II) sulfate pentahydrate as the catalyst. Peptide **11** was cyclised via ring closing metathesis (RCM). Fmoc protected, resin bound peptide (D(723)-W(739)) was treated with Grubbs second generation catalyst, lithium chloride, and heated using microwave irradiation[16]. Cyclisation was completed after two cycles of microwave irradiation. Deprotection of the N^{α}-Fmoc group from the resin-bound cyclic peptide and cleavage from the resin gave peptide **11**. The highest yield of peptide **11** was obtained when the RCM was performed on a 0.05 mmol scale, with lithium chloride added to the reaction mixture. Lithium chloride was used to disrupt unfavourable hydrogen bonding interactions that would be expected to hinder the formation of the complex between alkene groups of (S)-2-(((9H–fluoren-9-yl)methoxy)carbonylamino)-2-methyl-hept-6-enoic acid (S5) residues (Figure 3-2A) and the ruthenium catalyst[17, 18], which resulted in lower levels of cyclic product (Table 3-1, entry 4). Peptides **7**, **10** and **11** were characterised using mass spectrometry and their

purity determined to be greater than 95% using analytical HPLC (see 4.2.9and 4.3 for experimental details).

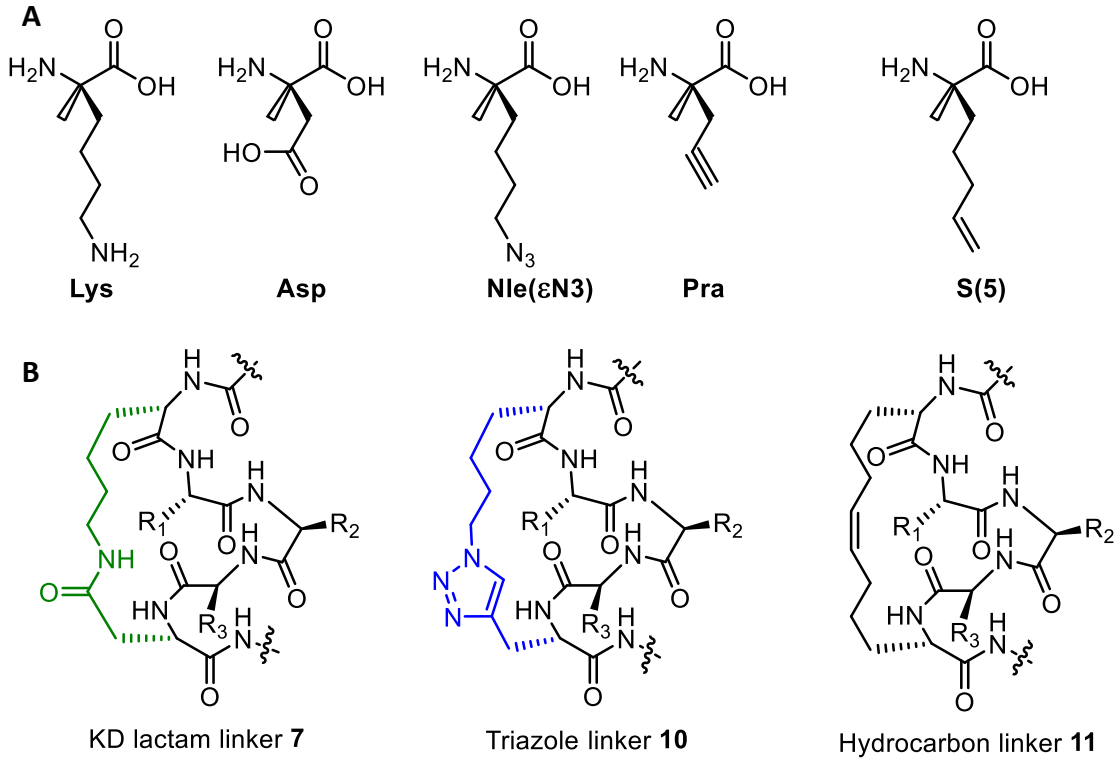


Figure 3-2 (A) Amino acids, Lys and Asp, Nle(εN3) and Pra, S(5) and S(5), incorporated at i and i+4th positions to allow for cyclisation to produce linked peptides **7**, **10** and **11** respectively. (B) The linked section of the peptides produced after cyclisation of i and i+4th residue side chains for peptides **7**, **10** and **11**, where R_1=Glu, R_2=Phe and R_3=Ala.

3.3.2 Structural Comparison of Triazole, All-Hydrocarbon and Lactam Linked Peptides **7**, **10** and **11**.

Circular dichroism spectra of triazole (**7**), all-hydrocarbon (**10**), and lactam (**11**) linked peptides were recorded in phosphate buffer (Figure 3-3) to determine which constraint induced the most helical structure. The experimental parameters used in CD measurements were obtained from previous work on the characterisation of short helical peptides reported by Shepherd et al.[19]. The maximum molar ellipticity was determined using the equation: $[\theta]_{max} = (-44\,000 + 250T)(1 - k/n)$, where temperature (T) was 298K, number of non-bonded carbonyls (k) was 4, and number of residues (n)

was 18. The ratio of calculated molar ellipticity to measured maximum molar ellipticity (λ = 222 nm) was used to determine percent helicity (detail provided in 4.1.2.2).

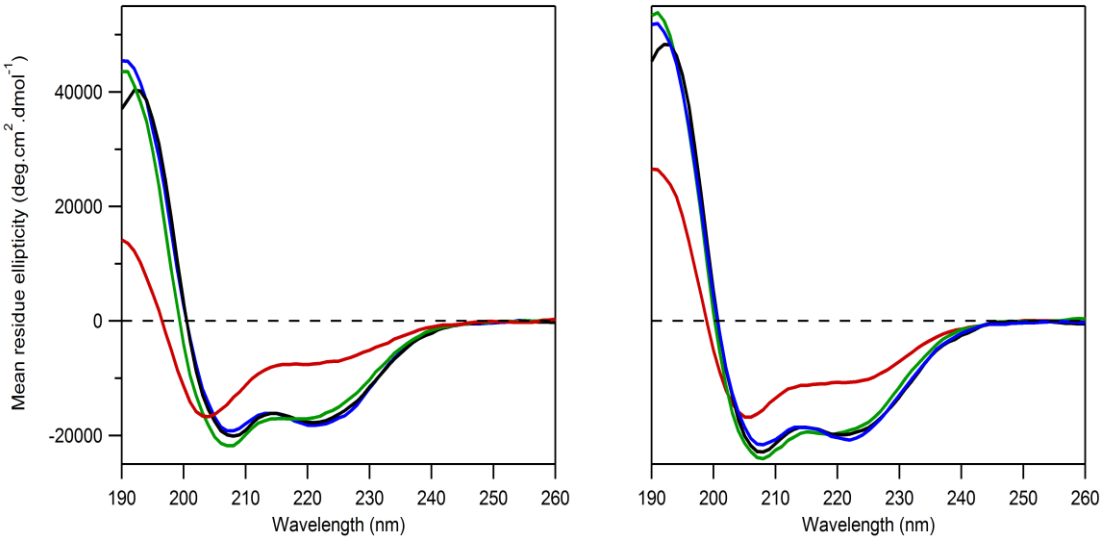


Figure 3-3: CD spectra for β3MP(722-739) (red), lactam linked **7** (green), triazole linked **10** (blue) and all-hydrocarbon linked **11** (black) illustrating linker effect on helical content (left: 10 mM phosphate buffer, pH 6.5, 298K, right: 10 mM phosphate buffer, 10% TFE, pH 6.5,

All linked peptides **7**, **10** and **11** displayed markedly greater percent α-helicity (58-62%) than the linear peptide, β3MP(722-739) (23%), indicating all linkers were able to stabilise helical structure as expected. α-Helical content of linked peptides increased in the order: lactam-linked **7** (57%) < all-hydrocarbon linked **11** (60%) < triazole linked **10** (62%) (Figure 3-3, left, Table 3-2), indicating the triazole linker induced the most helicity of all the linkers.

Table 3-2: Molar ellipticity [θ] (deg cm^2 $dmol^{-1}$ $residue^{-1}$) at 222 nm, percentage helicity and ratios of ellipticity at 222 nm/208 nm for β3MP(722-739) peptides, **7**, **10** and **11** determined from CD data acquired in 10 mM phosphate buffer (pH 6.5, 25°C).

Peptide	$[\theta]_{222}$[a]	Helicity[b]	$[\theta]_{222}/[\theta]_{208}$[c]
β3MP(722-739)	-8106.49	23%	0.55
7	-13121.9	57%	0.76
10	-18123.6	62%	0.95
11	-17622.2	60%	0.88

[a] Molar ellipticity [θ] (deg cm^2 $dmol^{-1}$ $residue^{-1}$) at 222 nm, acquired from CD data

[b] percent helicity

[c] ratios of ellipticity at 222 nm/208 nm

The ratio of molar ellipticity at 222 nm to that at 208 nm ($[\theta]_{222}/[\theta]_{208}$), is often used as a measure of helix ideality, with a value of ~1 indicating optimal α-helical structure[20]. The deeper minima at 208 nm compared to 222 nm, seen in the CD spectrum of peptide **7** in particular, suggests the peptides contain a mixture of α-helical structure and 3_{10} helical structure or unstructured regions. The greatest ratio of $[\theta]_{222}/[\theta]_{208}$ was 0.95 for the triazole-linked peptide **10,** making it the peptide closest to ideal α-helical structure. Next was, the all-hydrocarbon-linked peptide **11** (0.88), followed by the lactam peptide **7** (0.76) (Table 3) i.e. the same order observed for percent helicity.

To determine if the peptides were demonstrating their maximal helicity, 10% TFE was added to the solution as it is a known helix inducing solvent[21]. All peptides showed an increased helicity with addition of 10% TFE (Figure 3-3, right, Table 4-1). This suggests that helical content has not been maximised by any of these constraints, likely due to their long length (18 residues), and suggests that additional constraints along the length of the peptides may lead to a further increases in helical structure for all linker types.

The results of this study suggest peptide **10**, containing the triazole linker, displayed the greatest helical content, followed by **11** (all-hydrocarbon linker), and then **7** (lactam linker). In contrast, de Araujo et al.[8] study suggest the lactam linked peptides display the greatest helicity when compared to all-hydrocarbon and triazole linked peptide, for model pentapeptides and longer 18-mer peptides. The discrepancy between the 18-

mer peptide synthesise by de Araujo et al.[8] and the 18 residue peptides **7**, **10**, **11** is likely due to the differences in peptide sequence; the present study uses a biologically relevant sequence, consisting of charged and polar residue, while the de Araujo et al.[8] peptides were mostly comprised of alanine residues. This suggests that the specific peptide sequence plays an important role in determining the 'optimal' linker to nucleate helicity. Therefore, it cannot yet be generalised that either lactam or triazole linkers will nescessarily produce peptides with greatest helical content. More general conclusions require further study into sequence effects and linker types on the helical content of peptides. The integrin sequence may in fact favour a particular form of helical structure, which varies from the ideal α-helix, specific for its native protein-protein interaction with talin.

All three linkers, lactam, all-hydrocarbon and triazole, produce peptides with consistently high helicity in both this study and the study performed by de Araujo et al.[8]. Producing a peptide with the highest helical content possible may not be of utmost importance in all studies using helical peptides. Other favourable characteristics of the linked peptide may be required depending on their application. One such desirable quality is cellular uptake, which was therefore investigated in the following section, using fluorescently tagged peptides.

3.3.3 Synthesis of Fluorescent Peptides **F7**, **F10** and **F11** with Triazole, All-Hydrocarbon and Lactam Linkers

The 18-residue peptides derived from the integrin tail, containing lactam, triazole and all-hydrocarbon linkers, were tagged with carboxyfluorescein to produce Fβ3MP(722-739), **F7**, **F10** and **F11** (Table 3-1, entries 5-8, general structure shown in Figure 3-4), in order to address the influence of the linker, and induced helicity, on cellular uptake.

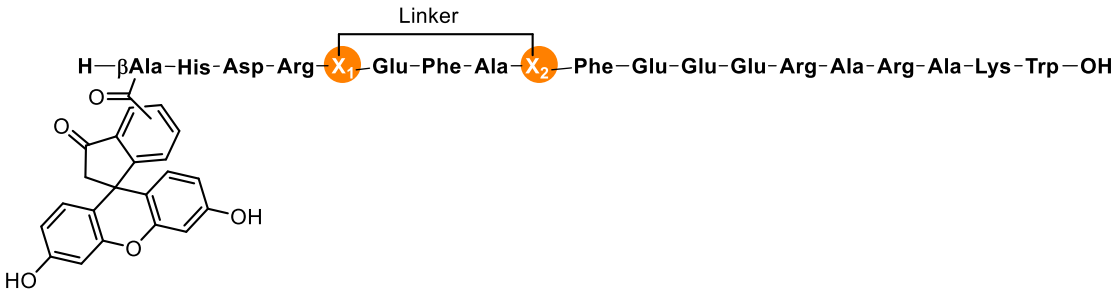


Figure 3-4: General structure for carboxyfluorescein tagged peptides, Fβ3MP(722-739) **F7**, **F10** and **F11**. Residues X_1 and X_2 identity as displayed in Table 3-1.

The fluorescently tagged peptides Fβ3MP(722-739) **F7**, **F10** and **F11** (Table 3-1, entries 5-8 for the sequences) were synthesised using Fmoc solid phase peptide synthesis protocols[15] incorporating carboxyfluorescein (Figure 3-4) as detailed in the experimental section (4.3). Linkers were produced via cyclisation as previously described for peptides **7**, **10** and **11** to give peptide **F7**, **F10** and **F11** respectively. Additionally, a positive control, CF-NYAD-1, was synthesised for the comparison of confocal microscope images in cellular uptake studies. This peptide was based on NYAD-1, a peptide previously used to compare cellular uptake of linked peptides by Spokoyny [12]. The addition of the carboxyfluorescein group was achieved by the coupling of Fmoc-β-alanine-OH followed by carboxyfluorescein to the *N*-terminus of the peptide synthesised as described by Zhang et.al[13]. These peptides were characterised by mass spectrometry and their purity confirmed to be greater than 95% by HPLC.

3.3.4 Cellular Uptake of Triazole, All-Hydrocarbon and Lactam Linked Peptides

To determine cellular uptake, tagged peptides Fβ3MP(722-739), **F7, F10** and **F11** as well as controls CF and CF-NYAD-1 (5 μM) were separately added to NIH 3T3 mouse embryonic fibroblast cells. The cells were incubated with peptides for 3 hrs, after which the cells were washed to remove peptide remaining in the media. Samples were trypsinised to remove any peptide on the cell surface. These cells were then fixed to glass slides and imaged using confocal microscopy. Confocal images of peptides **F10** and **F11** displayed clear, green punctate fluorescence, consistent with the peptides being localised in endosomes within cells (Figure 3-5). Lower levels of green punctate fluorescence are seen in images of Fβ3MP(722-739) and **F7**, suggest low levels of cellular uptake for both these peptides. Negative control cells, treated with

carboxyfluorescein (CF) alone, displayed low levels of background fluorescence. Positive control cells, treated with CF-NYAD-1, displayed punctate fluorescence similar to cells treated with **F10** and **F11**. However, cellular uptake was significantly greater for CF-NYAD-1, which displayed higher intensity fluorescence in confocal images.

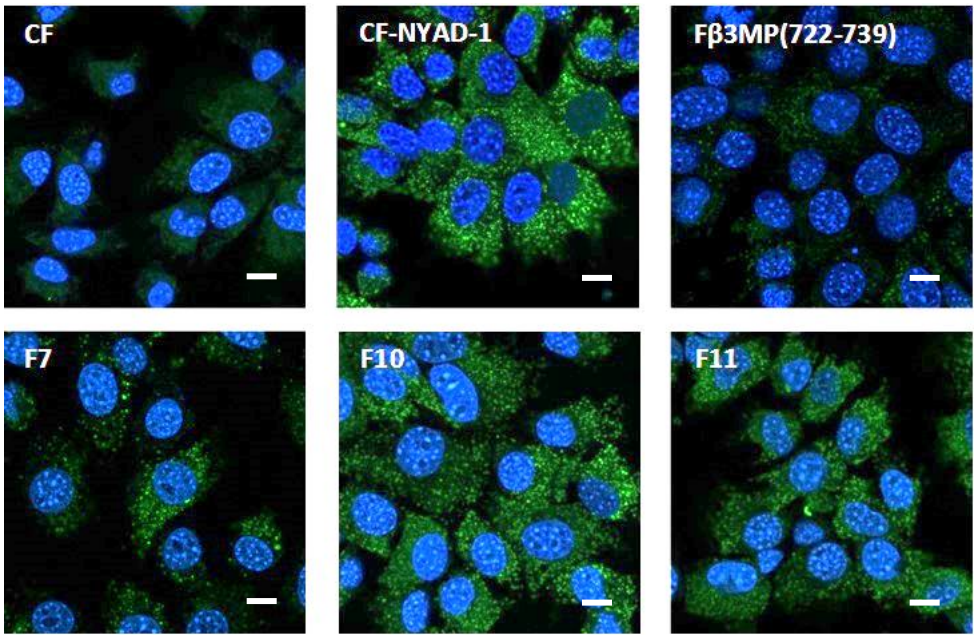


Figure 3-5: Fluorescent confocal microscopy images of mouse NIH3T3 cells treated with CF (carboxyfluorescein, green fluorescence), CF-NYAD-1, FβMP(722-739), **F7**, **F10** and **F11**. Cells were stained with DAPI nuclear stain (blue fluorescence). Scale bars represent 10µm.

To verify cellular uptake of the peptides was not limited to one cell line, cellular uptake was also investigated in human HeLa cells, an immortal cell line derived from human cervical cancer cells. Cells were treated with peptides **F10**, **F11**, **CF**, and CF-NYAD-1, using the same protocols as previously described for the cellular uptake studies using NIH 3T3 cells. Confocal images of treated HeLa cells showed similar results to those of mouse NIH 3T3 cells, however HeLa cell images displayed higher background fluorescence, as seen in images of the negative control cells treated with **CF** (Figure 3-6). Fluorescence in cells treated with peptides **F10** and **F11** was clearly above background, and was observed as green punctate fluorescence, once again likely due to the peptide being localised in cellular endosomes.

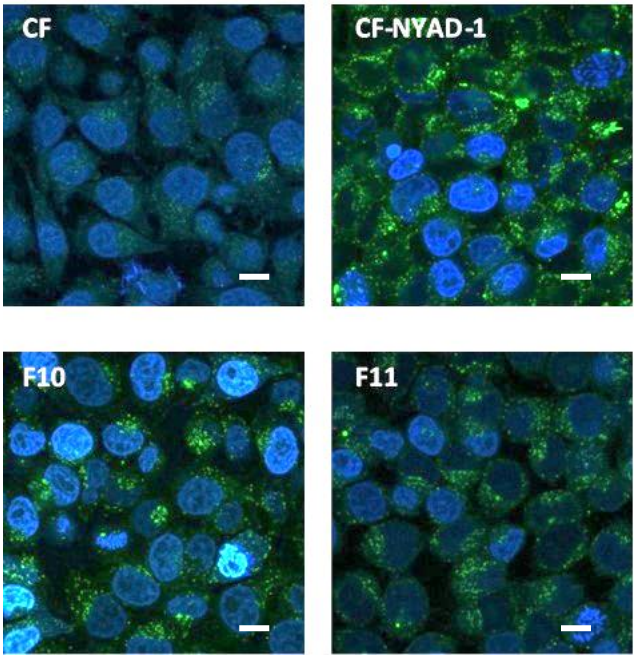


Figure 3-6: Fluorescent confocal microscopy images of HeLa cells treated with: CF (carboxyfluorescein), CF-NYAD-1, **F10** and **F11**. Scale bars represent 10µm.

The mean cellular fluorescence was used to quantify the extent of cellular uptake for each peptide (Figure 3-7) using fluorescence activated cell sorting (FACS), a specialised type of flow cytometry. Tagged peptides Fβ3MP(722-739), **F7, F10** and **F11** (5 µM) were each added to NIH 3T3 mouse embryonic fibroblast cells. The cells were incubated with the peptides for 3 hrs, after which the cells were washed to remove peptide remaining in the media. Samples were trypsinised as described above to remove any peptide on the cell surface. Peptide **F11** with the all-hydrocarbon linker was found to have the greatest mean fluorescence indicating highest cellular uptake. **F10** displayed lower cellular uptake than **F11**, but greater cellular uptake than **F7**, suggesting that linker type influences cellular uptake of peptides (Figure 3-7).

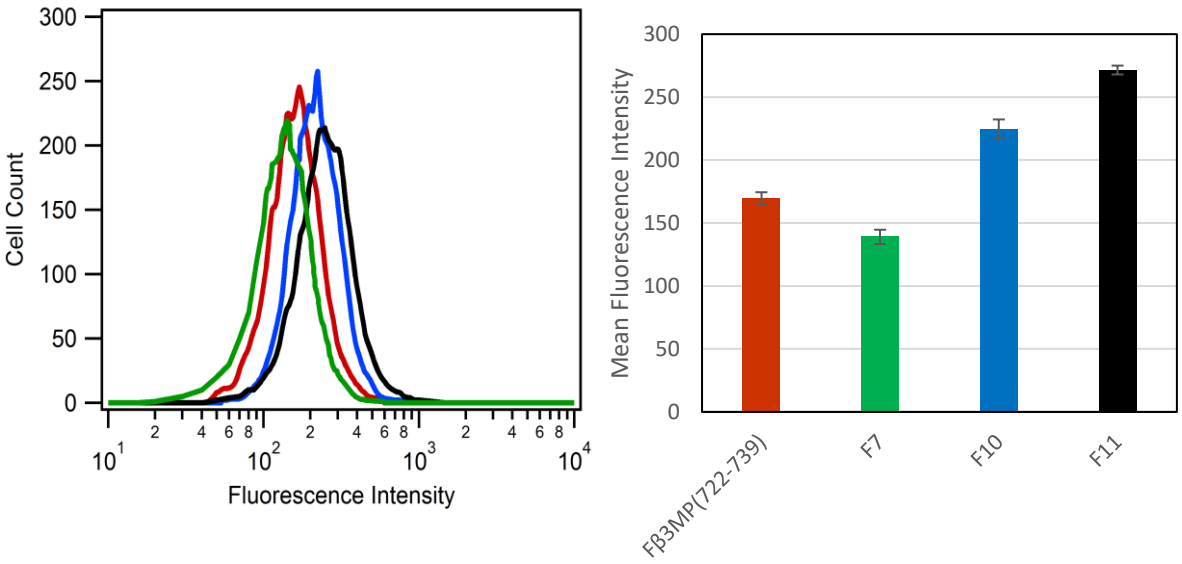


Figure 3-7: (Left) Representative graph of FACS analysis of NIH 3T3 cell treated with Fβ3MP(722-739) (red), **F7** (green), **F10** (blue), **F11** (black) (5µM), indicating the number of cells (cell count) with each level of fluorescence intensity, and (Right) Histogram displaying mean fluorescence from two experiments displaying errors bars indicating standard deviation.

The basis of the differences between the cellular uptake of **F11**, **F10** and **F7** is likely related to the linker properties and peptide helicity. In terms of helicity dictating the levels of cellular uptake, the triazole linked peptide (**F10**) should display the greatest cellular uptake followed by the all-hydrocarbon linked peptide (**F11**) and then the lactam linked peptide (**F7**) according to helical content as determined by CD spectroscopy of the untagged peptides. The lactam linked peptide **F7** does have the lowest cellular uptake of the three linked peptides, consistent with this peptide also having the lowest helical content, however **F11** displays greater cellular uptake than **F10**, with **10** displaying greater helical content than **11**. This suggests that the extent of helicity alone does not explain the observed cellular uptake trend. It is possible the polarity of the linker plays a role in dictating the levels of cellular uptake[22], with the all-hydrocarbon linker displaying the lowest levels of polarity, compared to the triazole and amide linkers. However the triazole is likely less hydrophobic than the amide[23], with the amide displaying a dipole moment of 3.5 D[24] while the triazole displays a dipole moment of 5 D[25]. Alone hydrophobicity of the linker and peptide helicity cannot explain the pattern of cellular uptake, with **F11** showing the greatest cellular uptake followed by

F10 and then **F7** (Figure 4), suggesting the linker polarity and peptide helicity both play a role in cellular uptake.

Previous studies suggest a positive correlation between peptide helicity and cellular uptake[26]. Interestingly, Fβ3MP(722-739) displayed greater cellular uptake than **F7**. This is unexpected as untagged peptide **7** with the lactam linker displayed the greater percentage helicity (57%) than the linear peptide β3MP(722-739) (23%), as identified in CD analysis and confirmed by NMR studies (2.3.5). The cause of lower levels of cellular uptake of **F7** in comparison to Fβ3MP(722-739) may be due to the difference in overall peptide charge. The original linear peptide sequence (overall charge +2) contains two free positively charged lysine residues, which are modified to neutral linkers in peptides **F7**, **F10** and **F11**. Increasing the number of positively charged residues in a peptide has previously been associated with greater cellular uptake of the peptides[11]. This result suggests the presence of the positively charged lysine residues in the native sequence, Fβ3MP(722-739) , had a greater impact on cellular uptake than the incorporation of a lactam bridge, despite **F7** having the greater helical content. As such the effect of charge of the peptide on cellular uptake was investigated in the following section.

3.3.5 Investigation into the Effect of Peptide Charge on Cellular Uptake of the β3 Integrin Tail

A panel of short 14-mer peptides derived from the β3 integrin sequence H(722)-A(735) (2.3.2) were synthesised with hydrocarbon linkers incorporated to investigate the effect of peptide charge on cellular uptake. Overall peptide charge was varied by altering the peptide sequence based on a study by Bernal et al.[27]. Bernal et al.[27] reported a short peptide with an all-hydrocarbon linker which did not display cellular uptake, which was attributed to the negative charge of the peptide. The authors synthesised a panel of peptides in which negatively charged residues were substituted with either neutral or positively charged side chains. Peptides with an overall neutral or positive charge displayed cellular uptake, while the original sequence (overall negative charge) did not[27].

The 14-mer β3 integrin sequence H(722)-A(735) was used as the basis for fluorescently tagged peptides **F13**, **F14**, **F15** and **F16** (Table 3-3). Peptide **F13** (Table 3-3, entry 1) has

an overall charge of -2 and consisted of the native residues, apart from residues at positions 725 and 729 (assigned the letter X in Table 3-3) that were modified to incorporate the hydrocarbon linker. Negatively charged residues were substituted with either neutral or positively charged residues for peptides **F14**, **F15** and **F16**, except for residues E726 and E733 as they are likely involved in charge related interactions within the peptide, or with talin (2.3). Peptide **F14** was designed with all non-critical negatively charged residues replaced with either glutamine or asparagine, resulting in a peptide with +1 charge. These residues in peptide **F15**, were substituted with arginine to further increase the overall charge to +4. To assess whether changes in peptide uptake were due to arginine residues being located at specific sites, **F16** was synthesised for comparison, where the arginine residues were added to the *N*-terminus and non-critical negatively charged residues in the sequence were replaced with either glutamine or asparagine, giving a peptide with the same +4 charge.

Peptides **F13**, **F14**, **F15** and **F16** were synthesised manually from A(735) to H(722) incorporating non-natural amino acid **S5** (Figure 3-2) in the *i* and *i*+4th positions, shown as residue 'X' in Table 3-3. Resin bound Fmoc protected peptides were cyclised by ring closing metathesis, as described for peptide **10**. After cyclisation, β-alanine and carboxyfluorescein were coupled to the resin bound peptide. Peptides were then cleaved from the resin to produce **F13-16**., which were characterised by mass spectrometry and analytical HPLC.

Table 3-3: Sequence and charge for peptides **F13**, **F14**, **F15** and **F16**. All-hydrocarbon linkers are incorporated, *C*-terminal amidated and *N*-terminal β-alanine CF tag.

Peptide	Sequence	Charge
F13	HDRXEFAXFEEERA	-2
F14	HNRXEFAXFQQERA	+1
F15	HRRXEFAXFRRERA	+4
F16	RRRHNRXEFAXFQQERA	+4

NIH 3T3 cells were treated with fluorescein labelled peptides, **F13**, **F14**, **F15** and **F16**, and incubated for 3 hrs before cells were rinsed and trypsinised to remove peptides bound to the cell surface. Cells were then analysed using FACS and results for mean fluorescence intensity were used to infer peptide levels within cells (Figure 3-8). Cells treated with negatively charged peptide **F13** (-2) and positively charged peptide **F14** (+1) displayed no significant difference in mean fluorescence intensities. This suggests that negative charge was not the major contributing factor to low cellular uptake of **F13**. Cells treated with **F16** had the highest mean fluorescence intensity of the 4 samples. Cells treated with **F15** displayed greater cellular uptake than either **F14** or **F13**, however mean fluorescence intensity was much less than that of **F16**. This result was unexpected as **F15** and **F16** had the same overall charge. These results suggest that location of positively charged residues affects cellular uptake of peptides, and that the positive charge grouped at the *N*-terminal was more effective for increasing cellular uptake of a peptide sequence than just increasing overall peptide charge through sequence modifications.

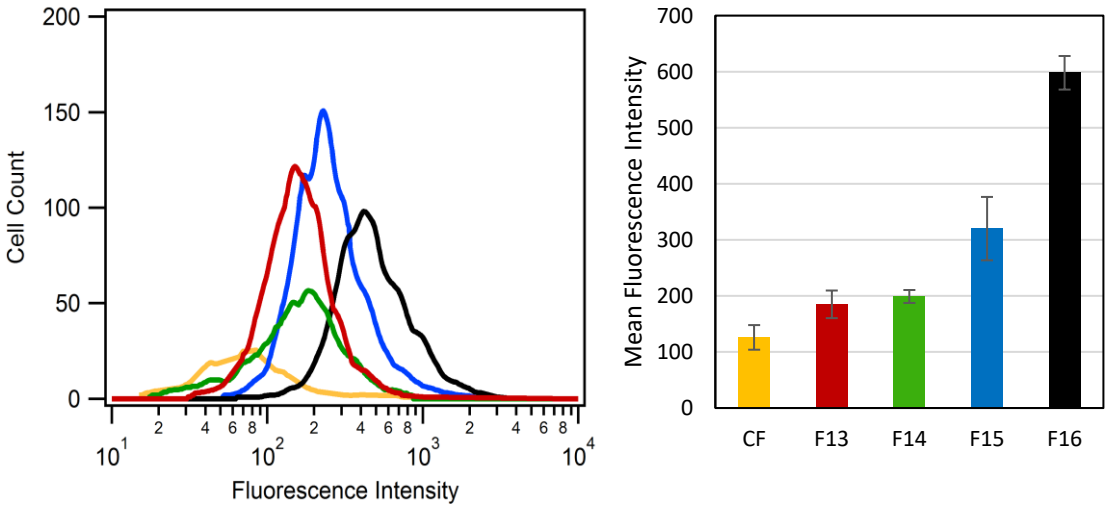


Figure 3-8: (Left) Representative graph of FACS analysis of NIH 3T3 cell treated with carboxyfluorescein (CF, yellow) **F13** (red), **F14** (green), **F15** (blue) and **F16** (black) (5µM), and (Right) Histogram displaying mean fluorescence from three experiments displaying errors bars indicating standard deviation. Comparison between all measurements using a two-tailed t-test indicated a $p>0.05$, except for F13 with F14.

Confocal images taken of NIH 3T3 cells treated with **F13**, **F14** and **F16** displayed green puncta fluorescence within cells, confirming peptides were entering the cells (Figure 7). The intensity of fluorescence in confocal images were consistent with the data from FACS analysis.

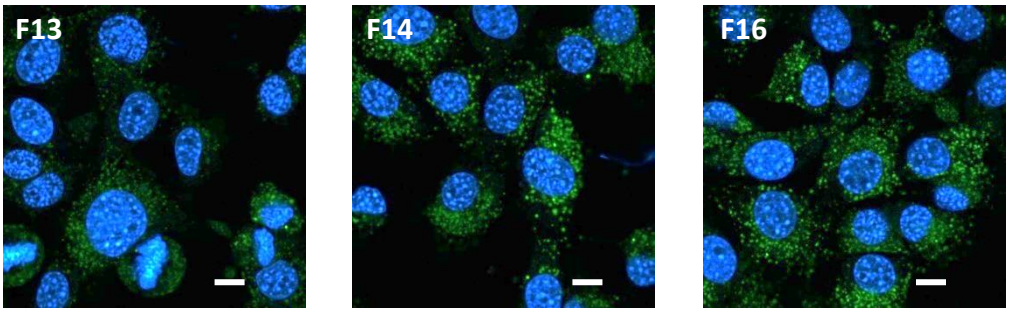
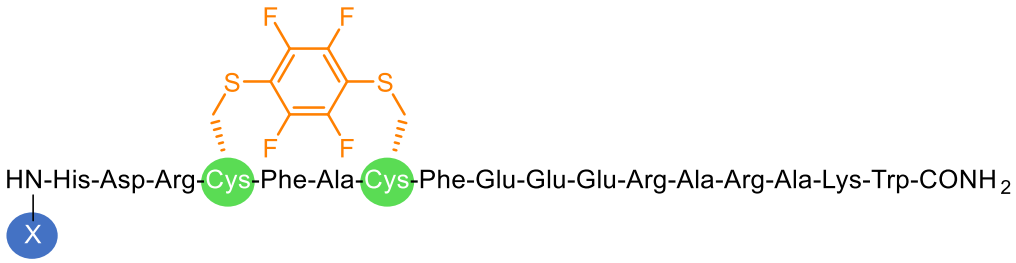


Figure 3-9: Representative fluorescent confocal microscopy images of mouse NIH 3T3 cells treated with **F13**, **F14** and **F16**. Scale bars represent 10µm.

3.3.6 Design and Synthesis of Perfluorobenzyl Thioether (PF benzyl) Linked Peptides

The final study aimed to determine the effect of PF benzyl linkers on integrin peptide helical content and cellular uptake, as PF benzyl linked peptides were identified as peptides displayed high cellular uptake in a report by Spokoyny et al. [12]. In the work by Spokoyny et al. [12] the PF benzyl linker was found to induce helical structure, and produced a peptide that displayed cellular uptake comparable to that of the equivalent all-hydrocarbon linked peptide. However, the subsequent study by de Araujo et al.[8] found a pentapeptide peptide containing the PF benzyl linker was not highly helical, when compared to lactam, all-hydrocarbon and triazole linked peptides.

PF benzyl linked peptides **12** and **F12** (Figure 3-10) were synthesised via Fmoc solid phase peptide synthesis using an automated peptide synthesiser and characterised by mass spectrometry and HPLC. Cysteine residues incorporated at positions 725 and 729 (*i* to *i*+4) were linked to produce the PF benzyl linkers in solution on treatment with hexafluorobenzene. Nucleophilic aromatic substitution at positions 1 and 4 of hexafluorobenzene was achieved via reaction with the cysteine residue sulfhydryl groups (Figure 3-10).



12: X=H **F12**: X=βAla-CF

Figure 3-10: PF benzyl linker incorporated in Peptides **12** where X=H and **F12** where X= βAla-CF. The PF benzyl linker is highlighted in orange.

The cyclisation reactions were low yielding (less than 2%) due to precipitation of the peptides on addition of tris base. The low solubility of the peptides in DMF could be overcome by addition of water, however this also stopped the reaction with hexafluorobenzene. Additionally, peptides were purified before and after cyclisation by semi-prep HPLC which likely further reduced peptide yields. In an attempt to improve the amount of peptide recovered after cyclisation, on resin cyclisation was attempted using a previously described protocol involving selective removal of cysteine protecting groups (Mmt) and cyclisation with 1,4-dibromo-2,3,5,6-tetrafluorobenzene and DIPEA in DMF[8]. However, after the on-resin cyclisation, masses for the expected cyclic products were not identified in the LCMS spectrum, indicating this cyclisation protocol was not effective for PF benzyl linked peptides.

3.3.7 Structural Analysis of Perfluorobenzyl Thioether Linked Peptide

The CD spectrum of peptide **12** (Figure 3-11, left) displays features characteristic of helical structure. A maximum around 190 nm is a clear indicator of helical structure. The minimum at 205 nm, is slightly shifted from the expected 208 nm, providing an indication of low levels of helical structure. The shallow minimum observed at 222 nm would normally suggest low helical content, however, Spokoyny et al. [12] identified an absorbance maximum around 222 nm in the CD spectrum[12] for peptides containing PF benzyl linkers due to the presence of the C_6F_4 group. Therefore, absorbance of the linker

will mask any minimum due to helicity, making alternative methods for determining helicity necessary.

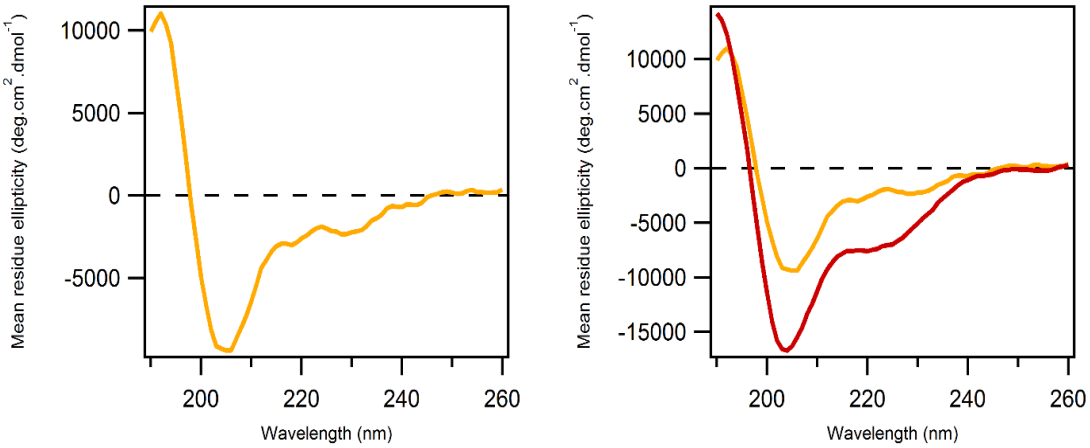


Figure 3-11: CD spectra of (left) **12** and (right) **12** overlayed with β3MP(722-739), illustrating the effect of the linker on helical content. (10 mM phosphate buffer, pH 6.5, 298K)

The original report by Spokoyny et.al made a semi-quantitative estimation of percent helicity for peptides containing the PF benzyl linkers[12] by correcting for the ellipticity of the perfluorobenzyl group at 222 nm. This estimation involves subtracting the mean residue absorbance of the 'control' CD spectrum (perfluorobenzyl group connected to two fully protected cysteine amino acids) from the measurement obtained for the full PF benzyl cyclised peptide. This method was applied to the mean residue ellipticity at 222 nm for peptide **12,** to obtain a 'corrected' value, which was used to determine the percent helicity**.** α-Helical content of **12** was estimated to be 74% using this method.

de Araujo et al[8] also investigated the effect of a PF benzyl linker in a short peptide using characteristics of the CD spectrum to determine relative helicity. In their study, CD spectra of helical peptides were compared to the spectrum of the PF benzyl linked peptides. The lack of maxima at 190 nm in particular suggests little α-helical structure for the short PF benzyl linked pentapeptide[8]. In the present study, a similar analysis of helicity was performed by comparing the CD spectrum of **12** with the linear peptide β3MP(722-739) (Figure 3-11, right). The CD spectrum of **12** displays a lower maximum at 190 nm than that of β3MP(722-739) suggesting lower levels of helical structure. Peptide **12** displays a minimum at 205 nm, while β3MP(722-739) displays a minimum at

204 nm, this slight shift toward 208 nm, indicative if helical structure, and suggests the PF benzyl linker increases helical content of the peptide. Linked peptides **7**, **10** and **11** (section 3.3.2), of the same sequence, display minima at 208 nm, suggesting the PF benzyl linked peptide has lower helical content. The similarities in the CD spectra of β3MP(722-739) and peptide **12** suggests that the structures of the two peptides are similar. As β3MP(722-739) had a percent helicity of 25%, peptide **12** is likely to also have approximately 25% helicity.

Two extremely different estimates of helical content were determined for peptide **12**, 74% and 25%, using methods previously applied to PF benzyl linked peptides. These results suggest that Spokoyny et.al[12] method of subtracting the mean molar ellipticity of PF benzyl linker at 222 nm provides a higher estimate of helicity than de Araujo et al[8]. This may give an over estimate of helical structure given the lack of other features expected in the spectrum. However, using the CD spectra characteristics to compare peptide helicities, similar to the method presented by de Araujo et al[8], may underestimate helical content due to the absorbance of the perfluorobenzyl group. These results suggest CD spectroscopy is not the best method for characterisation of PF benzyl linked peptide structure. Therefore, alternative methods, such as NMR spectroscopy, may give a more accurate determination of helicity for these peptides.

3.3.8 Cellular Uptake of Perfluorobenzyl Thioether Linked Peptide

Peptide **F12** was used to investigate cellular uptake of PF benzyl linked peptides in two cell lines. Fluorescein tagged peptide **F12**, positive control **CF-NYAD-1,** and negative control **CF** were added to mouse NIH T3T cells and human HEK 293 cells. Cells were incubated for 3 hrs, before being washed and trypsinised. Green punctate fluorescence was observed within the cytoplasm indicating significant uptake of **F12**, similar to the positive control CF-NYAD-1 (Figure 3-12). These results are consistent with those reported by Spokoyny et al which show the PF linked peptides have significant cellular uptake.[12]

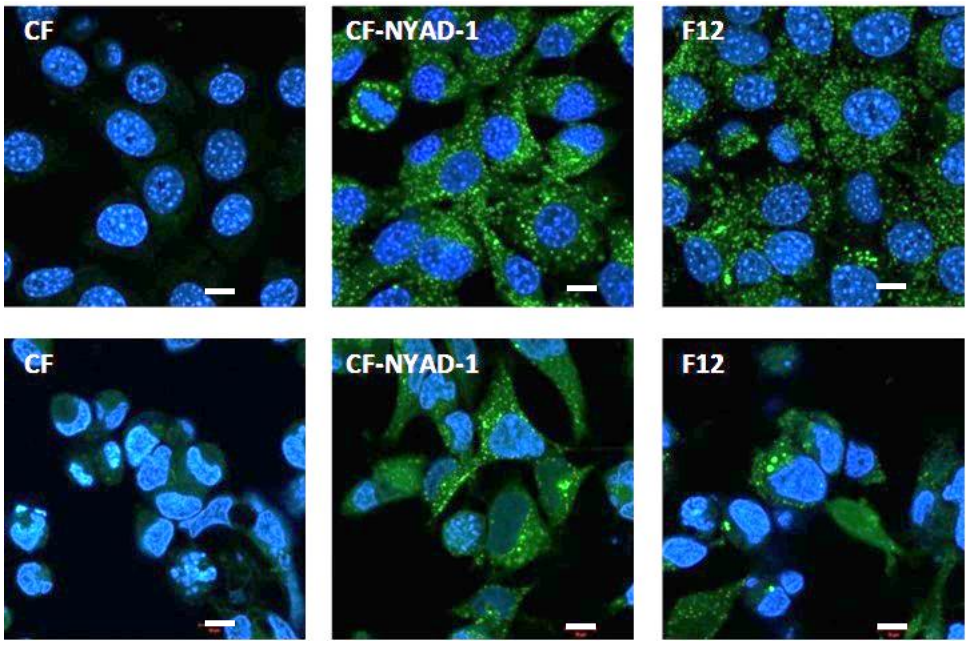


Figure 3-12: Confocal microscopy images of NIH 3T3 mouse cells (top) and human HEK 293 cells (bottom) treated with **CF**, CF-NYAD-1 and **F12**. Scale bars represent 10µm.

FACS analysis of cells treated with **F12** was not attempted due to the low yields acquired from cyclisation reactions. Further work is required to optimise cyclisation conditions, such that enough peptide can be obtained to quantify the cellular uptake and directly compare the cellular uptake of **F12** (PF benzyl), with **F7** (lactam), **F10** (triazole) and **F11** (all-hydrocarbon).

3.3.9 Conclusions

Four different linker types (lactam, triazole, all-hydrocarbon and PF benzyl) were incorporated in a peptide sequence derived from β3 integrin cytoplasmic tail. CD spectroscopy was used to compare peptides **7** (lactam), **10** (triazole) and **11** (hydrocarbon), revealing that all linkers gave rise to similar helical content. Peptide **10**, containing the triazole linker, had the greatest percentage helicity (62%) of the three peptides studied, displaying a $[\theta]_{222}/[\theta]_{208}$ ratio (0.95) closest to 1. These results differ from the study by de Araujo et al.[8] in which model alanine rich peptides with lactam linkers produced the greatest helical structure, when compared to triazole and all-hydrocarbon linked peptides. The difference between these two studies suggests that peptide sequence must be considered when determining which type of linker will create the largest helical content.

Cellular uptake studies of fluorescein tagged peptides **F7**, **F10** and **F11**, identified peptide **F11** (all-hydrocarbon) as the peptide that displayed the highest levels of cellular uptake. Peptide **F10** (triazole linker), displayed greater cellular uptake than the linear peptide (Fβ3MP(722-739)), whereas the lactam peptide **F7** resulted in lower levels of cellular uptake than Fβ3MP(722-739) as determined by FACS analysis. The basis of the differences between the cellular uptake of **F11**, **F10** and **F7** appears to be affected by both helicity of the peptide and the polarity of the side chain linker. The least polar of the linkers was the all-hydrocarbon linker, this gave rise to the peptide with the greatest cellular uptake. The amide linker of **F7** is less polar than the triazole linker of **F10**, which would suggest that **F7** should have greater cellular uptake, however this is not the case with **F10** displaying greater cellular uptake. These results may be due to **F10** having greater helical content than **F7**. In this case it appears the helicity of the peptide may be more significant than polarity in determining cellular uptake.

The effect of charge on cellular uptake of integrin-derived peptides was investigated using all-hydrocarbon linkers. This study suggests an increase in localised positive charge at the *N*-terminus of the peptide (**F15**), achieved by addition of three arginine

residues, resulted in a far greater increase in cellular uptake than distributed positive charge within the peptide sequence (**F15** vs **F16**).

The PF benzyl linker was incorporated into the 18-residue integrin-derived sequence, **12**, equivalent to the lactam, triazole and all-hydrocarbon linked peptides, **7**, **10** and **11**. Helicity could not be directly compared with **7**, **10** and **11** as the CD spectrum of **12** was complicated by absorption due to the perfluorobenzyl group, which masks any minimum at 222 nm used to determine helicity. Two previously reported approaches were used to determine the helicity of **12**; subtraction of the PF group absorption at 222 nm from the ellipticity of **12**, and comparison of the ellipticity of **12** with β3MP(722-739) at 195 and 208 nm. These approaches provided significantly different results, with no clear evidence that either method was more valid. Thus, in this case CD spectroscopy is not an appropriate method for quantifying helicity of PF benzyl linked peptides. NMR structural studies would likely provide a clearer indication of helical structure. Cellular uptake of the PF benzyl linked peptide was visualised using confocal microscopy indicating cellular uptake of the peptide. However, FACS analysis is required for a quantitative comparison to other peptide linkers.

3.4 References

1. Frankiewicz, L.; Betti, C.; Guillemyn, K.; Tourwé, D.; Jacquot, Y.; Ballet, S., Stabilisation of a short α-helical VIP fragment by side chain to side chain cyclisation: a comparison of common cyclisation motifs by circular dichroism. *Journal of Peptide Science* **2013,** *19* (7), 423-432.
2. Lau, Y. H.; de Andrade, P.; Wu, Y.; Spring, D. R., Peptide stapling techniques based on different macrocyclisation chemistries. *Chemical Society Reviews* **2014**.
3. Houston, M. E., Jr.; Gannon, C. L.; Kay, C. M.; Hodges, R. S., Lactam bridge stabilization of alpha-helical peptides: ring size, orientation and positional effects. *J Pept Sci* **1995,** *1* (4), 274-82.
4. Kawamoto, S. A.; Coleska, A.; Ran, X.; Yi, H.; 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. *Journal of Medicinal Chemistry* **2011,** *55* (3), 1137-1146.
5. Blackwell, H. E.; Grubbs, R. H., Highly Efficient Synthesis of Covalently Cross-Linked Peptide Helices by Ring-Closing Metathesis. *Angewandte Chemie International Edition* **1998,** *37* (23), 3281-3284.
6. Schafmeister, C. E.; Po, J.; Verdine, G. L., An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *Journal of the American Chemical Society* **2000,** *122* (24), 5891-5892.
7. Cantel, S.; Le Chevalier Isaad, A.; Scrima, M.; Levy, J. J.; DiMarchi, R. D.; Rovero, P.; Halperin, J. A.; D'Ursi, A. M.; Papini, A. M.; Chorev, M., Synthesis and Conformational Analysis of a Cyclic Peptide Obtained via i to i+4 Intramolecular Side-Chain to Side-Chain Azide−Alkyne 1,3-Dipolar Cycloaddition. *The Journal of Organic Chemistry* **2008,** *73* (15), 5663-5674.
8. de Araujo, A. D.; Hoang, H. N.; Kok, W. M.; Diness, F.; Gupta, P.; Hill, T. A.; Driver, R. W.; Price, D. A.; Liras, S.; Fairlie, D. P., Comparative α-Helicity of Cyclic Pentapeptides in Water. *Angewandte Chemie International Edition* **2014,** *53* (27), 6965-6969.
9. Verdine, G. L.; Hilinski, G. J., All-hydrocarbon stapled peptides as Synthetic Cell-Accessible Mini-Proteins. *Drug Discovery Today: Technologies* **2012,** *9* (1), e41-e47.
10. Milletti, F., Cell-penetrating peptides: classes, origin, and current landscape. *Drug Discovery Today* **2012,** *17* (15–16), 850-860.
11. Chu, Q.; Moellering, R. E.; Hilinski, G. J.; Kim, Y.-W.; Grossmann, T. N.; Yeh, J. T. H.; Verdine, G. L., Towards understanding cell penetration by stapled peptides. *MedChemComm* **2015,** *6* (1), 111-119.
12. Spokoyny, A. M.; Zou, Y.; Ling, J. J.; Yu, H.; Lin, Y.-S.; Pentelute, B. L., A Perfluoroaryl-Cysteine SNAr Chemistry Approach to Unprotected Peptide Stapling. *Journal of the American Chemical Society* **2013,** *135* (16), 5946-5949.
13. Zhang, H.; Zhao, Q.; Bhattacharya, S.; Waheed, A. A.; Tong, X.; Hong, A.; Heck, S.; Curreli, F.; Goger, M.; Cowburn, D.; Freed, E. O.; Debnath, A. K., A Cell-penetrating Helical Peptide as a Potential HIV-1 Inhibitor. *Journal of Molecular Biology* **2008,** *378* (3), 565-580.
14. Lau, Y. H.; de Andrade, P.; Wu, Y.; Spring, D. R., Peptide stapling techniques based on different macrocyclisation chemistries. *Chemical Society Reviews* **2015,** *44* (1), 91-102.
15. Sheppard, R., The fluorenylmethoxycarbonyl group in solid phase synthesis. *Journal of Peptide Science* **2003,** *9* (9), 545-552.

16. Roberts, K. S.; Sampson, N. S., Increased polymer length of oligopeptide-substituted polynorbornenes with LiCl. *J Org Chem* **2003,** *68* (5), 2020-3.
17. Paradis-Bas, M.; Tulla-Puche, J.; Albericio, F., The road to the synthesis of "difficult peptides". *Chemical Society Reviews* **2016,** *45* (3), 631-654.
18. Hossain, M. A.; Rosengren, K. J.; Zhang, S.; Bathgate, R. A. D.; Tregear, G. W.; van Lierop, B. J.; Robinson, A. J.; Wade, J. D., Solid phase synthesis and structural analysis of novel A-chain dicarba analogs of human relaxin-3 (INSL7) that exhibit full biological activity. *Organic & Biomolecular Chemistry* **2009,** *7* (8), 1547-1553.
19. Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P., Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *Journal of the American Chemical Society* **2005,** *127* (9), 2974-2983.
20. Wallimann, P.; Kennedy, R. J.; Miller, J. S.; Shalongo, W.; Kemp, D. S., Dual wavelength parametric test of two-state models for circular dichroism spectra of helical polypeptides: anomalous dichroic properties of alanine-rich peptides. *J Am Chem Soc* **2003,** *125* (5), 1203-20.
21. Sonnichsen, F. D.; Van Eyk, J. E.; Hodges, R. S.; Sykes, B. D., Effect of trifluoroethanol on protein secondary structure: an NMR and CD study using a synthetic actin peptide. *Biochemistry* **1992,** *31* (37), 8790-8798.
22. Klein, M. J.; Schmidt, S.; Wadhwani, P.; Bürck, J.; Reichert, J.; Afonin, S.; Berditsch, M.; Schober, T.; Brock, R.; Kansy, M.; Ulrich, A. S., Lactam-Stapled Cell-Penetrating Peptides: Cell Uptake and Membrane Binding Properties. *Journal of Medicinal Chemistry* **2017,** *60* (19), 8071-8082.
23. Kacprzak, K.; Skiera, I.; Piasecka, M.; Paryzek, Z., Alkaloids and Isoprenoids Modification by Copper(I)-Catalyzed Huisgen 1,3-Dipolar Cycloaddition (Click Chemistry): Toward New Functions and Molecular Architectures. *Chemical Reviews* **2016,** *116* (10), 5689-5743.
24. Hol, W. G., Effects of the alpha-helix dipole upon the functioning and structure of proteins and peptides. *Advances in biophysics* **1985,** *19*, 133-65.
25. Kolb, H. C.; Sharpless, K. B., The growing impact of click chemistry on drug discovery. *Drug Discovery Today* **2003,** *8* (24), 1128-1137.
26. Muppidi, A.; Li, X.; Chen, J.; Lin, Q., Conjugation of Spermine Enhances Cellular Uptake of the Stapled Peptide-Based Inhibitors of p53-Mdm2 Interaction. *Bioorganic & medicinal chemistry letters* **2011,** *21* (24), 7412-7415.
27. Bernal, F.; Tyler, A. F.; Korsmeyer, S. J.; Walensky, L. D.; Verdine, G. L., Reactivation of the p53 Tumor Suppressor Pathway by a Stapled p53 Peptide. *Journal of the American Chemical Society* **2007,** *129* (9), 2456-2457.

Chapter 4

Experimental

4.1.1 General Information

Solvents and reagents were of reagent grade or higher and used as supplied. 9-Fluorenylmethoxycarbonyl (Fmoc) protected L-amino acids, Fmoc-Trp(OtBu)-OH, Fmoc-Ala-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Asp(OtBu)-OH, Fmoc-Asp(Opip)-OH, Fmoc-Cys(Trt)-OH, Fmoc-Glu(OtBu)-OH, Fmoc-Glu(Opip)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Lys(Mmt)-OH, Fmoc-Phe-OH, Fmoc-Pra-OH and Fmoc-His(Trt)-OH, as well as Boc-L-Lys-OH, *N*-[(Dimethylamino)-1*H*-1,2,3-triazolo-[4,5-*b*]pyridin-1-ylmethylene]-*N*-methylmethanaminium hexafluorophosphate *N*-oxide (HATU), and O-(1H-6-chlorobenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HCTU) were purchased from Chem-Impex International (Wooddale, IL, USA). Dimethylformamide (AR grade) (DMF), *N*-methyl-2-pyrrolidone (NMP), dichloromethane (DCM), and piperidine were purchased from Merck (Darmstadt, Germany). Acetonitrile (HPLC grade) and methanol (HPLC grade) were purchased from Scharlau (Sentmenat, Barcelona, Spain). Trifluoroacetic acid (TFA), trifluoroethanol (TFE) (NMR grade), and diisopropylethylamine (DIPEA) were purchased from Alfar Aesar (Ward Hill, MA, USA). *N*-hydroxybenzotriazole (HOBt), (S)-N-Fmoc- ξ-(4'pentenyl) alanine, Fmoc-Lys(N_3)-OH, and Fmoc Rink amide resin were purchased from, GL Biochem (Shanghai, China), Okeanos, (Japan), ChemPep Inc. (Wellington, USA), and Agilent (Santa Clara, CA, USA) respectively .All other reagents were purchased from Sigma-Aldrich (USA, or Castle Hill, NSW, Australia).

4.1.2 Physical and spectroscopic techniques

4.1.2.1 Mass spectrometry (MS) and Nuclear magnetic resonance (NMR) spectroscopy characterisation.

High resolution mass spectra (HMRS) were acquired with an Agilent 1260 infinity TOF LC/MS 6230, running Agilent Mass Hunter software. The process of elongation was monitored during peptide synthesis using an electrospray ionization ion trap mass spectrometer (Bruker, HCT mass spectrometer).^{1}H NMR, ^{13}C NMR and 2D spectra were recorded using the Varian Oxford NMR 600 MHz, with Agilent DD2 console. Software

running on the spectrometers was Agilent VnmrJ 4.2. Chemical shifts are reported in parts per million (ppm) on the δ scale.

4.1.2.2 Circular dichroism (CD) analysis

Samples were prepared in 10 mM phosphate buffer, containing 0% or 10% TFE with the final peptide concentration of approximately 25 µM. The exact concentrations of peptides were determined by the UV absorption of tyrosine residue at 280 nm which was used to calculate molar ellipticity.

CD measurements for all peptides were performed using a Jasco J-815 circular dichroism spectropolarimeter (Jasco Corp., Japan), calibrated with *d*-10-camphorsulfonic acid. The quartz cell temperature of 25°C was stabilised using a Peltier temperature controller. Secondary spectra for all peptides were obtained between wavelengths of 190-260 nm under the same parameters, where scan speed was 10 nm/minute, bandwidth was 1.0 nm and the resolution was 1 nm with a 1 second response. A 0.1 mm quartz cuvette were used, and three repeat scans were compiled to generate the average spectra. The CD signal resulting from the solvent alone was subtracted from the spectrum of each peptide solution. The results were evaluated using the Jasco Spectra Manager and any remaining noise in the spectra was smoothed using Savitzky-Golay convolution (width: 9). Data were converted to ellipticity ($deg.cm^2.dmol^{-1}.residue^{-1}$) according to the equation 1[1].

The α-helical content of each peptide was determined from the mean residue molar ellipticity at 222 nm.

$$[\theta] = \theta / (10\ l\ c\ n) \qquad \textit{eq. 1}$$

Where [θ] is the mean residue molar ellipticity ($deg.cm^2.dmol^{-1}$), θ is ellipticity (mdeg), *l* is the path length (cm), c is the concentration (mol L^{-1}) and n is the number of residues.

Percent helicity was calculated from the ratio $[\theta]_{222}/[\theta]_{max}$, where

$$[\theta]_{max} = (-44\,000 + 250T)(1 - k/n) \qquad \textit{eq. 2}$$

Where $[\theta]_{max}$ is the maximum mean residue molar ellipticity (deg.cm^2.dmol^{-1}), T is the temperature (°C), k is the number of non-hydrogen-bonded peptide carbonyls. For free N-terminal peptide it has been suggested a value of 4, n is number of residues[1].

For our system we used $[\theta]_{max}$ = −23 400 for k = 4.0 and n = 18 (number of amino acid residues in the peptide).

Table 4-1: Molar ellipticity [θ] (deg cm2 dmol-1 residue-1) at 222 nm, percentage helicity and ratios of ellipticity at 222 nm/208 nm for β3MP(722-739) peptides, 7, 10 and 11 determined from CD data acquired in 10% TFE, 10 mM phosphate buffer (pH 6.5, 25oC).

Peptide	$[\theta]_{222}$[a]	Helicity[b]	$[\theta]_{222}/[\theta]_{208}$[c]
β3MP(722-739)	-10691.6	36%	0.69
7	-19073.7	65%	0.79
10	-20789.1	71%	1.18
11	-19744.9	67%	0.88

[a] Molar ellipticity [θ] (deg cm^2 dmol^{-1} residue^{-1}) at 222 nm, acquired from CD data

[b] percent helicity

[c] ratios of ellipticity at 222 nm/208 nm

4.1.3 Purification and isolation techniques

4.1.3.1 Purification and analysis by reverse phase high performance liquid chromatography (RP-HPLC)

Peptides were purified by RP-HPLC, performed using a Gilson semi prep system (GX-271 liquid handler, UV/VIS-155, and 322-pump) fitted with a Supleco (C18, 300 Å, 5 µm, 10 mm x 250 mm) column. Peptides were generally purified with a gradient of 20-45% ACN/H_2O with 0.1% TFA over 25 minutes Carboxyfluorescein labelled peptides were purified using a gradient of 25-50% ACN/H_2O with 0.1% TFA over 25 minutes. Fractions were collected where absorbance at 215 nm exceeded 50 mV.

Peptide purity was determined using analytical RP-HPLC, performed on an Agilent 1260 system fitted with an analytical (C18, 300 Å, 5 µm, 4.6 mm x 150 mm) column. The eluents were 0.1% aqueous TFA and 0.1% TFA in acetonitrile.

4.1.3.2 Lyophilisation

Synthetic peptides were lyophilised on a Christ Alpha 2-4 LSC freeze-dryer.

4.1.4 General Procedure I: Manual SPPS

Where stated, peptide elongation was performed manually on Fmoc-Rink Amide resin on a 0.1 mmol scale, using the Fmoc/*t*Bu strategy[2]. *Deprotection:* The *N*α-Fmoc group was removed by treatment of the resin with a solution of 20% v/v piperidine in DMF containing HOBt (0.1 M) for 10 minutes. *Amino acid coupling:* The *N*α-Fmoc protected amino acid (5 equiv.) was pre-activated with a solution of HATU in DMF (5 equiv., 0.45 M) and DIPEA (neat) (10 equiv.) for 2-10 minutes at rt. The activated solution was added to the resin and the mixture was left for 1 hr at rt. The solution was drained, and the resin was washed with DMF (3 x 5 ml), DCM (3 x 5 ml) and DMF (3 x 5 ml). Coupling reactions were monitored *via* the TNBSA test[3] for primary amines. *Unusual amino acid couplings: N*α-Fmoc protected *Fmoc-Lys(Mmt)-OH, Fmoc-Asp(Opip)-OH, Fmoc-Glu(Opip)-OH*), *Fmoc-Pra-OH, Fmoc-L-Nle (εN$_3$)-OH* were coupled as detailed below. *N*α-Fmoc protected amino acid (2.5 equiv.) was pre-activated with a solution of HATU in DMF (2.5 equiv.) and DIPEA (neat) (5 equiv.) for 2 minutes at rt. The activated solution was added to the resin and the mixture was left for 16 hrs at rt. The solution was drained, and the resin was washed with DMF (3 x 5 ml), DCM (3 x 5 ml) and DMF (3 x 5 ml). To confirm reaction had reached completion 10 mg of resin was cleaved (general method III), and the peptide analysed by HPLC and ESI-MS. If the reaction was found to be incomplete a second coupling was performed with fresh reagents.

4.1.5 General procedure II: Automated peptide synthesis

Where stated, automated peptide synthesis was carried out on a Liberty Microwave Peptide Synthesiser (CEM Corporation) using the Fmoc/*t*Bu strategy[2]. *Deprotection*: The *N*α-Fmoc group was deprotected before addition of *N*α-Fmoc protected amino acids by consecutive additions of 20% v/v piperidine in DMF for 30 seconds and 3 minutes. The maximum microwave power was set to 25 W and the maximum temperature was set to 75°C. Resin was then washed with DMF (4 x 7 ml). *Amino acid couplings:* The *N*α-Fmoc protected amino acid in DMF (5 equiv., 0.2 M), HATU in DMF (4.5 equiv., 0.45 M) and DIPEA in NMP (10 equiv., 2 M) were added to the rink amide resin and subjected to 25 W microwave irradiation for 6 minutes. Each coupling was performed twice with a maximum temperature of 50°C for both couplings. The exception was Fmoc-Arg(Pbf)-OH, which was coupled for 1500 seconds at rt, followed by coupling for 300 seconds at 25 W microwave irradiation and the maximum temperature of 75°C.

4.1.6 General Procedure III: Cleavage of Peptides from Rink Amide Resin

Cleavage of the peptide from the resin was performed using 2.5% triisopropylsilane (TIPS), 2.5% 2,2'-(ethylenedioxy)diethanethiol (DODT) and 2.5% water in TFA (1 ml per 50 mg) for 2 hrs. Half the TFA was removed under nitrogen, and peptide was precipitated by addition of diethyl ether. The suspended peptide was isolated by centrifugation and residue dissolved in 30% acetonitrile in water and lyophilised.

4.1.7 General Procedure IV: Peptide cyclisation for lactam bridged peptides

Fully protected resin bound peptide (0.1 mmol) was successively washed with DMF (5 x 5 ml) and DCM (5 x 5 ml). A solution of DCM, TFA, TIPS (94:1:5, 5 ml) was added to the resin to selective remove the side chain protecting groups, methoxytrityl (Mmt) and 2-phyenyl isopropyl oxy (Opip) of lysine (Lys) and aspartic acid (Asp)/ glutamic acid (Glu) respectively. The solution was stirred for 2 minutes, and then washed with DCM (5 x 5 ml). This procedure was repeated 10 times. The resin was then washed with DCM

(5 x 5 ml) and DMF (5 x 5 ml). The resin bound peptide was treated with PyBOP (52 mg, 0.1 mmol,) and DIPEA (86 µl, 0.5 mmol,) in DMF. After 18 hrs the reagents were removed by filtration, and a sample of the resin (10 mg) was cleaved (general method III) and analysed by HPLC and ESI-MS. If the cyclisation had not reached completion, the coupling was repeated with fresh reagents. The overall cyclisation reaction varied from 1 to 3 days as shown in each example below.

4.1.8 General Procedure V: Cyclisation of Hydrocarbon Linked Peptides.

Dried resin bound linear peptide precursors (0.05 mmol) was swollen in anhydrous DCM (6 ml). Lithium chloride dissolved in anhydrous DMF (4M, 0.4 ml) was added to solution, followed by the addition of 2nd generation Grubbs' catalyst (11 mg, 12.5 µmol). The reaction mixture was then heated at 100 W µwave, with the maximum temperature set at 80°C for 1 hr.

4.1.9 General Procedure VI: Carboxyfluorescein Labelling

Resin bound peptide (0.1 mmol) was swollen in DCM: DMF (1:1, 5 ml). Fmoc- β-alanine (155 mg, 0.5 mmol) dissolved in DMF (2 ml), was pre-activated by addition of HCTU in DMF (1 ml, 0.5M) and DIPEA (174 µl) for 2 minutes. The resin was washed with DMF (5 x 5 ml) followed by addition of pre-activated Fmoc- β-alanine, and allowed to react for 1 hr. The resin was then drained, followed by a second addition of pre-activated Fmoc-β-alanine to the resin which was allowed to react for 1 hr. *Na*-Fmoc protecting group was removed via treatment of the resin bound peptide with piperidine (20%) and HOBt (0.1 M) in DMF (5 ml) for 10 minutes. The resin was then washed with DMF (5 x 5 ml), DCM (2 x 5 ml) and DMF (2 x 5 ml). 5(6)-Carboxyfluorescein (95 mg, 0.25 mmol) was dissolved in NMP (0.5 ml) and added to PyBOP (130 mg, 0.25 mmol) dissolved in NMP (0.5 ml). DIPEA (87 µl, 0.5 mmol) was then added and the reaction mixture was allowed to react for 2 minutes. The reaction mixture was then added to the resin and allowed to react for 18 hrs. Reagents were removed by filtration, and a sample of the resin

(10 mg) was cleaved (general method III) and analysed by HPLC and ESI-MS. If the cyclisation had not reached completion, the coupling was performed with fresh reagents.

4.2 Experimental work described in Chapter 2

4.2.1 Peptide β3MP(722-735) (NH_2-HDRKEFAKFEEDRA-$CONH_2$)

Resin bound peptide was synthesised on rink amide resin on a 0.1 mmol scale. The initial deprotection and addition of the *C*-terminal Fmoc-Ala-OH to the resin was performed manually according to general procedure I. Automated peptide synthesis was used to elongate the sequence from R(734) to R(724), following methods described in general procedure II. Manual addition of final residues D(723) and H(722) was performed according to general procedure I. The peptide with free *N*-terminal was then cleaved according to general procedure III, and the crude peptide was purified by RP-HPLC. Pure fractions were collected and lyophilised to afford β3MP(722-735) as a white powder. LCMS (ESI+) $[M+2H]^{+2}$ calculated for $[C_{78}H_{119}N_{25}O_{24}]$:895.9509;observed: 895.9496 m/z.

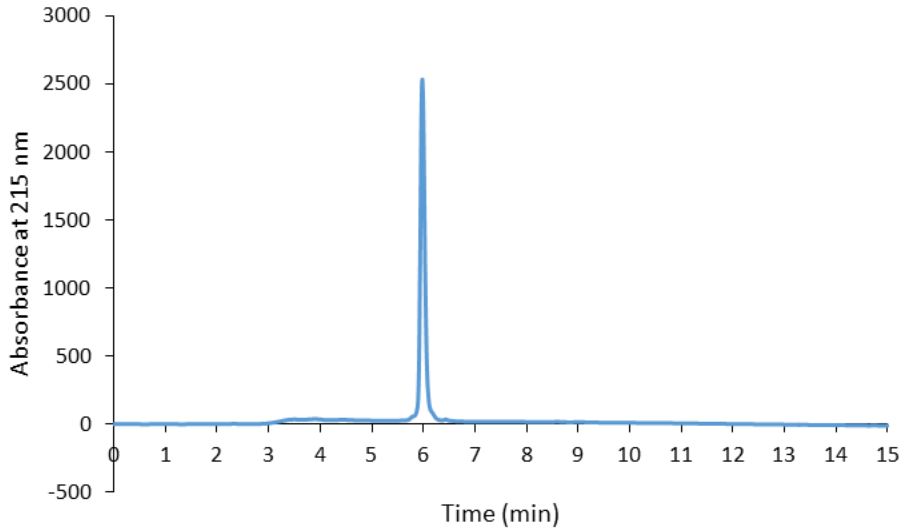


Figure 4-1: RP-HPLC trace of purified β3MP(722-735) monitoring at a wavelength of 215 nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.2.2 Peptide **1** (NH_2-HDRKEFAK*FEED*RA-$CONH_2$)

Resin bound peptide was synthesised on rink amide resin on a 0.1 mmol scale according to methods described in the general procedure I from A(735)-D(723). Uncommon amino acids Fmoc-Asp(Opip)-OH and Fmoc-Lys(Mmt)-OH were coupled at positions 733 and 729, in accordance with coupling uncommon amino acids. Cyclisation of the resin bound Fmoc protected peptide was performed according to general procedure IV, followed by the addition of the final residue H(722) following methods described in general procedure I. The completed peptide was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **1** as a white powder. LCMS (ESI+) m/z calculated for $C_{76}H_{112}N_{24}O_{24}$: 1757.8598 (average isotopes); observed: m/z 879.93 ($[M+3H]^{+3}$) 586.95 ($[M+4H]^{+4}$).

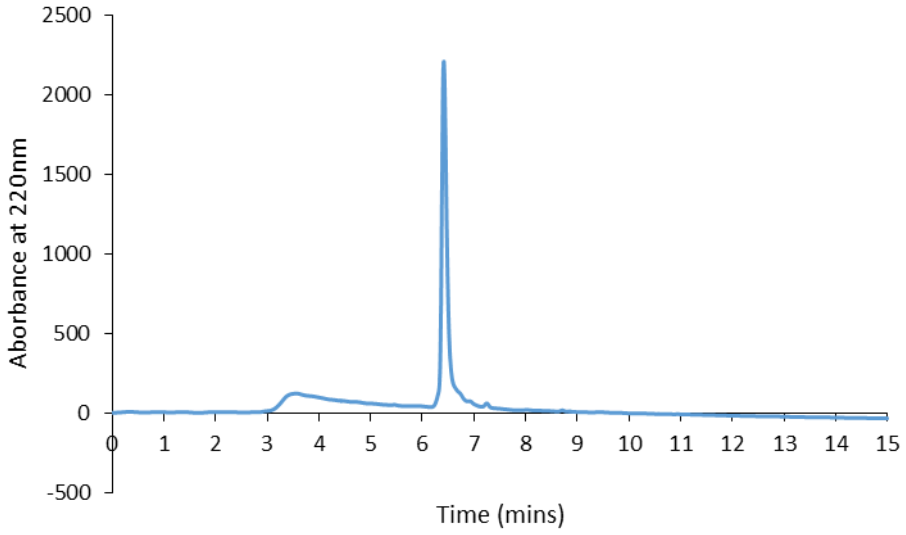


Figure 4-2: RP-HPLC trace of purified **1** monitoring at a wavelength of 215nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.2.3 Peptide **2** (NH_2-HDRKEFAK*FEEE*RA-$CONH_2$)

Resin bound peptide was synthesised on rink amide resin on a 0.1 mmol scale according to methods described in the general procedure I from A(735) to D(723). Fmoc-Glu(Opip)-OH and Fmoc-Lys(Mmt)-OH were coupled at positions 733 and 729 in accordance with coupling uncommon amino acids. Cyclisation of the Fmoc protected peptide was performed using general procedure IV followed by the manual addition of final residue H(722) following methods described in general procedure I. The completed peptide was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **2** as a white powder. LCMS (ESI+) m/z calculated for $C_{78}H_{117}N_{25}O_{23}$: 1771.8754 (average isotopes); observed: m/z 886.94 ($[M+3H]^{+3}$) 591.63 ($[M+4H]^{+4}$).

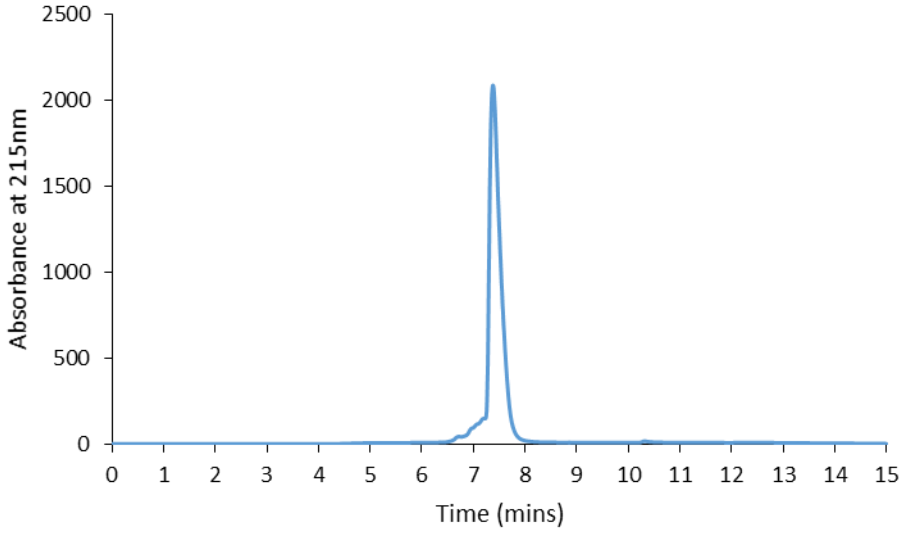


Figure 4-3: RP-HPLC trace of purified **2** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.2.4 Peptide **3** (NH_2-HDRK*EFAD*FEEERA-$CONH_2$)

Resin bound peptide was synthesised on rink amide resin on a 0.1 mmol scale. The initial deprotection of the resin and addition of the *C*-terminal Fmoc-Ala-OH was performed manually according to general procedure I. Automated peptide synthesis was used to elongate the sequence from R(734) to F(730), following general procedure II. Manual elongation of the peptide up to D(723) was performed using methods described in general procedure I, including coupling of uncommon amino acids Fmoc-Asp(Opip)-OH and Fmoc-Lys(Mmt)-OH at positions 729 and 725. Cyclisation of the Fmoc protected peptide was performed using general procedure IV, followed by the addition of final residue H(722) following methods described in general procedure I. The completed peptide, with N_α-terminal Fmoc deprotected, was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **3** as a white powder. LCMS (ESI+) m/z calculated for $C_{75}H_{110}N_{24}O_{24}$: 1730.8125 (average isotopes); observed: m/z 866.41 ($[M+3H]^{+3}$) 577.94 ($[M+4H]^{+4}$).

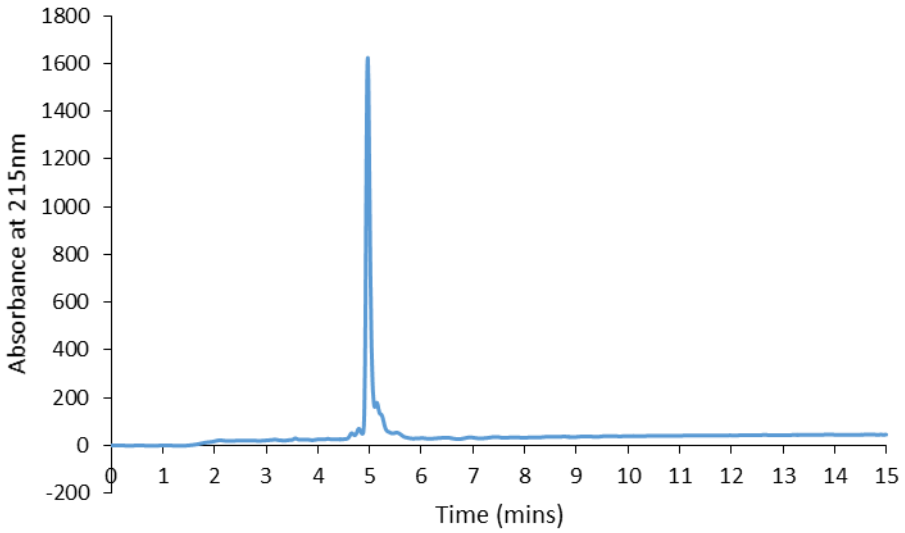


Figure 4-4: RP-HPLC trace of purified **3** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.2.5 Peptide **4** (NH_2-HDRE*EFAK*FEEERA-$CONH_2$)

Resin bound peptide was synthesised on rink amide resin on a 0.1 mmol scale. The initial deprotection of the resin and addition of the *C*-terminal Fmoc-Ala-OH was performed according to general procedure I. Automated peptide synthesis was used to elongate the sequence from R(734) to F(730), following general procedure II. Manual elongation of the peptide to D(723) was performed using methods described in general procedure I, including coupling of uncommon amino acids Fmoc-Glu(Opip)-OH and Fmoc-Lys(Mmt)-OH in positions 729 and 735. Cyclisation of the Fmoc protected peptide was performed using general procedure IV, followed by the addition of final residue H(722) following methods described in general procedure I. The completed peptide, with N_α-terminal Fmoc deprotected, was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **4** as a white powder. LCMS (ESI+) m/z calculated for $C_{76}H_{112}N_{24}O_{24}$: 1744.8281 (average isotopes); observed: m/z 873.41 ($[M+3H]^{+3}$) 582.61 ($[M+4H]^{+4}$).

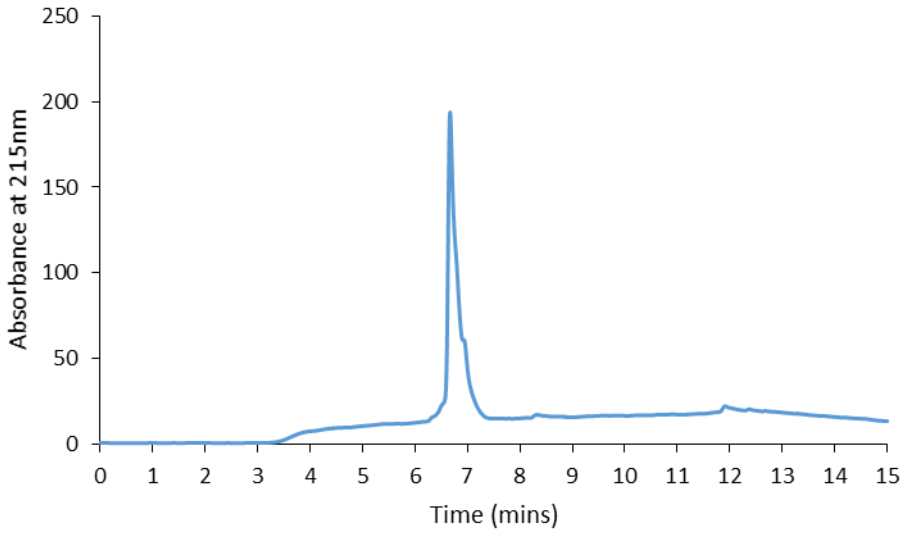


Figure 4-5: RP-HPLC trace of purified **4** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes

4.2.6 Peptide **5** (NH_2-HDRD*EFAK*FEEERA-$CONH_2$)

Resin bound peptide was synthesised on rink amide resin on a 0.1 mmol scale. The initial deprotection of the resin and addition of the *C*-terminal Fmoc-Ala-OH was performed according to general procedure I. Automated peptide synthesis was used to elongate the sequence from R(734) to F(730), following general procedure II. Manual elongation of the peptide to D(723) was performed using methods described in general procedure I, including coupling of uncommon amino acids, Fmoc-Asp(Opip)-OH and Fmoc-Lys(Mmt)-OH at positions 729 and 725. Cyclisation of the resin bound Fmoc protected peptide was performed according to general procedure IV, followed by the addition of final residue H(722) according to methods described in general procedure I. The completed peptide with free N_α-terminal, was cleaved from the resin following general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **5** as a white powder. LCMS (ESI+) m/z calculated for $C_{75}H_{110}N_{24}O_{24}$: 1730.81248 (average isotopes); observed: m/z 866.41 ($[M+3H]^{+3}$) 577.92 ($[M+4H]^{+4}$).

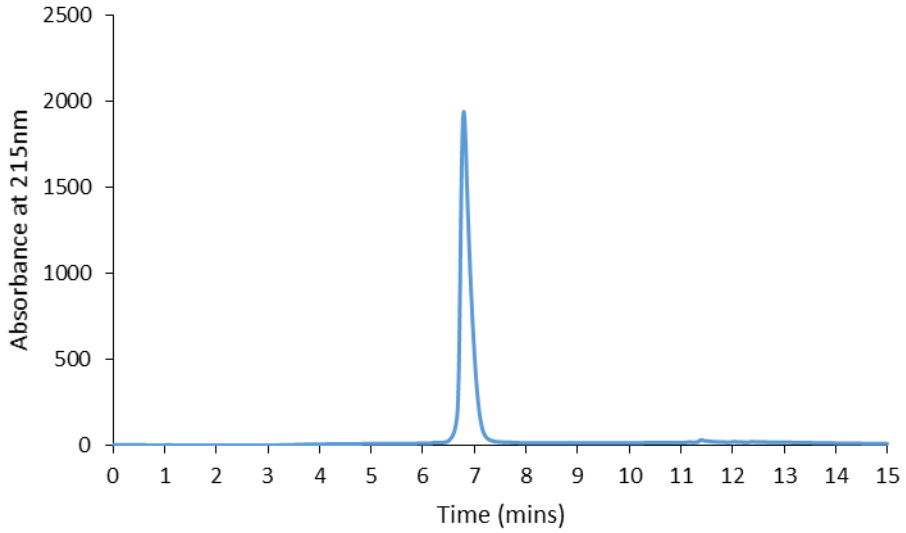


Figure 4-6: RP-HPLC trace of purified **5** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes

4.2.7 Peptide β3MP(722-739) (NH_2-HDRKEFAKFEEDRARAKW-$CONH_2$)

The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin. The initial deprotection and coupling of *C*-terminal Fmoc-Trp(OtBu)-OH to the resin was performed manually according to methods described in general procedure I. Automated peptide synthesis was used to elongate the sequence from K(738) to R(724) according to general procedure II. Manual addition of final residues D(723) and H(722) was then completed in accordance with general procedure I. The desired peptide, with free N_α-terminal, was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **β3MP(722-739)** as a white powder. HRMS (ESI+) $[M+3H]^{+3}$ calculated for $[C_{104}H_{158}N_{34}O_{28}]$:778.0741;observed: m/z 778.0712.

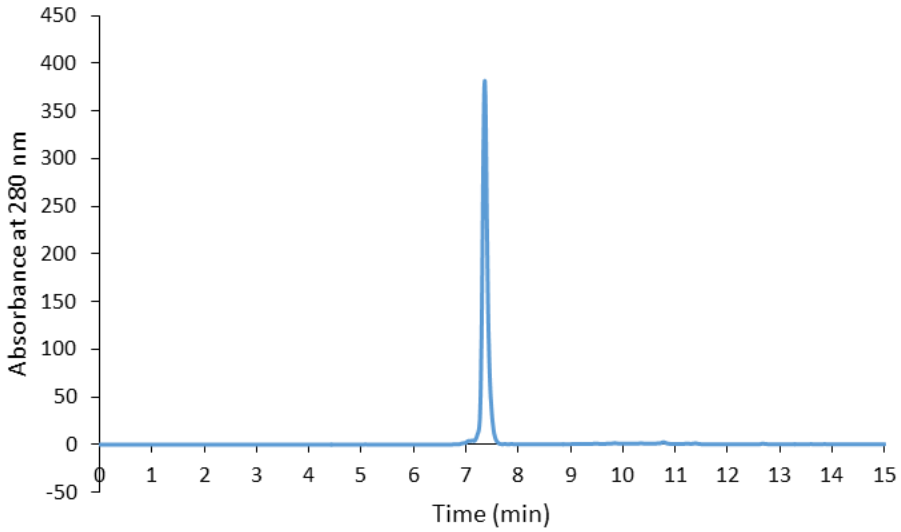


Figure 4-7: RP-HPLC trace of purified β3MP(722-739) monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.2.8 Peptide **6** (NH_2-HDRKEFAK*FEED*RARAKW-$CONH_2$)

The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin. The initial deprotection of the resin and addition of the *C*-terminal Fmoc-Trp(OtBu)-OH was performed according to general procedure I. Automated peptide synthesis was used to elongate the sequence from K(738) to R(734) according to general procedure II. The peptide was then elongated manually according to general procedure I, including coupling of uncommon amino acids, Fmoc-Asp(Opip)-OH and Fmoc-Lys(Mmt)-OH at the positions 733 and 729. Cyclisation of the resin bound Fmoc protected peptide was performed using general procedure IV, followed by the addition of final residue H(722) following methods described in general procedure I. The completed peptide, with free N_α-terminal Fmoc, was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **6** as a white powder. HRMS (ESI+) $[M+3H]^{+3}$ calculated for $[C_{103}H_{154}N_{34}O_{27}]$: 767.3987 ;observed: m/z 767.3983.

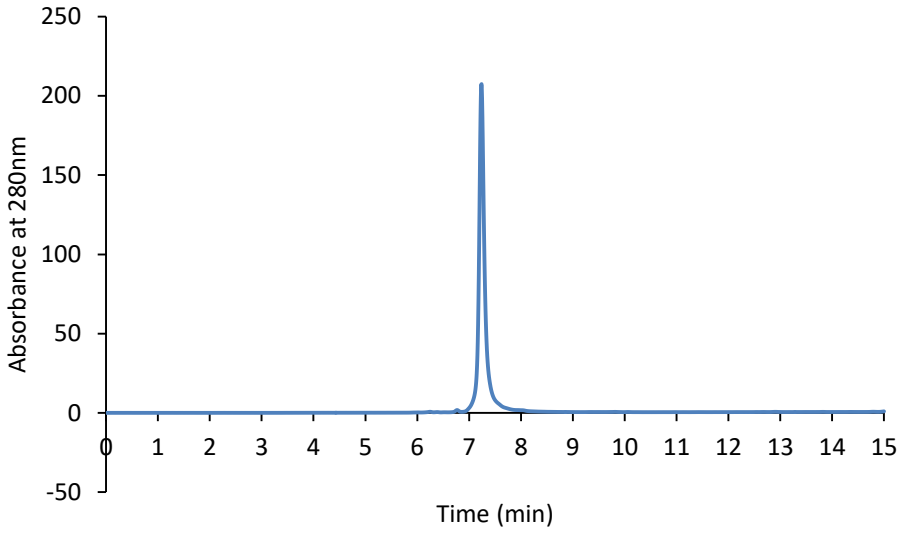


Figure 4-8: RP-HPLC trace of purified **6** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.2.9 Peptide **7** (NH_2-HDRK*EFAD*FEEERARAKW-$CONH_2$)

The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin. The initial deprotection of the resin and addition of the *C*-terminal Fmoc-Trp(OtBu)-OH was performed according to general procedure I. Automated peptide synthesis was used to elongate the sequence from K(738) to F(730), according to general procedure II. The peptide was then elongated manually to D(723) according to general procedure I, including coupling of uncommon amino acids, Fmoc-Asp(Opip)-OH and Fmoc-Lys(Mmt)-OH at positions 729 and 725. Cyclisation of the resin bound Fmoc protected peptide was performed according to general procedure IV followed by the addition of final residue H(722) in accordance with methods described in general procedure I. The completed peptide, with free N_α-terminal, was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **7** as a white powder. HRMS (ESI+) $[M+4H]^{+4}$ calculated for $[C_{102}H_{149}N_{33}O_{29}]$:576.0379 ;observed: m/z 576.0371.

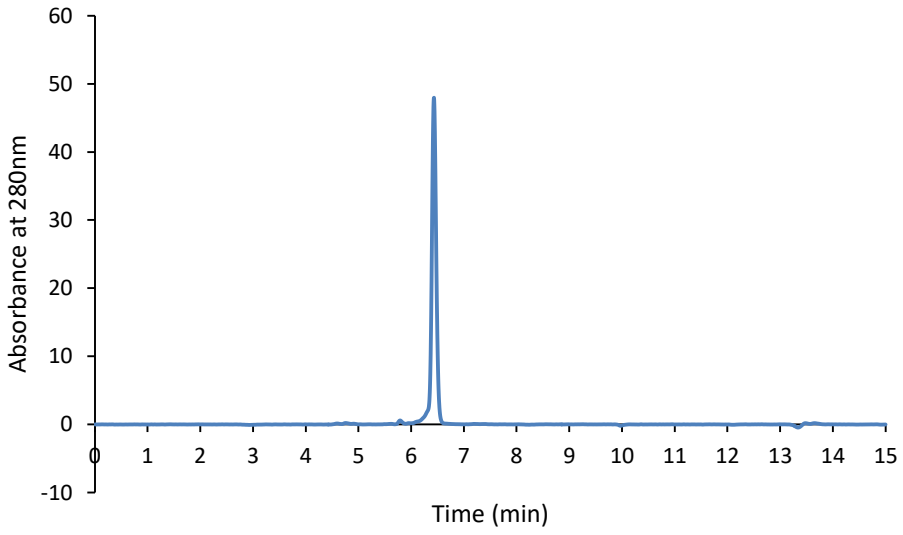


Figure 4-9: RP-HPLC trace of purified **7** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.2.10 Peptide **8** (NH_2-HDRKEFKK*FED*ERARAKW-$CONH_2$)

The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin. The initial deprotection of the resin and addition of the *C*-terminal Fmoc-Trp(OtBu)-OH was performed according to general procedure I. Automated peptide synthesis was used to elongate the sequence from K(738) to E(733) was performed according to general procedure II. The peptide was then elongated manually to D(723) according to general procedure I, including coupling of uncommon amino acids, Fmoc-Asp(Opip)-OH and Fmoc-Lys(Mmt)-OH at the positions 734 and 729. Cyclisation of the resin bound Fmoc protected peptide was performed using general procedure IV followed by the addition of final Fmoc- residue H(722) according to methods described in general procedure I. The completed peptide, with free N_α-terminal, was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **8** as a white powder. HRMS (ESI+) $[M+4H]^{+4}$ calculated for $[C_{106}H_{161}N_{35}O_{27}]$:590.0655 ;observed: m/z 590.0660.

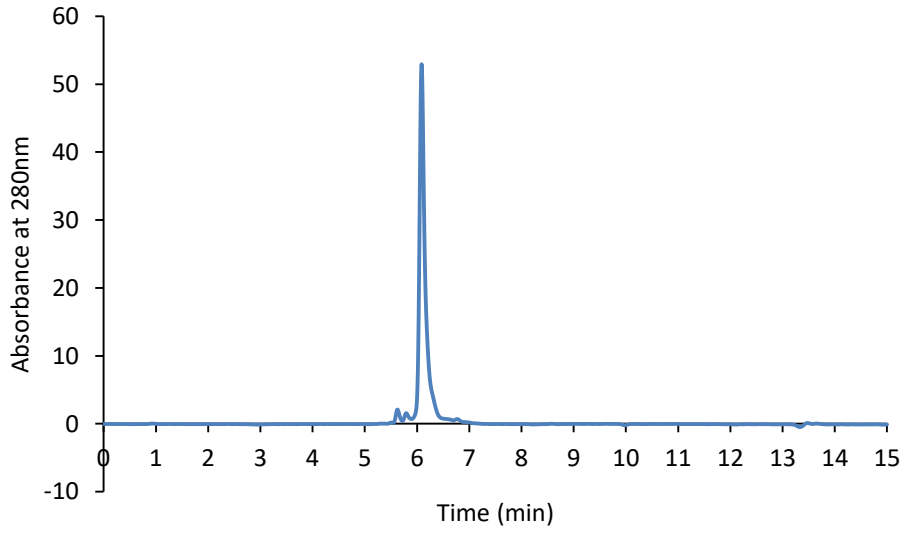


Figure 4-10: RP-HPLC trace of purified **8** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.2.11 Peptide **9** (NH_2-HDK*KEFD*KFEEERARAKW-$CONH_2$)

The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin. The initial deprotection of the resin and addition of the *C*-terminal Fmoc-Trp(OtBu)-OH was performed according to general procedure I. Automated peptide synthesis was used to elongate the sequence from K(738) to F(730) according to general procedure II. The peptide was then elongated manually to D(723) according to general procedure I, including coupling of uncommon amino acids, Fmoc-Asp(Opip)-OH and Fmoc-Lys(Mmt)-OH at positions 728 and 724. Cyclisation of the Fmoc protected peptide was performed using general procedure IV followed by the addition of final Fmoc-His-OH residue. The completed peptide, with N_α-terminal Fmoc deprotected, was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **9** as a white powder. HRMS (ESI+) $[M+4H]^{+4}$ calculated for $[C_{105}H_{156}N_{32}O_{29}]$:583.3008 ;observed: m/z 583.3020.

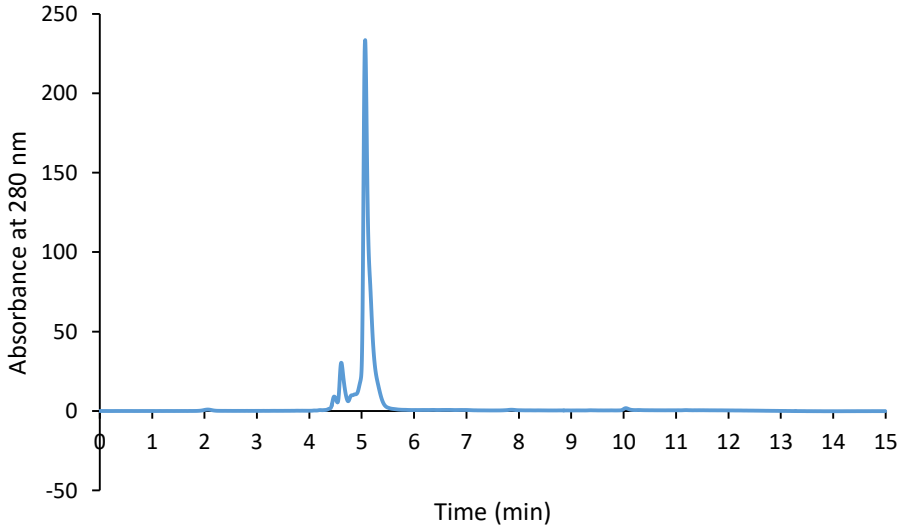


Figure 4-11: RP-HPLC trace of purified **9** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

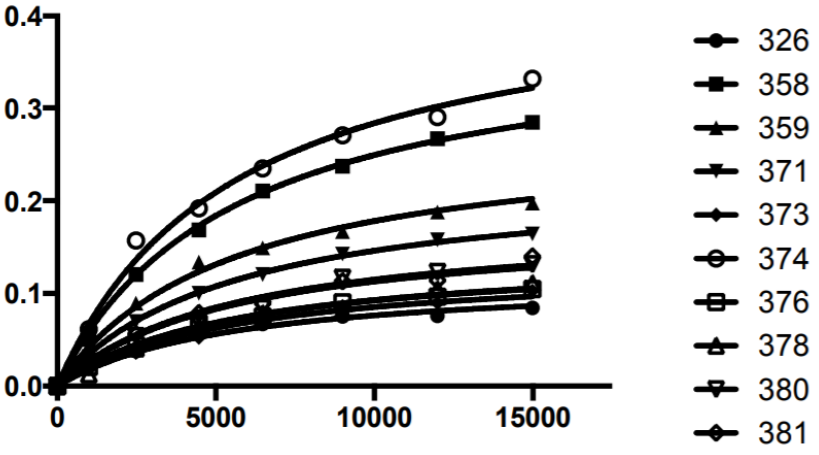
4.2.12 NMR Titrations

^{15}N-^{1}H HSQC titrations of U-^{15}N-labeled talin F3 domain (0.1 mM), with linear peptide β3MP(722-739) and linked peptides **6**-**9** were performed by Dr Kate Wegener according to protocols described in Keeling et.al.[4]

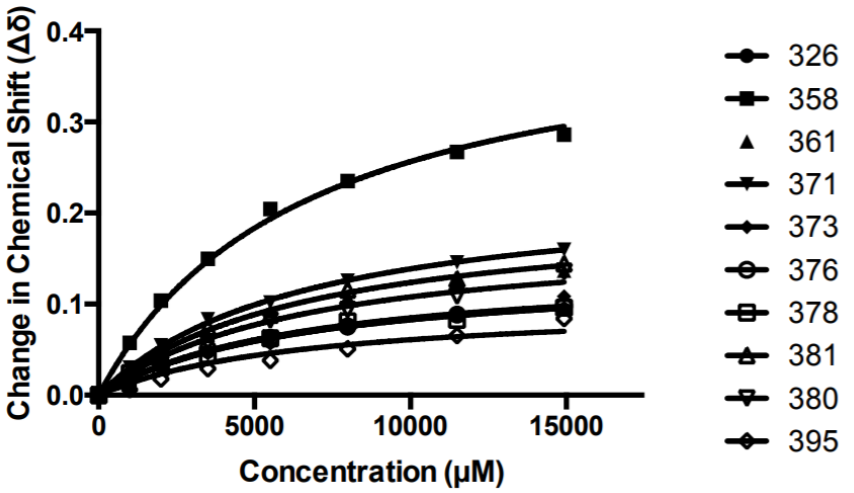
4.2.13 NMR structure determination

Peptide structures of linear peptide β3MP(722-739), and linked peptides **6**-**9** were calculated by Dr Kate Wegener using a version of ARIA 2.3, modified to deal with lactam linked peptides, described in Keeling et.al.[4]

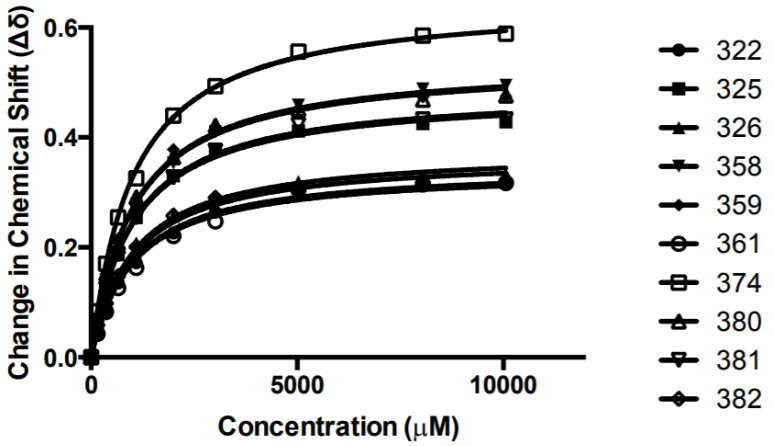
β3MP(722-739)



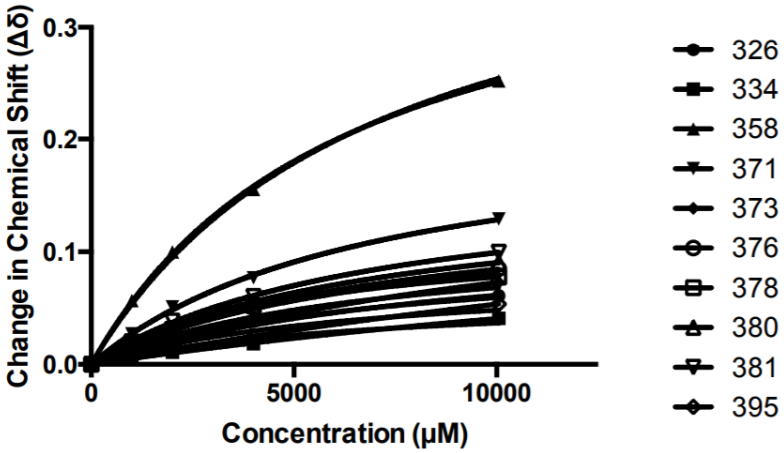
Peptide **6**



Peptide **7**



Peptide **8**



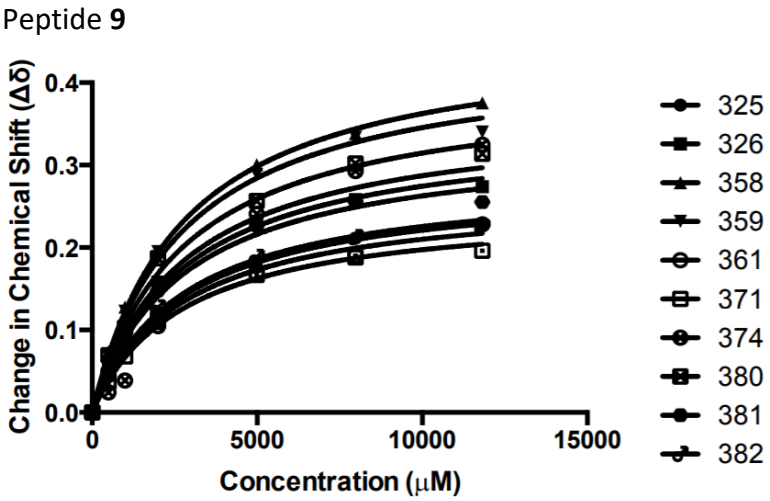
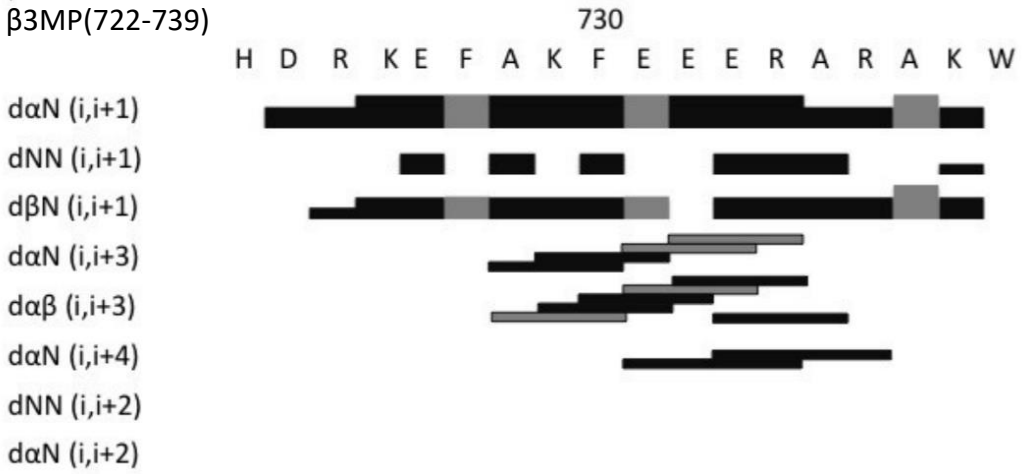
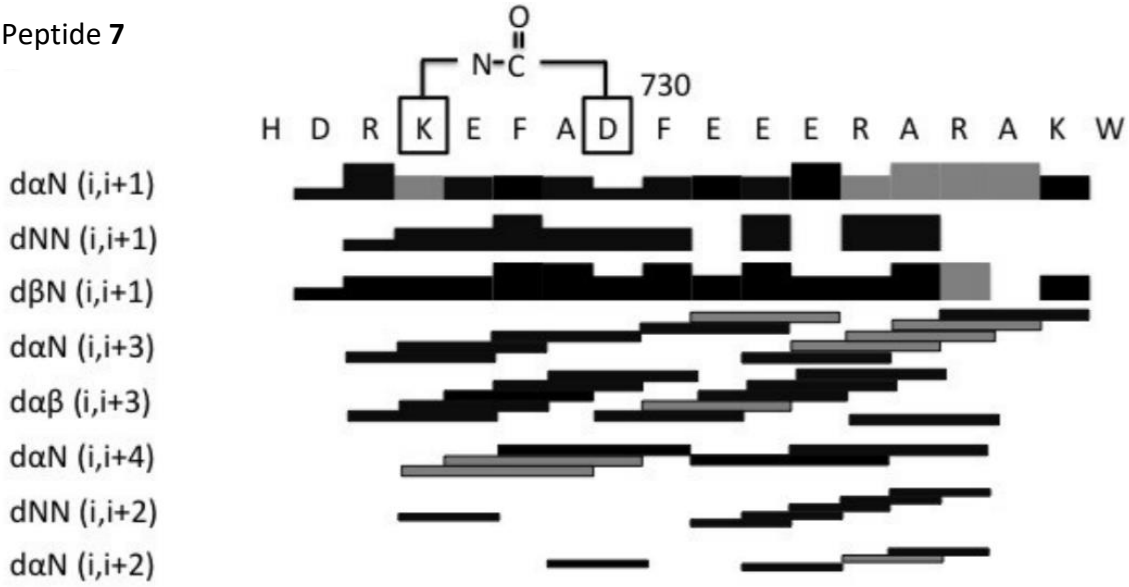


Figure 4-12: Binding curves displaying the change in chemical shift (Δδ) with increasing peptide ligand concentration (μM), for the ten most shifted peaks in each HSQC spectrum. Residue identities are indicated to the right of each chart.



Peptide **6**

Peptide **7**



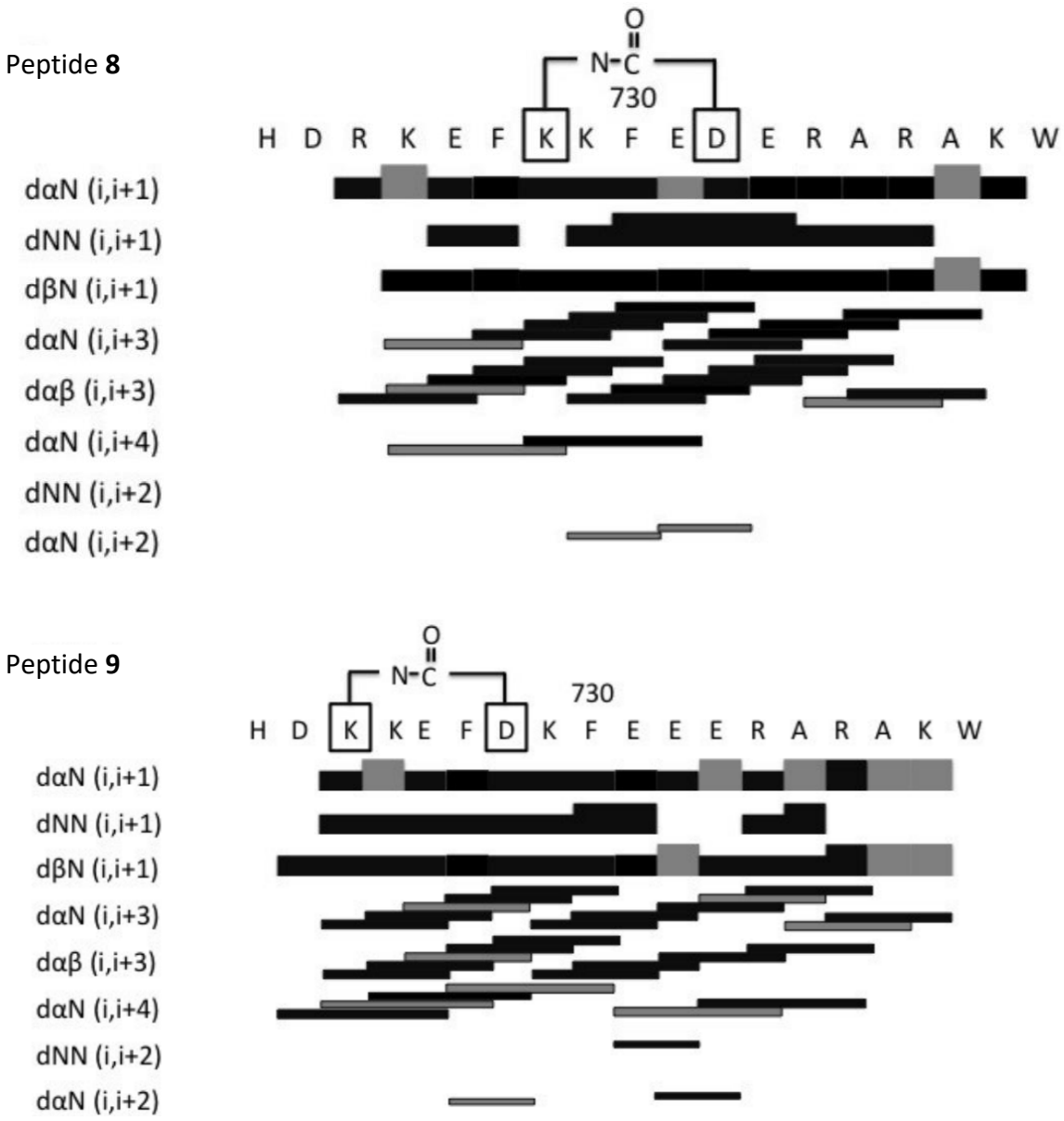


Figure 4-13: NOE connectivity maps. Ambiguous NOEs are indicated in grey

4.2.14 Proteolysis

Proteolysis experiments were carried out in collaboration with Dr Okki Cho, (The University of Adelaide, Australia).

Purified peptide β3MP(722-739), 7,and 8 (110 µM, 50 µl) was combine with of α-chymotrypsin (Sigma Aldrich, 500 nM, 5 µl) and incubated it in Tris HCl (50 mM), pH 7.5, at rt, for the times specified for each peptide. The reaction was quenched by addition of 20% TFA in water (3.3 µl), and the samples of digested peptide were frozen to prevent

further degradation. Samples were defrosted for analysis and aliquots (50 µl) were injected on eclipse XDB-C18 column (4.6 x 150 mm, 5 µm, Agilent System 1260 Infinity) coupled to a mass spectrometry (Bruker HCT, Germany). The samples were eluted with a gradient of 1-100% ACN/H_2O with 0.1% formic acid over 20 minutes. Spectra were analysed using Hystar software (3.1) and the remaining substrate was quantified using extracted ion mass chromatogram. Relative protein concentrations were determined from the area under the curve (Table 4-2, Table 4-3, and Figure 4-14)

Table 4-2: Raw data collected for proteolysis of β3MP(722-739) by chymotrypsin.

Time (minutes)	**% intact peptide remaining**	**% intact peptide remaining**	**% intact peptide remaining**
0	1	1	1
10	0.552	0.729	0.481
20	0.335	0.504	0.406
30	0.112	0.323	0.265
45	0.117	0.146	0.061
60	0.018		0.087
90	0.025		0.018
120	0.033		

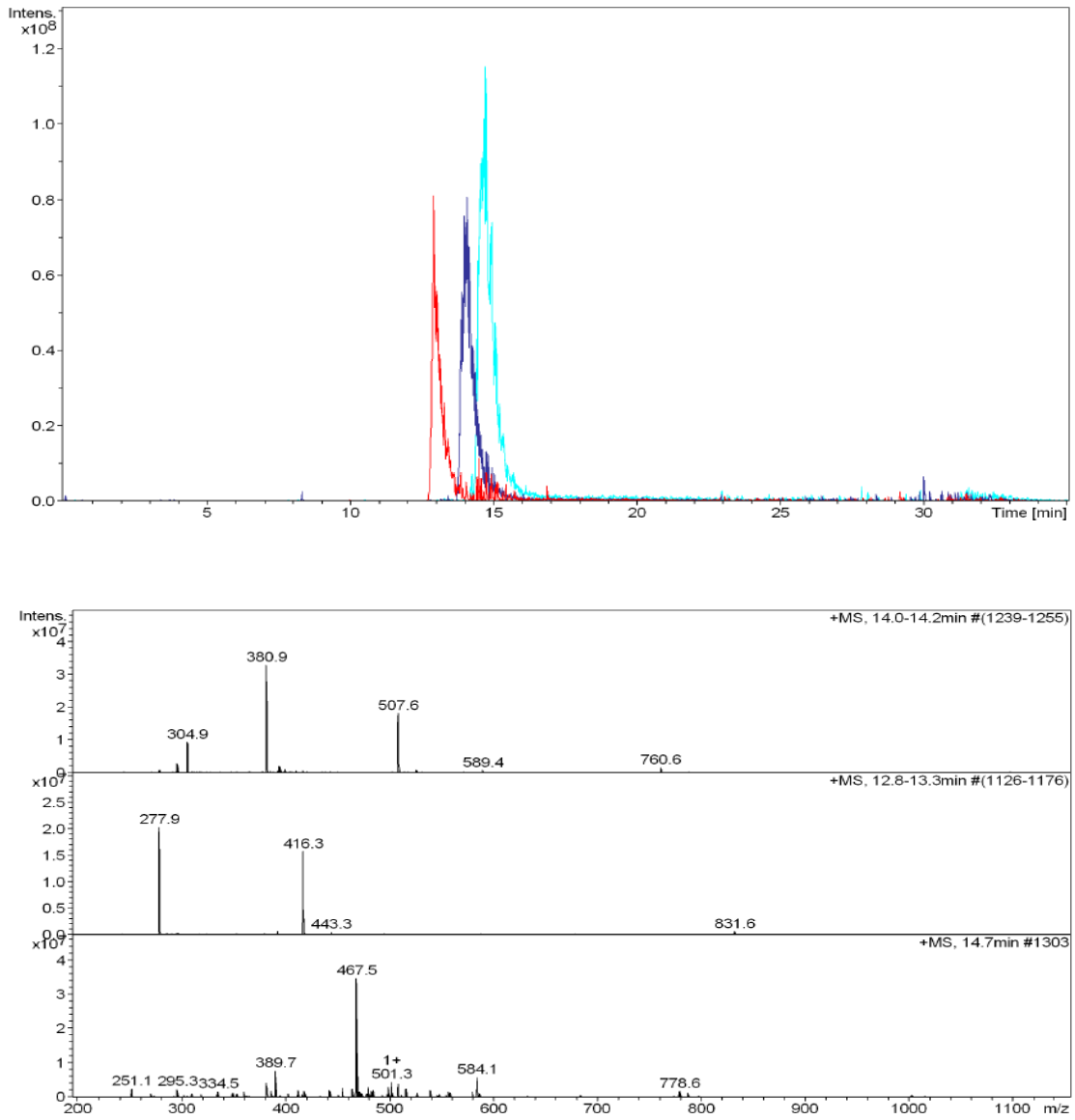


Figure 4-14: Left, the extracted ion chromatogram for β3MP(722-739) (light blue), and hydrolysis products, shown in blue and red. Right, Mass spectrum of β3MP(722-739) 778 amu $[M+4]^{+4}$, 584 amu $[M+5]^{+5}$, 467 amu $[M+6]^{+6}$ (bottom), and hydrolysis products (middle and bottom).

Table 4-3: Raw data collected for proteolysis of peptide **7** by chymotrypsin

Time (minutes)	% intact peptide remaining	% intact peptide remaining	% intact peptide remaining
0	1	1	1
45	1.087	0.974	0.956
60	1.033	0.952	0.899
90	1.063	0.942	0.909
180	1.061	1.016	0.978
240	0.901	0.839	0.9
1440		0.906	0.82

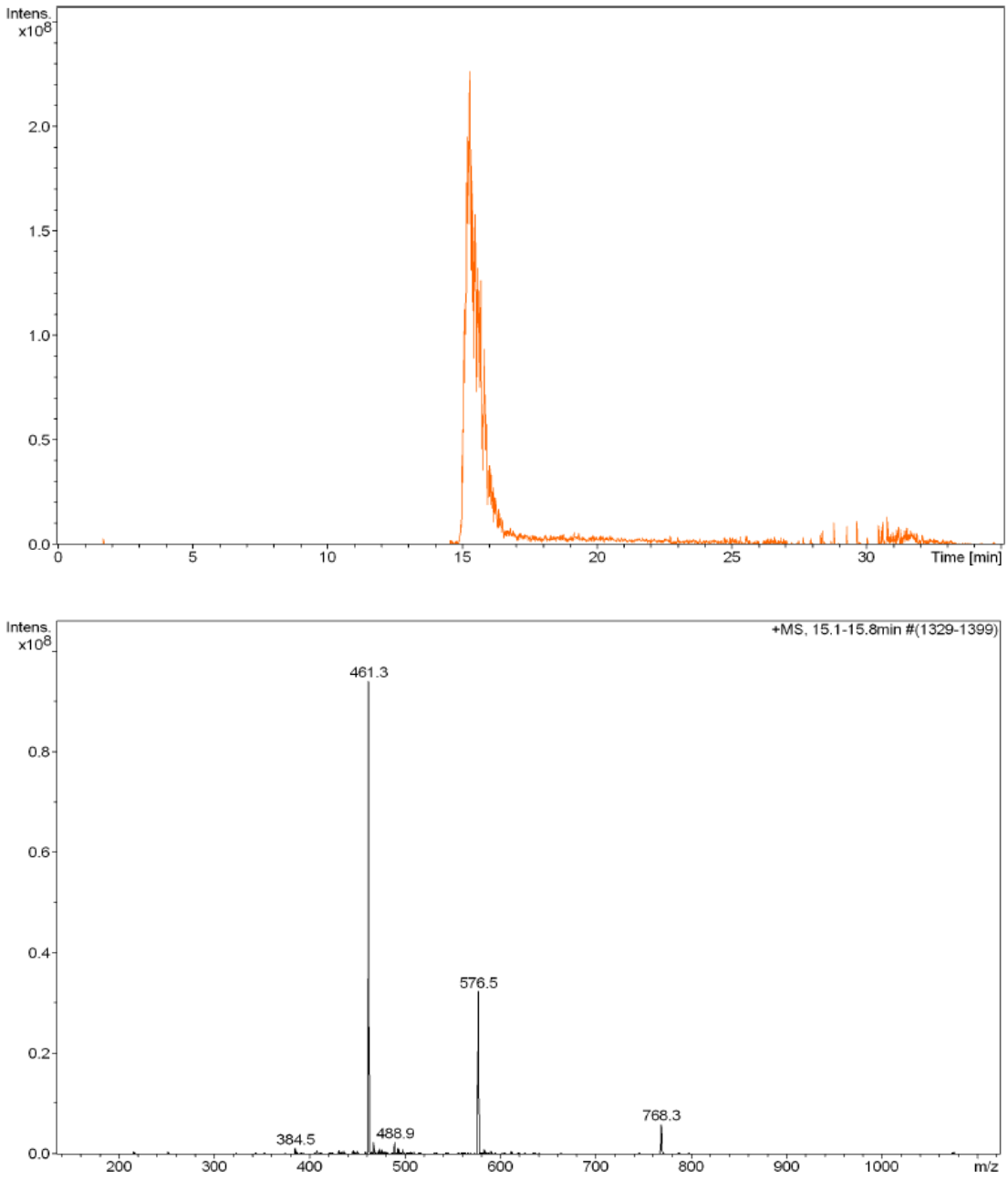


Figure 4-15: Left, the extracted ion chromatogram for peptide **7**. Right, Mass spectrum of peptide **7** 768 amu $[M+4]^{+4}$, 576 amu $[M+5]^{+5}$, 461 amu $[M+6]^{+6}$.

Table 4-4: Raw data collected for proteolysis of peptide **6** by chymotrypsin

Time (minutes)	% intact peptide remaining	% intact peptide remaining	% intact peptide remaining
0	1	1	1
10		0.857	0.766
30		0.858	0.662
45	0.589	0.757	0.597
60	0.447		0.548
90	0.615	0.675	0.51
120	0.367	0.696	0.201
180	0.342	0.57	
240	0.347	0.586	
1080		0.341	
1440		0.336	

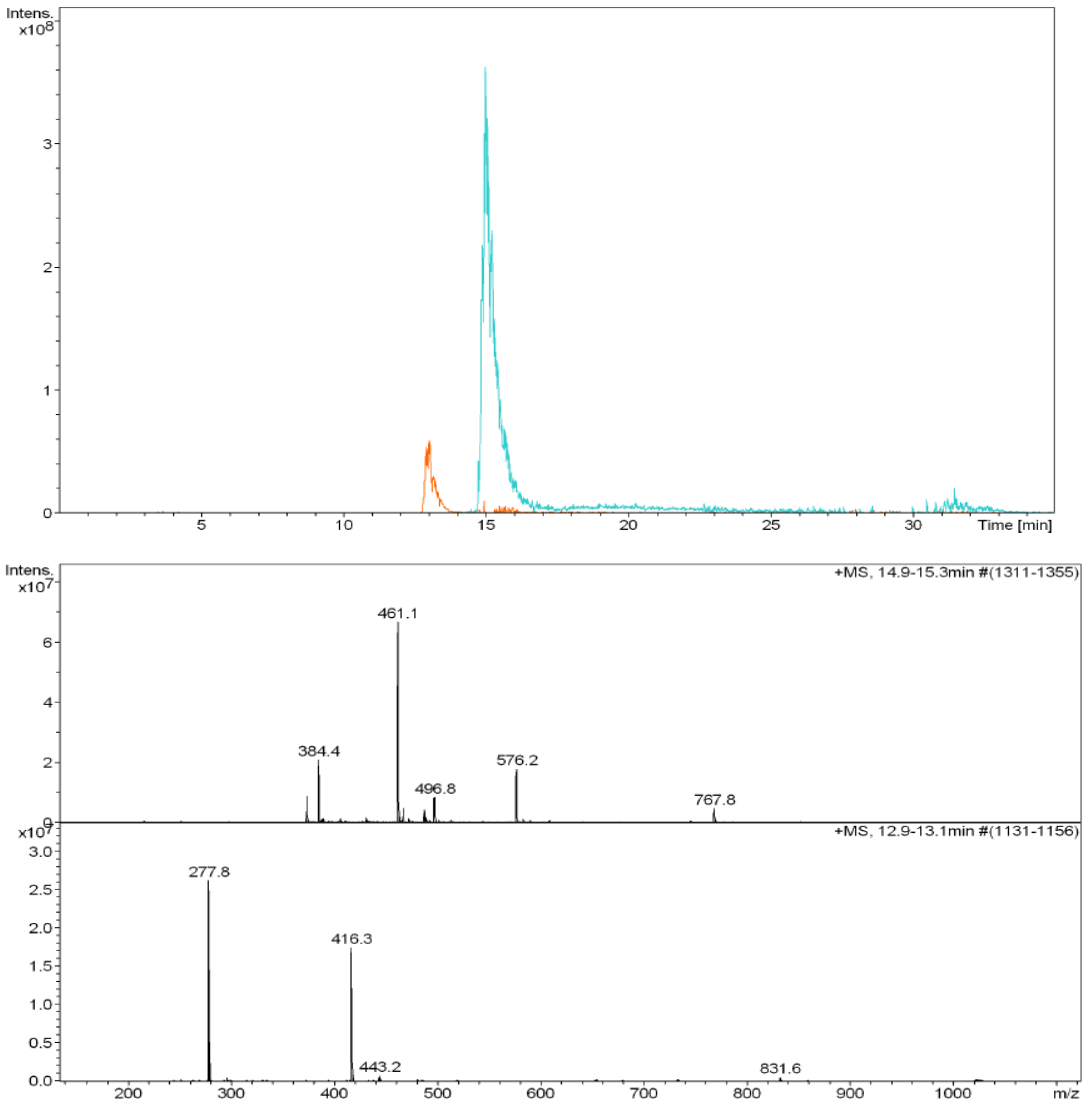


Figure 4-16: Left, the extracted ion chromatogram for peptide **6** (light blue) and hydrolysis product (orange). Right, Mass spectrum of peptide **6** 767 amu $[M+4]^{+4}$, 576 amu $[M+5]^{+5}$, 461 amu $[M+6]^{+6}$(top) and hydrolysis product (bottom).

4.3 Experimental work described in Chapter 3

Peptide β3MP(722-739) (NH_2-HDRKEFAKFEEDRARAKW-$CONH_2$)

β3MP(722-739) was synthesised as described previously (See 4.2.1,)

Peptide **7** (NH_2-HDRK*EFAD*FEEDRARAKW-$CONH_2$)

Peptide **7** was synthesised as described previously (4.2.9)

4.3.1 Peptide **10** (NH_2-HDR[Nle(εN_3)]*EFA[Pra]*FEEDRARAKW-$CONH_2$)

The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin. The initial deprotection and coupling of the *C*-terminal Fmoc-Trp(OtBu)-OH were performed manually according to general procedure I. Automated peptide synthesis was used to elongate the sequence from K(738) to F(730) according to general procedure II. The peptide was then elongated manually from Pra(729) – H(722) according to general procedure I, which included coupling of uncommon amino acids Fmoc-L-Pra-OH and Fmoc-L- Nle(εN_3)-OH in the positions 729 and 727. The completed peptide, with N_α-terminal Fmoc deprotected, was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford the linear peptide precursor of **10** as a white powder. Purified linear peptide precursor was dissolved in water and tert-butanol (2:1, 1 mg/ml), followed by addition of ascorbic acid (4.4 eq) and $CuSO_4$ (4.4 eq). The reaction mixture was stirred for 18 hrs. The cyclised peptide was purified by RP-HPLC. Pure fractions were lyophilised to afford **10** as a white powder. LCMS (ESI+) m/z calculated for $C_{103}H_{149}N_{35}O_{28}$: 2324.1311; observed: m/z 775.7145 ($[M+3H]^{+3}$) (expected: 775.7104).

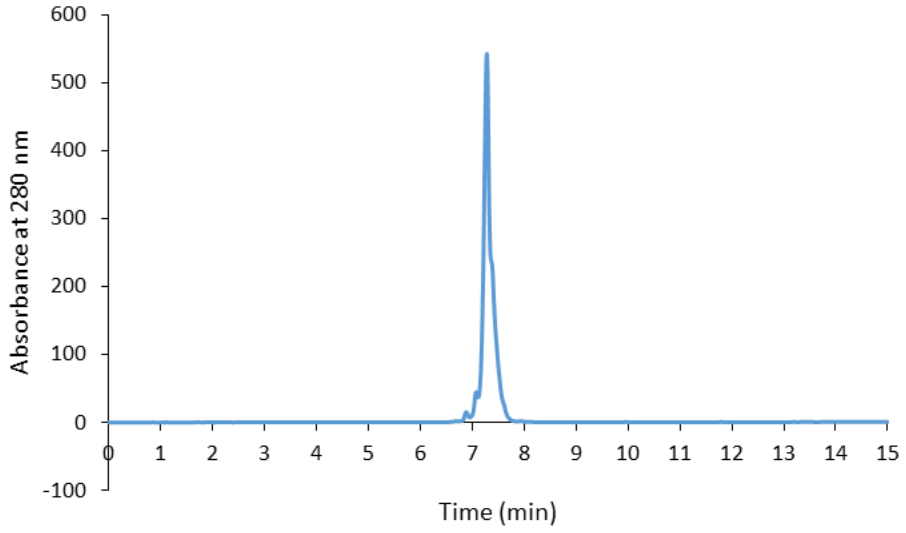


Figure 4-17: RP-HPLC trace of purified **10** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.2 Peptide **11** (NH_2-HDRX*EFAX*FEEDRARAKW-$CONH_2$)

The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin. The initial deprotection and addition of the *C*-terminal Fmoc-Trp(OtBu)-OH to the resin was performed according to general procedure I. Automated peptide synthesis was used to elongate the sequence from K(738) to F(730) according to general procedure II. The peptide was then elongated manually according to general procedure I, including coupling of uncommon amino acids (S)-N-Fmoc- ξ-(4'pentenyl) alanine (X) at positions 729 and 727. Cyclisation of the Fmoc protected peptide was performed following the general procedure V. The N_α-terminal Fmoc group was deprotected as described in general procedure I and the completed peptide was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **11** as a white powder. LCMS (ESI+) m/z calculated for $C_{106}H_{156}N_{32}O_{28}$: 2325.1767 (average isotopes); observed: m/z 776.0670 ($[M+3H]^{+3}$).

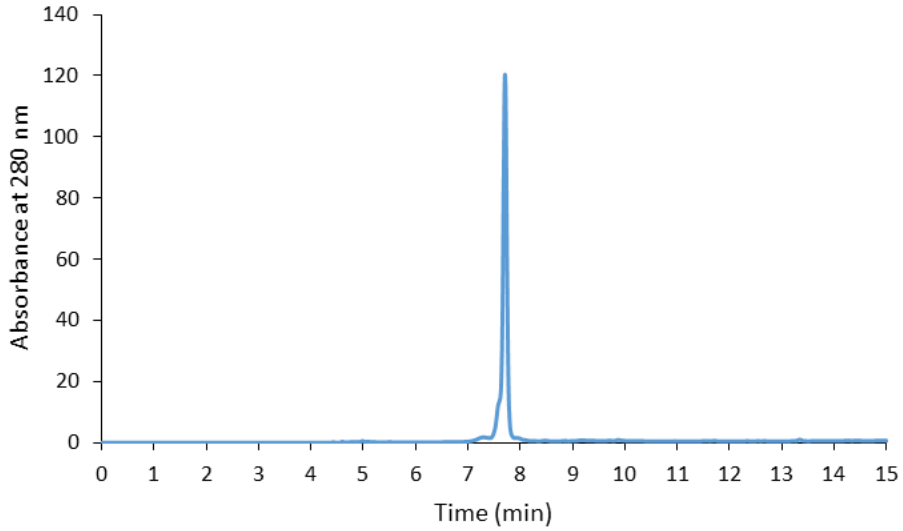


Figure 4-18: RP-HPLC trace of purified **11** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.3 Peptide **12** (NH_2-HDRC*EFAC*FEEDRARAKW-$CONH_2$)

The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin. The initial deprotection and addition of the Fmoc-Trp(OtBu)-OH to the resin was performed

according to general procedure I. Automated peptide synthesis was used to elongate the sequence K(738) to R(724) according to general procedure II. The final residues D(723) and H(722), were incorporated according to general procedure I. The N_{α}-terminal Fmoc group was deprotected as described in general procedure I and the peptide was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford the linear precursor of **12** as a white powder. The linear precursor (1 eq), hexafluorobenzene (25 eq), 2-amino-2-(hydroxymethyl)-1,3-propanediol (Tris base) (13 eq) and LiCl (5 eq) were dissolved in DMF (1 mg/ml) and sonicated at room temperature for 30 minutes, the reaction mixture was then allowed to react for a further 18 hrs. Solvent was removed in vacuo and the residue was dissolved in water with 0.1% TFA. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **12** as a white powder. LCMS (ESI+) m/z calculated for $C_{104}H_{142}F_4N_{32}O_{28}S_2$: 2427.0049; observed: m/z 810.0081 ($[M+3H]^{+3}$).

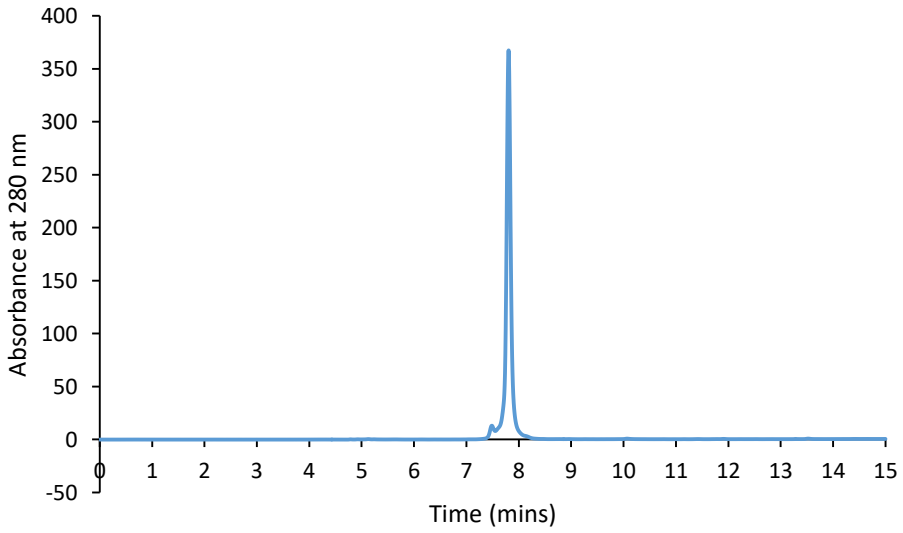


Figure 4-19: RP-HPLC trace of purified **12** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.4 Peptide CF–NYAD-1 (CF-βAla-ITFX*DLLX*YYGP -$CONH_2$)

CF–NYAD-1 peptide was based on the NYAD peptide synthesised by Zhang.et.al[5]. The resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin according to methods described in the general procedure I, including couplings of uncommon amino acid (S)-N-Fmoc- ξ-(4'pentenyl) alanine (X). Cyclisation of the Fmoc protected peptide was performed following the general procedure V. The N_α-terminal Fmoc was then deprotected, and the peptide was labelled as described in the general procedure VI. The peptide was then cleaved from the resin and the crude peptide purified by RP-HPLC. Pure fractions were collected and lyophilised to afford CF–NYAD-1 as a yellow powder. LCMS (ESI+) m/z calculated for $C_{98}H_{121}N_{13}O_{25}$: 1863.8647 (average isotopes); observed: m/z 932.93 ($[M+2H]^{+2}$).

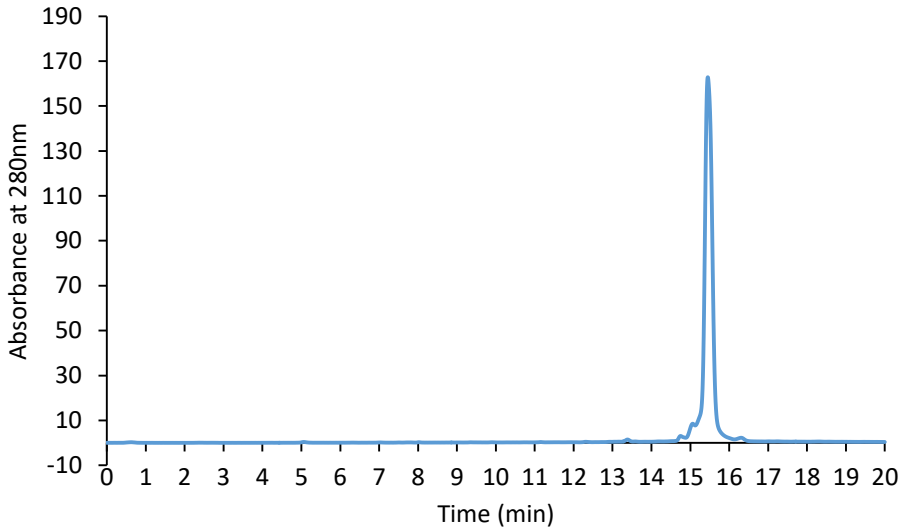


Figure 4-20: RP-HPLC trace of purified CF-NYAD-1 monitoring at a wavelength of 280nm with a gradient of 0-100% ACN/H_2O with 0.1% TFA over 20 minutes.

4.3.5 Peptide Fβ3MP**(722-739)** (CF-βAla-HDRKEFAKFEEDRARAKW-$CONH_2$)

Resin bound peptide was synthesised as described previously for resin bound peptide β3MP(722-739). The peptide was fluorescently labelled as described in the general procedure VI, before the peptide was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford Fβ3MP(722-739) as a yellow powder. LCMS (ESI+) m/z calculated for $C_{128}H_{173}N_{35}O_{35}$: 2761.2906 (average isotopes); observed: m/z 921.4330 ($[M+3H]^{+2}$).

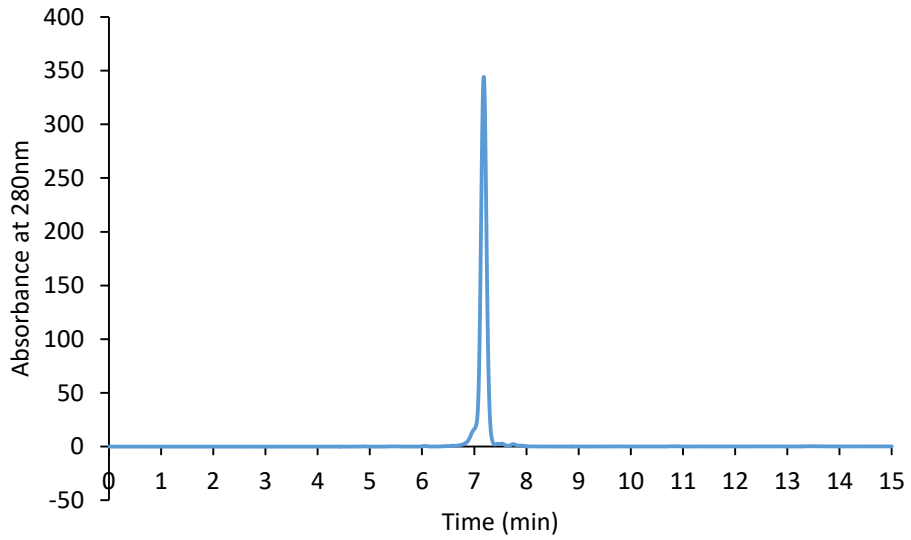


Figure 4-21: RP-HPLC trace of purified Fβ3MP(722-739) monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.6 Peptide **F7** (CF-βAla-HDRK*EFAD*FEEDRARAKW-$CONH_2$)

Resin bound peptide was synthesised as described previously for resin bound peptide **7**. The peptide was labelled as described in the general procedure VI, before the peptide was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **F7** as a yellow powder. LCMS (ESI+) m/z calculated for $C_{126}H_{164}N_{34}O_{36}$: 2761.2906 (average isotopes); observed: m/z 911.0764 ($[M+3H]^{+2}$).

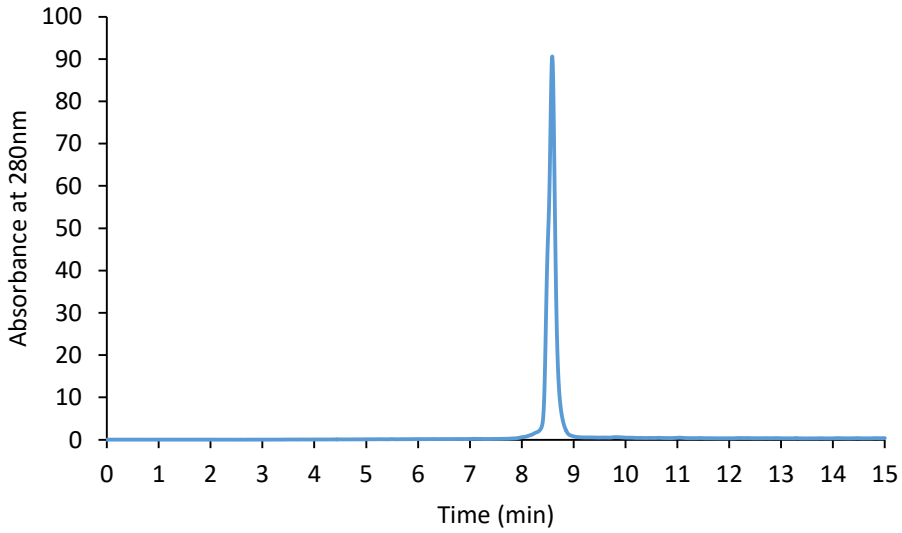


Figure 4-22: RP-HPLC trace of purified **F7** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.7 Peptide **F10** (CF-βAla-HDR[Nle(εN$_3$)]*EFA[Pra]*FEEDRARAKW-CONH$_2$)

Resin bound peptide was synthesised as described previously for resin bound peptide **10**. The peptide was labelled as described in the general procedure VI before the peptide was cleaved from the resin according to general procedure III. Linear precursor of F10 was given as a yellow powder. LCMS (ESI+) m/z calculated for $C_{126}H_{164}N_{34}O_{36}$: 2761.2906 (average isotopes); observed: m/z 911.0764 ($[M+3H]^{+2}$).

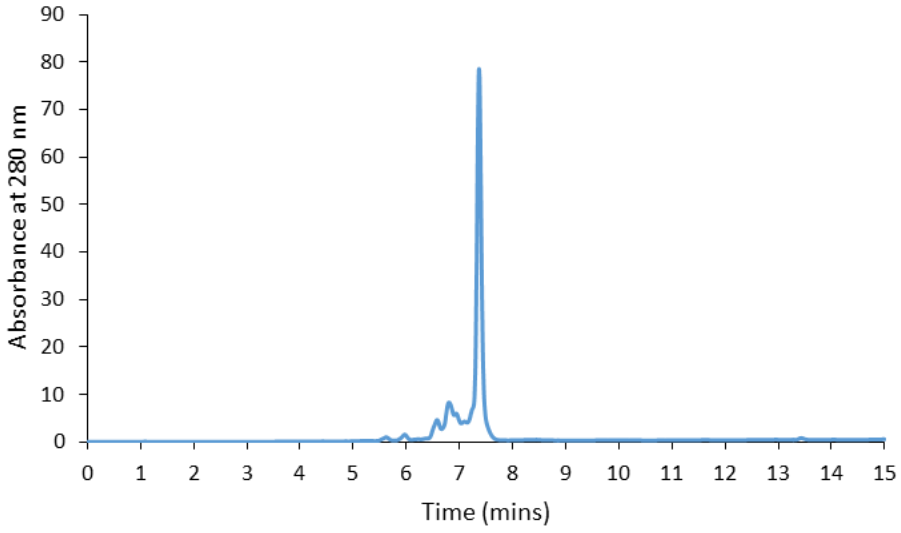


Figure 4-23: RP-HPLC trace of purified **F10** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.8 Peptide **F11** (CF-βAla-HDRX*EFAX*FEEDRARAKW-$CONH_2$)

Resin bound peptide was synthesised as described previously for resin bound peptide **11**. The peptide was labelled as described in the general procedure VI, before the peptide was cleaved from the resin according to general procedure III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **F11** as a yellow powder. LCMS (ESI+) m/z calculated for $C_{130}H_{171}N_{33}O_{35}$: 2754.2615 (average isotopes); observed: m/z 689.5734 ($[M+4H]^{+2}$).

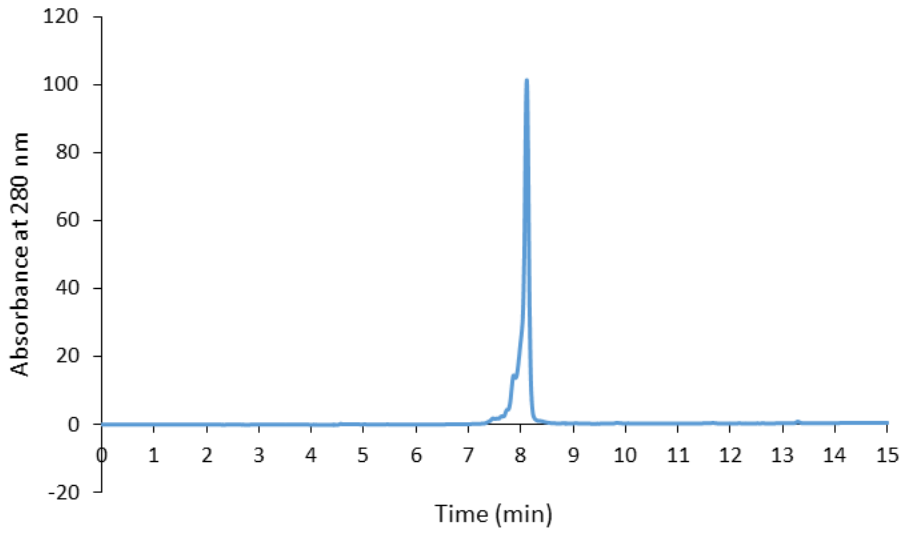


Figure 4-24: RP-HPLC trace of purified **F11** monitoring at a wavelength of 280nm with a gradient of 10-70% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.9 Peptide **F12** (CF-βAla-HDRC*EFAC*FEEDRARAKW-$CONH_2$)

Linear peptide precursor of **F12** was synthesised as described previously for resin bound linear peptide precursor of **12**. The peptide was labelled as described in the general procedure VI, before the peptide was cleaved from the resin according to general procedure III. Linear peptide precursor of F12 was given as a yellow powder. Linear peptide precursor of **F12**, hexafluorobenzene, Tris base and LiCl were dissolved in DMF and sonicated at room temperature for 30 minutes, then allowed to react for a further 18 hrs. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford F12 as a yellow powder. $C_{128}H_{157}N_{33}O_{35}S_2$: 2856.0897 (average isotopes); observed: m/z 952.0252 ($[M+3H]^{+2}$).

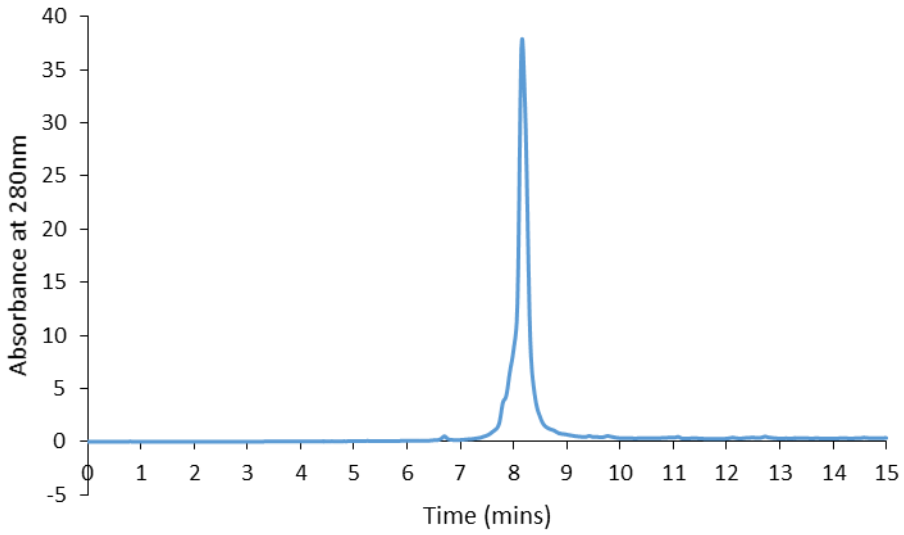


Figure 4-25: RP-HPLC trace of purified **F12** monitoring at a wavelength of 280nm with a gradient of 0-100% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.10 Peptide **F13** (CF-βAla-HDRX*EFAX*FEEERA-$CONH_2$)

Resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin, from A(735) to D(723), according to methods described in the general procedure I; including couplings of uncommon amino acid, (S)-N-Fmoc- ξ-(4'pentenyl) alanine (X). Cyclisation of the Fmoc protected peptide was performed according to general procedure V. H(722) was incorporated according to general method I. The peptide with a free *N*-terminal was labelled following the general procedure VI. The peptide was then cleaved from the resin according to methods described in general method III. The crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **F13** as a yellow powder. $C_{104}H_{132}N_{24}O_{31}$: 2212.9490 (average isotopes); observed: m/z 738.6584 ($[M+3H]^{+2}$).

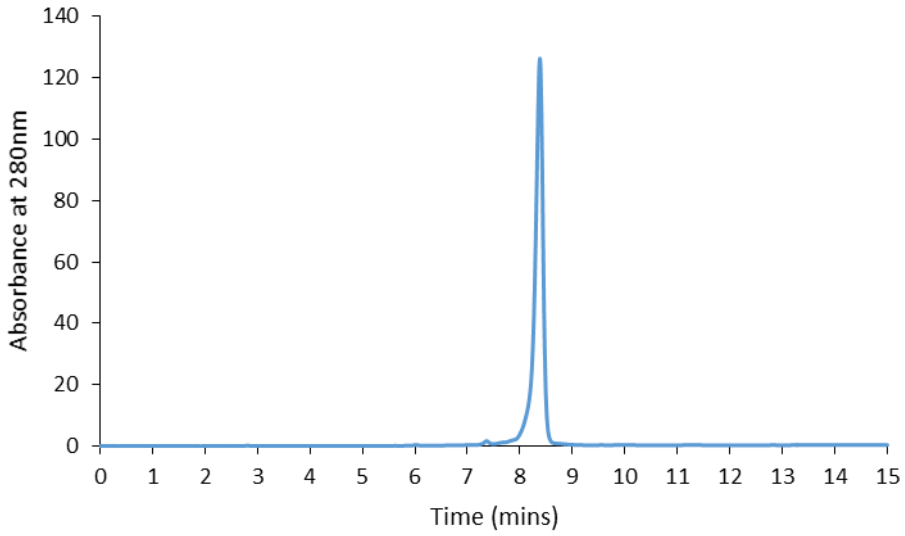


Figure 4-26: RP-HPLC trace of purified **F13** monitoring at a wavelength of 280nm with a gradient of 0-100% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.11 Peptide **F14** (CF-βAla-HNRX*EFAX*FQQERA-CONH$_2$)

Resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin, from A(735) to N(723), according to methods described in the general procedure I; including couplings of uncommon amino acid (S)-N-Fmoc- ξ-(4'pentenyl) alanine (X). Cyclisation of the Fmoc protected peptide was performed using general procedure V. H(722) was incorporated according to general method I. The peptide with a free *N*-terminal was labelled following the general procedure VI. The peptide was then cleaved rom the resin following general procedure III, and the crude peptide was purified by RP-HPLC. Pure fractions were collected and lyophilised to afford **F14** as a yellow powder. LCMS (ESI+) m/z calculated for $C_{104}H_{135}N_{27}O_{28}$: 2209.9970 (average isotopes); observed: m/z 737.67 ($[M+3H]^{+3}$) 554.50 ($[M+4H]^{+4}$).

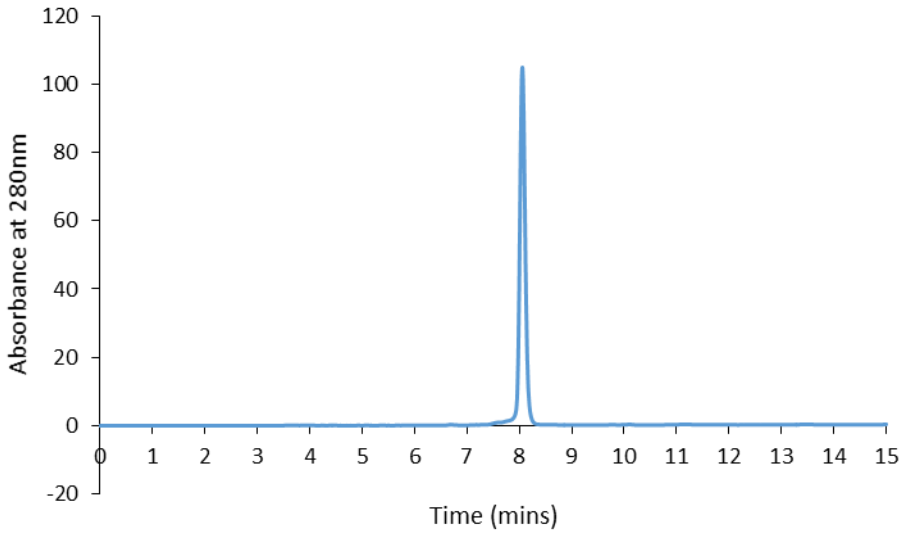


Figure 4-27: RP-HPLC trace of purified **F14** monitoring at a wavelength of 280nm with a gradient of 0-100% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.12 Peptide **F15** (CF-βAla-HRRX*EFAX*FRRERA-$CONH_2$)

Resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin, from A(735) to N(723), according to methods described in the general procedure I; including couplings of uncommon amino acid (S)-N-Fmoc- ξ-(4'pentenyl) alanine (X). Cyclisation of the Fmoc protected peptide was performed using general procedure V. H(722) was incorporated according to general method I. The peptide with a free *N*-terminal was labelled following general procedure VI. The peptide was then cleaved from the resin and the crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **F15** as a yellow powder. LCMS (ESI+) m/z calculated for $C_{121}H_{169}N_{39}O_{31}$: 2678.3003 (average isotopes); observed: m/z 893.77 ($[M+3H]^{+3}$) 671.58 ($[M+4H]^{+4}$).

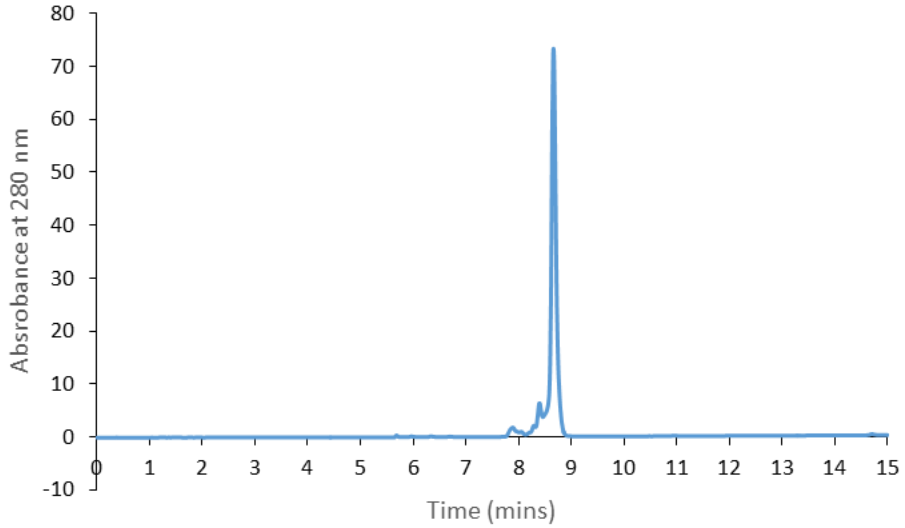


Figure 4-28: RP-HPLC trace of purified **F15** monitoring at a wavelength of 280nm with a gradient of 0-100% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.13 Peptide **F16** (CF-βAla-RRRHNRX*EFAX*FQQERA-$CONH_2$)

Resin bound peptide was synthesised on a 0.1 mmol scale on rink amide resin, from A(735) to N(723), according to methods described in the general procedure I; including couplings of uncommon amino acid (S)-N-Fmoc- ξ-(4'pentenyl) alanine (X). Cyclisation of the Fmoc protected peptide was performed using general procedure V. H(722) and three arginine residues were incorporated at the *N*-terminal according to general method I. The peptide with a free *N*-terminal was labelled following general procedure VI. The peptide was then cleaved from the resin and the crude peptide was purified by RP-HPLC, pure fractions were collected and lyophilised to afford **F16** as a yellow powder. LCMS (ESI+) m/z calculated for $C_{121}H_{169}N_{39}O_{31}$: 2678.3003 (average isotopes); observed: m/z 893.77 ($[M+3H]^{+3}$) 671.58 ($[M+4H]^{+4}$).

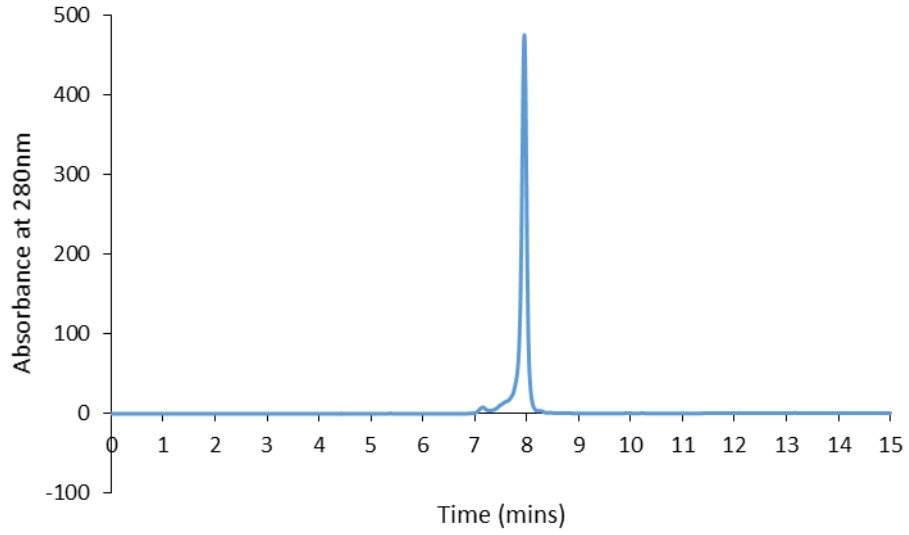


Figure 4-29: RP-HPLC trace of purified **F16** monitoring at a wavelength of 280nm with a gradient of 0-100% ACN/H_2O with 0.1% TFA over 15 minutes.

4.3.14 Confocal Microscopy

Proteolysis experiments were carried out in collaboration with Dr Briony Gliddon, (Molecular Signalling Laboratory Centre for Cancer Biology, SA Pathology, Adelaide, Australia).

Confocal microscopy was used to visualise the cellular uptake of peptides Fβ3MP(722-739), **F7**, **F10** and **F11** NIH3T3 cells ($1.5x10^5$) were seeded in a poly-L-lysine coated 12-well plate overnight. A confluent monolayer of cells (70-80%) was used for experiment. The growth medium was removed, and the cells were washed with PBS followed by serum free medium. Serum free medium was then added to the cells for 3 hrs. The CF labelled peptides Cellular uptake of peptides Fβ3MP(722-739), **F7**, **F10** and **F11** (5 μM) were treated and incubated for 2 hrs at 37 °C. Cells were then washed with PBS (3x) and fixed by treatment with 4% paraformaldehyde. Fixed cells were washed with PBS (x2). The cells were scanned and imaged under a Zeiss LSM510 laser scanning confocal microscope (Zeiss).

4.3.15 Fluorescence-activated cell sorting (FACS)

FACS analysis used to compare cellular uptake of peptides were carried out in collaboration with Dr Okki Cho, (The University of Adelaide, Australia).

Cellular uptake of the peptides Fβ3MP(722-739), **F7**, **F10**, **F11** and **F13-F16**, were investigated in NIH3T3 cells (mouse embryonic fibroblast, kindly provided by Jansen laboratory, The University of Adelaide). The cells were cultured in minimum essential medium supplemented with foetal bovine serum (10% v/v), L-glutamine (2 mM), sodium pyruvate (1 mM), penicillin (100 units/ml), and streptomycin (100 μg/ml). The cells ($4x10^5$) were seeded in a 6-well plate overnight, and a confluent monolayer of cells (70-80%) was used for the experiment. The growth medium was removed, and the cells were washed twice with PBS and serum free medium, then a serum free medium was added for 3 hrs. Stock solutions of CF (carboxyfluorescein) labelled peptides (250 μM) in DMSO and water were diluted to give a peptide concentration of 5 μM, and final

DMSO concentration of 0.5% in medium. Cells were treated with the peptide and incubated for 2 hrs at 37 C. The cells were then washed with ice cold PBS and treated 1 mg/ml trypsin for 10 minutes to remove any peptide bounded to the surface of the cells. After washing with ice cold PBS (x2), the cells were collected by centrifugation (1,500xg for 4 minutes). Cellular penetration of the peptides was measured immediately using FACS LSRII flow cytometer (BD Bioscience). The fluorescence of viable cells were gated and displayed 10,000 events. The data were analysed by using FlowJo Software.

4.4 References

1. Shepherd, N. E.; Hoang, H. N.; Abbenante, G.; Fairlie, D. P., Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *Journal of the American Chemical Society* **2005,** *127* (9), 2974-2983.
2. Sheppard, R., The fluorenylmethoxycarbonyl group in solid phase synthesis. *Journal of Peptide Science* **2003,** *9* (9), 545-552.
3. Antoni, G.; Presentini, R.; Neri, P., A simple method for the estimation of amino groups on insoluble matrix beads. *Analytical Biochemistry* **1983,** *129* (1), 60-63.
4. Keeling, K. L.; Cho, O.; Scanlon, D. B.; Booker, G. W.; Abell, A. D.; Wegener, K. L., The key position: influence of staple location on constrained peptide conformation and binding. *Org Biomol Chem* **2016,** *14* (41), 9731-9735.
5. Zhang, H.; Zhao, Q.; Bhattacharya, S.; Waheed, A. A.; Tong, X.; Hong, A.; Heck, S.; Curreli, F.; Goger, M.; Cowburn, D.; Freed, E. O.; Debnath, A. K., A Cell-penetrating Helical Peptide as a Potential HIV-1 Inhibitor. *Journal of Molecular Biology* **2008,** *378* (3), 565-580.